Evaluating achievement on mathematics and science problems: The role of global and local processing

Suzanne Catherine Morris

Department of Psychology and Human Development

University College London

Institute of Education

Thesis submitted for the degree of Doctor of Philosophy

The work presented in this thesis was supported by a Bloomsbury studentship.

I confirm that the work presented in this thesis is my own.

Suzanne Catherine Morris

[redacted]

Abstract

This thesis sought to clarify relationships between whole-part constructs; where responses are thought to reflect a focus on the whole stimulus or context, or on individual elements (the parts). Children aged 5 to 10 years completed a number of tasks allowing developmental changes to be measured on a cross-sectional and longitudinal basis. Global and local processing tasks (Navon tasks: free choice, selective attention, and divided attention) revealed a change in response patterns between the ages of 6 and 7 years, the precise nature of which varied depending on the attentional demands of the task. Field independence tasks (Children's Embedded Figures Test (CEFT) and Design Organisation Test (DOT)) revealed that children became more field independent with increasing age. A parental questionnaire measuring systemizing (the analysis or construction of a rule-based system) was administered at a single timepoint and revealed no cross-sectional age-related changes.

Behavioural tasks and eye-tracking technology were employed to understand possible mechanisms underlying field independence performance. Visuospatial IQ and working memory explained variation on both field independence tasks. Higher accuracy on the CEFT reflected fewer and shorter fixations on distractor areas as well as longer and more fixations on target areas. Better *response* inhibitory control related to higher disembedding accuracy, while better *semantic* inhibitory control related to fewer and shorter fixations on distractor areas. Together, these explain how domain-general factors contribute to performance on the CEFT.

The second part of this thesis examined associations between whole-part constructs and mathematics and science achievement. Global and local processing and systemizing revealed few significant associations with the academic scores. Better field independence was associated with higher scores on both mathematics and science. After controlling for age and domain-general factors, the field independence tasks explained additional variance on specific mathematics and science tests, which likely reflect common processes involved in the tasks.

Impact Statement

The impact of this thesis falls into two broad categories – the contribution to the field of whole-part processing, and implications of individual differences in whole-part processing for mathematics and science achievement.

The work presented in this thesis is important for clarifying and developing our understanding of whole-part constructs and how they relate to each other. Much research to date has involved adult participants or comparisons between neurotypical and atypical populations. There is little evidence relating to whole-part processing in neurotypical children, especially where individuals complete a number of tasks. Furthermore, this is the first global and local processing study which has a longitudinal element. A strength of this research is that rather than focussing purely on describing patterns of association, this thesis endeavours to understand domain-general factors which explain variance on the different tasks. The behavioural tasks are supported by an eye-tracking study, which allows for the examination of how patterns of attention across a stimulus relate to performance. Field independence has not previously been measured in children using eye-tracking technology.

The second area of impact relates to examining whole-part correlates with mathematics and science. This is the first time that global and local processing has been related to mathematics and science achievement. There are previous studies identifying associations between field independence and academic subjects, however these rarely measure overlapping variance with domain-general tasks, and therefore can only describe rather than explain these relationships. Work presented in this thesis has identified the unique contributions made by two field independence tasks to different mathematics and science tasks after covarying domain-general factors.

There were a number of novel tasks designed for this thesis. The visuospatial working memory task (an example of a reverse spatial span) was designed to be engaging and accessible for children aged 5- to 10-years-old. This task has already been used in other studies, including the UnLocke project (http://unlocke.org), a study examining attention and creativity (Massonnié, Rogers, Mareschal, & Kirkham, 2019), and a study examining mental imagery (Bates & Farran, 2019). There was no appropriate general science task for use with children aged 5 to 10 years, so a novel task was designed based on the UK national curriculum. In order to measure associations with different subsets of mathematics (Number, Word, Shape, and Graph) in Year 5 children, a novel mathematics task was designed based on the layout of national tests, which ensured that the design would be familiar to the children. Finally, a short, multiple choice misconceptions task was designed to measure responses to common counterintuitive concepts in children aged 5 to 10 years, which has been adapted for use by the UnLocke team. These novel tasks have expanded the available tests suitable for use with children of primary school age.

The work presented in **Chapter 5** was published in the following paper:

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, *13*(4), 268–278. https://doi.org/10.1111/mbe.12214

Acknowledgements

I would like to thank my two wonderful supervisors, Emily Farran and Iroise Dumontheil, for their unwavering support throughout my PhD years, and during my writing-up year. I have learnt an enormous amount from you both, and I truly value all your input as well as your patience and encouragement. I could not have wished for better supervisors! Thank you also to my fellow students in lab groups and the CEN. I have enjoyed widening my understanding through hearing about your research, as well as having the opportunity to share my research and hear your feedback. I would particularly like to thank Annie Brookman-Byrne for all your encouragement and for answering all my questions with undiminishing enthusiasm and patience.

I would like to extend my thanks to everyone who helped me to collect data for my research. Thank you to all the staff, parents, and children at all the schools I worked with in Crowborough and Tunbridge Wells – without your flexibility and enthusiasm, I would not have been able to carry out my research. Thank you also to Jackie Phelps for giving up your time to help with my data collection. Your dedication and commitment were invaluable, and enabled me to collect a large dataset from 135 children in about 3 months – your help is very much appreciated. I also owe an enormous thank you to Tim Smith for lending me an eye tracker for one of my studies, and to his team who offered me training and support.

I would like to thank all my friends, work colleagues, and my family. Thanks to my most wonderful friend Julie Peckham for the coffee breaks, wine, and chats, and to my colleagues at Kent Fire and Rescue Service for all your support and encouragement. I would like to thank my parents, Jean and Colin, for all your help in enabling me to balance study, work, and family responsibilities. Finally, my thanks and love go to my family – husband Ade, and children Nathan and Charlotte. Thank you for your patience and understanding throughout my PhD, and for being my most favourite people. I am so proud of you, as I know you are of me; your love, laughter, and support has given me the belief, motivation and determination to keep going. You are, and always will be, my whole world.

Table of Contents

Abbreviations

ANOVA	analysis of variance
CEFT	Children's Embedded Figures Test
D-EFT	2D and 3D Embedded Figures Test
DOT	Design Organisation Test
DV	dependent variable
EEG	electroencephalogram
EF	executive functions
EFT	Embedded Figures Test
EQ	Empathizing Quotient
EQ-C	Empathizing Quotient for children
FI	field independence
GEFT	Group Embedded Figures Test
IC	inhibitory control
IQ	Intelligence Quotient
IV	independent variable
L-EFT	Leuven Embedded Figures Test
LISAS	linear integrated speed-accuracy score
LMM	linear mixed model
M-EFT	Meaningful Embedded Figures Test
M	mean
RT	response time
SD	standard deviation
SE	standard error
SEND	special educational needs and disabilities
SQ	Systemizing Quotient
SQ-C	Systemizing Quotient for children
STEM	science, technology, engineering, mathematics
T1	time 1
T2	time 2
WM	working memory
WIAT	Wechsler Individual Achievement Test

List of figures

List of tables

Chapter 1: Introduction to the thesis

1.1 Introduction

This thesis explores the associations between global and local processing, and science and mathematics achievement in neurotypical children aged 5-10 years. On the surface, the commonalities between visual perception in general, and specifically global and local processing, and mathematics and science are not immediately obvious. This is reflected by the lack of current research which investigates these specific relations. However, as discussed below, there are several disparate strands of research which, taken together, compellingly support the existence of such associations. Thus, the aim of this thesis is to explore and clarify these associations, and in doing so, make a novel contribution to our understanding of global and local processing and its relevance to wider fields of study.

1.2 Visual perception

The human perceptual system continually encounters complex stimuli in often quickly changing environments, yet it is able to make sense of this multi-layered incoming sensory information in surprisingly consistent ways (Milne & Szczerbinski, 2009; Scherf, Behrmann, & Luna, 2009). Visual perception, the organisation and interpretation of input from the environment, has interested psychologists and psychophysicists for decades, particularly in relation to how incoming information is initially grouped into objects and backgrounds, outlines and details, and foreground and background (Vetter & Newen, 2014; Wagemans et al., 2012a). For example, a single object may be derived from a multitude of colours and shapes, or may even be partially hidden by other features of the scene, and yet must be grouped and segregated appropriately to be meaningfully understood (Brooks, 2015). This is achieved by perceptual processes which structure and interpret the raw light data received at the retina (Kimchi, Hadad, Behrmann, & Palmer, 2005; Milne & Szczerbinski, 2009). This ability to make sense of the environment from raw sensory data is vital for tasks such as navigation, remembering locations, object recognition and manipulation (Tzuriel & Egozi, 2010), as well as social interactions (Van der Hallen, Evers, Brewaeys, Van den Noortgate, & Wagemans, 2015).

One important role of the visual system is to interpret the whole stimulus as well as detailed elements or parts. There are a number of different constructs relating to the interplay between the perception of the whole and the perception of the elemental parts, collectively described as ‘whole-part’ constructs. These are termed ‘global and local processing', which refers to the processing of the global hierarchical level (whole) and its constituent local elements (parts) (e.g. Kimchi, 2015; Navon, 1977; Wang, Mottron, Peng, Berthiaume, & Dawson, 2007); 'field independence', which describes the ability to isolate a target shape (part) embedded within a complex background (whole) (e.g. Goodenough & Witkin, 1977; Van der Hallen, Chamberlain, De-Wit, & Wagemans, 2018; Zhang, 2004); and 'Gestalt processing', which describes the automatic grouping (whole) of individual features (parts) (e.g. Bölte, Holtmann, Poustka, Scheurich, & Schmidt, 2007; Han & Humphreys, 1999; Wagemans et al., 2012a). A further construct which relies on an understanding of parts, although not in the visual domain, is 'systemizing', which describes a tendency to create and understand rule-based systems (whole) using an understanding of details (parts) to group, characterise, and understand information (e.g. Baron-Cohen, Ashwin, Ashwin, Tavassoli, & Chakrabarti, 2009; Escovar, Rosenberg-Lee, Uddin, & Menon, 2016; Focquaert, Steven, Wolford, Colden, & Gazzaniga, 2007). Despite the fact that they share a common focus on whole-part interactions, and appear to describe somewhat overlapping processes, there are also some important distinctions. For example, field independence describes the ability to separate a target from its context, while Gestalt processing describes the principles which result in a cohesive perception of a stimulus. Therefore, they should not be used interchangeably, and will be treated as separate entities throughout this thesis.

1.2.1 Visual attention

The sensory system receives a huge amount of information, and if it were to interpret and use all the available information, it could quickly become overloaded. Therefore, the incoming information is filtered by attentional processes which enable goal-related relevant elements, and not irrelevant inputs, to reach short-term memory and visual awareness (Amso & Scerif, 2015; Itti & Koch, 2001; Kimchi, Yeshurun, Spehar, & Pirkner, 2016; May, Rinehart, Wilding, & Cornish, 2013; Pinto, van der Leij, Sligte, Lamme, & Scholte, 2013). It is also possible to make judgements about objects outside the focus of attention, but the information is far less detailed and less accurate (Itti & Koch, 2001). It is thought that there are three subtypes of attentional processes – sustained attention (also referred to as 'alerting') which is when attention is maintained in readiness for a response; selective attention (also referred to as 'orienting') which is when attention is focussed on a particular part of the stimulus thought to be related to the task; and executive attention or divided attention, which involves the inhibition of distractors and resolution of conflict (Peng & Miller, 2016; Petersen & Posner, 1990; Steele, Karmiloff-Smith, Cornish, & Scerif, 2012). These are likely to have unique, but overlapping, underlying neural systems (Steele et al., 2012).

The relationship between visual attention and perception is important for optimising the encoding of stimuli in short-term memory (Pinto et al., 2013; Shimi, Nobre, Astle, & Scerif, 2014). Attention can be shifted towards, or away from, a location (known as spatial attention), or a feature (known as feature-based or object-based attention) (Noudoost, Chang, Steinmetz, & Moore, 2010), using a combination of bottom-up and top-down processes. Bottom-up attentional processes are stimulus-driven, quick and involuntary (Katsuki & Constantinidis, 2014; Pinto et al., 2013). Stimulus features, such as a highlighted word or a flashing light, automatically direct a viewer's attention towards more salient parts of a visual input, even if they are irrelevant to the task (Buschman & Miller, 2007). Top-down modulation is a slower, more active process, with attention being directed depending on object expectation, prior knowledge, and behavioural goals (Corbetta & Shulman, 2002; Katsuki & Constantinidis, 2014). Visual perception can therefore be modulated by both bottom-up and top-down attentional processes (Kimchi et al., 2016). This is important to note when studying behaviour responses to visual perceptual tasks. It may be unclear the extent to which responses derive from an individual's processing advantage of the whole or detailed level of a stimulus, or their ability to appropriately direct attention to meet the task goals, thus inconsistencies in response patterns may arise from differing attentional demands of the task.

1.2.2 Mathematics and science

The importance of science, technology, engineering and maths (STEM) to the economy as a whole, and to the lives of individuals, is undeniable. At a national level, engineering accounted for 27.1% of the total UK Gross Domestic Product (GDP) in 2014, employing 18.2% of the national workforce (Centre for Economics and Business Research (CEBR), 2015), while mathematical science research accounted for approximately 10% of UK jobs in 2010 (Deloitte, 2013). However, employers are currently experiencing an issue with recruitment; 43% of STEM vacancies are difficult to fill due to applicants not having the requisite skills or experience for the role (Bennett, 2016), and only 24% of STEM graduates are working in STEM occupations 6 months after graduating (National Audit Office, 2018). At an individual level, mathematics achievement at school is associated with earnings in adulthood (Walker & Zhu, 2013) and early problems encountered by children in both mathematics and science have been shown to persist through school and beyond (Cragg & Gilmore, 2014; Morgan, Farkas, Hillemeier, & Maczuga, 2016). Therefore, understanding the factors affecting mathematics and science performance through development is important for improving outcomes in individuals as well as maintaining a significant contribution to the national economy.

Mathematics achievement is associated with higher IQ (Deary, Strand, Smith, & Fernandes, 2007; Kyttala & Lehto, 2008), higher scores on general cognitive control processes known as executive functions (EF) (Alloway & Alloway, 2010; Alloway & Passolunghi, 2011; Cragg & Gilmore, 2014; Formoso et al., 2018; Merkley, Thompson, & Scerif, 2016) and better spatial abilities (Gilligan, Flouri, & Farran, 2017; Gilligan, Hodgkiss, Thomas, & Farran, 2018). Although science achievement has received far less attention, studies have identified associations with EF (Gropen, Clark-Chiarelli, Hoisington, & Ehrlich, 2011; Nayfeld, Fuccillo, & Greenfield, 2013; Vosniadou et al., 2018; Zaitchik, Iqbal, & Carey, 2014) and spatial abilities (Hodgkiss, Gilligan, Tolmie, Thomas, & Farran, 2018). These associations vary across development (Alloway & Passolunghi, 2011; De Smedt et al., 2009; Männamaa, Kikas, Peets, & Palu, 2012; Mazzocco & Kover, 2007) and with expertise (Bull & Scerif, 2001).

1.2.3 Rationale for examining whole-part associations with mathematics and science achievement

1.2.3.1 Common processes

Mathematics and science both involve sets of rules to be followed, pattern detection and analysis, and concepts which can be applied across many contexts (Bressan, 2018; X. Wei, Yu, Shattuck, McCracken, & Blackorby, 2013), and mathematics is an activity which requires an ability to restructure a problem to find a solution (Mousavi, Radmehr, & Alamolhodaei, 2012). These characteristics are also evident in whole-part tasks. Additionally, an understanding of the relationships between wholes and parts is more overtly important for mathematical tasks such as identifying missing terms in an equation (Baroody, 2000). This suggests that an ability to focus on details or to separate a target from its context may be advantageous to mathematics and science performance. It has also been suggested that the integration of parts into wholes represents a process of abstraction which is important across many domains including integrating letters to form words (De-Wit & Wagemans, 2015), and for conceptual understanding in mathematics and science (Ferrari, 2003; Mitchelmore & White, 2007; Nersessian, 1989). These characteristics represent common themes between visual and cognitive processes, suggestive of a relationship between mathematics and science and the whole-part constructs of global / local processing, field independence, and systemizing.

1.2.3.2 Atypical populations

1.2.3.2.1 Autism.

This thesis focusses on neurotypical populations, however, there has been a substantial volume of literature examining atypical perceptual processing using whole-part tasks, particularly in autistic individuals. Autism is a neurodevelopmental condition with impairments in social interaction and social communication, as well as repetitive and restricted behaviours, activities, and interests (American Psychiatric Association, 2013). Researchers have recognised that visuospatial abilities are often a strength in autistic participants, sometimes described as an 'islet of ability' (Shah & Frith, 1983).

1.2.3.2.2 Visual processing in autistic individuals

Studies comparing autistic participants with control groups have found relatively strong performances in tasks which may benefit from an ability to focus on details and to ignore the distracting context (Almeida, Dickinson, Maybery, Badcock, & Badcock, 2014; Happé, 1999; Happé & Frith, 2006; Jarrold, Gilchrist, & Bender, 2005; Pellicano, Maybery, Durkin, & Maley, 2006; Plaisted, O'Riordan, & Baron-Cohen, 1998). However, some studies report a lack of significant difference with neurotypical individuals on global and local processing (Hayward et al., 2012; Iarocci, Burack, Shore, Mottron, & Enns, 2006; Ozonoff, Strayer, McMahon, & Filloux, 1994), field independence (Brian & Bryson, 1996; De Jonge, Kemner, Naber, & Van Engeland, 2009; White & Saldaña, 2011), and Gestalt tasks (Hadad & Ziv, 2015). Further, some studies have only identified poorer performance by autistic individuals when there is additional task complexity such as inhibiting or switching (Mann & Walker, 2003; White, O'Reilly, & Frith, 2009), and a meta-analysis using datasets from 56 studies revealed no significant detail-level advantage in autism, although they did find slower responses compared with neurotypical individuals to the global level (Van der Hallen et al., 2015).

There is substantial variation across studies, which is likely to result from the different stimuli and methodology involved, the heterogeneity of the autistic group, and differing ways of matching to a control group. Despite this, in certain task conditions, autistic individuals have greater success on tasks that require a focus on the detail than neurotypical controls.

1.2.3.2.3 Mathematics and science in autistic individuals

A number of studies have examined the academic subject choices preferred by autistic students. There is a higher proportion of autistic students on physics, engineering and mathematics university courses than on non-science courses (Baron-Cohen, Wheelwright, Burtenshaw, & Hobson, 2007; Baron-Cohen et al., 1998).

Additionally, although the proportion of autistic individuals applying to college was low, they were more likely to select a mathematics or science course (Wei et al., 2013). A preference for mathematics- and science-based courses and professions has also been found in parents of autistic children (Briskman, Happé, & Frith, 2001), and in those who have a family member with autism (Baron-Cohen et al., 1998). These studies indicate that those with autism or autistic traits are more likely to choose to study or work in mathematics or science-related fields. It should be noted that although the greater involvement in mathematics and science subjects may be due to strengths associated with autistic traits, it could also be explained by a desire to be involved with more structured interests, rather than the chaos which may be associated with social interactions (James, 2010).

Studies have revealed a more mixed picture of mathematical ability in autistic youth. A study of adolescents with autism revealed more individuals with a 'peak' than a 'dip' in arithmetic ability, suggesting that overall, mathematics is an area of strength for autistic individuals (Jones et al., 2009). However, other research has found that mathematical achievement in autism is in line with expectation based on IQ levels and is therefore relatively poorer than average (May et al., 2013; Mayes & Calhoun, 2003). Other studies have identified deficits in mathematical outcomes for autistic children and adolescents across a number of different tasks including word problem-solving tasks (Bae, Chiang, & Hickson, 2015; Wei, Christiano, Yu, Wagner, & Spiker, 2015), numerical operations and calculation tasks (Griswold, Barnhill, Myles, Hagiwara, & Simpson, 2002; Wei et al., 2015), and overall mathematical achievement (Aagten-Murphy et al., 2015). However, this poorer performance in autistic individuals without an intellectual impairment may be related to deficits in attention switching (May et al., 2013; May, Rinehart, Wilding, & Cornish, 2015), working memory (Pellizzoni & Passolunghi, 2017), processing speed, and fine motor skills (Assouline, Foley Nicpon, & Dockery, 2012).

1.2.3.3 Summary

The nature of mathematics and science may lead to advantages for those who are detail-focussed, and who are able to separate a target from its context. Additionally, despite the lack of consensus about mathematical and scientific ability in autistic individuals, the research suggests that some characteristics associated with autism may be positive assets for mathematics and science. Certainly, there is a high prevalence of autism amongst those with exceptional mathematical ability (Bressan, 2018; Chiang & Lin, 2007). This could indicate a possible advantage of a detail-focus on achievement in mathematics and science tasks (Baron-Cohen, 2007), which may also be evident in neurotypical individuals.

1.2.4 Chapter overview

This chapter will introduce the key strands of research which form the basis for this thesis. **Sections 1.3** to **1.6** will clarify the whole-part constructs of interest here, namely global and local processing (**Section 1.3**), field independence (**Section 1.4**), Gestalt processing (**Section 1.5**), and systemizing (**Section 1.6**). Each of these sections explores factors which influence performance on tasks, changing responses which occur due to development, and the associations with mathematics and science. Finally, **Section 1.7** brings together the different research threads to set the scene for this thesis.

1.3 Global and local processing

An important role of the visual system is to identify which elements of a visual stimulus belong together to form a whole object (global processing) and which are details or textural features (local processing). Global and local levels exist in environments which are hierarchical in nature – where individual properties may be considered as structurally more global or local than other properties (Farran, Jarrold, & Gathercole, 2003; Harrison & Stiles, 2009; Poirel, Pineau, & Mellet, 2006). For example, a whole forest can be described as being more global than an individual tree which in turn would be more global than a single leaf. The processing of these different hierarchical levels enables us to understand the structure and context of visual scenes and make rapid sense of what we are seeing (Bruce, Green, & Georgeson, 2003; D'Souza, Booth, Connolly, Happé, & Karmiloff-Smith, 2016).

Global processing refers to the act of acquiring the gist, while local processing refers to the acquisition of detailed information (Koldewyn, Jiang, Weigelt, & Kanwisher, 2013; Nayar, Franchak, Adolph, & Kiorpes, 2015). Global processing involves two distinct operations; grouping (or element clustering) to determine which local elements belong together, and shape formation (or element configuration) to determine the shape of the whole. Grouping requires few attentional resources and enables us to quickly perceive information such as the position and size of an object (Trick & Enns, 1997). Shape formation is achieved through using element configuration to perceive the boundary of grouped elements, separating them from the background and surrounding information (Behrmann & Kimchi, 2003; Razpurker-Apfeld & Kimchi, 2007). Shape formation is more likely to require attention than grouping (Behrmann & Kimchi, 2003). Global processing is therefore primarily concerned with the spatial relationships between local elements, so they can be integrated into a recognisable global object. In contrast, local processing involves the separation of elements which might otherwise have been perceived together; a process known as segmentation or individuation (Kimchi, 2015; Scherf et al., 2009). Once elements are grouped or separated, attentional resources and the retrieval of information from visual long-term memory enable object identification (Gerlach & Poirel, 2018; Kimchi, 2015; Poirel, Pineau, & Mellet, 2008).

Global and local processing are two distinct processes, rather than opposite extremes of a single visual processing continuum (Milne & Szczerbinski, 2009). Navon (1977) proposed that human perception was organised with a global-to-local order of processing, although these processes are likely to overlap rather than occurring in sequence (Navon, 1981).

1.3.1 Global and local processing tasks

1.3.1.1 Typical task design

To research global and local processing, Navon (1977) designed hierarchical stimuli so that the global and local levels can be compared fairly. A typical example of a Navon stimulus would be a large H (global stimulus) made of small S's (local stimuli) (see **Figure 1.1A**). Both levels are equally recognisable, and the content at each level is independent and cannot be predicted from one another (Kimchi, 1992; Navon, 1977). Navon tasks, therefore, can measure both global and local processing independently without the confounds of pre-understood inter-level relationships which would be present in everyday scenes (Kinchla & Wolfe, 1979). When encoding a Navon stimulus, processing advantage is inferred through comparing accuracy and response time (RT) data between global and local responses, as well as through studying congruency effects. Congruency effects are recognised by faster RT and higher accuracy when the stimuli at global and local levels are identical (congruent trials) than when the levels include inconsistent information (incongruent or neutral trials) (**Figure 1.1B**). Interference is greater when irrelevant information is presented at an individual's preferred level, than at their non-preferred level (Farran et al., 2003).

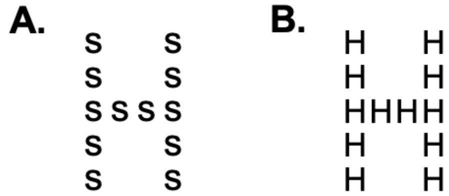


Figure 1.1: *Navon figure using letters.* ***A.*** *Incongruent example with H at the global level and S at the local level.* ***B.*** *Congruent example with H at both the global and local levels.*

Global and local processing is usually studied in the visual domain, but can also be measured in other domains. In the verbal domain, individual letters can be considered to the local elements of the whole word, or words the local elements of a sentence or paragraph. Also, the context of the sentence can determine which words might be more appropriate. Tasks measuring whole-part relations in the verbal domain include sentence completion using the context of the sentence (Booth & Happé, 2010), selecting the most appropriate missing word (Norbury & Nation, 2011), selecting the correct homograph for the context (Frith & Snowling, 1983; Happé, 1997), and error-spotting tasks (Norbury & Nation, 2011). A key problem with using verbal tasks is that there is a confound when reading and verbal abilities vary within the sample. This is particularly pertinent in developmental studies where semantic experience, reading level, and comprehension will vary widely between children within the same age group, and across age groups. Therefore, verbal whole-part tasks were not considered appropriate measures for this thesis.

1.3.1.2 Global advantage and global-to-local interference

In neurotypical adults, studies have revealed a global advantage as well as global-to-local interference (Kimchi, 2015). Recent research has suggested that these are manifestations of two distinct mechanisms. The former results from the earlier processing of the global level compared with the local level, known as global precedence (Navon, 1977, 1981). This is due to sensory mechanisms of the visual system which create a rough global representation before a fine-detailed representation. An initial coarse-grained representation of the global level, limited to determining the presence of an object and basic categorisation, can be achieved 150 ms after the onset of the stimulus (Hegdé, 2008), and is processed by the quicker dorsal pathway (Gerlach & Poirel, 2018; Liu, Wang, Zhou, Ding, & Luo, 2017; Poirel, Pineau, & Mellet, 2008). This then provides the structure of the visual scene to guide subsequent local processing of finer-grained details along the slower ventral pathway (Flevaris, Martinez, & Hillyard, 2014; Liu et al., 2017; Thomas, Kveraga, Huberle, Karnath, & Bar, 2012). The dorsal visual pathway automatically creates an integrated global representation, even when the global level is irrelevant to the task which is important for the quick recognition of objects and for making early predictions (Gerlach & Poirel, 2018; Kveraga, Boshyan, & Bar, 2007; Liu et al., 2017). This faster global processing takes place automatically and without modulation by object recognition (Poirel, Pineau, & Mellet, 2008), resulting in the global precedence effect identified by Navon (1977).

In contrast, global-to-local interference effects are likely to result from attentional processes involved in object identification as well as conflict resolution processes. This interference effect demonstrates the dominance of the global level over the local level, rather than simply the differences in the speed at which each level is processed (Kimchi, 1992). In studies comparing effects of meaningful and non-meaningful stimuli, global-to-local interference was only evident when the Navon stimuli comprised recognisable letters and objects (Beaucousin et al., 2011; Poirel, Pineau, & Mellet, 2008). This suggests that processes involved in object identification drive the global-to-local interference effect. Object recognition relies on the coarse-grained initial global outline to be matched with representations stored in long-term memory; more efficient use of the initial global outline for matching objects in long-term memory results in a greater global advantage in response time as well as greater global-to-local interference (Gerlach & Poirel, 2018; Poirel, Pineau, & Mellet, 2008). Studies comparing global and local processing using electroencephalogram (EEG) data, have revealed that the global level automatically captures attention at an early stage of processing leaving fewer resources for processing the local level. However, when the global level contains a meaningless object, it does not capture attention and therefore attentional resources are available to process the local level, resulting in no interference effect (Beaucousin et al., 2011, 2013).

In sum, the global precedence effect and the global-to-local interference effect identified by Navon are likely to result from different processes. The former is due to automatic sensory mechanisms in the visual system while the latter is due to cognitive mechanisms associated with object recognition and identification. Additionally, attentional resources are involved in resolving the conflict between levels that results from incongruent information being presented. Attentional processes enable individuals to focus on relevant elements, and ignore irrelevant inputs (Pinto et al., 2013). Where information differs at each level, inhibitory control is employed to resolve conflict and control the allocation of attentional processes (Amso & Scerif, 2015).

1.3.1.3 Exceptions to global precedence

One focus of research has been to establish how, and indeed whether, people can be categorised as being more inclined towards global or local processing. Navon's original research with hierarchical figures in neurotypical adults demonstrated a global advantage (Navon, 1981), however, subsequent research has indicated that this is not universal. In the adult population, global precedence varies across different neurodevelopmental disorders (Bölte et al., 2007; D'Souza et al., 2016; Plaisted, Swettenham, & Rees, 1999), special-interest groups such as artists and musicians (Chamberlain, Van der Hallen, Huygelier, Van de Cruys, & Wagemans, 2017; Stoesz, Jakobson, Kilgour, & Lewycky, 2007), and cultures (Davidoff, Fonteneau, & Goldstein, 2008; Lao, Vizioli, & Caldara, 2013; Oishi et al., 2014). It also varies depending on participants' emotional state (Fredrickson & Branigan, 2005; Srinivasan & Hanif, 2010), and can be manipulated through priming (Huttermann, Bock, & Memmert, 2014; Poirel et al., 2014). This suggests that any observed visual processing advantage towards either the global or local level is malleable rather than static.

1.3.2 Key influences on global and local responses

1.3.2.1 Attentional demands

When completing a Navon task, participants can be asked to select which of two presented choices is most similar to a target stimulus (a free choice task) (Dukette & Stiles, 1996; Harrison & Stiles, 2009); to name the stimuli at a pre-specified level (a selective attention task) (Wang et al., 2007); or to identify whether a particular shape or letter is present or where a target is located (a divided attention task) (Katagiri, Kasai, Kamio, & Murohashi, 2013; Plaisted et al., 1999). These tasks require participants to attend to a single level (selective attention Navon tasks) or to switch their attention between two levels (divided attention Navon tasks) (Katagiri et al., 2013). As such, the degree to which an individual demonstrates a global or local advantage is not solely due to processes associated with object perception, but can also depend on the attentional demands of the task (Caparos, Linnell, Bremner, de Fockert, & Davidoff, 2013; Dale & Arnell, 2013). In fact, increased attentional demands of a task and an increase in task difficulty can magnify any pre-existing perceptual preferences and modulate performance (Mondloch, Geldart, Maurer, & de Schonen, 2003; Navon, 2003).

1.3.2.2 Effects of stimuli and task design

Global and local responses can also be modulated by the characteristics of the stimuli, including visual angle, density, exposure time, form, background, and saliency of elements.

There is a bias towards global responses in stimuli with a smaller visual angle and towards local responses with a larger visual angle (Kinchla & Wolfe, 1979; Lamb & Robertson, 1990; Wang et al., 2007). The point at which this change occurs varies by stimulus and task design, but ranges between 7^{o} and 10^{o} (Kimchi, 2015). One sensory mechanism which could explain this finding is that the visual system is most adept at processing stimuli of a particular size, regardless of whether they present at the global or local level (Kinchla & Wolfe, 1979).

Global advantage is also modulated by the density of local elements. The grouping of many, smaller elements requires less effort than grouping fewer, larger elements, so stimuli with a lower density of local elements elicit a reduction in global advantage (Dukette & Stiles, 1996; Harrison & Stiles, 2009; Kimchi, 2015; Kimchi et al., 2005). Conversely, individuated sparse local elements are easier to segment perhaps reflecting the increased likelihood that they are distinct objects rather than textural elements of a larger object (Kimchi et al., 2005). This pattern was also evident in a study which required children aged between 4 and 8 years to draw Navon stimuli of different densities; more accurate representations of the global or local level depended on the number and size of the local elements (Dukette & Stiles, 2001).

Exposure time of the stimuli can also impact responses, whereby very short presentation times are more likely to elicit responses commensurate with a greater global advantage (Kimchi, 2015; Scherf, Behrmann, & Luna, 2009; Wang et al., 2007). This likely reflects the quicker, coarse-grained processing of an initial global percept along the dorsal stream.

Global advantage can be modulated by the form of the stimuli, and is reduced if a novel object is presented at the global level and a more familiar object is positioned at the local level (Poirel et al., 2006; Poirel, Pineau, & Mellet, 2008). This is likely to be driven by attentional differences between the automatic processing of meaningful stimuli and greater resources required for recognising and categorising novel objects (Harrison & Stiles, 2009; Poirel et al., 2006). In a study with children which used Navon stimuli comprising of objects and non-objects, participants had to judge whether pairs of stimuli were the same or different. Inter-level interference only occurred in the older age groups of 6- and 9-year-olds when the non-target level was meaningful and therefore more salient, whereas interference had a constant effect in the younger, and therefore more inexperienced, children aged 4- and 5 years old (Poirel, Mellet, Houdé, & Pineau, 2008). A further influence relating to the form of the stimuli is whether they involve a closed shape (for example, a circle) or a non-closed shape (for example, a 'x'). Local stimuli are harder to segment when they are a non-closed shape, resulting in a greater global advantage (Han & Humphreys, 2002, 1999; Kimchi, 1994).

Studies have also identified an effect of background, such that Navon stimuli presented on a background of crosses reduces the global advantage in comparison to a plain background. The distracting background disrupts the grouping processes of the local elements as it is harder to discern which elements are similar to each other and therefore belong together (Han & Humphreys, 2002, 1999). The background colour of a stimulus can also affect global advantage. Studies examining sensory processing in the visual system have indicated that specific colours can suppress activity in the magnocellular cells in the faster dorsal stream, which could therefore disrupt early processing of the global level (Awasthi, Williams, & Friedman, 2016; Hugrass, Verhellen, Morrall-Earney, Mallon, & Crewther, 2018; Michimata, Okubo, & Mugishima, 1999). However, some studies have found no significant effect of background colour on performance (Dore, Dumani, Wyatt, & Shepherd, 2018), and at present, the variations between study and stimulus design make it challenging to fully comprehend the influence of colour on global advantage (Vidal-López & Romera-Vivancos, 2009).

Finally, a single local element in a contrasting colour to the rest of the figure reduces the global advantage in a divided attention paradigm as attention is drawn to the local element. The distraction of the salient local element needs to be overcome in order to process the global level effectively. However, there is no effect of the salient local element in global selective attention paradigm, indicating that the bottom-up influence of the local element can be overcome by top-down selective attention processes (Han & Humphreys, 2002).

1.3.2.3 Effects of gender

In additional to stimuli and task effects, gender differences have been identified in responses to Navon tasks. Using a free choice Navon task, Scheuringer and Pletzer (2016) observed a similar proportion of global choices between genders, but faster responses to local than global matches in females and faster global matches in males than females. Similarly, in a divided attention task, Roalf, Lowery and Turetsky (2006) found that females responded more quickly when the target was presented at the local level than the global level, but males revealed no difference in RT. Furthermore, where behavioural differences are absent, studies have nevertheless identified differences in neural activation, mediated by sex hormones, suggesting unique gender-related global or local neural strategies (Pletzer & Harris, 2018; Razumnikova & Volf, 2011). These gender differences are also revealed in other spatial tasks including mental rotation and navigation, where males use a more holistic or global strategy while females use a more detailed or local strategy (Pletzer, Petasis, & Cahill, 2014; Tzuriel & Egozi, 2010). However, there are also studies providing no evidence of gender effects. In a study assessing performance on Navon stimuli designed with varying line orientation and shape closure, there were no differences in responses depending on level, although females revealed enhanced discrimination of shape properties compared to males and poorer local responses with incongruent line orientation stimuli (Kimchi, Amishav, & Sulitzeanu-Kenan, 2009). A further study also revealed no difference in global and local responses between gender in a task assessing both global and local matching preferences and accuracy in judging the orientation of lines. However, males were better at judging line orientation than females, and correlations between the two tasks revealed a small positive association between global matches and line orientation judgement (Basso & Lowery, 2003). Although these studies did not identify a gender effect, the findings related to line orientation do lend support for the aforementioned positive relationship between mental rotation ability and global processing.

Studies comparing performance between genders in childhood have revealed fairly mixed outcomes. In two studies using the same free choice task, males chose a higher proportion of global matches than females in a sample of 4- to 12-year-olds (Kramer, Ellenberg, Leonard, & Share, 1996) and in a sample of 6- to 7-year-olds (Tzuriel & Egozi, 2010), consistent with the adult studies where a gender effect was identified. In contrast, using a different free choice paradigm, 4-year-old and 6-year-old females made a higher proportion of global matches than males, but this effect disappeared when geometric shapes were used instead of letters (Dukette & Stiles, 1996). There are also studies where no gender difference has been detected, such as in a study where children aged 3- to 9-years-old judged which of four choices was most similar to the target Navon stimulus (Vinter, Puspitawati, & Witt, 2010), and in a free choice matching task for children aged 7 to 10 years (Harrison & Stiles, 2009).

In sum, there is some evidence of a greater global advantage in males than females, however there does not seem to be a consistent pattern of gender differences, even across similar tasks or similar age ranges.

1.3.3 Global and local development

1.3.3.1 Importance of a developmental approach

In order to gain a fully-rounded understanding of behavioural responses in adults, it is important to consider the developmental perspective (Karmiloff-Smith, 1994). This can provide a deeper understanding of global and local processing through characterising children's performance across tasks developmentally. Further, developmental changes in performance may affect cross-task comparisons such that associations between tasks may differ with age (Thomas et al., 2009).

A global advantage has been identified in neurotypical adults, however the processes associated with global and local processing may develop along different trajectories, resulting in changing relative strengths across development. Although the visual system undergoes extensive changes in the first year of life, the development of perceptual processing continues until adolescence (Leat, Yadav, & Irving, 2009). For example, while adult-level visual acuity is generally reached by about 7-years of age (Mondloch et al., 2003), neural developmental processes associated with visual attention occur at a slower rate (Mondloch et al., 2003).

There is general agreement in the literature that a local-to-global processing change occurs with development. This shift is thought to result from later maturation of processes associated with grouping elements relative to the processes involved with segmenting elements from a whole, particularly when the task parameters (e.g., duration) and stimulus design (e.g., density of local elements) leads to more effortful grouping and easier segmenting (Kimchi et al., 2005; Scherf, Luna, Kimchi, Minshew, & Behrmann, 2008). However, there is a lack of consensus on the timescale of this change. This is largely due to the high volume of cross-sectional studies and lack of longitudinal studies. Although cross-sectional studies can give an indication of developmental changes, the most effective method is to follow-up these studies with a longitudinal study (Thomas et al., 2009).

1.3.3.2 Development of performance on free choice tasks

Cross-sectional studies using free choice tasks have revealed an increasingly global advantage with age. Dukette et al. (1996) found that although both 4- and 6-year olds made a higher proportion of global matches than local matches using 'standard' stimuli, 4-year olds, but not 6-year olds, made more local choices when presented with sparser stimuli, which suggests an improvement in the grouping processes between the ages of 4 and 6 years. Using a similar task design, Harrison et al. (2009) found that children aged 9-years-old and younger made fewer global matches than older participants when the stimuli were presented with no bias towards a particular level. Both Dukette et al. (1996) and Harrison et al. (2009) presented the target and choice stimuli together, however only Harrison et al. (2009) included a maximum RT, which may have led to the later developmental shift observed in this study. In both studies, some stimuli were designed to encourage a response towards a particular level, and participants of all ages then responded more globally or locally depending on the bias. This suggests that grouping and segmenting process were available to children and adults when there was less competition between the response choices.

Further studies using a free choice matching paradigm, revealed a higher proportion of global matches in 5-year-olds (Vinter et al., 2010), 7-year-olds (Poirel et al., 2011), and 7- to 12-year-olds (Kramer et al., 1996) compared with younger children. Consistent with these findings, Booth (2006) also identified a local-to-global pattern of development in a free choice Navon task. However, the change occurred in adolescence rather than childhood: significantly fewer global matches were made in the free choice Navon task in children and young adolescents (8-10 and 11-13 years) than mid-adolescents and young adults (14-16 and 17-25 years). A possible explanation for this more delayed transition to a global advantage is that each target stimulus was only visible for 250ms, so participants had to hold information in working memory and to make judgements after only a brief presentation of the information.

In summary, these results suggest that as the task demands increased, through using sparser stimuli, shortening the presentation time, or increasing memory load, the age at which responses became more global in standard free-choice design tasks increased. Despite this, the studies generally indicated that changes in global and local responses to free choice tasks occurred between the ages of 5- and 7-years-old.

1.3.3.3 Development of performance on other global/local tasks

Changes in global and local processing in childhood have also been observed at approximately 5 to 7 years using other task designs. On a task where participants had to identify whether two stimuli were the same or different, 4- to 5-year-olds made fewer errors on the local level than the global level, while 9-year-olds made fewer errors on the global level (Poirel et al., 2008). Using a drawing task where children had to recreate a given Navon stimulus, the local level dominated children's drawings until the age of 5 years, at which point levels began to be more integrated (Vinter et al., 2010).

In a small-scale selective attention Navon study comparing performance of children (8- to 13-year-olds), adolescents (14- to 17-year-olds), and adults (aged at least 18 years), a local advantage was identified in both children and adolescence and a global advantage in adults (Scherf et al., 2009). This later transition may be due to the lack of maximum response time, which could indicate that the local level is dominant in the final percept when enough time to process both the global and local level has been allowed. However, this seems at odds with the study by Dukette et al. (1996) which also had no time limit but recorded an earlier transition to a global advantage. A key difference between these tasks was that the former was a selective attention task while the latter was a free choice matching task, additionally highlighting the impact of task design on responses.

Contrary to most findings, Mondloch et al. (2003) observed later local than global processing development in a same-different task paradigm. Adult-like global responses were observed in 10-year-olds, but adult-like local responses were not observed until 14 years. The very short presentation time of only 50 ms may explain this difference, such that participants access only the early perceptual processes of basic shape perception along the quicker dorsal visual pathway (Hebart & Hesselmann, 2012; Zachariou, Klatzky, & Behrmann, 2014). Further, in a divided attention study with three hierarchical levels where participants had to identify whether a square was present or absent in trials with zero to five distractor stimuli, the 5-year olds responded most quickly and accurately to the global level, whereas the 9-year olds and adults responded equally accurately to both the global and intermediate levels. All age groups made most errors to the local level. This global advantage in the youngest age group may result from poorer attention control when faced with more complex stimuli, resulting in an inability to disengage from the more salient (global) level (Krakowski et al., 2016).

1.3.3.4 A developmental shift between 5 and 7 years of age

There are several suggestions as to why a developmental shift may be observed in this age range. A study with two groups of 6-year olds, identified group-level differences in grey matter volume in the right hemisphere; the group responding with a global advantage had reduced grey matter volume than the group with a local advantage on a Navon matching task. This may reflect group differences in maturation and specialisation in these regions, such that those with a global advantage had experienced neural changes commensurate with a perceptual and attentional shift towards global processing (Poirel et al., 2011).

Developmental changes in EF and attention may explain these childhood changes in global and local processing. A transition around this age has been observed in behavioural studies examining the development of inhibitory control and attention. In a study examining congruency effects in a flanker inhibition task, there was a reduction in the stimulus interference from incongruent information between the ages of 7 and 10 years (Cragg, 2016). Similarly, response inhibition in an adapted Go/No-Go task, improved between the ages of 7 to 9 years (Cragg & Nation, 2008). In a study testing sustained attention, children aged 5 to 6 years were slower and made more errors than the 11- to 12-year olds (Betts, Mckay, Maruff, & Anderson, 2006). Improved response inhibition, selective attention, and sustained attention may allow children to better selectively attend to global or local information and inhibit a dominant response towards a preferred level.

Additionally, it has been suggested that school-based activities could have an impact on global and local responses. Reading ability has been shown to correlate positively with performance on visual search tasks, which may result from changes in the allocation of attention while learning to read such that the focus moves from individual graphemes to whole words (Krakowski et al., 2016). This is supported by a study comparing Navon responses in dyslexic children and neurotypical readers, whereby the dyslexic group exhibited a poorer global advantage than neurotypical readers. After a training intervention using action video games, the dyslexic group's responses to the global level improved along with an increase in reading speed (Franceschini, Bertoni, Gianesini, Gori, & Facoetti, 2017). This raises the possibility that the improvements in reading ability over the first few years of school may contribute to the changes in global and local responses observed using the Navon tasks.

1.3.4 Global / local processing and mathematics / science

The association between global and local processing and mathematics and science in neurotypical individuals has not previously been explored directly. There is, however, some research looking at mathematical ability in autistic individuals, as discussed in **Section 1.2.3.2.3**. The greater detailed focus exhibited by autistic individuals compared with neurotypical individuals, may be the reason for a relatively high proportion of individuals with autism or autistic traits electing to study or work in mathematics and science compared with other fields. Where there is lower achievement on mathematics tasks in autistic individuals who do not have an intellectual impairment, this may be related to deficits domain-general abilities including working memory (Pellizzoni & Passolunghi, 2017) and processing speed (Assouline et al., 2012). However, there is not conclusive evidence of an association between enhanced local processing relative to global processing and mathematics or science performance, so further research is required including a consideration of mechanism.

1.3.5 Summary

Global and local processing describes the relationship between the whole (gist) and elemental parts (details) and is typically measured using Navon stimuli. In adults, these tasks reveal congruency effects, a global precedence, and global-to-local interference. In the vast majority of studies with children, a local to global processing change has been observed with development. The main difference between studies is the age at which this transition takes place. Several studies have identified a change occurring between the ages of 5 and 7 years which may relate to parallel development of attentional systems and EF, as well as school-related activities such as reading which encourage a global focus.

In both adults and children, the design of the task and stimuli, such as number and size of the local elements, length of presentation, and use of meaningful stimuli can modulate responses. Responses to Navon tasks are not only related to perceptual processes, but are also influenced by attentional processes. Global advantage can also vary between genders, cultures, and an individual's emotional state.

Currently, there are no studies which have examined the relationship between global / local processing and mathematics and science in neurotypical populations.

1.4 Field independence

Field independence (FI) measures the extent to which individuals are able to separate a target from its context, and therefore the extent to which individuals are distracted by a particular context. It was originally conceptualised by Witkin and Asch (1948) with a rod and frame test, where participants were faced with a tilted rectangular frame and a differently tilted rod. The task required participants to twist the rod into an upright position, ignoring the distracting frame. Witkin and Asch (1948) noticed that performance on this task associated with other tasks requiring a target to be separated from its field (Evans, Richardson, & Waring, 2013). Individual differences on the tasks were therefore interpreted as reflecting a general ability to disembed a focal stimulus from its surroundings in order to make sense of it without distractions. Those who could disembed more successfully, as reflected in higher accuracy or quicker RT, are described as field independent, while those more distracted by the context are described as field dependent. However, one criticism levelled at the tasks measuring FI is that they only measure field independence. There is no active measure of field dependence; it is simply denoted by lower scores on the FI task (Evans et al., 2013; Richardson & Turner, 2000). The concern is that there may be reasons, other than a difficulty in overcoming the influence of the context, that result in lower scores on the task and which are unmeasured. However, there are few suggestions as to what these factors might be, apart from a misunderstanding of task demands and poor concentration, which one could argue are also factors in any cognitive task. Nonetheless, for clarity, this thesis will refer to participants as 'more' or 'less' field independent, to reflect the ability being actively measured in these tasks.

The literature sometimes describes FI in more perceptual terms, whereby the task relates to differences in processes of perceptual analysis, and ability to overcome perceptual grouping (De-Wit, Huygelier, Van der Hallen, Chamberlain, & Wagemans, 2017; Miyake, Witzki, & Emerson, 2001). However, it is also described in broader, cognitive terms, whereby FI describes the extent to which individuals are able to analyse and restructure a given stimulus to solve a problem that requires details to be decontextualised (Pithers, 2002; Rémy & Gilles, 2014; Witkin, Moore, Goodenough, & Cox, 1977). FI associates with global and local processing (Chamberlain et al., 2017; Poirel, Pineau, Jobard, & Mellet, 2008), spatial thinking (Rémy & Gilles, 2014), formal operational reasoning (Lawson, 1976), and orientation abilities (Boccia, Piccardi, Di Marco, Pizzamiglio, & Guariglia, 2016), as well as general cognitive abilities (Flexer & Roberge, 1980; Miyake et al., 2001). This has resulted in a lack of clarity about what the construct of FI actually represents (Busch, Watson, Brinkley, Howard, & Nelson, 1993; Evans et al., 2013). Witkin was originally more interested in theorising FI as a cognitive style (De-Wit & Wagemans, 2015). Cognitive styles are general strategies of representing and processing information which an individual applies consistently across cognitive, perceptual, and even personality systems (Davey, 1990; Davies & Graff, 2006; Witkin et al., 1977). In keeping with FI as a cognitive style, studies have identified general characteristics and traits which characterise those who are more and less field independent. Those who are more field independent tend to be more analytical and reflective (Arnup, Murrihy, Roodenburg, & McLean, 2013; Brosnan et al., 2011), are more intrinsically motivated, and favour independent working (Witkin, Moore, Oltman, et al., 1977). In contrast, those who are less field independent demonstrate a preference for social interactions, and prefer more structured learning environments (Evans et al., 2013; Onyekuru, 2015; Witkin & Goodenough, 1977). Although these studies have identified interesting associations with personality traits, they have little explanation of the mechanisms which are responsible for these observed patterns.

1.4.1 Field Independence tasks

FI is typically measured using the Embedded Figures Test (EFT) (Brosnan, Gwilliam, & Walker, 2012; Hao et al., 2013), although the Block Design task (Globerson, Weinstein, & Sharabany, 1985) and Navon figures (Billington, Baron-Cohen, & Bor, 2008) have also been used. In the EFT, participants must locate a simple target shape within a complex stimulus. When a target is embedded within a context, it becomes hidden due to shared lines, overlapping colours, and distracting patterns, and is therefore perceived as part of the whole (De-Wit et al., 2017; Mumma, 1993; Poirel, Pineau, Jobard, et al., 2008). This contrasts with visual search paradigms and matching tasks where a target is surrounded by distractor elements, but is not embedded within a context (Li, Wu, Zhu, & O'Boyle, 2014; Wolfe & Horowitz, 2017). The process of identifying a target embedded within a context is known as disembedding. There are several variants of the EFT, tailored for particular age groups and types of testing. These include the Group EFT (GEFT), a pen-and-paper version of the task which is administered to participants in a group situation, and the Children's EFT (CEFT), which uses colourful meaningful images and only two simple target shapes across the task. More recently, further variants have been created to systematically examine factors affecting the disembedding process. The Leuven EFT (L-EFT) examined the effects of the number of lines forming the target shape outline, the number of lines shared with the background, and target shape symmetry; the D-EFT contrasted performance between 2D scenes and 3D scenes; and the M-EFT compared meaningful and non-meaningful stimuli. The L-EFT, D-EFT, and M-EFT revealed that task performance was poorer in adults when the target shape was composed of fewer lines, shared more lines with the background, and was asymmetrical, as well as when the stimulus design was 2D rather than 3D, and when it contained non-meaningful information (Chamberlain et al., 2017; De-Wit et al., 2017; Huygelier, Van der Hallen, Wagemans, De-Wit, & Chamberlain, 2018; Van der Hallen et al., 2018). Together, this suggests that target shapes are harder to disembed when they are more integrated into the background, and when the shape is less unique, and therefore harder to distinguish from similar distractor shapes in the stimulus (De-Wit et al., 2017).

In the Block Design task, participants must partition a pattern into a grid of squares, and then match each segment of the grid with a choice of patterns. The pattern becomes harder to segment when the design spreads over a grid boundary, as the large-scale pattern has to be overcome to identify the individual elements. This process is known as segmenting (Schorr, Bower, & Kiernan, 1982; Shah & Frith, 1993). It differs somewhat from disembedding because it is guided by fixed gridlines, whereas disembedding involves a search process to identify the boundaries of the target. Some research has suggested that success on the Block Design task requires a suppression of the overall distracting figure in order to identify the smaller elements creating the pattern (De Jonge et al., 2009). However, this may vary depending on the strategy being used to complete the task (Rozencwajg & Corroyer, 2002). Performance may reflect use of the overall figure if participants change the elemental blocks until the pattern matches the design of the target, through a 'trial and improvement' strategy. Conversely, responses may reflect an individual's focus on details, if they employ an analytic strategy of mentally separating individual elements of the pattern and match with the choice elements. Participants may even use a combination of these approaches, for example, when their chosen approach varies depending on the design of the pattern. When the edges of individual elemental blocks are clearly discernible, an analytic approach may be more appropriate, whereas when the pattern includes a recognisable object like an arrow, an approach which aims to match the whole image using a 'trial and improvement' strategy may be more successful (Schorr et al., 1982). Some variants of the Block Design task may encourage a particular strategy, for example, pen-and-paper versions do not allow an opportunity for trial and improvement strategies.

To identify the local parts of a Navon stimulus, the participant has to overcome the distracting global level by segmenting the global image into parts. As the elemental parts of the Navon figures are already quite distinct, the challenge of segmenting is more straightforward than in the Block Design task.

Although there are two key processes involved in these FI tasks (disembedding and segmenting), both test the ability to extract target information from its context.

1.4.2 Influences on field independence

1.4.2.1 Effect of gender

Studies have identified gender effects in FI, with males achieving higher accuracy in the GEFT than females in adult groups (Jantan, 2014; Onyekuru, 2015), and in adolescence (Flexer & Roberge, 1980; Witkin et al., 1977). Conversely however, there was no significant gender difference in an undergraduate student sample using a forced-choice online version of the EFT (Billington et al., 2007), which may indicate a role of task design in the measurement of disembedding. There is also evidence of a small male advantage on Block Design tasks. One study examining intelligence over a number of tasks found that the Block Design task had one of the larger between-gender differences, although the overall conclusion was of no overall gender effect on intelligence (Colom, García, Juan-Espinosa, & Abad, 2002). This is supported by a study examining changes in Block Design responses over the adult life-span which also identified a small gender effect with higher scores in males compared with females (Rönnlund & Nilsson, 2006).

Some research has identified a similar male advantage for FI using measures of disembedding in childhood over a range of age groups, including 10- to 11-year-olds (Jantan, 2014), 7-to 8- and 10-to 11-year-olds (Amador-Campos & Kirchner-Nebot, 1997), as well as 8-year-olds and 11-year-olds (Cairns, Malone, Johnston, & Cammock, 1985). This gender difference has also been identified in adolescent groups aged 10 to 17 years (Witkin, Goodenough, & Karp, 1967) and 11 to 14 years (Flexer & Roberge, 1980). One study using a Block Design task identified a male advantage in children aged 8 to 10 years, but no gender effect in age groups 11 to 25 years (Booth, 2006). Some studies using disembedding tasks have not identified a male dominance in FI, including a longitudinal study of 11- to 14-year-olds (Flexer & Roberge, 1983) and a cross-sectional study of children aged 5 to 10 years (Bigelow, 1971), while one study which administered the Block Design task to children aged 4 to 8 years identified no gender effect (Akshoomoff & Stiles, 1996). A study with 5-year-olds even identified greater FI in females than males (Coates, 1974). Overall, there appears to be better performance on FI tasks in males than females, but where these are identified the absolute differences tend to be small, which may explain the inconsistencies between studies.

1.4.2.2 Associations with IQ and Executive Functions

Success on FI tasks is dependent on perceptual factors such as the ability to parse a visual stimulus into parts, as well as cognitive factors including the ability to organise and analyse a stimulus, and to keep a target in mind whilst searching the distracting context (Guisande, Rodríguez, Almeida, Tinajero, & Páramo, 2008; Huygelier et al., 2018; Mumma, 1993). It is therefore not surprising that individual differences in IQ and EF associate with performance on FI tasks.

IQ is a measurement of general cognitive abilities, which includes many elements including problem solving, reasoning, processing speed, and adaptability (Duncan, Chylinski, Mitchell, & Bhandari, 2017; Richardson, 2002; Sternberg, 2008). It can be measured using verbal or non-verbal tests. There are a number of studies that report a positive association between FI and IQ (Flexer & Roberge, 1980; Swyter & Michael, 1982). In fact, some studies have even concluded that the GEFT is indistinguishable from a general intellectual measure (Leo-Rhynie, 1985; Vernon, 1972). Although the Block Design task can be considered to be a test of FI, it also forms a sub-test of the Wechsler intelligence scale, and therefore is considered to be a measure of visuospatial ability (Killgore & Gogel, 2014). However, a study which grouped participants according to performance on an EFT observed no group difference on a non-verbal IQ task, indicating that the association with IQ is not always related to FI (Li et al., 2014). Overall, the relationship between IQ and FI is likely to reflect the common task demands of segmenting and abstracting target information from the complete stimulus (Duncan et al., 2017; Goodenough & Karp, 1961).

Studies have also revealed an association between FI and EF, particularly working memory (WM) and inhibitory control (IC). EF is an umbrella-term for top-down control processes which are required to problem solve, achieve task goals, and adapt responses in a novel situation (Diamond, 2013; Hughes, 2011). These processes can be grouped into three correlated but distinct factors: Working memory, inhibitory control, and cognitive flexibility or shifting (Carlson, Zelazo, & Faja, 2014; Miyake et al., 2000). WM describes the fixed, limited capacity where information is held and manipulated on a short-term basis in order to complete a task (Gathercole, Pickering, Knight, & Stegmann, 2004), and can be measured in both the visuospatial and verbal domains. IC refers to the ability to ignore irrelevant distractors or to stop a prepotent but incorrect response (Diamond, 2013; Miyake et al., 2000). It is not considered to be a unitary entity, and generally, it is divided into two categories (Brookman-Byrne, Mareschal, Tolmie, & Dumontheil, 2018; Verbruggen, Liefooghe, & Vandierendonck, 2004). Semantic inhibition refers to the suppression of meaning, while response inhibition refers to the suppression of a motor response. WM and IC are thought to be aspects of similar yet distinct mental processes, and most cognitive tasks draw on a combination of both processes; the maintenance of WM processes requires that IC prevents actions inconsistent with task aims and reduces the impact of distractions (Luna, Velanova, & Geier, 2008). Cognitive flexibility describes the ability to adjust behaviour depending on changes in the environment, and usually refers to task switching whereby participants differentially respond to stimuli based on pre-determined criteria (Dajani & Uddin, 2015). EF continues to develop into adulthood although individual processes follow different developmental trajectories. For example, there is substantial development of IC in preschool, whereas the development of WM and cognitive flexibility is more protracted (Best & Miller, 2010). Cognitive flexibility is less distinct in children, possibly because efficient task and set shifting relies on appropriately reallocating attention as well as inhibiting previous information, and therefore these processes need to be developed first (Dajani & Uddin, 2015; St Clair-Thompson & Gathercole, 2006). Individual differences in EF have far-reaching consequences, from variation in academic success (Mariëtte Huizinga, Baeyens, & Burack, 2018) to emotional regulation (Hughes, 2011; Hughes & Ensor, 2011) to health matters and the likelihood of a criminal conviction (Moffitt et al., 2011).

In order to successfully complete the EFT, participants have to hold the target shape in mind, whilst they monitor the complex figure and ignore distractor shapes, so they can isolate items of importance from the context. Research has identified a positive association in adults and children between WM and EFT in the verbal (Guisande et al., 2008) and visuospatial (Miyake et al., 2001) domains, as well as with IC (Imanaka, Kakigi, & Nakata, 2017). IC is important for appropriately directing attention on the EFT, so that the overall complex figure is not a distraction (Rittschof, 2010). By focussing on the most relevant information in the stimulus and filtering out distractor information, field independent individuals are less likely to become overwhelmed with information and have a reduced memory load, so performance on the EFT is improved (Alamolhodaei, 2002; Evans et al., 2013; Jia, Zhang, & Li, 2014). Studies with adults have tended to use one or two EF tasks to assess performance, such as a forwards corsi memory span (Huygelier et al., 2018), a verbal 2-back memory span (Miyake et al., 2001), or auditory and somatosensory Go/No-Go tasks (Imanaka et al., 2017). Currently, therefore, there is a lack of clarity about how FI may differentially relate to verbal or visuospatial WM, or to different types of IC processes.

In order to be successful on the Block Design task, participants inhibit the effect of the overall pattern (although, as discussed earlier, this may depend on their chosen strategy), and they need to keep the design of the option blocks or the segmented element in mind so that they can be matched. There are very few studies that have examined associations between EF and the Block Design task. However, one study revealed that the Block Design correlated with tasks measuring WM, but not with tasks measuring IC or cognitive flexibility (Friedman et al., 2006). This suggests that the processes of disembedding (as measured by the EFT) and segmenting (as measured by the Block Design task) may rely on different EF, although both have an association with WM.

1.4.3 Field independence development

Research has consistently shown a developmental trend for individuals to demonstrate greater FI with age (Akshoomoff & Stiles, 1996; Amador-Campos & Kirchner-Nebot, 1997; Busch et al., 1993; Glynn & Stoner, 1987; Goodenough & Eagle, 1963). This development is not necessarily linear however, with one study identifying no significant difference between children aged 5-6 and 6-7 years, but a large difference between the age groups of 7-8 and 8-9 years (Bigelow, 1971), and performance appears to plateau around the age of 17 years for both disembedding and segmenting (Amador-Campos & Kirchner-Nebot, 1997; Booth, 2006; Witkin et al., 1967). Although there may be large variability within each age group (Bigelow, 1971), this developmental change has been replicated in both cross-sectional (Amador-Campos & Kirchner-Nebot, 1997; Cairns et al., 1985; Flexer & Roberge, 1980) and longitudinal studies (Flexer & Roberge, 1983). It is worth noting that Witkin et al. (1967) collected longitudinal data, but then excluded these data from their analyses as the longitudinal group performed better than the age-equivalent group in their cross-sectional study. Only a subset of participants in the cross-sectional study returned for the longitudinal follow-up, so it is possible that this apparent practice effect was due in part to differences between this returning subset of participants and the main group.

Studies have revealed stable responses to FI tasks over time in adults on the EFT (Kepner & Neimark, 1984). Similar patterns have been revealed using the Block Design task although here, there was some decline in age groups over 60 years (Rönnlund & Nilsson, 2006).

1.4.4 Comparison with global and local processing

Both global and local processing, and FI describe relationships between wholes and parts, however there are some key differences between the two constructs. In the Navon task, the global and local levels are clearly distinguishable from each other and can be processed independently. There is often inter-level competition such that a response requires the non-target level to be ignored. In contrast, results from a recent study with adapted versions of the EFT has suggested that the whole is not ignored when disembedding the target shape (Chamberlain et al., 2017). In fact, the configurations of the lines and shapes of the whole may be used to successfully locate the target shape. However, this association between global processing and EFT is not consistent in the literature, and other studies have found that local processing relates to better FI and global processing associates with poorer FI (Milne & Szczerbinski, 2009; Poirel, Pineau, Jobard, et al., 2008). This difference may be due to the contrasting task demands; the Navon task in the Chamberlain study required participants to identify whether the letter at the directed level was a vowel or a consonant, which adds an additional complexity before a response can be made. Therefore, responses here may have revealed different patterns from tasks when the response required a more straightforward identity of the letter or a same/different judgement.

1.4.5 Field independence and mathematics / science.

A consistent picture has emerged from the literature of a positive association between FI and academic success in children and adolescents, particularly in mathematics (Alamolhodaei, 2002; Azari, Radmehr, & Mohajer, 2013; Buriel, 1978; Tinajero & Paramo, 1997). The majority of studies have used a general mathematics measure (Roberge & Flexer, 1983; Tinajero & Páramo, 1998) but some focussed on specific subsections of mathematics. In a study with university students, Zhang (2004) found that FI only associated with the geometry subsections of their mathematics measure. This may indicate that FI is restricted to visuospatial processes and therefore only associates with cognitive tasks with similar visual inputs. Other studies have found associations with word problems in adolescents (Alamolhodaei, 2002; Azari et al., 2013). Word problems are characterised by embedding the abstract mathematical idea into a real-world context. Those who can separate relevant conceptual information from a context are more successful than those who are more distracted by the context.

Fewer studies have investigated associations between FI and science achievement. Higher scores on the EFT have been achieved by those studying science compared with non-scientific disciplines at the undergraduate level (Billington, Baron-Cohen, & Wheelwright, 2007; Derussy & Futch, 1971) and in adolescents (Leo-Rhynie, 1985; Tinajero & Paramo, 1997). There have also been differences detected according to topic, e.g., geology being more associated with FI than evolution in undergraduates (Lawson, 1983), and there was a significant relationship in natural sciences but not social sciences in adolescents (Tinajero & Paramo, 1997). This suggests that the common underlying processes in FI and science vary in strength between subject topics, and may reflect differing requirements to decontextualise information. Differences may also depend on the style of questioning, with enhanced performance in more field independent individuals when questions require the application of knowledge and in essay assignments, contrasting with no FI advantage being identified on multiple choice questions (Lawson, 1983). This indicates that the development of conceptual understanding and fact retrieval is not limited in those less field independent, but the application of these facts into different contexts is more of a challenge.

Although there is evidence for a positive relationship between FI and academic achievement, there is limited understanding about the cognitive processes which underpin this association. It is possible that any association between FI and academic measures may be explained by other factors such as IQ and EF. As described in **Section 1.4.2.2**, studies have identified positive associations between FI and IQ, WM, and IC. These domain-general factors are also important in mathematics and science performance, therefore some shared variance is likely to be explained by these domain-general factors.

1.4.6 Summary

FI describes the extent to which individuals are able to separate a target from the context into which it has been embedded. The two processes associated with FI tasks are disembedding and segmenting. Both require the overall stimulus to be reorganised and analysed in order to isolate the smaller elements of interest. Task success is dependent on the visual features of the overall stimulus, which can make a target harder to disembed or a whole harder to segment. These visual features encourage the visual system to create a whole through early visual sensory processes. This will be discussed further in **Section 1.5**.

There is an improvement in FI with age which stabilises at around 17-years-old. This is usually examined cross-sectionally rather than longitudinally due to possible practice effect confounds. Although there are some inconsistencies in the literature, males may be slightly more field independent than females in childhood.

1.5 Gestalt processing

A third theory of perceptual organisation is Gestalt processing, which describes the grouping and integration of elements into distinct objects (De-Wit & Wagemans, 2015; Wagemans et al., 2012a). Gestalt processing is thought to be involved in the early preattentive stage of visual perception, when the visual field is segmented so that figures can be distinguished from the background, known as figure-ground organisation. There are a number of perceptual laws which determine which elements of a scene should be perceived as a group, and these are known as Gestalt principles of perceptual organisation (Wagemans, 2015; Wagemans et al., 2012a). These include proximity (**Figure 1.2A**), where elements located close to each other are grouped together; similarity (**Figure 1.2B**), where elements which look the same are grouped together; closure, where a complete shape is perceived even when information is missing (**Figure 1.2C**); and good continuation, where a line is perceived as following a continuous direction (**Figure 1.2D**) (Brooks, 2015; Han & Humphreys, 1999). As grouping occurs early in the perceptual process, the whole, or Gestalt, is often grouped before individual parts enter consciousness (Wagemans et al., 2012b). The Gestalt psychologists who observed and described the stimulus features which lead to the perception of wholes, considered that the whole is other than the sum of its parts, emphasising that the Gestalt was not simply an additive result of its elemental parts (Wagemans, 2015). The configuration of context of the individual elements create the whole, and in turn, the perception of these individual elements is influenced by the whole.

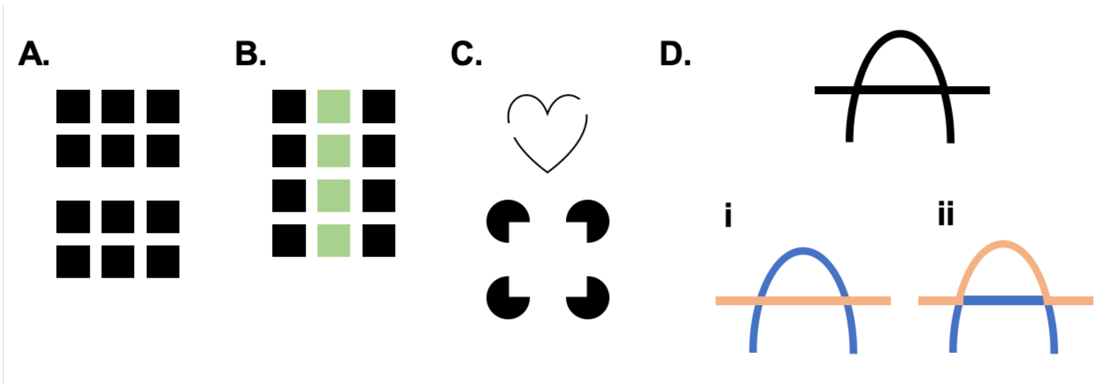


Figure 1.2: *Examples of Gestalt principles of perceptual organisation.* ***A.*** *Proximity. The stimulus is perceived as two groups of small squares.* ***B.*** *Similarity. The stimulus is perceived as three vertical columns.* ***C.*** *Closure. Lines are continued so a whole heart-shape is perceived, and a central white rectangle can be perceived even though there is no outline to the shape.* ***D.*** *Good continuation. Lines are perceived as continuous, so the stimulus will be perceived as a line and an arch (i), rather than as lines with abrupt changes in direction (ii).*

1.5.1 Tasks measuring Gestalt processing.

Gestalt processing can be measured using several tasks. These measure specific Gestalt laws, such as the law of closure as measured by Kanizsa triangles and fragmented pictures (Bölte et al., 2007; Nayar et al., 2015; Panis, De Winter, Vandekerckhove, & Wagemans, 2008), and sensitivity to the Gestalt using visual illusions (Van der Hallen et al., 2015). The common theme in these tasks is the ability to make sense of the whole, and the extent to which the whole distracts from understanding detailed elements. For example, Kanizsa stimuli (**Figure 1.2C**) lead to the perception of a central shape despite the fact that this shape is not actually present. The positioning and shape of the four individual elements drive the visual perception system to create a meaningful whole (Bölte et al., 2007), which in this case is a white square overlaying four black circles. A task requiring participants to perceive this white square can measure an individual's Gestalt processing using the principle of closure. Optical illusions are designed to confound the visual system by making use of Gestalt laws in unexpected ways. For example, an image of two face silhouettes in profile making a vase-shape between them interrupts the figure-ground organisation. Both parts of the stimulus could be the focus, and therefore the figure-ground law is violated. Similarly, Titchener circles, designed by Ebbinghaus, confuse the visual system where the perception of the size of the central circle is modulated by the size of the circles surrounding it. If participants isolate the central circles from their context, they can more accurately compare the size, demonstrating a reduced influence of the Gestalt (Kovacs, 2000; Walter, 2007).

Early Gestalt theory was criticised for being too descriptive with little consideration of the mechanistic explanations of perceptual integration (Wagemans et al., 2012b). Although there is still little focus on the mechanism of Gestalt processing, studies continue to measure the behavioural associates of Gestalt features, adding to the quantitative nature of the field (Jäkel, Singh, Wichmann, & Herzog, 2016). However, one problem with quantifying Gestalt processing in everyday situations is that typically there are several types of grouping occurring simultaneously. Processes can be isolated in individual research tasks, but there also needs to be an understanding of how these processes interact when presented together (Jäkel et al., 2016; Quinlan & Wilton, 1998).

1.5.2 Comparison with global processing

Global processing and Gestalt processing are both constructs describing the automatic drive for integration in visual perception, however there is a key difference. Global processing uses the configural relationships of the local elements for the whole to be understood. This means that local processing elements have a 'place' relationship with global processing, such that the global level can still be perceived regardless of the form of the local elements. For example, in Navon stimuli, a global 'H' can be recognised regardless of whether the local elements are S's or T's. In contrast, Gestalt processing additionally utilises the structural or relational properties of the local elements. This means that Gestalt local elements have a 'nature' relationship with the global Gestalt such that the form as well as position of the local elements are integral to the processing of the stimulus as a whole (Bölte et al., 2007; Brosnan, Scott, Fox, & Pye, 2004; Kimchi, 2015; Quinlan & Wilton, 1998; Wagemans et al., 2012a). The perception of the Gestalt can be altered by changing its local elements. For example, the squares in **Figure 1.2B** are equidistant from each other and are arranged in a rectangle configuration. However, by shading the middle column of the squares, the figure is not perceived as a whole grid, but as columns of squares by the law of similarity; thus, the change in their form with the same configuration results in a different understanding of the whole.

Despite this accepted distinction between the two constructs, it should be noted that a global advantage on Navon stimuli can be affected by a change in the form of local elements through differing object familiarity (Poirel et al., 2006), the use of colour (Han & Humphreys, 2002), or classifying by shape closure / non-closure (Kimchi, 1994). One explanation for this influence of form, is that it creates an imbalance in attention between the global and local levels, resulting in the local level becoming comparatively more salient. Additionally, as the processing of the global level of Navon stimuli is facilitated by the Gestalt laws of similarity and proximity, the global advantage will be reduced if these laws are disrupted, as revealed with an increased local advantage in Navon stimuli with fewer, larger local elements compared with many, smaller local elements.

1.5.3 Comparison with field independence

In order to focus on the details of the FI stimuli, the Gestalt principles which strengthen the perception of the whole need to be overcome (Poirel, Pineau, Jobard, et al., 2008; Shah & Frith, 1993). The L-EFT explored this more systematically by varying the target and background features according to the Gestalt grouping principles of symmetry, shape closure, and good continuation. Target shapes which were embedded into the background using a greater level of Gestalt principles were harder to disembed (De-Wit et al., 2017). This also indicates that there is an influence of early, automatic visual processing on disembedding. FI and Gestalt processing might therefore be expected to negatively correlate, but in fact this does not appear to be the case (Bölte et al., 2007; Milne & Szczerbinski, 2009) suggesting that other key processes are involved in these tasks.

Similarly, the Block Design task is harder to segment where lines, patterns and colours continue across block boundaries so that visual grouping occurs across the position where the pattern in fact needs to be segmented. The Gestalt influence can be further increased by using a pattern of stripes in the complex figure, rather than shapes such as diamonds (Rozencwajg & Corroyer, 2002). This Gestalt influence is illustrated by differences in accuracy between Block Design tasks and segmented Block Design tasks (Shah & Frith, 1993). In the latter, the Gestalt is interrupted at the boundary and therefore it is more straightforward to segment the whole and match the elemental parts to the choices.

1.5.4 Gestalt processing and mathematics / science

Gestalt principles of grouping describe the automatic organisation of visual stimuli into perceptual wholes relying primarily on bottom-up, stimulus-based processes (Kimchi et al., 2016; Marini & Marzi, 2016). It is important to have a clear understanding of Gestalt processing as performance on visual whole-part tasks may require Gestalt principles to be overcome in order to successfully segment or disembed elemental targets (White & Saldaña, 2011). However, it is unlikely that Gestalt processing will associate with academic performance as it reflects the automatic sensory processes of the visual system. For this reason, Gestalt processing is considered to be outside the scope of this thesis, and will not form part of the analysis into associations with mathematics and science.

1.5.5 Summary

Gestalt processing describes the features of a visual stimulus which determine how individual elements are integrated automatically into wholes. There are several examples of Gestalt principles of perceptual organisation, including proximity, similarity, closure, and good continuation. Together, they describe a heuristic account of how the most parsimonious grouping of a stimulus is achieved. Gestalt principles may explain varying performance on whole-part stimuli, but are unlikely to be related to mathematics and science achievement.

1.6 The systemizing-empathizing theory

A further construct which warrants examination as it also explores a detailed focus, is the systemizing-empathizing theory. This is a slightly different theory from those discussed above, as it is not limited to visual perception.

The systemizing-empathizing theory was initially proposed by Baron-Cohen (2002) to describe and explain broader characteristics of autism. It incorporates two dimensions: Systemizing describes an individual’s understanding of the physical world, and empathizing describes an individual’s understanding of the social world (Riekki, Salmi, Svedholm-Häkkinen, & Lindeman, 2018). Systemizing measures the drive to explore, create and understand systems (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003; Caldwell-Harris & Jordan, 2014). These systems follow rules, enabling categorisation as well as predictions to be made, so systemizers tend to isolate detailed features, rather than combine information together (Baron-Cohen, 2009). This makes it an important construct to consider in this thesis as it suggests an analytical approach including a focus on elemental parts. Empathizing measures the extent to which individuals can recognise and appropriately respond to other people's emotions (Baron-Cohen, 2002). Systemizing and empathizing scores have a weak negative correlation in neurotypical adults, but in childhood, studies have found either no correlation (Escovar et al., 2016), a small positive correlation (Wakabayashi, 2013), or a small negative correlation which actually became non-significant when only neurotypical children were examined (Auyeung & Wheelwright, 2009). Overall, this suggests that they are measuring unique constructs rather than being two extremes of a single continuum (Greenberg, Warrier, Allison, & Baron-Cohen, 2018; Wakabayashi et al., 2006; Wheelwright & Baron-Cohen, 2006). When these scores are combined, people can be described as having either a systemizing, a balanced, or an empathizing brain (Baron-Cohen, 2002).

Systemizing and empathizing are typically measured using questionnaires with a Likert format. For adults, there is the self-report systemizing quotient (SQ) and empathizing quotient (EQ) (Baron-Cohen, 2002; Wakabayashi et al., 2006), and for children, parents complete the questionnaires (SQ-C and EQ-C) on the child's behalf (Auyeung & Wheelwright, 2009). For each statement, the responder selects one of four possible levels of agreement from 'definitely agree' to 'definitely disagree'. The use of a Likert scale means the responses can be easily quantified and gives an indication of the strength of agreement compared with a simple yes/no response. However, particularly in the case of the SQ-C and EQ-C, the responses may simply reflect individuals' perception of themselves or their children (Bressan, 2018; MacLeod, 2008).

There is some question about whether the SQ is measuring a single systemizing dimension, as a single factor has been shown to be a poor fit in factor analyses (Baron-Cohen et al., 2003; Ling, Burton, Salt, & Muncer, 2009). However, a recent large-scale study with adult students found that the measure did indeed represent a single entity of systemizing, with subgroups of stronger inter-item correlation being driven by the common contexts of the statements (Allison, Baron-Cohen, Stone, & Muncer, 2015).

1.6.1 Influences on systemizing and empathizing

1.6.1.1 Effect of autistic traits

As the origin of the systemizing-empathizing theory was to explain the characteristics associated with autism, there are many studies which have compared the responses of autistic individuals with neurotypical controls. Autistic individuals tend to score highly on the systemizing scale and poorly on the empathizing scale, such that their combined score is located at the extreme systemizing end of the continuum (Auyeung & Wheelwright, 2009; Baron-Cohen, 2009). In fact, one study with over

670,000 participants revealed that the greatest predictor (by far) of variation in autistic traits was the difference between their systemizing and empathizing scores (Greenberg et al., 2018). However, some studies with children (Evers, Steyaert, Noens, & Wagemans, 2015; Vanegas & Davidson, 2015) and adults (Carroll & Chiew, 2006; Vanmarcke et al., 2016) have found that, contrary to expectation, autistic groups do not excel in the SQ although they do score poorly compared with neurotypical individuals on the EQ. This suggests that the negative correlation between EQ and autistic traits is stronger than the positive correlation between SQ and autism, which is supported by more detailed analysis in the Greenberg et al. (2018) study.

1.6.1.2 Effects of gender

There are several studies which have identified gender differences in systemizing and empathizing quotients, with males scoring more highly on systemizing measures and females more highly on empathizing measures (Auyeung & Wheelwright, 2009; Billington et al., 2007; Bressan, 2018; Carroll & Chiew, 2006; Cook & Saucier, 2010; Ling et al., 2009; Wakabayashi et al., 2006; Wheelwright & Baron-Cohen, 2006; Wright & Skagerberg, 2012). This pattern is slightly different in autistic groups, where both males and females are more likely to score higher on the systemizing quotient (Greenberg et al., 2018). Autistic individuals have therefore been described as having an 'extreme male brain' (Baron-Cohen, 2002). However, as this term may lead to incorrect assumptions about the incidence of autism in females, or the variability within the male population in terms of systemizing-empathizing, this thesis will refer to individuals at the extremes of the continuum as 'hyper-systemizers' and 'hyper-empathizers' (Baron-Cohen et al., 2009).

Consistent with the gender comparison in adults, children's responses reveal higher systemizing scores in males, and higher empathizing scores in females (Auyeung & Wheelwright, 2009; Escovar et al., 2016; Wakabayashi, 2013). The difference between genders tends to be greater on the empathizing measure than on the sytemizing measure.

1.6.1.3 Effects of general cognitive abilities

Studies have found no association between IQ and systemizing and empathizing (Lai et al., 2012; Ling et al., 2009), but some associations with spatial abilities have been identified. A study with 88 neurotypical adults found that both systemizing and empathizing predicted performance on a 3D mental rotation task, however systemizing did not significantly relate to mental rotation scores directly and empathizing was a greater (negative) predictor in the regression model (Cook &

Saucier, 2010). Other studies with adults have shown that mental rotation did indeed associate with SQ scores (Brosnan, Daggar, & Collomosse, 2010; Ling et al., 2009), although one study identified a significant relationship only in the non-rotational aspects of the task such as encoding and comparing two stimuli (Brosnan et al., 2010). This was rather surprising, as there would appear to be a large overlap between systemizing and mental rotation ability; Baron-Cohen (2008) suggested that mental rotation tasks involve systemizing by considering the transformation of each feature in the stimulus and making predictions according to common rules. The inconsistent findings may indicate that it is not translation abilities that relate to systemizing, but a focus on the details enabling accurate mental representations to be created. This is supported by a study finding a negative relationship between systemizing and visual illusions in an adult sample, which is suggestive of a focus on the relevant details of the task whilst ignoring the distracting context (Walter, 2009). Both findings would require a featural comparison of the two presented stimuli to ascertain whether they were the same or different.

1.6.2 Development of systemizing / empathizing

Changes in responses to systemizing and empathizing across development have only been examined in one study. In a sample of 6- to 15-year-olds, there was no correlation between systemizing or empathizing and age, and no interaction between age group and gender (Wakabayashi, 2013). This suggests that these questionnaires are measuring stable attributes already present in 6-year-olds, rather than processes that develop throughout childhood and adolescence.

1.6.3 Relations with other whole-part constructs

One of the features of systemizing traits is the ability to break concepts or processes into smaller components in order to gain a better understanding of their unique features (Baron-Cohen, 2009). It is perhaps unsurprising then that studies with adults have found a positive correlation between SQ and local processing (Billington et al., 2008). Positive associations have also been revealed between the SQ and EFT in neurotypical groups (Billington et al., 2007) and in autistic groups (Brosnan et al., 2012), while a negative association between EQ and the EFT has been identified in a sample of undergraduates (Brosnan et al., 2011). There have been no further studies carried out which directly compare performance on the SQ with other whole-part measures.

1.6.4 SQ / EQ and mathematics / science

There are a number of studies identifying systemizing and empathizing differences at a subject level at university. Higher SQ scores have been identified amongst science students, whereas humanities students score more highly on the EQ (Billington et al., 2007; Focquaert et al., 2007; Kidron, Kaganovskiy, & Baron-Cohen, 2018; Svedholm-Häkkinen & Lindeman, 2016; Wakabayashi et al., 2006). This association varied still further depending on the specific scientific courses (Focquaert et al., 2007; Wheelwright & Baron-Cohen, 2006), suggestive of unequal systemizing demands across science subjects; for example, students on engineering and physics courses revealed greater systemizing traits than those on chemistry and mathematics courses. A recent study with adults has found that systemizing predicted performance on mathematical tasks, and that in fact, it was a better predictor of mathematics ability than gender (Bressan, 2018). Together, these studies support the notion that the ability to identify patterns and follow a rule-based system are predictive of better mathematics and science ability.

In a study with mainly female psychology students, Morsanyi et al. (2012) found no association between SQ and mathematics, nor between SQ and a systemizing measure (Bennett mechanical comprehension test) after controlling for spatial thinking styles. This appears to contradict other research, however there are possible explanations. The high proportion of females in the sample may result in differing results from a more mixed sample. The mathematics task tested factual knowledge which may not rely on a systemizing ability to the same extent as problem-solving tasks (Bressan, 2018). Additionally, spatial abilities may be a key component of both systemizing and mathematics, resulting in the lack of association after controlling for spatial thinking. Performance on a mental rotation spatial measure has been found to correlate with SQ scores in university students (Ling et al., 2009), adding some support for this suggestion.

A study with neurotypical children aged 7 to 12 years revealed no association between SQ and calculation, but a positive association between SQ and problem-solving. However, SQ did not independently predict performance on the problem-solving task once IQ and reading ability were accounted for. In contrast, EQ had a negative association with calculation and remained a significant predictor after accounting for the domain-general abilities, but it had no relationship with problem-solving (Escovar et al., 2016). This relationship between EQ and calculation performance may result from reacting to classroom distractions, as social awareness was found to mediate the relationship, particularly in girls. As this was the first study to compare EQ with mathematics performance, this association warrants further investigation.

1.6.5 Summary

The systemizing-empathizing theory combines an individual's drive to isolate details in order to understand rule-based systems (systemizing), and their emotional intelligence (empathizing). Autistic individuals are often categorised as hyper-systemizers on account of their high systemizing score and low empathizing score. Even though there is no visuospatial element of the SQ measure, studies have identified associations between systemizing and both local processing and FI.

Systemizing and empathizing generally do not correlate with each other, which suggests they are independent measures. Patterns of responses remain stable across childhood, and gender differences indicate a larger female advantage on empathizing than the male advantage on systemizing.

1.7 Thesis overview

Mathematics and science achievement have an influence on an individual's life outcomes, as well as making a significant contribution to the national economy. However, STEM industries have identified falling levels of engagement and skills in these subjects. There is a substantial body of evidence identifying an association between IQ and EF and both mathematics and science achievement, although the latter to a lesser extent. However, there is also evidence to suggest a relationship with whole-part tasks and mathematics and science. Firstly, although a direct association between global / local processing and mathematics and science has not been explored as yet, there is some evidence to suggest that autistic individuals, who have atypical responses to whole-part tasks, are more likely to work in mathematical or scientific fields. This association is fairly weak however, partly due to the variability in global and local response patterns and mathematics ability in the autistic population. Secondly, FI positively associates with mathematics and science, although it is unclear the extent to which this may be due to overlapping IQ and EF processes. FI is often a relative strength in autistic individuals which may result from reduced interference of the Gestalt or an enhanced detail-focus. Thirdly, systemizing positively associates with involvement in mathematics and science professions and academic courses. Autistic populations often score highly on the SQ and poorly on the EQ, which again indicates a preference for analysing and focussing on the detailed level. Combined, these separate strands of evidence indicate a potential association between whole-part tasks and mathematics and science, which warrants further exploration. Hence, these gaps in the literature have fed into the research questions for my thesis.

The main difficulty in drawing firm conclusions and predictions about these relationships is the variability between studies due to differing participant groups, stimuli, and tasks parameters. Further, even when associations with mathematics and science have been identified, there has rarely been a consideration of mechanistic processes or an acknowledgement of common underlying factors which may explain the relationships. This is particularly important where similar behaviour on a task may be due to different underlying mechanisms, especially in respect to differences between the neurotypical and autistic populations. For example, in the EFT, neurotypical participants may need to overcome the influence of the distracting context through inhibitory control, focussed attention, or breaking the whole into parts, while autistic participants, who may be less distracted by the whole, reveal comparatively greater activation in the visual cortex (Booth, 2006; Lee et al., 2007; Manjaly et al., 2007). Although there is some research which has identified an association between FI and academic achievement, it is unclear what is driving this, and it is possible that tasks are in fact simply reflecting the common IQ and EF elements in both types of task. Further, much of the current research has been undertaken with adult participants, and therefore a gap in our understanding is how the relationship between visual perception and mathematics and science changes over development. Clearly then, there are significant gaps in our understanding which has resulted in a lack of clarity about whole-part visual perception itself, as well as its relevance to other, more cognitive, domains.

The overarching aim of this thesis is to examine the associations between two visual perception constructs (global and local processing, and FI), as well as the construct of systemizing / empathizing, and mathematics and science achievement in 5-10 year olds. This age range encompasses the UK school Years 1, 3 and 5, and therefore will allow for any implications to be applicable to key stage 1 and 2 children at primary schools. It also covers a key period of development of domain-general processes and mathematical understanding, which means that we may expect associations to vary over this time period. To gain a clearer understanding about visual perceptual development across this age range, **Chapter 2** uses cross-sectional and longitudinal data to examine global and local processing and FI development. This chapter will also examine systemizing and empathizing development using cross-sectional data. These data will then be used to assess the extent to which tasks within each construct reveal common factors, and also any associations in performance across constructs. A more nuanced examination of the factors contributing to performance on FI tasks can be found in **Chapter 3**. This chapter considers the overlapping variance of IQ and EF in FI achievement through a behavioural study and a study using eye-tracking technology. The focus of the thesis then shifts towards the academic tasks. The developmental patterns of performance on the mathematics and science tasks and their relation to IQ and EF is examined in **Chapter 4**. Finally, the associations between whole-part performance and mathematics and science achievement is examined in **Chapter 5**. The discussion in **Chapter 6** will summarise the key findings, discuss limitations of the studies, and identify potential opportunities for future research. In sum, this thesis provides a systematic examination of global and local processing development and associations with other whole-part theories in neurotypical children, as well as an investigation into their associations with mathematics and science achievement.

The data in this thesis were collected with two participant samples: Sample 1 completed a behavioural study with tasks undertaken at two time-points, and sample 2 completed a behavioural and eye-tracking study. Data from sample 1 feature in all four experimental chapters, and data from sample 2 appear in **Chapters 3, 4,** and **5**. Details about each task will be included in the Method section of the Chapter in which they first appear, and thereafter they will only be briefly summarised. A summary of the tasks included in each Chapter can be found in each Method section.

Chapter 2: Developmental changes in whole-part processing

2.1 Introduction

Research reviewed in **Chapter 1** revealed key differences between the three whole-part constructs which form the focus of this thesis. In each construct, a number of factors have been identified as contributing to response variation on different tasks. One of these factors is age. This chapter examines developmental changes over childhood in first global and local processing, then field independence, and finally systemizing, empathizing, and autism traits. Comparisons are then made across the whole-part tasks to identify any variance associated with common underlying processes.

2.1.1 Changes over development

The vast majority of cross-sectional studies have revealed a local-to-global shift in processing advantage during childhood (e.g. Dukette & Stiles, 1996; Harrison & Stiles, 2009; Poirel, Mellet, Houdé, & Pineau, 2008). However, much of the research using Navon stimuli has been conducted with adult samples (e.g. Poirel, Pineau, & Mellet, 2006; Roalf, Lowery, & Turetsky, 2006) or has compared responses between neurotypical and atypical children and adolescents (e.g. Rinehart, Bradshaw, Moss, Brereton, & Tonge, 2000; Wang, Mottron, Peng, Berthiaume, & Dawson, 2007). There is limited research which has investigated neurotypical responses to Navon tasks from a developmental perspective, or where children have completed a battery of global and local tasks with varied attentional demands. Crucially, there are no prior longitudinal studies using Navon stimuli to assess developmental changes.

Developmental changes in disembedding have been identified in cross-sectional studies (Amador-Campos & Kirchner-Nebot, 1997; Cairns et al., 1985; Flexer & Roberge, 1980) as well as a longitudinal study (Flexer & Roberge, 1983), and reveal an improvement in field independence (FI) with age up to approximately 17 years of age (Amador-Campos & Kirchner-Nebot, 1997; Witkin et al., 1967). Developmental changes in segmenting have been examined cross-sectionally in children and adolescents using the Block Design task (Akshoomoff & Stiles, 1996; Booth, 2006), but the only longitudinal study using the Block Design task was conducted with adult participants (Rönnlund & Nilsson, 2006). As stated, FI studies have largely examined responses on the Embedded Figures Test (EFT) and its variants; there have been no prior studies with children comparing FI development across tasks.

Systemizing and empathizing has often been evaluated with large, usually adult, samples (e.g. Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003; Greenberg, Warrier, Allison, & Baron-Cohen, 2018; Wakabayashi et al., 2006). One study has looked at cross-sectional differences in response across 6- to 15-year-olds, and this identified no developmental changes in either systemizing or empathizing (Escovar et al., 2016).

To summarise, research to date examining developmental changes in the three whole-part constructs above, over childhood, has been reasonably sparse, and certainly few studies have compared development in a number of tasks within each construct.

2.1.2 Comparisons across constructs

The uniqueness of each of the whole-part constructs is apparent in studies which have presented a number of visual tasks to the same sample and found little correlation in performance across the tasks. Milne and Szczerbinski (2009) collected responses from neurotypical adults on 14 tasks designed to measure global and local processing, FI, and Gestalt processes. The resultant factor analysis revealed only one significant factor, which had significant loadings from the Block Design task, EFT accuracy and response time (RT), as well as substantial loadings from three other tasks which involved copying, finding a hidden pattern, and spotting differences between stimuli. This factor was therefore considered to represent 'disembedding' as all the tasks involved identifying and separating a target from its complex array. A more recent study found little overlap in response patterns between a Navon selective attention task and the EFT (Chamberlain et al., 2017). In a study with university students, there were no significant correlations between selective and divided Navon tasks, although there was a trend association between global advantage on the shape and letters divided attention Navon tasks (Pletzer, Scheuringer, & Scherndl, 2017). A further study which tested adults on a selective attention Navon task, a free choice paper-and-pencil Navon task, and a global / local hybrid face-matching task, found no inter-task correlation. The factor analysis revealed three factors which simply represented the three distinct tasks (Dale & Arnell, 2013). This pattern is not specific to adult participants as illustrated by a smaller-scale study with children and adolescents which also revealed a lack of correlations between the varied global and local measures (Van Eylen, Boets, Steyaert, Wagemans, & Noens, 2018). In sum, studies which have used a range of tasks with the same participants, have not been able to identify behaviour synonymous with common underlying whole-part processes across task. This lack of association between tasks which purport to measure common factors, may indicate that there are many processes involved in the tasks and that the factor of interest is not the most prominent. Equally, it may suggest that the grouping and segmentation processes are implemented differently in each task, such that no commonality is identified when comparisons are made across tasks (De-Wit & Wagemans, 2015). By strengthening our understanding of the unique features of each construct, it may be possible to better explain the lack of consistency in responses across the tasks.

2.1.3 Current study

There is a gap in our current understanding of whole-part development, particularly with regards to comparisons between tasks designed to measure similar processes. The studies examined in this chapter aim to address this gap, and establish a clear starting point, which will then enable comparisons of these constructs with mathematics and science in neurotypical children to be made (see **Chapter 5**).

As discussed in **Chapter 1**, there are many factors which can modulate responses on global and local tasks, including stimuli design where, for example, larger, sparser local elements reduce global advantage (Dukette & Stiles, 2001; Harrison & Stiles, 2009; Kimchi, 2015), as well as task design where, for example, responses vary according to the attentional demands of the task (Dale & Arnell, 2013; Van Eylen et al., 2018). Here, global and local processing was studied across three types of Navon tasks in order to evaluate and compare developmental changes with varying attentional task demands when stimuli design remained constant. In this way, it was possible to evaluate the impact of domain-general factors associated with completing the tasks, rather than observed effects being due to visual differences in the stimuli which are likely to be driven by automatic Gestalt processes (**Section 1.5**). Similarly, FI will be measured in two tasks, enabling comparisons to be made between task processes associated with disembedding and those with segmenting. Both sets of data were studied across three different age groups, using a combination of cross-sectional and longitudinal data. Systemizing, empathizing, and autism traits were calculated from parental questionnaire responses from one time point only.

The first aim was to determine how performance on each task varies in neurotypical children between the ages of 5 and 10 years. Based on the studies discussed in **Chapter 1**, it was hypothesised that there would be a shift from a local to global advantage over time in each of the Navon tasks, and increasing FI with increasing age. In contrast, there were unlikely to be any developmental changes identified in the responses to the questionnaires about systemizing, empathizing, and autism traits.

Second, the data were used to compare performance across tasks within each whole-part construct, to understand the impact of different task demands on responses. The hypothesis here was that different patterns of global and local responses and their changes during development would be observed between the Navon tasks because of the impact of task design on responses. In contrast, it was expected that responses on the two FI tasks would reveal a more consistent developmental change from less to more field independent, due to the development of domain-general factors thought to contribute to success on these tasks (**Section 1.4.2.2**). This may also result in significant correlations between the two tasks. It was expected that systemizing and empathizing would not significantly associate with each other, as previous studies with children revealed these measures to be distinct from each other. However, systemizing was expected to positively associate with autism traits while empathizing would negatively associate with autism traits, reflecting the current literature. This study will additionally be an opportunity to test whether associations revealed in larger-scale studies with neurotypical and autistic children (Auyeung & Wheelwright, 2009; Wakabayashi, 2013) can be replicated in a smaller, neurotypical sample of children.

Finally, comparisons were made across all of the tasks to determine whether common factors relating to whole-part processing could be identified. Based on cross-task studies undertaken primarily with adults (Chamberlain et al., 2017; Dale & Arnell, 2013; Milne & Szczerbinski, 2009), the expectation was that common underlying whole-part factors would not be identifiable due to the interference of additional processes involved in each of the tasks, or because the tasks measure abilities and processes too distinct from each other.

2.2 Method

2.2.1 Participants

The study involved participants from sample 1 and was approved by the local ethics committee. It included cross-sectional data, as well as longitudinal data collected two years after the initial study. Participants were recruited from a single community primary school. Parents and carers were given the opportunity to opt-out at the first time-point (T1) and to opt-in at the second time-point (T2) two years later. All children provided verbal consent to take part in the study at each time-point. There were marginal gender differences between Years at T1, $\chi^2(2) = 5.75$, $p = .056$, and significant gender differences between Years at T2, $\chi^2(1) = 4.29$, $p = .038$. This reflected the gender distribution in the school itself. The inclusion criterion was that participants did not have a statement of special educational needs and disabilities (SEND). Sixty-two percent of Year 1 children undertook the longitudinal follow-up when they were in Year 3, while 71% of Year 3 children undertook the longitudinal follow-up when they were in Year 5 (**Table 2.1**).

Table 2.1: *Participant demographics.*

	Year 1	**Year 3**	**Year 3**	**Year 5**	**Year 5**
	T1	**T2**	**T1**	**T2**	**T1**
Number of participants	45[a]	28[a]	45[b]	32[b]	45
Males : females	16 : 29	10 : 18	27 : 18	20 : 12	24 : 21
Mean age in months (SD)	68.6 (3.3)	92.2 (3.6)	92.8 (3.5)	114.2 (3.6)	115.4 (3.6)
Mean age in years; months	5; 8	7; 6	7; 6	9; 6	9; 7
Raven's CPM mean (mean percentile)	18 (57)		26 (72)		29 (50)
BPVS mean (mean percentile)	85 (52)		105 (52)		125 (34)

[a] Longitudinal cohort in Year 1 at T1 and Year 3 at T2
[b] Longitudinal cohort in Year 3 at T1 and Year 5 at T2
Notes: CPM: Coloured Progressive Matrices; BPVS: British Picture Vocabulary Scale

Non-verbal and verbal abilities were measured using standardised tests at T1. Non-verbal IQ was measured using Raven's Coloured Progressive Matrices (Raven, Raven and Court, 1998) and receptive vocabulary was measured using the British Picture Vocabulary Scale III (Dunn, Dunn, Styles and Sewell, 2009). Both were administered following the standard guidelines, and raw scores were recorded for each (see **Section 3.2.4** for further details about administration and recording). This enabled comparisons to be carried out to determine whether the longitudinal sub-groups differed at T1 from the whole sample in verbal or non-verbal IQ scores. No significant differences were observed using independent t-tests (Year 1: verbal IQ, $t(71) = -.181$, $p = .857$; non-verbal IQ, $t(71) = .468$, $p = .641$; Year 3: verbal IQ, $t(75) = -.251$, $p = .802$; non-verbal IQ, $t(75) = -.284$, $p = .777$).

2.2.2 Summary of tasks

A summary of the tasks completed by participants can be found in **Table 2.2.**

Table 2.2: *Summary of tasks and measures.*

Study	Measure	Task
Study 1	Global / local responses	Free choice Navon task
		Selective attention Navon task
		Divided attention Navon task (Big-small)
		Divided attention Navon task (Yes-no)
	Response switching	Big-small flying saucer task
	Disembedding	Children's Embedded Figures Test (CEFT)
	Segmenting	Design Organisation Test (DOT)
	Systemizing	Systemizing quotient for children (SQ-C)
	Empathizing	Empathizing quotient for children (EG-C)
	Autistic traits	Autism quotient (AQ-10)
	General intelligence	Visuospatial: Raven's Coloured Progressive Matrices
		Verbal: British Picture Vocabulary Scale 3 (BPVS-3)

2.2.3 Navon tasks

2.2.3.1 Navon stimuli

Three types of Navon task were included in the test battery: A free choice task, a selective attention task, and divided attention tasks. All tasks used simple geometric shapes (squares, circles, triangles, and trapeziums), rather than letters or digits, which minimised the effect of differential familiarity with the alphabet or numbers across individuals and across Years (Dukette & Stiles, 1996). Each global and local stimulus occupied a square-shaped area. The global stimulus was 30 mm wide, which subtended a visual angle of 3.44°, based on an unrestrained viewing distance of 50 cm. The local stimulus size was 5 mm wide, which subtended a visual angle of 0.57°. These sizes were in line with other research which have used Navon stimuli (Hayward et al., 2012; Katagiri et al., 2013; Mondloch et al., 2003; Scherf et al., 2009, 2008; Volberg & Hübner, 2007; Wang et al., 2007) and have been shown to elicit faster global than local responses in neurotypical adults (Hayward et al., 2012; Wang et al., 2007).

The Navon stimuli were presented via MATLAB (R2010b), using the Cogent toolbox (http://www.vislab.ucl.ac.uk/cogent_2000.php), on a 12-inch Dell laptop. Responses were made using laptop keys which were identified using stickers.

Participants responded with a key press using their left hand to press the left key and their right hand to press the right key. The responses associated with each key were counterbalanced across participants within each age group. To distinguish the separate Navon tasks, different pale background colours were used. Although some studies have found that background colour affects performance on Navon tasks (e.g., Michimata, Okubo, & Mugishima, 1999), others have found no significant effects on performance (e.g., Dore, Dumani, Wyatt, & Shepherd, 2018). The background colours used in the current study are lighter and paler than the region of the visible light spectrum suggested to influence responses (Dore et al., 2018). The inter-trial interval was 500 ms in all tasks. Sequences included equal numbers of different stimuli types. For details about the task instructions, see the **Appendix**.

2.2.3.2 Free choice Navon task

An incongruent target stimulus was presented above a line. Underneath the line were two further incongruent stimuli; one of which matched the target at the local level while the other matched the target at the global level (**Figure 2.1**). Children were asked to select the figure that was most similar to the target shape, forcing them to choose a response consistent with a focus either on the global or local similarities. There was no time limit, ensuring all children had the opportunity to respond according to their preferred level. There were only four trials, which minimised the effect of priming response choices (Koldewyn et al., 2013). The position of the global and local choices was counterbalanced across the four trials, with the same order of the four trials for all participants. The number of matches at each level were recorded.

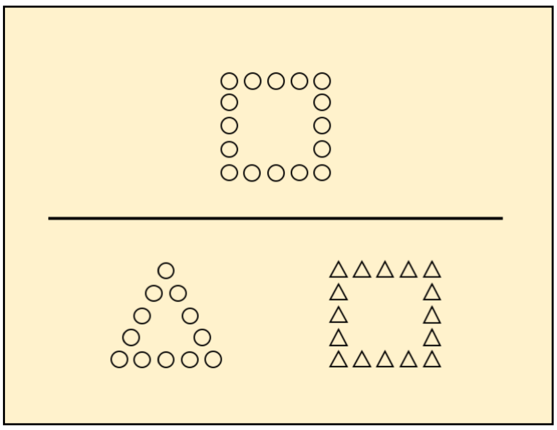


Figure 2.1: *Example trial from the free choice Navon task. Here, the bottom left-hand figure matches the top target at the local level, while the right-hand figure matches the target at the global level.*

2.2.3.3 Selective attention Navon task

Participants were directed to attend to either the global or local level for one block of 36 trials each. There were six different stimuli, each presented six times in each block. Participants had to indicate whether the shape at the attended level was a triangle or a square. Participants were given a maximum of 2000 ms to respond, which is longer than some studies due to the younger sample tested in the present study. Each block had an equal number of congruent stimuli (where there was no interference), incongruent stimuli (where a triangle was presented at one level and a square at the other), and neutral stimuli (where a triangle or square was presented at one level, and a circle at the other) (**Figure 2.2**). The circle shape was neutral as it was not associated with a key press whereas there was possible direct interference from the non-attended level in the incongruent stimuli. The stimuli within each block were presented in a fixed pseudo-random sequence, with no repeats of the same response more than three times in succession. The pseudo-random sequence provided control over the order of stimuli whilst ensuring responses were not predictable.

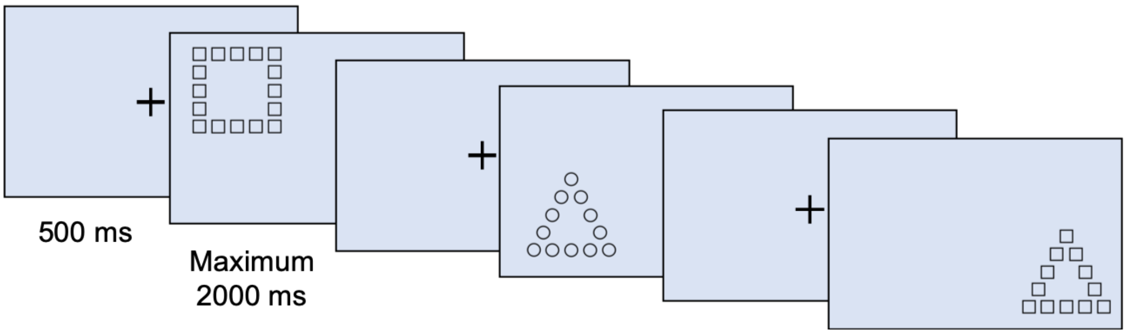


Figure 2.2: *Example trials from the global block of the selective attention task, where the responses would be square (congruent stimulus), then triangle (neutral stimulus), and then triangle (incongruent stimulus). The stimuli have all been enlarged relative to the crosshair here, so the detail can be seen more easily.*

As participants were focussing on one hierarchical level at a time, it is possible that the attentional field would narrow or widen according to the task demands, and only the relevant level would be actively perceived (Gerlach & Krumborg, 2014; Wilkinson, Halligan, Marshall, Büchel, & Dolan, 2001). To minimise this, the stimuli moved around the screen in a pseudo-random order, such that an equal number of stimuli appeared in each of four locations (**Figure 2.2**). The offset was 140 pixels on both the y and x axes, equivalent to a distance from the central cross of 6°. RT and accuracy data were recorded.

2.2.3.4 Divided attention Navon tasks

Participants had to identify whether a circle was featured at the global or local level, and had a maximum of 3000 ms to respond. The instruction was simplified for the children by asking whether the circle was the big or small shape (big-small divided attention task). This meant that both the global and local levels could be examined, and therefore attention was likely divided between the hierarchical levels. There were two blocks of 37 trials separated by a short break. To assess possible priming effects and switching cost (Katagiri et al., 2013; Wilkinson, Halligan, Marshall, Büchel, & Dolan, 2001), responses where the target appeared at the same level were presented in groups of two, three, or four consecutive stimuli (**Figure 2.3A**). At the end of each sequence of repeat trials, there was a switch trial where the response switched from global-to-local, or local-to-global. Within each block, there were six global-to-local and six local-to-global switches. The trials were presented in a pseudo-random order in each block, with all participants presented with the same order of stimuli, which allowed control over the trial order. RT and accuracy data were recorded.

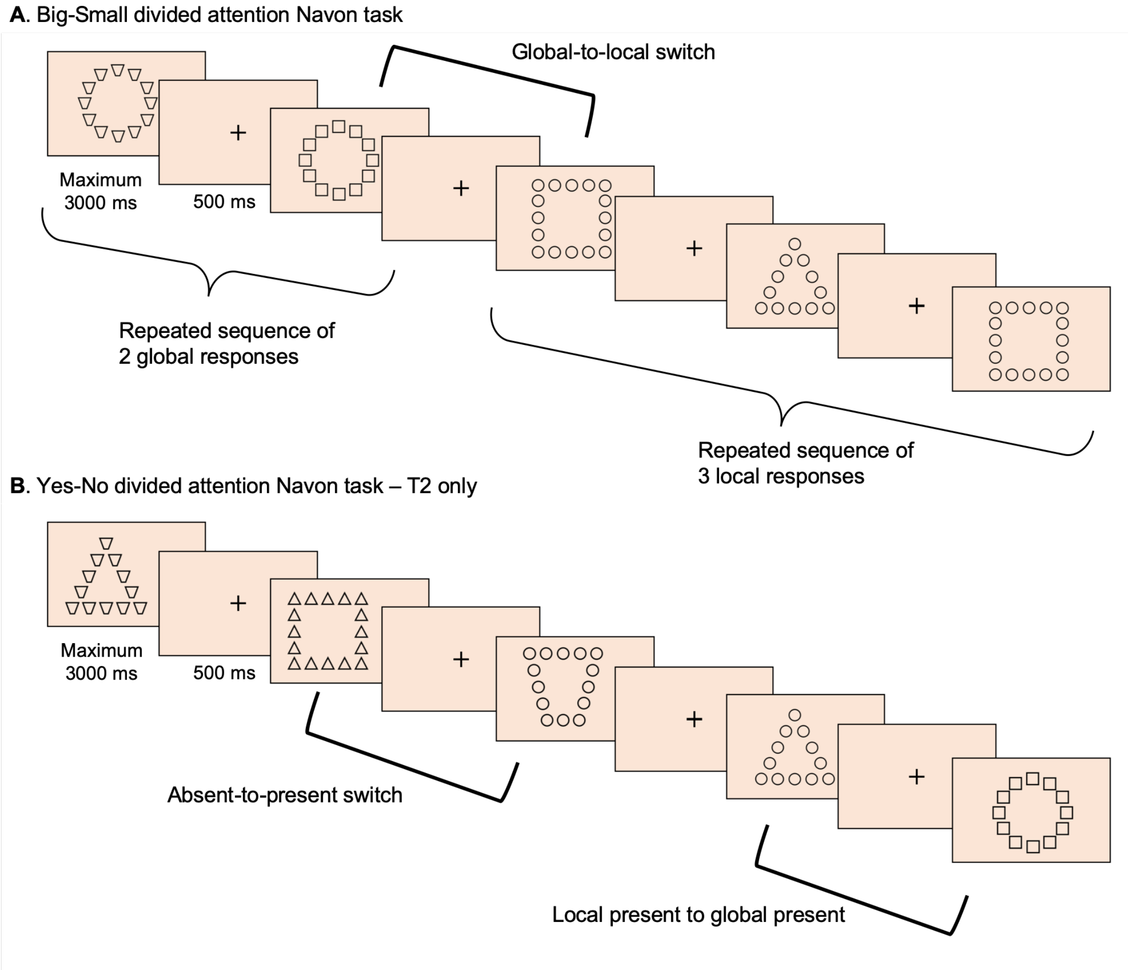


Figure 2.3: *Example trials from the divided attention Navon tasks. The stimuli in have been enlarged relative to the crosshair here, so the detail can be seen more easily.* ***A****. A sequence of repeated trials and a global-to-local switch trial from the big-small divided attention Navon task.* ***B****. A sequence of trials from the yes-no divided attention task introduced at Time 2.*

Although the original divided attention task was included to measure responses when children's attention was split across both levels, it is possible that children were inferring the presence or absence of a circle while attending to a single level. To investigate this possibility, a divided attention task variant was added at T2. In this task, participants were asked to identify whether a circle was present or absent, increasing the need for both global and local level to be examined before responding (**Figure 2.3B**). The instruction was simplified for the children by asking whether there was a circle, with response buttons indicating 'yes' or 'no' (yes-no divided attention task). There were two blocks of 36 trials separated by a short break. In half of the trials the target was absent; in the other half the target was present either at the local level or at the global levels, with equal frequency of both trial types. Participants had up to 3000 ms to respond and the trials were presented in a pseudo-random order in each block. All participants were presented with the same order of stimuli in each block. RT and accuracy data were recorded. The two divided attention tasks were separated in the battery at T2 to reduce confusion arising from the different task instructions.

2.2.3.5 Response switching task

This task measured the speed and accuracy of responses when participants had to switch between big and small stimuli without interference from different hierarchical levels. By comparing differences in performance between this task and the big-small divided attention task, a true indication of the cost of inter-level interference and switching between global and local hierarchical levels can be revealed.

Participants at T1 and T2 were asked to respond according to whether a flying saucer image was big or small. The images were selected to be distinct from the stimuli used in other tasks, and to be equally plausible when presented as a big or small image. The design of the task replicated that of the big-small divided attention task, with a series of repeated responses followed by a switch trial of either big-to-small or small-to-big (**Figure 2.4**). There were two blocks of 37 trials each, with three differently-coloured stimuli repeated 12 times in each block – six times as a large image and six times as a small image. As size is a relative characteristic, the big and small stimuli were presented together during the instruction stage so that the experimental trials could be immediately recognisable as being either big or small.

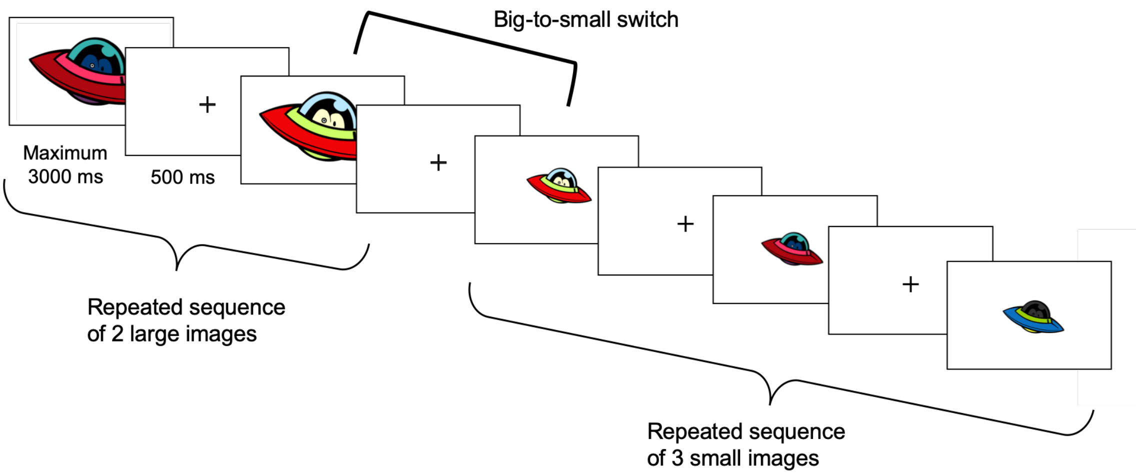


Figure 2.4: *An example sequence of repeated trials and a big-to-small switch trial from the response switching task. The stimuli have been enlarged relative to the crosshair here, so the detail can be seen more easily.*

2.2.4 Field independence tasks

2.2.4.1 Children's Embedded Figures Test (CEFT)

The Children's Embedded Figures Test (CEFT; Karp & Konstadt, 1963) was used to measure disembedding. The test comprises 25 colourful pictures which are generally recognisable to children, including a clock, boat, and a pram. One of two target shapes is embedded within each stimulus; a tent (triangle) is the target shape in 11 trials, and a house is the target shape in 14 trials. Children were asked to find the target shape in the complex image (**Figure 2.5**).

Figure 2.5: *Practice trial example of the Children's Embedded Figures Test (CEFT) with the tent (triangle) target shape alongside the complex image where the tent shape is embedded within.*

In order for accuracy data to be comparable across all the participants, the same sub-set of 13 images was presented to all participants. The stimuli were selected according to accuracy data from a study in Spanish primary school children (Amador-Campos & Kirchner-Nebot, 1997). Mean accuracy from the Amador-Campos et al. (1997) study was averaged across the age groups of interest and sorted into a descending order. Every second stimulus was selected from the list, to ensure that the test was of a manageable length for all ages. A total of six tent stimuli (T1, T3, T7, T6, T10, T11) and seven house stimuli (H3, H6, H4, H10, H8, H7, H11) were presented, so the test included a mix of difficulty and of target shape.

Contrary to the manual instructions, the target shape was visible throughout the task to reduce the influence of WM (Booth, 2006; Huygelier et al., 2018). Participants indicated that they had found the embedded shape by pointing to it. Both accuracy and RT were recorded, however only accuracy data were used in the analyses as children may have employed various response strategies, which could modulate their true RT. For example, some may respond based on their initial instinct, while others may re-check the stimulus before responding. Participants had a maximum of 30 seconds to respond, consistent with other studies with child participants (Pellicano et al., 2006). The maximum score was 13; scores were converted to percentage accuracy.

2.2.4.2 Design Organisation Test (DOT)

The Design Organisation Test (DOT; Killgore & Gogel, 2014) was used to measure segmenting. This is a pen-and-paper version of the Block Design subtest of the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997). Participants had to identify which numbered pieces fitted together to create a larger pattern, and then write those numbers in an empty grid (**Figure 2.6**). Based on a pilot study, participants were given three minutes to complete as much of the sheet as possible, rather than the two minutes allocated for adult participants (Killgore & Gogel, 2014). A mark was given for each correctly identified small square, with a maximum score of 56.

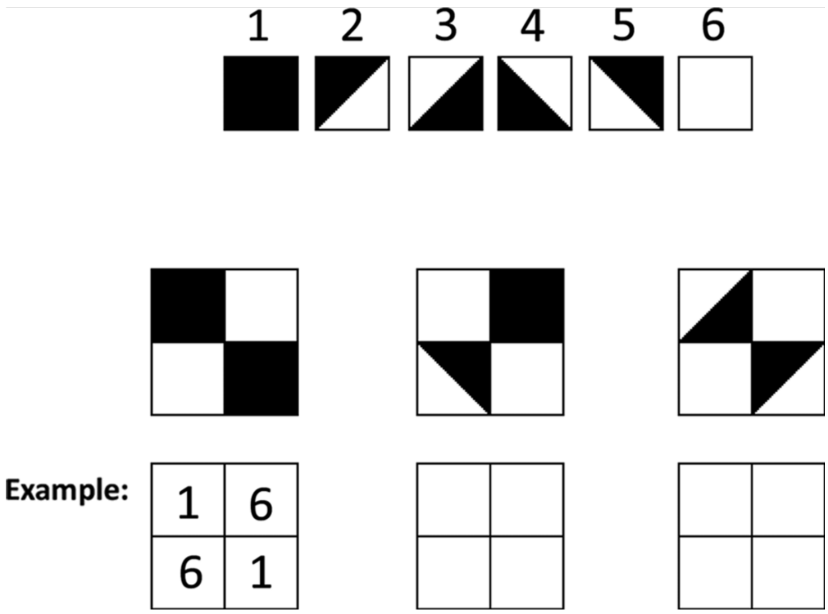


Figure 2.6: *Practice trials from the Design Organisation Test (DOT) with the six numbered option blocks, the pattern stimuli, and the response boxes.*

Children were asked to complete sheet A of the task, pattern by pattern. They worked across the page left to right, completing the five 2x2 patterns first, before moving on to the four 3x3 patterns, and were not permitted to segment the patterns by adding grid lines. This removed the opportunity for some children to achieve higher scores through the use of appropriately selected strategies, increasing the likelihood that observed variance was due to segmenting ability rather than strategy selection.

2.2.5 Systemizing and Empathizing Quotients

These data were collected at T1 only. The children's versions of the systemizing (SQ-C) and empathizing (EQ-C) quotients (Auyeung & Wheelwright, 2009) were combined with a short autism quotient (AQ-10) questionnaire designed to give an initial assessment about whether a child should be referred for diagnosis (Allison, Auyeung, & Baron-Cohen, 2012). The AQ-10 gave an indication of individual differences in autism traits. Combined, this created a single parental questionnaire with the SQ-C and EQ-C questions first and the AQ-10 questions at the end. There was nothing to distinguish the tasks from each other and the order of the questions was not changed. All the questionnaires used the same four-step Likert scale, where parents responded, 'definitely agree', 'slightly agree', 'slightly disagree', or 'definitely disagree'.

A raw score for systemizing, empathizing, and autism was calculated, along with a D-score (difference) to determine relative strengths between systemizing and empathizing (Auyeung & Wheelwright, 2009; Goldenfeld, Baron-Cohen, Wheelwright, Ashwin, & Chakrabarti, 2007). The maximum score was 56 on the SQ-C, 54 on the EQ-C, and 10 on the AQ-10. The D-score was calculated as D-score = (S-Standardised – E-Standardised) / 2, with S-Standardised = (SQ – group mean SQ) / 56 and E-Standardised = (EQ – group mean EQ) / 54 (Auyeung & Wheelwright, 2009; Goldenfeld, Baron-Cohen, Wheelwright, Ashwin, & Chakrabarti, 2007).

2.2.6 Procedure

At both time-points, testing took place at the participants' primary school in quiet, well-lit rooms during lesson times. The data collected were part of a larger test battery which included maths and science misconception questions at both time points, plus executive functions (EF), mathematics and science tasks at T1 only (for further details, see **Chapters 3 and 4**). Children completed these tasks within several sessions lasting 30 minutes each. The order of the whole battery was counterbalanced within each age group, with children completing the activities in one of two sequences. The free choice task was always the first test performed to eliminate priming effects from other tasks, and the global and local selective attention blocks were separated in the test battery to avoid confusion. For the selective and divided attention Navon tasks, and the response switching task, the order of blocks was counterbalanced across participants within each age group.

Each task was explained to the participants immediately prior to them attempting the activity. Participants started with six practice trials for the selective and divided attention Navon tasks and the response switching task to ensure that they understood the instructions. If they scored less than four on the practice trials, they repeated the practice up to a maximum of five times. With the exception of one participant in the divided attention task, all participants passed the practice trials. Participants did not have practice trials for the free choice Navon task, as there were no correct or incorrect responses. Before commencing the CEFT, a series of practice cards were shown to participants to check they were able to match the simple shapes and could locate the shape when embedded within an image. There were also practice trials for the DOT, with a reduction in support given by the researcher across the three practice trials so the participants could independently complete the final practice stimulus.

2.2.7 Statistical analyses and outliers

2.2.7.1 Statistical analyses

Accuracy and RT data were analysed using a series of linear mixed models (LMMs) combining cross-sectional and longitudinal data. LMMs allow for differing numbers of repeated measures and time points between participants, such as instances where participants were tested at T1 but not at T2 (Molenberghs & Verbeke, 2001). It also allows for the fact that observations by the same individual at different time points are not independent (Shek & Ma, 2011; West, 2009). LMMs have been shown to be robust when data violate the normal distribution assumptions (Schielzeth et al., 2020). This was important here where a Kolmogorov-Smirnov test revealed some data were negatively skewed; for example, selective Navon task global accuracy, D(135) = .171, $p < .001$, and local accuracy, D(135) = .214, $p < .001$, as well as the divided Navon task global accuracy, D(134) = .226, $p < .001$, and local accuracy, D(134) = .235, $p < .001$. A compound symmetry covariance structure was used, which assumes constant variance and correlations between the time points, as the gap between each Year group was two years and the gap between each time point was also two years. Once the variables which are repeated measures for each participant have been specified, the between-subject and within-subject factors which form the model can be selected, with an opportunity to examine both fixed and random effects and interactions. For every LMM, all main effects and interactions between factors were modelled as fixed effects. Testing session was included in the models as a repeated measures variable but was not entered as a factor in the models, as the focus was on the effect of Year across both time points. Mean RTs included data from correct responses only. Estimated means (*M*) and standard error (*SE*) were reported in all analyses. All follow-up paired comparisons used Bonferroni-corrected alphas.

The data for each task were analysed separately, identifying distinct response patterns relating to gender and Year group. Additionally, the effect of congruency was examined in the selective attention task, and the cost of switching from one level to the other was examined in the divided attention task. A second divided attention task was added at T2 to further investigate potential sources of task-related differences in response. Correlational analyses were carried out to compare individual performance across the tasks.

To address any differences in speed-accuracy trade-offs across groups, the selective attention and divided attention tasks analyses were also run with linear integrated speed-accuracy scores (LISAS) (Vandierendonck, 2017, 2018). This integrates RT, proportion of errors (PE) and standard deviations (SD) of each individual participant for each measure, to create RT estimates which account for differences in accuracy using the following calculation: $LISAS = RT + ((SD_{RT} / SD_{PE}) \times PE)$. This calculation has been used to compute a combined speed and accuracy measure in a number of studies (e.g. Hasenäcker, Verra, & Schroeder, 2019; Pereg & Meiran, 2018; Ritchie & de Beeck, 2019; Schuch & Pütz, 2019).

2.2.7.2 Outliers

Outliers are datapoints that appear to be non-typical compared with the rest of the dataset. It is important to be aware of outliers and respond to them effectively, in order to reduce the possibility of Type I or Type II errors; in other words, to ensure hypotheses are not accepted or rejected incorrectly (Aguinis, Gottfredson, & Joo, 2013; Leys, Delacre, Mora, Lakens, & Ley, 2019). There may be a multitude of reasons why a datapoint has been recognised as an outlier, including an error in recording or administering the task, a loss of concentration by the participant, or a misunderstanding of the task goals by the participant. However, outliers may also the represent true performance of an individual on a task, and can be informative for the researcher by highlighting further moderators that may influence responses (Aguinis et al., 2013). Despite the importance of identifying and responding to outliers, there is relatively little clear guidance given to researchers in the literature, and there is wide variability in how outliers are dealt with, including removal, transformation of data, or replacing data with average datapoints (Leys et al., 2019; Pollet & Meij, 2017). Outliers which exist due to errors in recording should be corrected where possible, but otherwise be removed from the sample (Leys et al., 2019). If outliers are not due to recording errors or administration errors, it can be difficult to determine what sort of outliers they might be. However, if it is possible to identify outliers which are interesting (i.e. which may reveal further factors moderating a participant’s responses), they should be explored in greater depth (Aguinis et al., 2013). One method of responding to outliers which may be influential on the dataset can be to run the analyses both including and excluding the outlier to reveal the level of influence the outlier(s) is exerting on the analyses (Aguinis et al., 2013; Cousineau & Chartier, 2011; Leys et al., 2019; Pollet & Meij, 2017).

In this thesis, outliers will be identified as datapoints ± 3.29 standard deviations (*SD*) from the mean (Field, 2013). Most of the tasks included in this thesis can only be started after successful completion of practice trials. These practice trials aim to ensure that all participants understand the task goals and are able to demonstrate an ability to complete the task prior to commencement. Therefore, the responses on the task should reflect the participants' abilities on that task, but it is possible that loss of attention or additional cognitive processes which are not the target of the task, may also influence their responses. It can be difficult to explain why non-typical or outlier responses are present in a dataset; whether the datapoint represents a participant's true response on that task, or whether the measurement has been influenced by additional unidentified cognitive processes, or whether the task was too challenging for the participant. Simply removing the outliers may reduce the generalisability of any findings to the wider population. Throughout this thesis, if outliers have resulted from errors in recording, they will be excluded from the analyses. If not, the analyses will be run with and without the outliers. Any changes to main effects or interactions will be identified and reported so that the influence of the outliers on the analyses can be revealed, reducing the potential for Type I or Type II errors. In the free choice Navon task, high scores represent a global preference and low scores a local preference, therefore this data-set contains no outliers.

2.3 Results

All children attempted all tasks, however, one Year 1 male participant failed the practice session of the divided attention Navon task at T1 and was not included in the analysis of this task.

2.3.1 Global and local processing development

For each task, the data collected at T1 and T2 were plotted on the same figure for comparison purposes (**Figures 2.7 A-E**). It should be noted that those participating in the longitudinal study therefore have their T1 data plotted twice – once as part of the cross-sectional mean, and once as T1 of the longitudinal data. There was no evidence of practice effects; the T2 responses in Year 3 were not more accurate or quicker than the T1 Year 3 responses.

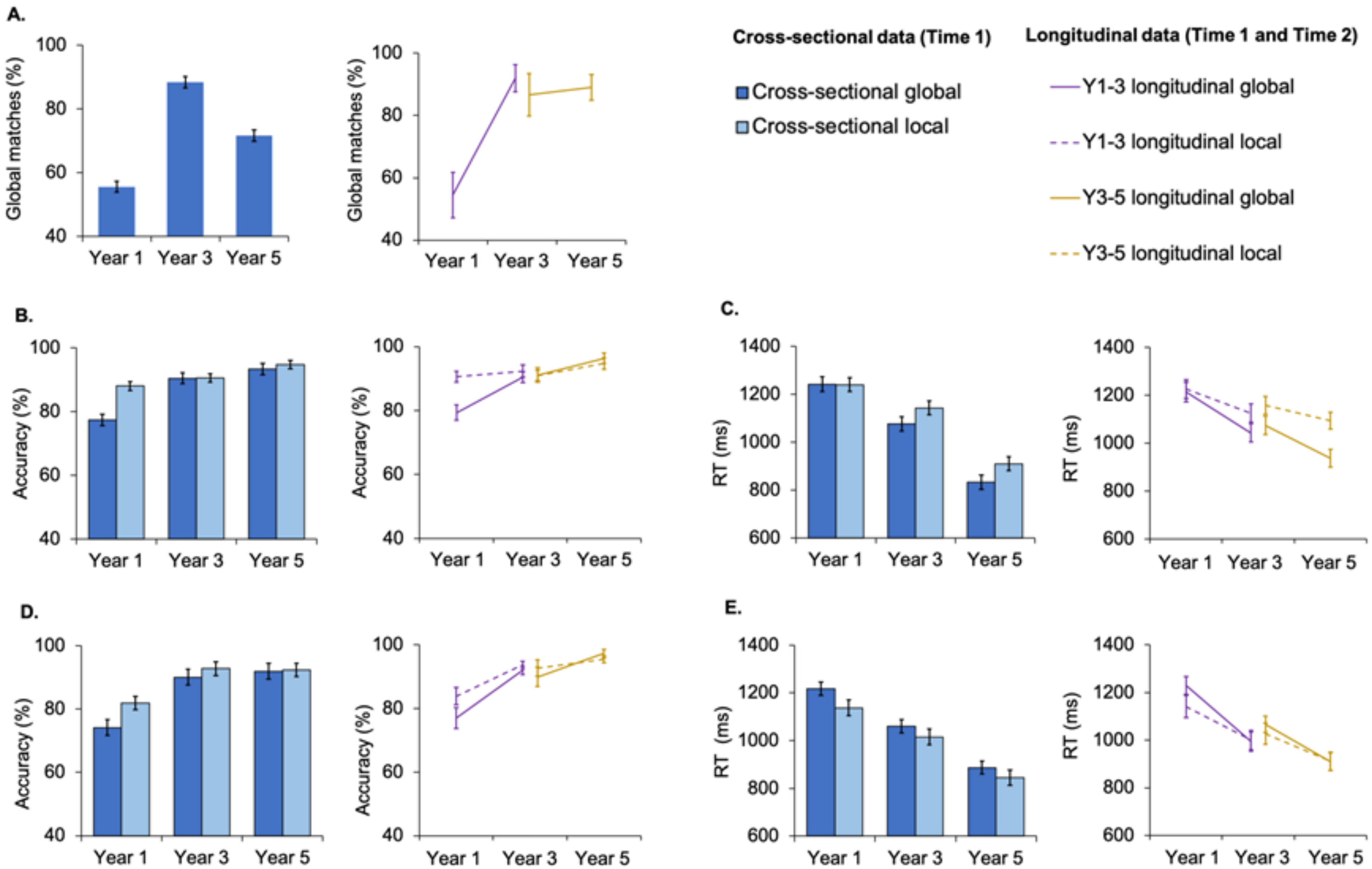


Figure 2.7: *Mean cross-sectional (bar charts) and longitudinal (line charts) performance on each Navon task. All data are plotted as a function of Year.* ***A.*** *Free choice Navon.* ***B.*** *Selective attention Navon task accuracy.* ***C.*** *Selective attention Navon task response time (RT).* ***D.*** *Big-Small divided attention Navon task accuracy.* ***E.*** *Big-Small divided attention Navon task RT. All results are plotted as M ± SE.*

2.3.1.1 Free choice Navon task

A linear mixed model (LMM) was carried out with the proportion of global responses as the dependent variable (DV) with Year and gender as between-subject factors. There was no main effect of gender (*p* = .083). There was a significant main effect of Year, *F*(2,126.2) = 16.13, *p* < .001, with higher proportions of global matches in Years 3 and 5 than in Year 1 (*p*'s < .001), but no difference between Years 3 and 5 (*p* = .239) (**Figure 2.8**).

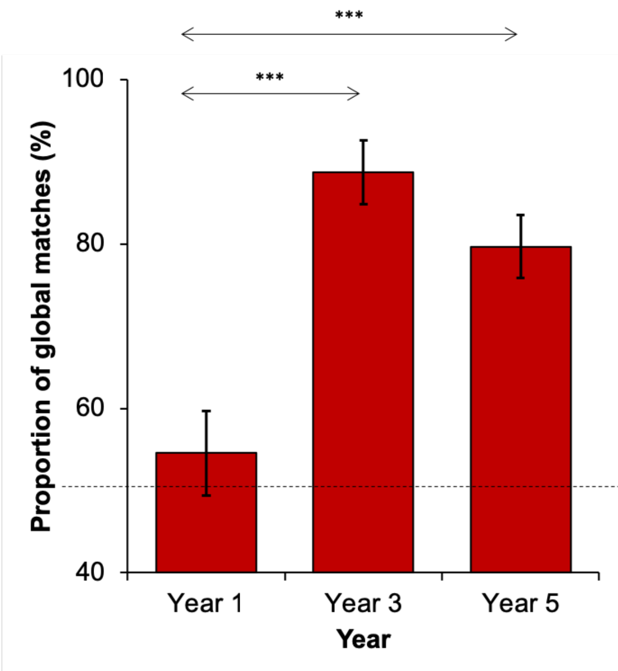


Figure 2.8: *Proportion of global-level matches as a function of Year (estimated mean ± SE). Dashed line at 50% represents no bias towards global or local choices. * p < .05, ** p < .01, *** p < .001.*

A one-sample t-test was computed to determine whether global choices in each Year and at each time-point were significantly different from 50%. This revealed that Year 1 children at T1 did not demonstrate either a global or local advantage, *t*(44) = 0.875, *p* = .386. All other Years made significantly more global than local choices, Year 3 T1: *t*(44) = 9.183, *p* < .001; Year 5 T1: *t*(44) = 3.872, *p* < .001; Year 3 T2: *t*(27) = 9.033, *p* < .001; Year 5 T3: *t*(31) = 10.072, *p* < .001. Overall, this indicates a developmental change from no preference to a preference for matching at the global level between the ages of 6 and 7 years.

2.3.1.2 Selective attention Navon task

LMMs were carried out with accuracy, RT, or LISAS as the DV, with level (global, local) and congruency (congruent, neutral, incongruent) as within-subjects factors, and Year and gender as between-subjects factors.

All main effects are reported in **Table 2.3**. To summarise, there was a main effect of congruency such that participants were faster and more accurate in congruent and neutral trials than in incongruent trials (*p*'s < .001), while performance did not differ between congruent and neutral trials (*p*'s > .220). Participants were also faster and more accurate with each increasing Year (*p*'s < .001). For the main effects of level, participants were faster but less accurate in global trials, and slower but more accurate in local trials. There was no difference in accuracy between genders, but males responded more quickly than females.

Table 2.3: *Linear mixed model main effects of selective attention task performance with accuracy, response time (RT), and linear integrated speed-accuracy scores (LISAS). Estimated marginal means and standard error (SE) are reported. Significant effects in bold.*

Main effect	Measure	Statistics	Estimated means (SE)
Level	Accuracy	***F*(1,994.2) = 33.96, *p* < .001**	Global: 87.0% (0.7) Local: 91.3% (0.7)
	RT	***F*(1,997.0) = 27.32, *p* < .001**	Global: 1051 ms (15) Local: 1110 ms (15)
	LISAS	*p* = .065	
Congruency	Accuracy	***F*(2,994.2) = 80.68, *p* < .001**	Congruent: 93.2% (0.8) Incongruent: 82.6% (0.8) Neutral: 91.6% (0.8)
	RT	***F*(2,997.0) = 11.64, *p* < .001**	Congruent: 1057 ms (16) Incongruent: 1119 ms (16) Neutral: 1066 ms (16)
	LISAS	***F*(2,995.2) = 27.40, *p* < .001**	Congruent: 1133 ms (19) Incongruent: 1262 ms (20) Neutral: 1158 ms (19)
Year	Accuracy	***F*(2,628.2) = 45.16, *p* < .001**	Year 1: 82.5% (1.1) Year 3: 90.2% (0.8) Year 5: 94.6% (0.8)
	RT	***F*(2,906) = 70.53, *p* < .001**	Year 1: 1219 ms (21) Year 3: 1077 ms (16) Year 5: 947 ms (16)
	LISAS	***F*(2,825.8) = 98.28, *p* < .001**	Year 1: 1396 ms (25) Year 3: 1169 ms (20) Year 5: 988 ms (20)
Gender	Accuracy	*p* = .229	
	RT	***F*(1,135.4) = 6.28, *p* = .013**	Males: 1046 ms (20) Females: 1116 ms (19)
	LISAS	*p* = .054	

The main effects in the accuracy data were modulated by two-way interactions between congruency and Year, *F*(4,994.2) = 7.42, *p* < .001, between congruency and level, *F*(2,994.2) = 6.05, *p* = .002, and between level and Year, *F*(2,994.2) = 22.27, *p* <

.001. These interactions were further modulated by a three-way interaction between congruency, level, and Year, $F(4,994.2) = 5.88$, $p < .001$. The three-way interaction was followed up by running separate LMMs in each Year, revealing a significant interaction between congruency and level in Year 1 only, $F(2,215.0) = 7.95$, $p < .001$ (Years 3 and 5: *F*'s <1). Further follow-up LMMs were run for each type of congruency in Year 1 which indicated that accuracy was higher in local than global trials when stimuli were incongruent ($p < .001$) and neutral ($p = .001$), but there was no effect of level in congruent trials ($p = .400$). The 3-way interaction was therefore driven by the fact that Year 1 children, but not older children, showed relatively poorer performance in global than local incongruent and neutral trials (**Figure 2.9A**). This interaction also drove the significant two-way interactions.

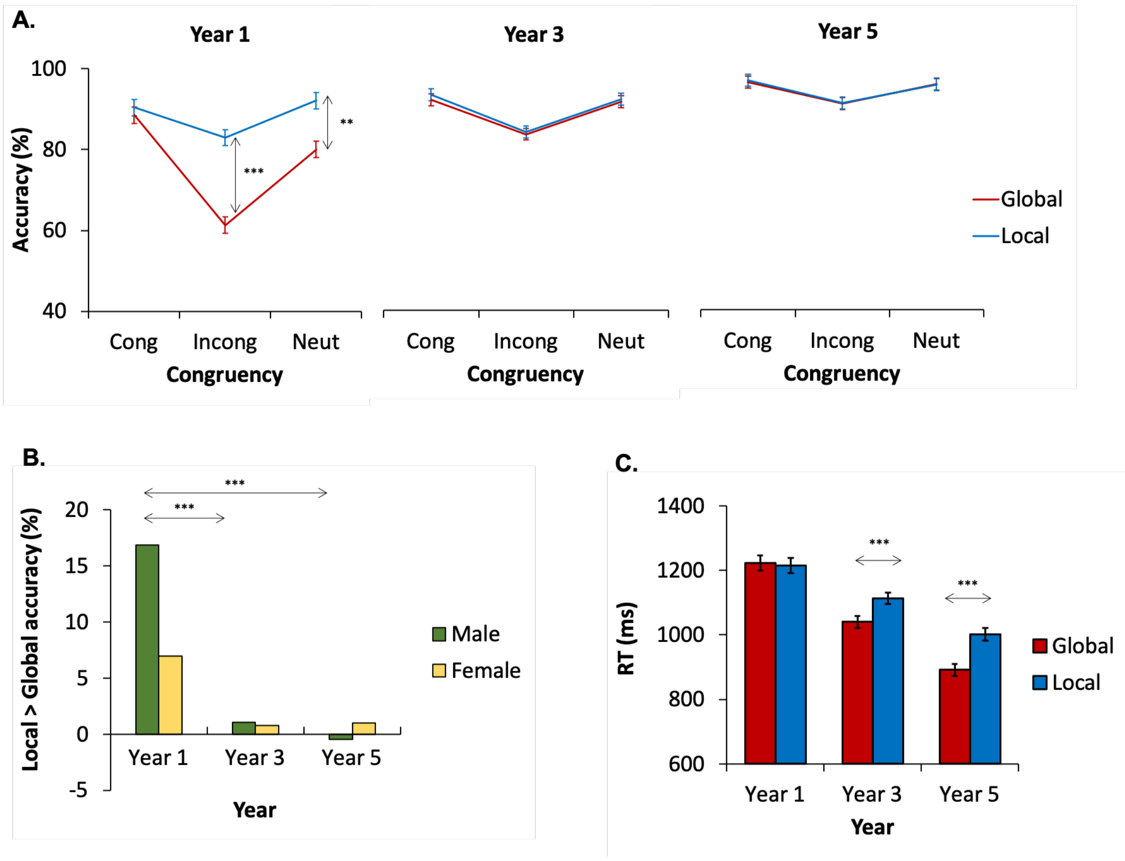


***Figure 2.9:** Performance on the selective attention Navon task (estimated mean ± standard error). **A.** Accuracy as a function of Year, congruency (Cong: congruent, Incong: incongruent, Neut: neutral) and level. **B.** Accuracy difference between global and local responses as a function of gender and Year. **C.** Response time (RT) as a function of Year and level. *p < .05; **p < .01; ***p < .001.*

Accuracy also showed a significant three-way interaction between level, Year, and gender, $F(2,994.2) = 3.45$, $p = .032$. As other follow-up analyses were inconclusive regarding the source of the three-way interaction, LMMs were performed on the difference in accuracy between global and local responses (local - global) for each gender separately and revealed a significant effect of Year in males (p < .001) but not in females (p = .063). Pairwise comparisons in males revealed a greater difference between global and local responses in Year 1 than in Years 3 and 5 (p's < .001). There was no difference between Year 3 and 5 (p > .99) (**Figure 2.9B**).

There was an interaction between level and Year in the RT data, F(2,997.0) = 8.18, p < .001. LMMs run for each Year revealed no effect of level in Year 1 (p = 812), but significantly quicker responses to the global level than the local level in Years 3 and 5 (p's < .001) (**Figure 2.9C**).

The above analyses indicated differing levels of speed-accuracy trade-offs across groups, emphasising the importance of running the LMM with LISAS scores to account for both accuracy and RT (**Table 2.2**). Consistent with the accuracy and RT analyses, participants had quicker LISAS scores in congruent and neutral trials than incongruent trials (p's < .001), with no difference between congruent and neutral trials (p = .481). Additionally, LISAS scores improved with each older Year group (p's < .001). There was no main effect of level, but there was a significant level by Year interaction, F(2,995.1) = 16.55, p < .001. Follow-up LMMs for each Year group revealed quicker LISAS scores at the local level than the global level (p = .011) in Year 1, but conversely, quicker LISAS scores at the global level than the local level (p's < .002) in Years 3 and 5 (**Figure 2.10**).

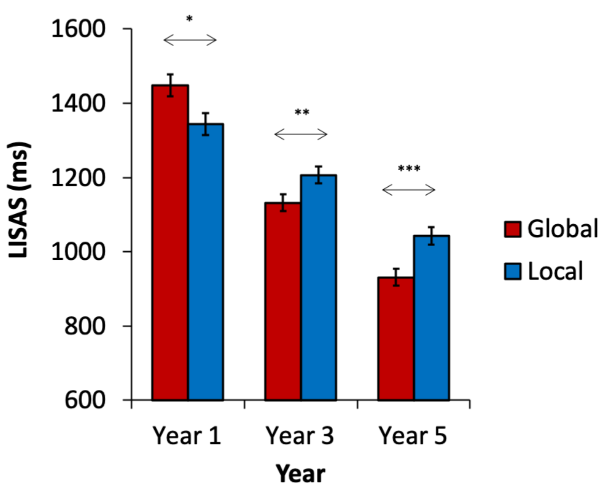


Figure 2.10: *Linear integrated speed-accuracy score (LISAS) as a function of Year and level on the selective attention Navon task (estimated mean ± standard error). *p < .05; **p < .01; ***p < .001*

Results on this combined measure are consistent with the lower accuracy in global than local trials observed in Year 1, and the faster RTs to the global than local level in Years 3 and 5. All other main effects and interactions were non-significant, which suggests that the gender and congruency effects observed in the accuracy and RT analyses were due to differing levels of speed-accuracy trade-off between genders and congruency.

Using the criteria explained in **Section 2.2.7.2**, outliers were identified from each selective attention level (**Table 2.4**) and removed from across the task at that time point. The global and local selective attention tasks were administered separately in the test battery, however the data were entered into the same analyses to examine the effect of level as a single task. Therefore, outlier data were removed from both global and local datasets. The same series of LMMs were then run with accuracy, then RT, then LISAS as the DV.

Table 2.4: *Number of outliers removed from the selective attention Navon data, by Time, Year, and level.*

	Global			**Local**		
	Year 1	Year 3	Year 5	Year 1	Year 3	Year 5
Time 1	2	0	0	1	1	0
Time 2	-	1	0	-	1	1

The accuracy data without outliers revealed an additional significant interaction between level and gender accuracy, $F(1,972.4) = 4.38$, $p = .037$, which was further modulated by the three-way interaction between level, Year, and gender as described above. The LISAS analysis without the outliers revealed an additional significant main effect of gender, $F(1,138.5) = 4.66$, $p = .033$, with better performance by males than females. There was also an additional main effect of level, $F(1,968.5) = 4.61$, $p = .032$, with better performance on local than global trials. This was further modulated by the two-way interaction with Year described above. All other significant patterns of the main effects and interactions remained as before. This reveals that although the outliers had some influence on the models, their removal did not impact on the overall story of the data.

In summary, for the selective attention Navon task, participants were overall less accurate and slower in incongruent than congruent and neutral trials, and more accurate but slower at the local than the global level. The difference in accuracy between the local and global levels was driven by Year 1 children, particularly in incongruent trials. When accuracy and RT were both accounted for, the only remaining significant interaction was between level and Year. Overall, this indicates a developmental change from a local advantage in Year 1 to a global advantage in Years

3 and 5 in the selective attention task. Females achieved poorer overall scores than males when accuracy and RT data were combined and outliers removed.

2.3.1.3 Big-small divided attention Navon task

Research with adults has identified a switching cost in divided attention Navon tasks, but this has not been examined in children. An initial LMM was carried out to assess whether there was a difference between stay and switch trials in the present study. Level (global, local) and trial type (switch, stay) were entered as within-subject variables, and Year and gender as between-subjects factors. There were no main effects or interactions of trial type in either the accuracy data (p's > .245), nor the RT data (p's > .240). Switch/stay trial type was therefore removed from the LMM analysis. Analyses examining the effect of the number of stay trials before a switch trial, and comparisons with the response switching task (which was designed to act as a stay-switch control without hierarchical levels) were not undertaken. Significant statistical results of the LMM analysis without trial type are presented in **Table 2.5**.

Table 2.5: *Linear mixed model main effects of the big-small divided attention task performance with accuracy, response time (RT), and linear integrated speed-accuracy scores (LISAS). Estimated marginal means and standard error (SE) are reported. Significant effects in bold.*

Main effect	Measure	Statistics	Estimated means (SE)
Level	Accuracy	***F*(1,622.5) = 10.47, *p* = .001**	Global: 86.2% (1.0)
			Local: 89.0% (1.0)
	RT	***F*(1,633.5) = 6.00, *p* = .002**	Global: 1033 ms (15.4)
			Local: 1001 ms (15.4)
	LISAS	***F*(1,629.0) = 9.63, *p* = .002**	Global: 1169 ms (21)
			Local: 1109 ms (21)
Year	Accuracy	***F*(2,658.1) = 52.88, *p* < .001**	Year 1: 78.0% (1.5)
			Year 3: 90.4% (1.1)
			Year 5: 94.4% (1.2)
	RT	***F*(2,651.3) = 83.27, *p* < .001**	Year 1: 1187 ms (22)
			Year 3: 1002 ms (17)
			Year 5: 862 ms (17)
	LISAS	***F*(2,596.3) = 101.03, *p* < .001**	Year 1: 1411 ms (30)
			Year 3: 1100 ms (23)
			Year 5: 906 ms (23)

Participants were quicker and more accurate when responding to local trials than global trials, and both accuracy and RT improved with ascending Year group (*p*'s < .003). The main effects in accuracy were modulated by a two-way interaction between Year and level, *F*(2,622.5) = 5.64, *p* = .004. Follow-up LMMs run in each Year indicated higher accuracy in local than global trials in Years 1 and 3 (*p*'s < .010), but no effect of level in Year 5 (*p* = .269) (**Figure 2.11A**).

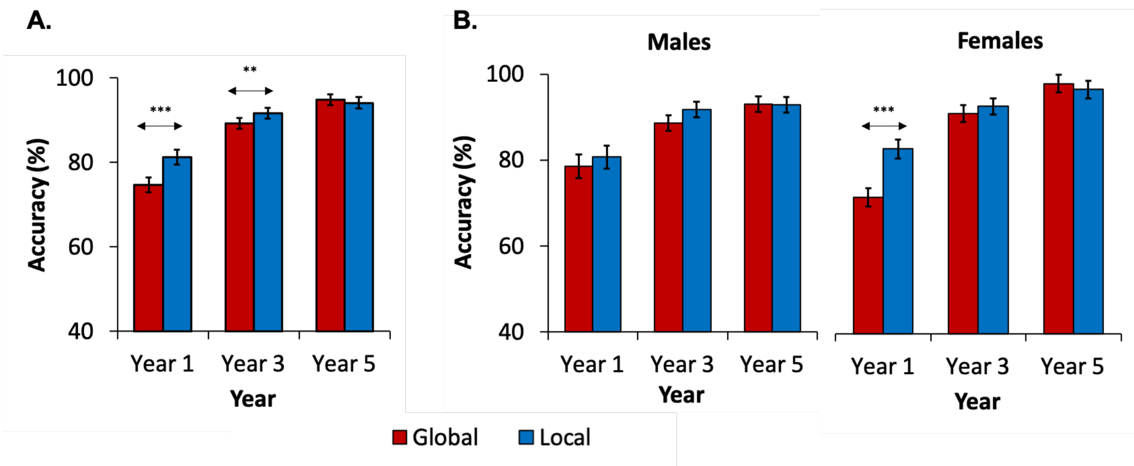


Figure 2.11: *Performance on the big-small divided attention Navon task (estimated mean ± SE).* ***A.*** *Accuracy as a function of Year and level.* ***B.*** *Accuracy as a function of gender, level and Year. *p < .05; **p < .01; ***p < .001.*

While there was no main effect of gender (*F*'s < 1), there was a significant three-way interaction in accuracy between Year, level, and gender, $F(2,622.5) = 3.36$, $p = .035$. LMMs run for each gender revealed an interaction between level and Year in females, $F(2,314.7) = 9.34$, $p < .001$, but not in males ($p = .362$). Further follow-up LMMs were run for each Year for females and showed higher accuracy in local trials than global trials in Year 1, $F(1,86.0) = 16.53$, $p < .001$, but no effect of level in Years 3 and 5 (*p*'s > 115). The three-way interaction was therefore driven by the fact that Year 1 females, but not older females or any males, showed relatively poorer performance in global than local trials (**Figure 2.11B**).

Although the data did not indicate group differences in speed-accuracy trade-offs, for consistency with the analysis of the selective attention task, an LMM was run with LISAS data. This supported the patterns identified in the accuracy and RT data, with quicker LISAS scores on local trials than global trials, and improving performance with ascending Year (**Table 2.5**). The interaction between Year, level, and gender observed in the accuracy data was not significant, suggesting that this was due to differing levels of speed-accuracy trade-off between genders. The interaction between level and Year which was significant for accuracy was marginal for LISAS data, $F(2,628.6) = 2.82$, $p = .060$.

Once outlier data-points had been removed (**Table 2.6**), the divided attention LMM was re-run to assess the impact of outlier datapoints.

Table 2.6: *Number of outliers removed from the divided attention Navon data, by Time and Year.*

	Year 1	Year 3	Year 5
Time 1	2	0	0
Time 2	-	1	0

As before, there were no main effects or interactions with the stay-switch variable (*p*'s > .215). A further series of LMMs as detailed above were run with accuracy, then RT, then LISAS as the DV. The accuracy model revealed the same main effects and interactions as the model with the full set of data, with two exceptions. The three-way interaction between Year, gender, and level became non-significant (p = .207), and an additional two-way significant interaction between Year and gender, $F(1,662.7) = 7.94$, $p < .001$, was identified (**Figure 2.12**).

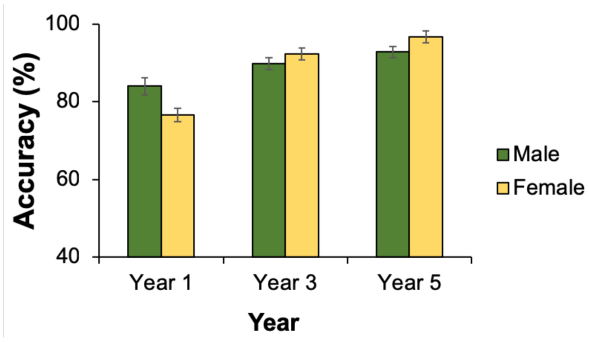


Figure 2.12: *Accuracy as a function of Year on the big-small divided attention Navon task without outliers.*

Follow-up LMMs did not reveal the source of this interaction. There were no significant differences of gender between Years, and both males and females revealed the same patterns of significance whereby Year 1 had significantly lower accuracy than Year 3, but there were no differences between Years 3 and 5. However, **Figure 2.12** suggests that the accuracy pattern across the Years is more pronounced in females than males.

All main effects and interactions in the RT and LISAS data remained as before, although the interaction between Year and level in LISAS data became significant (p = .005) where previously it had been just outside significance. Follow-up LMMs were run for each Year to compare level effects. This revealed a significant effect of level in Year 1 children ($p < .001$) with quicker LISAS scores in local than global trials, but no significant level effect in Years 3 and 5 (*p*'s > .147) (**Figure 2.13**).

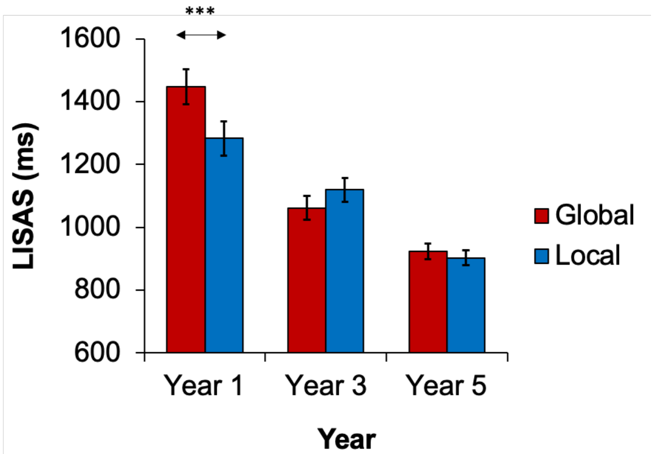


Figure 2.13: *Linear integrated speed-accuracy score (LISAS) as a function of Year and level on the big-small divided attention Navon task (estimated mean ± standard error). *p < .05; **p < .01; ***p < .001.*

This indicates that the outliers had a small influence on observed differences in accuracy between genders in different Years, and on Year and level differences in the combined LISAS score.

In summary, for the divided attention task, participants became faster and more accurate with age. While responses were overall faster and more accurate for local trials than global trials, there were differences in accuracy between level as a function of Year and gender, however this appears to be influenced by outliers. The LISAS calculation revealed a significant level by Year interaction when outliers were removed, suggesting a change in global and local responses occurring between Years 1 and 3, with a local processing advantage in Year 1 but no advantage in Years 3 and 5.

2.3.1.4 Yes-No divided attention Navon task at Time 2 only

Responses to the trials where the target shape was present were analysed using repeated measures ANOVAs. Trials where the target shape was absent were not informative on global or local processing development and were therefore excluded from these analyses. Accuracy, RT, and LISAS data were the DVs, with level (global, local) as the within-subjects factor and Year (3, 5) and gender as between-subjects factors. Main effects are presented in **Table 2.7**.

Responses were quicker and more accurate when the target appeared at the local compared with the global level, and fewer errors were made by the older Year group. The local advantage and improvement in performance with Year remained significant with the LISAS data. This indicates a local processing advantage across Years 3 and 5, which was only observed to be significant in Year 1 in the big-small divided attention task. There were no main effects of gender or any interactions.

Table 2.7: *Main effects of the yes-no divided attention task at Time 2 with accuracy, response time (RT), and linear integrated speed-accuracy scores (LISAS). Estimated marginal means and standard error (SE) are reported. Significant effects in bold.*

Main effect	**Measure**	**Statistics**	**Estimated means (SE)**
Level	Accuracy	***F*(1,56) = 4.34, *p* = .042, η^2_p = .072**	Global: 85.2% (2.2) Local: 90.9% (1.6)
	RT [(a)]	***F*(1,55) = 8.91, *p* = .004, η^2_p = .139**	Global: 1131 ms (27) Local: 1077 ms (28)
	LISAS [(a)]	***F*(1,55) = 10.20, *p* = .002, η^2_p = .156**	Global: 1297 ms (41) Local: 1177 ms (35)
Year	Accuracy	***F*(1,56) = 4.51, *p* = .038, η^2_p = .075**	Year 3: 85.1% (2.0) Year 5: 90.9% (1.9)
	RT [(a)]	*p* = .172	
	LISAS [(a)]	***F*(1,55) = 4.29, *p* = .043, η^2_p = .072**	Year 3: 1305 ms (49) Year 5: 1169 ms (45)

[(a)] One Year 3 participant did not respond correctly to any global level trials in the yes-no divided attention task.

With the removal of two outlier data-points (both Year 3 females, one global and one local), the main effect of Year on accuracy became non-significant (*p* = .079). An additional main effect of gender was observed, $F(1,54) = 1.32$, $p = .040$, $\eta^2_p = .076$, with higher accuracy in females than males. This demonstrates that these two outlier data-points were lowering the Year 3 female accuracy in the original analysis. The main effect of Year observed in the LISAS data became non-significant (*p* = .063), due to the narrower difference in LISAS scores between Year 3 and 5 after removal of the two outlier Year 3 scores. There were no further changes in the main effects or interactions for the accuracy, RT or LISAS data. Therefore, without outliers, the local processing advantage remained in evidence. While there was no difference in response by Year, females were more accurate than males.

2.3.1.5 Responses across global and local tasks

The analysis thus far has concentrated on group differences within each task, which does not allow for a comparison of individual performance across the Navon tasks. Correlations were therefore analysed between free choice global matches, and the LISAS global and local scores for the selective and big-small divided attention tasks. Additionally, the proportional difference between each participant’s global and local LISAS scores was calculated for each task, as (Global – Local) / Global. This gives an indication of within-task differences in global and local responses, whilst accounting for individual differences in processing speed.

Pearson's correlations were carried out with data from T1 with outliers removed. Correlations between tasks may be modulated by age- and gender-related individual differences across tasks so partial correlations were also carried out, controlling for these factors (**Table 2.8**)**.** One Year 1 participant achieved an overall big-small divided attention accuracy rate which did not fall outside the outlier boundaries. However, their global accuracy was only 2.7% (their local accuracy was 83.8%) and therefore a LISAS score could not be calculated for their global trials or their composite score. As a result, there are different numbers of participants in the global and local big-small divided attention correlations.

Table 2.8: *Pearson's correlations between percentage of global choices on the free choice Navon task and linear integrated speed-accuracy scores (LISAS) on the selective (Sel. Attn.) and big-small divided (Div. Attn.) attention Navon tasks. Data from Time 1 only. Partial correlations controlling for age in months and gender are shown below the diagonal. *p < .05; **p < .01; ***p < .001. Significant correlations are highlighted in bold.*

	Free Choice global %	Sel. Attn. Global	Sel. Attn. Local	Sel. Attn. composite	Div. Attn. Global	Div. Attn. Local	Div. Attn. Composite
Free Choice global % [(a)]		-.155	-.016	-.154	**-.179 ***	**-.174 ***	.010
Sel. Attn. Global [(b)]	-.061		**.520 *****	**.602 *****	**.572 *****	**.475 *****	.157
Sel. Attn. Local [(b)]	.079	**.224 ***		-.327 ***	**.424 *****	**.473 *****	-.086
Sel. Attn. Composite [(b)]	-.108	**.657 *****	**-.534 *****		**.193 ***	.085	**.217 ***
Div. Attn. Global [(b)]	-.106	**.291 ****	.104	.111		**.718 *****	**.331 *****
Div. Attn. Local [(c)]	-.113	**.232 ****	**.261 ****	-.003	**.607 *****		**-.389 *****
Div. Attn. Composite [(b)]	.037	.092	**-.202 ***	**.194 ***	**.317 *****	**-.526 *****	

[(a)] N = 135, df = 131
[(b)] N = 131, df = 127
[(c)] N = 132, df = 128

The partial correlations controlling for age and gender effects revealed clear within-task associations, whereby global and local scores were positively correlated within the selected attention task and, to a greater extent, within the big-small divided attention task (**Figure 2.14A and B**). Positive associations were also evident between the global trials and the local trials across the selective and big-small divided attention tasks (**Figure 2.14C and D**), as well as between the selective attention global trials and big-small divided attention local trials. Importantly, there was a significant, but small, positive association between the proportional difference scores of the selective and big-small divided attention tasks indicating that the relative advantage in responding to the global or local level was somewhat consistent across task (**Figure 2.14E**). There were no significant associations with the free choice Navon task.

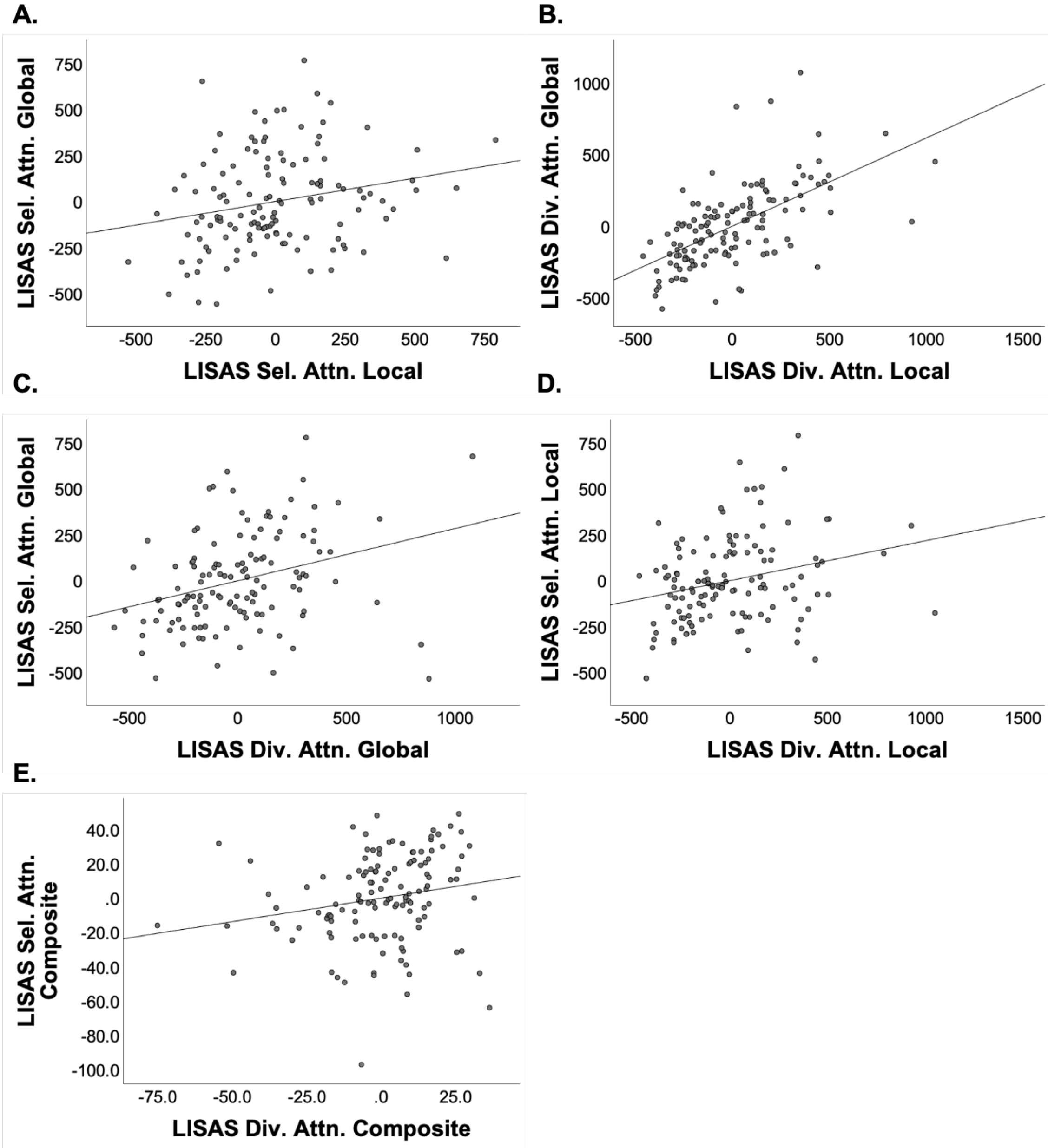


Figure 2.14: *Plots using Time 1 linear integrated speed-accuracy (LISAS) Navon data, illustrating the partial correlations controlling for age in months and gender.* ***A.*** *Selective attention global and local responses.* ***B.*** *Big-small divided attention global and local responses.* ***C.*** *Global responses from the selective and big-small divided attention tasks.* ***D.*** *Local responses from the selective and big-small divided attention tasks.* ***E.*** *Composite responses from selective and big-small divided attention tasks.*

2.3.1.6 Performance across the divided attention tasks

A summary of responses on the two divided attention tasks at T2 can be viewed in **Table 2.9**, enabling a comparison across task. Although a statistical analysis cannot be carried out to compare the two distinct tasks, the responses on the Yes-No task are generally slower and less accurate than on the Big-Small task.

Table 2.9: *Accuracy, RT, and linear integrated speed-accuracy scores (LISAS) on the two divided attention tasks at Time 2. Estimated marginal means and standard error (SE) are reported.*

		Year 3		Year 5	
		Global	**Local**	**Global**	**Local**
Accuracy	**Big-small**	92.0% (1.4)	93.7% (1.2)	97.3% (1.3)	95.4% (1.1)
	Yes-no	81.9% (3.1)	88.9% (2.2)	88.7% (2.9)	92.7% (2.0)
RT	**Big-small**	995 ms (40)	1001 ms (41)	911 ms (38)	912 ms (38)
	Yes-no	1155 ms (39)	1087 ms (40)	1088 ms (36)	1032 ms (37)
LISAS	**Big-small**	1097 ms (47)	1093 ms (49)	955 ms (44)	976 ms (45)
	Yes-no	1348 ms (59)	1203 ms (50)	1218 ms (54)	1110 ms (47)

Comparisons of LISAS scores on global and local trials and the proportional differences between global and local trials were made between the two divided attention tasks using T2 data only, to assess commonalities in performance across the tasks (**Table 2.10**). Pearson's correlations were carried out with outliers excluded, and partial correlations were run controlling for age and gender.

Table 2.10: *Pearson's correlations between linear integrated speed-accuracy scores (LISAS) on the two divided attention Navon tasks at Time 2. Composite measures were calculated as (global-local)/global. Partial correlations controlling for age in months and gender are shown below the diagonal. *p < .05; **p < .01; ***p < .001. Significant correlations are highlighted in bold.*

	Big-small global	**Big-small local**	**Big-small composite**	**Yes-no global**	**Yes-no local**	**Yes-no composite**
Big-small global (a)		**.764 *****	**.326 ***	**.317 ***	**.311 ***	.054
Big-small local (a)	**.753 *****		**-.344 ****	**.292 ***	**.401 ****	-.060
Big-small composite (a)	**.320 ***	**-.364 ****		.047	-.160	.218
Yes-no global (b)	**.297 ***	**.298 ***	.006		**.529 *****	**.448 *****
Yes-no local (b)	**.276 ***	**.386 ****	-.192	**.504 *****		**-.468 *****
Yes-no composite (b)	.080	-.034	.216	**.460 *****	**-.482 *****	

(a) N = 59, df = 55
(b) N = 58, df = 54

After controlling for age and gender, positive correlations were identified between levels within each task (**Figures 2.15A and B**). Weaker, but significant, correlations were also identified on global and local trials across task (**Figures 2.15C and D**), however, the composite measures did not significantly correlate between tasks. These results provide further support for the existence of considerable task demand influences on children's global and local responses.

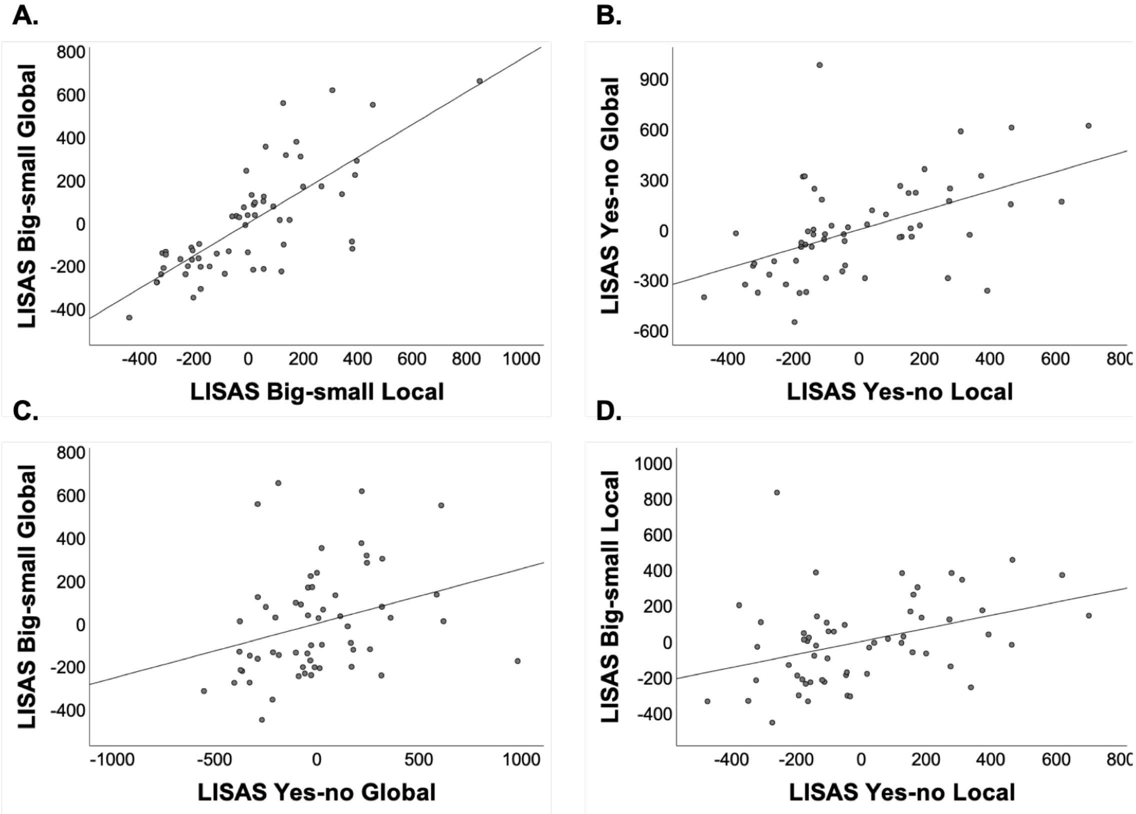


Figure 2.15: *Plots using Time 1 linear integrated speed-accuracy (LISAS) Navon data, illustrating the partial correlations controlling for age in months and gender.* ***A.*** *Big-small global and local responses.* ***B.*** *Yes-no global and local responses.* ***C.*** *Global responses from the big-small and yes-no divided attention tasks.* ***D.*** *Local responses from the big-small and yes-no divided attention tasks.*

2.3.2 Field independence development

For each task, the data collected at T1 and T2 were plotted on the same figure for comparison purposes (**Figures 2.16A and B**). As for the Navon tasks, it should be noted that those participating in the longitudinal study therefore have their T1 data plotted twice – once as part of the cross-sectional mean, and once as T1 of the longitudinal data.

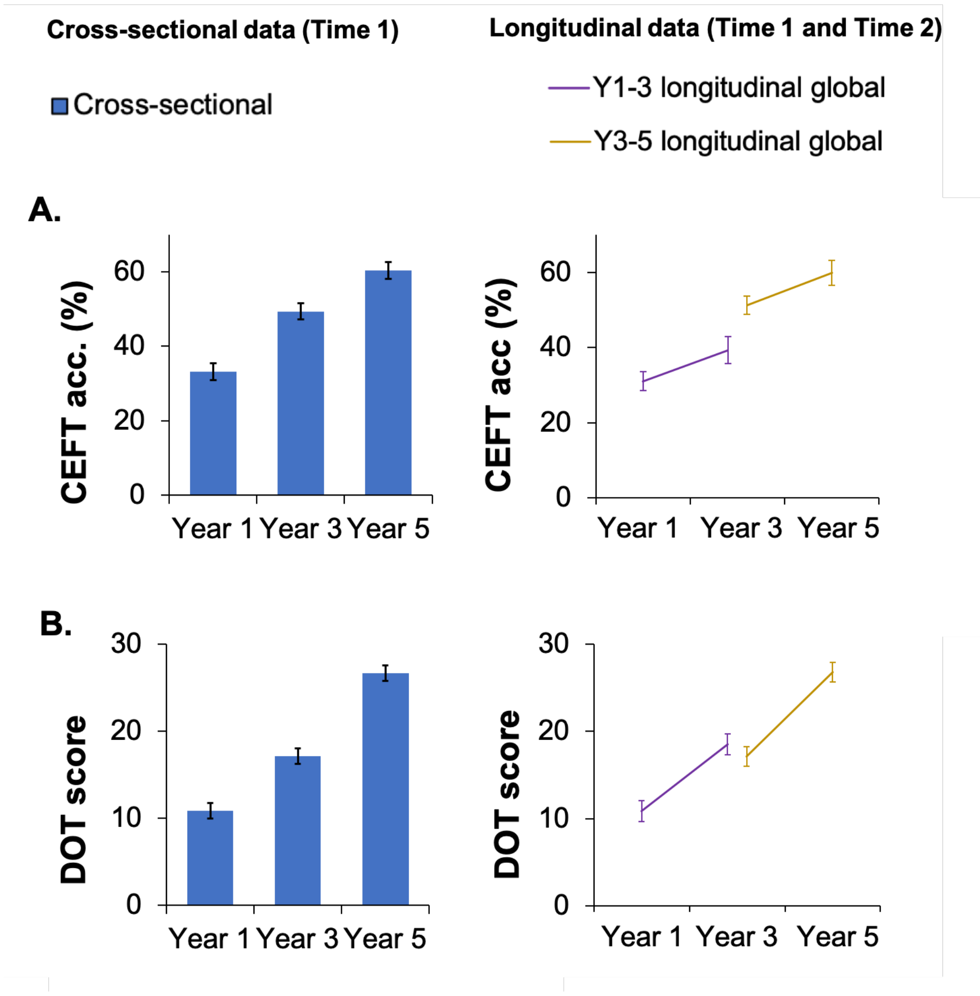


Figure 2.16: *Mean cross-sectional and longitudinal accuracy (acc.) on the field independence tasks. Data are plotted as a function of Year.* ***A.*** *Children's Embedded Figures Test (CEFT).* ***B.*** *Design Organisation Test (DOT). All results are plotted as estimated mean ± standard error, with cross-sectional data plotted as bar charts on the left and longitudinal data plotted as line charts on the right.*

There were no apparent practice effects in either task. In fact, the reverse appears to be the case in the CEFT, where the scores were lower in Year 3 for the T2 data compared with T1. There were no outlier data-points identified in either the CEFT or the DOT.

2.3.2.1 Children's Embedded Figures Test (CEFT)

An LMM was run with accuracy as the DV with trial type (tent, house) as the within-subjects factor and Year and gender as the between-subjects factors. Main effects are presented in **Table 2.11**. In line with expectation, participants achieved higher scores in trials with the tent target shape compared with the house target shape. As predicted, there was also a main effect of Year, with higher scores in each ascending Year group (*p*'s < .001). There were no further main effects or interactions.

Table 2.11: *Analyses of field independence task performance with Children's Embedded Figures Test (CEFT) accuracy and Design Organisation Test (DOT) score. Estimated marginal means and standard error (SE) are reported.*

Main effect	Measure	Statistics	Estimated means (SE)
Trial Type	CEFT	***F*(1,252.8) = 79.77,**	Tent: 55.6% (1.6)
		***p* < .001**	House: 38.9% (1.6)
Year			
	CEFT	***F*(2,331.5) = 34.77,**	Year 1: 35.7% (2.4)
		***p* < .001**	Year 3: 46.5% (1.8)
			Year 5: 59.6% (1.8)
	DOT	***F*(2,125.4) = 120.06,**	Year 1: 10.3 (0.9)
		***p* < .001**	Year 3: 17.7 (0.7)
			Year 5: 26.7 (0.7)

2.3.2.2 Design Organisation Test (DOT)

A further LMM was run with DOT score as the DV, and Year and gender as the between subjects factors. Main effects are presented in **Table 2.11**. There was a main effect of Year, with higher scores in each ascending Year group (*p*'s < .001). There were no further main effects or interactions.

2.3.2.3 Performance across field independence tasks

Pearson's correlations were run between CEFT accuracy and DOT scores from Time 1, revealing a significant positive association, $r = .624$, $p < .001$ (**Figure 2.17A**). A partial correlation was also run, controlling for age in months and gender, which remained significant, $r_p = .317$, $p = .015$ (**Figure 2.17B**).

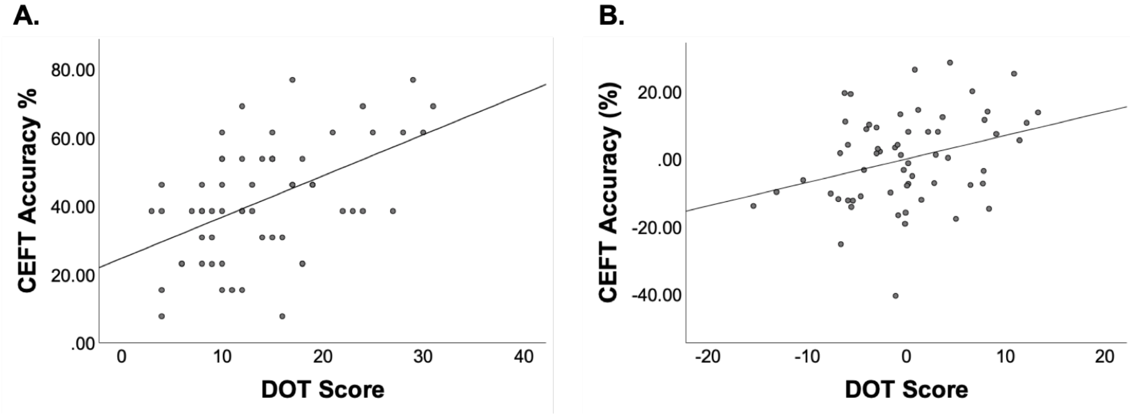


Figure 2.17: *Plots of correlations between Children's Embedded Figures Test (CEFT) accuracy and Design Organisation Test (DOT) score using Time 1 data.* ***A.*** *Correlation between CEFT and DOT.* ***B.*** *Partial correlation between CEFT and DOT, controlling for age in months and gender.*

2.3.3 Systemizing, empathizing, and autism quotients

2.3.3.1 Participants

Although the questionnaire was distributed to all parents of children who took part in the study at T1, only a subset returned the completed questionnaire (**Table 2.12**). The gender split differed from the main sample, but there were no significant group-level gender differences in the subset $\chi^2(2) = 1.50$, $p = .471$.

Table 2.12: *Questionnaire demographics.*

	Year 1	**Year 3**	**Year 5**
Number of participants (%)	22 (48.9%)	20 (44.4%)	23 (51.1%)
Males : females	8:14	11:9	11:12
Mean age in months (SD)	68.5 (3.3)	91.7 (3.3)	116.9 (3.6)
Mean age in years, months	5, 8	7, 8	9, 9

2.3.3.2 SQ-C, EQ-C, and AQ-10 development

A series of ANOVAs were carried out with SQ-C, EQ-C, or AQ-10 scores as the DV, and Year and gender as between-subject factors. There were no main effects or interactions (*p*'s > .110). A further ANOVA was run with D-scores as the DV, which again revealed no main effects or interactions (*p* > .080). Therefore, there were no differences in questionnaire scores across Years nor between genders. There were no outlier data-points identified in any of the variables.

2.3.3.3 Performance across SQ-C, EQ-C, and AQ-10

Pearson's correlations were run between the questionnaire scores and the D-score to examine relationships between systemizing, empathizing, and autism trait scores (**Table 2.13**). Although there were no Year effects in the analyses detailed above, partial correlations were run co-varying age in months and gender in case these factors modulated direct associations.

Table 2.13: *Pearson's correlations of systemizing (SQ-C), empathizing (EQ-C), autism trait scores (AQ-10), and combined systemizing and empathizing (D-score). Partial correlations controlling for age in months and gender are presented below the diagonal. *p < .05; **p < .01; ***p < .001. Significant correlations are highlighted in bold.*

	SQ-C	**EQ-C**	**AQ-10**	**D-score**
SQ-C		.167	.006	**.527** ***
EQ-C	.149		**-.444** ***	**-.749** ***
AQ-10	-.004	**-.471** ***		**.382** ***
D-score	**.546** ***	**-.747** ***	**.391** **	

N = 65, df = 61

The covariates of age and gender had little impact on the correlations. There was a negative correlation between EQ-C and AQ-10, whereby those who scored higher on the autism trait questionnaire achieved lower empathizing scores (**Figure 2.18A**). Also, those who achieved higher SQ-C scores than EQ-C scores, tended to score more highly on the AQ-10 (**Figure 2.18B**). There was no significant association between SQ-C and EQ-C, or between SQ-C and AQ-10.

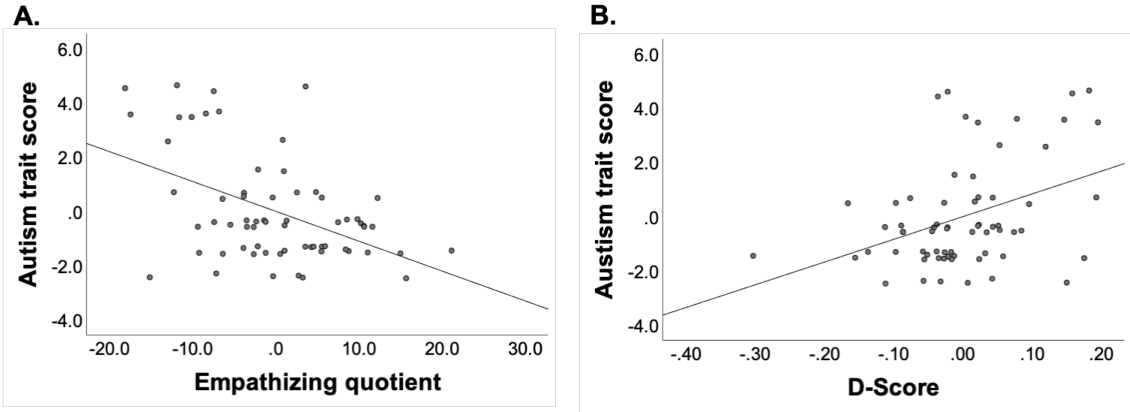


Figure 2.18: *Plots of partial correlations controlling for age in months and gender.* ***A****. Negative correlation between empathizing quotient and autism traits.* ***B****. Positive correlation between autism traits and composite of systemizing and empathizing quotients. Higher D-scores denote higher systemizing scores relative to empathizing scores.*

2.3.4 Associations across constructs

A correlation was run with T1 data from the Navon tasks, the FI tasks, and the parental questionnaires, to examine correlations across the whole-part constructs (**Table 2.14**). Any outlier datapoints were excluded from the analyses. Partial correlations were carried out controlling for age in months and gender.

Table 2.14: *Pearson's correlations (r) and partial correlations (r_p) controlling for age in months and gender, comparing free choice global choices, linear integrated speed-accuracy scores (LISAS) for selective attention (Sel. Attn.) and big-small divided attention (Div. Attn.) tasks, Children's Embedded Figures Test (CEFT), Design Organisation Test (DOT), and systemizing (SQ-C), empathizing (EQ-C), a composite measure (D-score) and autism traits (AQ-10) from questionnaires. Data from time 1 only. *p < .05; **p < .01; ***p < .001. Significant correlations are highlighted in bold.*

		Field Independence (FI)				**Systemizing - Empathizing**							
		CEFT [a]		**DOT** [a]		**SQ-C** [d]		**EQ-C** [d]		**D-score** [d]		**AQ-10** [d]	
		r	r_p	*r*	r_p	*r*	r_p	*r*	r_p	*r*	r_p	*r*	r_p
Global / local processing	**Free choice** [a]	.085	-.018	.141	.033	.040	.011	-.066	-.134	.083	.119	-.067	-.076
	Sel. Attn. Global LISAS [b]	**-.492 *****	-.150	**-.566 *****	-.150	-.119	.016	-.025	.134	-.061	-.097	-.183	-.201
	Sel. Attn. Local LISAS [b]	**-.443 *****	-.132	**-.555 *****	**-.214 ***	-.193	-.090	-.072	.001	-.069	-.056	-.181	-.191
	Sel. Attn. composite [b]	-.081	.040	-.094	.062	-.043	-.012	.032	.113	-.059	-.105	-.102	-.098
	Div. Attn. Global LISAS [b]	**-.398 *****	-.048	**-.574 *****	**-.234 ****	-.045	.098	.013	.152	-.044	-.058	-.044	-.196
	Div.Attn. Local LISAS [c]	**-.266 ****	.049	**-.494 *****	**-.220 ***	-.195	-.112	.102	.222	-.226	**-.267 ***	**-.292 ***	**-.307 ***
	Div. Attn. composite [b]	-.156	-.095	-.122	-.035	.105	.135	-.133	-.106	.188	.185	.111	.119
FI	**CEFT** [a]					.149	.034	.210	.136	-.077	-.094	.052	.028
	DOT [a]					.238	.135	.043	-.134	.131	.205	.057	.031

[a] N = 135, df = 131
[b] N = 131, df = 127
[c] N = 132, df = 128
[d] N = 65, df = 61

The correlations and partial correlations revealed few significant associations between constructs. There were significant negative associations between the DOT and the local trials on the selective and big-small divided attention tasks (**Figure 2.19A and B**), as well as global trials on the big-small divided attention task (**Figure 2.19C**). This revealed that those achieving a higher score on the DOT, had faster LISAS scores on those Navon tasks, after accounting for age and gender. There were no significant correlations between the FI tasks and the questionnaire responses. In fact, the only significant association with the questionnaire responses was a negative correlation after covarying age and gender between the big-small divided attention local responses and both D-score (**Figure 2.19D**) and AQ-10 (**Figure 2.19E**).

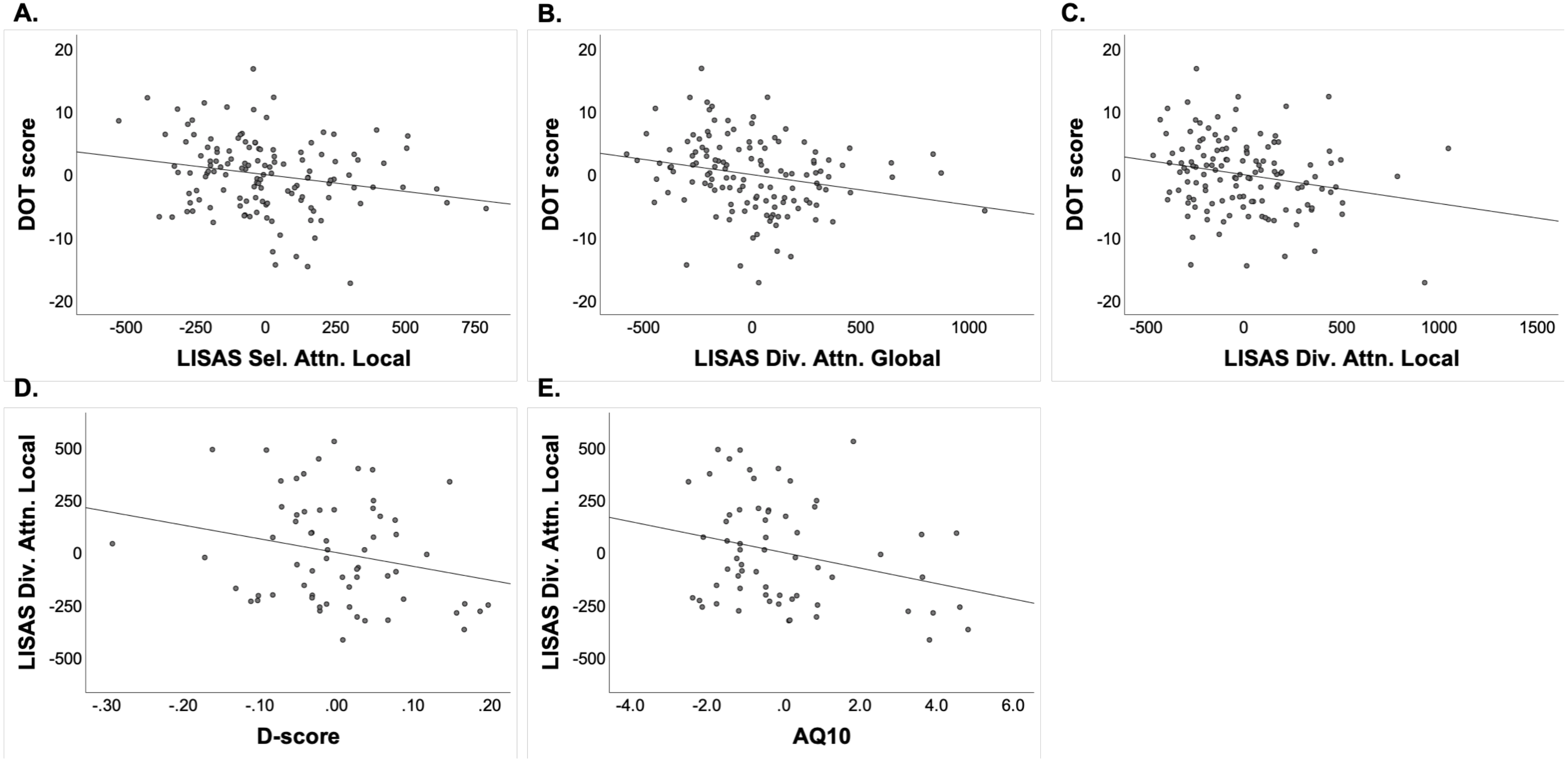


Figure 2.19: *Plots of partial correlations covarying age and gender.* ***A-C****. Negative correlations between Design Organisation Test (DOT) scores and linear integrated speed-accuracy scores (LISAS) selective attention (Sel. Attn.) Navon local trials (****A****), LISAS big-small divided attention (Div. Attn.) Navon global trials (****B****) and LISAS big-small Div. Attn. Navon local trials (****C****).* ***D-E.*** *Negative correlations between LISAS big-small Div. Attn. local trials and the relative strength in systemizing over empathizing (D-scores) (****D****) and autism trait questionnaire (AQ-10) (****E****).*

2.4 Discussion

The purpose of this chapter was to examine developmental changes in whole-part constructs, and to assess the influence of differing task demands on performance both within and between constructs. This was achieved through assessing individuals' performance on three Navon tasks and two FI tasks using a combination of cross-sectional and longitudinal data. Additionally, systemizing, empathizing, and autism traits were assessed at a single time-point. This is the first study to longitudinally examine global and local processing development in neurotypical children, as well as the first to comprehensively examine children's responses to whole-part tasks across the three whole-part constructs of global and local processing, FI, and systemizing. This is important for understanding the extent to which these three constructs share overlapping variance, and how this changes with development between the ages of 5 and 10 years.

2.4.1 Global and local processing

2.4.1.1 Global and local processing development

Performance on the three Navon tasks did not show a consistent pattern of responses, indicating that the demands of the tasks modulated the global and local performance. There were distinct patterns of responses as a function of Year depending on the task demands, with a shift from no preference to a global preference in the free choice paradigm, from a local advantage to a global advantage in the selective attention task, and from a local advantage to no effect of level in the big-small divided attention task. This supports prior observations that Navon tasks are not only measuring variation in global and local processing, but also a number of other processes such as attention, conflict resolution, and maintenance of task goals (De-Wit & Wagemans, 2015).

In the Free Choice task, Year 1 children had no preference in their matching choice, but those in Years 3 and 5 were more likely to select the global match. This largely reflects previous studies which found that neurotypical children make more global than local matches on free choice Navon tasks (Deruelle, Rondan, Gepner, & Fagot, 2006; Koldewyn et al., 2013; Wang et al., 2007), and supports studies where children from age 6- to 7-years-old begin to match the global level more consistently than the local level (Dukette & Stiles, 1996; Harrison & Stiles, 2009).

In the selective attention Navon task, there was a change in response pattern between Years 1 and 3. Year 1 children responded equally quickly to both levels, but were more accurate on local than global trials, while older children were equally accurate across level but responded more quickly to global than local trials. When RT was adjusted to account for accuracy differences, removing any trade-off in speed and accuracy, Year 1 children revealed a local advantage while Years 3 and 5 children showed a global advantage. Overall, this demonstrates a developmental change from a local advantage to a global advantage between Years 1 and 3 on the selective attention Navon task. In a selective attention study of 5- to 12-year-olds, Koldewyn et al. (2013) found no significant difference between global and local accuracy, and marginally quicker responses to the global than local level. However, there was no analysis of developmental differences, so it is unclear whether there were age-specific patterns similar to those observed in the present study. In contrast, Scherf et al. (2009) identified higher accuracy and quicker responses in local trials than global trials during a selective attention task in children (8-13 years) and adolescents (14-17 years). However, this more persistent local advantage may have resulted from smaller global stimuli making the local level comparatively salient.

In the big-small divided attention Navon task, higher accuracy was observed on local than global trials in Years 1 and 3, and no difference in accuracy was revealed between levels in Year 5. Across all Years, RT was quicker in local than global trials. When accuracy and RT data were combined using the LISAS calculation, there was a significant interaction between level and Year once outliers were removed, reflecting a local advantage in Year 1 and no difference in responses between levels in Years 3 and 5. The additional divided attention task, performed at T2 only, revealed a local advantage across Years 3 and 5. The results here contrast with previous divided attention studies, where participants had to identify whether a target was present or absent. In one study, children aged 6 to 14 years were quicker and more accurate when a letter was presented at the global level than the local level (Plaisted et al., 1999). Similarly, in another study, local responses were slower and less accurate than intermediate and global responses in 5- and 6-year-olds (Krakowski et al., 2016). In a further study Farran et al. (2003) found no significant difference in accuracy or RT between target stimuli presented at the global or local level in 6- to 7-year-olds. Differences between these studies and the present study may be due to contrasting saliency balances between the levels, whereby the global stimuli were larger and the local stimuli were smaller and more numerous (Krakowski et al., 2016; Plaisted et al., 1999) than in the present study. Differences may also be due to the use of letters rather than shapes (Farran et al., 2003) which has been shown to induce a larger global advantage in children and adults (Dukette & Stiles, 1996; Pletzer et al., 2017).

In summary, although the specific patterns of global and local responses differed between the tasks, this study suggests there are significant changes occurring between the ages of 6- and 7-years-old that led to shifts in the balance of global and local processing. A contributing factor to this developmental change may be the design of the Navon stimuli themselves. In Navon tasks, the global representation is not a solid shape as it involves a group of individual elements. It is possible that this contributes to the apparent later development of grouping processes compared with segmenting processes. This suggestion is supported by studies forming Navon stimuli with unfamiliar shapes, which have found that the automatic identification of a familiar object attracts attentional resources and influences responses (Poirel et al., 2006). If the global shape is less similar to younger children's pre-existing representation of that shape, it may be less salient than the local level. As children develop, their wider experiences of shape may result in a more varied internal representation and, therefore, quicker shape identification. This requires further examination to understand if this is indeed a modulating factor in responses of younger children to the global level of Navon tasks.

2.4.1.2 Global and local congruency effects

Congruency effects found in the present study showed slower and less accurate responses in incongruent trials than in congruent and neutral trials, which supports previous findings (Bouvet, Rousset, Valdois, & Donnadieu, 2011; Weinbach & Henik, 2014). This indicates that only distractors which were task-relevant interfered with performance and that the longer RTs are likely due to inhibitory control or conflict resolution processes involved in suppressing the irrelevant stimulus and response. Again, Year 1 children revealed distinct congruency effects in accuracy. However, when a single measure of accuracy and RT was used, these effects became non-significant and therefore likely reflected differences in speed-accuracy trade-off between congruency and Year.

2.4.1.3 Navon task: differences between stay and switch trials

Studies with adults have identified an accuracy or RT cost on switch trials compared with stay trials while attentional resources are shifted from one level to the other (Katagiri et al., 2013; Wilkinson et al., 2001). Surprisingly, this effect was not evident here in the divided attention Navon task, with no significant differences in accuracy or RT between stay and switch trials. This suggests that, in children, there are no lingering effects of the target level beyond an individual trial. There are two possible explanations for this lack of effect. First, children may examine both levels before making a decision when their attention is split, which leads to an accuracy and RT cost regardless of whether the trial is a stay or switch trial. Second, children may attend to only one level and respond according to whether the circle is present or absent. This would lead to no advantage of a stay trial over a switch trial as a single level is attended to throughout the task.

The yes-no divided attention task was introduced at T2 to compare accuracy and RT when children had to examine both levels before responding. This revealed distinct patterns of responses between divided attention task, with no significant differences in responses to level in Years 3 and 5 on the big-small task, but quicker and more accurate responses to local trials than global trials on the yes-no task. Additionally, responses on the yes-no task were generally slower and less accurate than on the big-small task, although this has not been statistically tested. This may indicate that the response strategies differed between task, which may suggest that both levels were not examined in each trial of the big-small divided attention task. Certainly, it emphasises once more the role that task characteristics can have on global and local responses.

2.4.1.4 Global and local gender differences

There were no gender effects evident in the free choice Navon task, nor in the divided attention tasks for LILAS scores. After outliers were excluded, a gender effect was revealed in LISAS scores on the selective attention task, where males achieved quicker LISAS scores than females. There were also some gender effects revealed in accuracy data in the selective and big-small divided attention tasks. A larger local advantage was found in Year 1 males than females in the selective attention task and, contrastingly, Year 1 females exhibited a local advantage in the big-small divided attention task but not Years 3 or 5 females nor males. Further, females achieved higher accuracy than males in the yes-no divided attention task when outliers were removed. As these effects were non-significant for LILAS scores, it is fair to assume that these gender effects were impacted by differing speed-accuracy trade-offs between genders.

Previous research has broadly suggested that males have a more global advantage than females (Kramer et al., 1996; Pletzer, Petasis, & Cahill, 2014; Scheuringer & Pletzer, 2016; Tzuriel & Egozi, 2010). However, some research has identified no gender effect (Kimchi et al., 2009) or even a more global advantage in females than males (Dukette & Stiles, 1996). This heterogeneity in gender differences in the literature is likely to result not only from the different experimental designs employed, but also from the fact that RT and accuracy measures are rarely combined. As is evident in this study, the integrated analyses revealed different gender effects from the RT or accuracy data.

2.4.1.5 Individual differences across Navon tasks

After controlling for age and gender, the correlation analyses revealed a closer association between global and local responses within the big-small divided attention task than the selective attention task. This may be partly explained by the fact that the divided attention global and local responses were derived from a single task, but the selective attention global and local levels were presented as separate, but similar, tasks. When assessing the consistency of responses across task, significant associations were evident between the global responses and local responses on the selective and big-small divided attention tasks. Further, there was a weak, but significant, positive correlation between the relative differences in responses to global and local trials of the selective and divided attention Navon tasks. This indicates that children who performed relatively better on global than local trials on the selective attention task, also performed relatively better on global than local trials on the big-small divided attention task. The comparison of scores on the two divided attention tasks at T2 revealed closer within-task global and local correlations than between-task correlations. The within-task correlations were greater in the big-small task than the yes-no task, despite both being presented as single tasks. The lack of a significant association between the two composite measures indicates that the relative advantage in global or local processing was not consistent across tasks. This suggests that even when attentional demands are broadly constant (i.e. both tasks were described as requiring attention to be divided between the levels), other task demands can affect participants' global and local responses.

There were no significant associations with the free choice Navon task, after covarying age and gender. A study of children aged 5 to 12 years, which used a free choice and a selective attention paradigm, also found no relationship between responses on each task (Koldewyn et al., 2013). This may be due to the very different design of the free choice task, with only four trials and no time limit on response. The task included only a small number of trials to elicit participants' decision about which was most similar to the target, without the effect of priming. However, it is possible that four trials is not sufficient to measure a global or local advantage, resulting in a lack of response overlap with the other tasks. The task also differed from the selective and divided attention tasks in that participants were asked to make a judgement about similarity, rather than respond according to a clear goal with correct and incorrect responses.

In summary, individual responses were weakly consistent in terms of global or local advantage across the selective and big-small divided attention Navon tasks. When comparing two tasks purporting to require similar types of attentional demand, in this case the two divided attention tasks, closer associations were observed on global and local responses within each task than between the tasks. Therefore, inconsistent global and local responses between tasks may relate not only to the differing attentional demands of the task (i.e. selective or divided attention), but also differing strategies adopted by individuals to complete the task.

2.4.2 Field independence

2.4.2.1 Field independence development

Accuracy on both FI tasks improved with Year, supporting previous studies with children in this age range (Amador-Campos & Kirchner-Nebot, 1997; Glynn & Stoner, 1987; Goodenough & Eagle, 1963). This development was significant between Years 1 and 3, and between Years 3 and 5 for both tasks, which contrasts with one study identifying no difference in CEFT performance between Years 1 and 3 (Bigelow, 1971). The main difference between this study and the study by Bigelow et al. (1971) is that here, the CEFT was administered in an adapted way to enable comparison across Years. As demonstrated in the study by Amador-Campos et al. (1997), the stimuli in the CEFT are not presented in an ascending order of difficulty, which may have led to the observed poorer performance in Year 1 children than older children in the study by Bigelow et al. (1971). This is the first study using the DOT with children but previous studies using the Block Design task have identified improvements in performance with age across 4.5- to 8.9-years-old (Akshoomoff & Stiles, 1996) and 8- to 16-years old (Booth, 2006).

As predicted, and in support of a previous study (Amador-Campos & Kirchner-Nebot, 1997), scores were higher on the tent stimuli than the house stimuli. This did not change with development. In a series of newly-designed EFT tested on adults (De-Wit et al., 2017; Huygelier et al., 2018), target shapes with a greater number of sides were found to be easier to disembed. Based on this, it might be expected that higher scores would be achieved on the CEFT stimuli with the house target shape. However, the same study also found that asymmetrical target shapes were harder to disembed, so it is possible that poorer performance on the CEFT house stimuli was due to the asymmetric nature of the target stimulus. This requires further exploration with child participants, to examine whether responses to target stimuli of different shapes and different levels of line continuation into the complex figure, elicit similar patterns of responses to adults.

2.4.2.2 Field independence gender differences

There were no gender differences identified in either task, supportive of previous studies using the CEFT with this age group (Bigelow, 1971) and with early adolescents (Flexer & Roberge, 1983). However, it contradicts other studies with children using the CEFT (Amador-Campos & Kirchner-Nebot, 1997; Cairns et al., 1985; Jantan, 2014) and the Block Design task (Booth, 2006). Significant gender differences in these studies were small, and were in fact only observed in some of the age groups within the same study (Amador-Campos & Kirchner-Nebot, 1997; Booth, 2006). This suggests that the effect of gender in FI in children is small, and only sometimes reaches significance.

2.4.2.3 Individual differences across field independence tasks

Correlations between the two FI tasks were significant and remained so after controlling for age and gender. This indicates that the two tasks were measuring common processes, despite being quite different in design and measuring two distinct aspects of FI, namely disembedding and segmenting. It is not possible to conclude from the current analyses whether the overlapping variance represents FI processes of separating a simple target from its complex surroundings. The association may also be related to domain-general processes of IQ and EF, which have been found to associate with FI tasks (Guisande et al., 2008; Imanaka et al., 2017; Leo-Rhynie, 1985; Miyake et al., 2001). This will be explored further in **Chapter 3**.

2.4.3 Systemizing, empathizing, and autism

2.4.3.1 Systemizing, empathizing, and autism development and gender differences

There were no differences in systemizing, empathizing, or autism traits by Year or by gender. This contrasts with other studies with children which have identified higher empathizing scores in females and higher systemizing scores in males (Auyeung & Wheelwright, 2009; Escovar et al., 2016; Akio Wakabayashi, 2013). It is possible that the sample size here (N = 65) was not sufficiently large to reveal significant gender effects.

2.4.3.2 Individual differences across SQ, EQ, and AQ

As predicted, systemizing and empathizing did not correlate with each other, indicating that they are separate dimensions, supportive of the current literature (Auyeung & Wheelwright, 2009; Escovar et al., 2016). However, contrary to expectation, higher systemizing scores also did not significantly associate with higher autism trait scores. This conflicts with some previous studies where autistic children and adults scored highly on the systemizing questionnaire (Auyeung & Wheelwright, 2009; Baron-Cohen, 2002; Greenberg et al., 2018; Wakabayashi et al., 2006). Here, higher empathizing scores related to lower scores on the autism trait questionnaire, in line with previous studies (Auyeung & Wheelwright, 2009; Baron-Cohen, 2009).

This overall pattern of association, where autism scores negatively associated with empathizing but had no association with systemizing, was also observed in other studies with children between the ages of 7 and 11 years (Evers et al., 2015; Vanegas & Davidson, 2015). In a study with adults, Greenberg et al. (2018) found that the negative correlation between autism traits and empathizing is stronger than the positive correlation between autism traits and systemizing. This may explain why the association between systemizing and autism has sometimes not been observed.

2.4.4 Comparison across constructs

The correlations across the whole-part constructs revealed very little overlapping variance after controlling for age in months and gender. Nonetheless, some commonality was identified.

Participants who were better at segmenting a pattern, also achieved better scores on both the global and local trials of the divided attention task, as well as the local trials of the selective attention task. The correlations between the DOT and the two local Navon measures may be due to common segmenting task demands, whereby both require participants to isolate and identify a local element from the context of the whole. However, the correlations are fairly weak which may be a consequence of differing levels of challenge; the local elements in Navon tasks are easier to segment than in the DOT, as they are already slightly separated from each other. The correlation between the DOT and the global trials of the divided attention task may indicate that children used a 'virtual' trial-and-error strategy where they imagined combining different elements together to match the whole pattern, as per a study comparing different strategies in completing the Block Design task (Rozencwajg & Corroyer, 2002). Alternatively, the correlations between the DOT and both levels of the divided attention task may reflect common inhibitory control processes required to overcome the distracting whole in the DOT and the non-target level in the divided attention task.

There were two significant cross-construct associations involving the questionnaire measures. First, those who achieved a quicker LISAS score on the divided attention local trials scored relatively higher on the SQ-C than the EQ-C. This partially supports a previous study with adults which identified a positive correlation between systemizing and local processing (Billington et al., 2008), however this direct association was non-significant in the present study. Second, those who achieved a quicker LISAS score on the divided attention local trials scored higher on the autism trait questionnaire. This is somewhat supportive of studies which identified a local advantage in autistic participants (Koldewyn et al., 2013; Plaisted et al., 1999; Wang et al., 2007), although two of these studies used selective attention paradigms, which did not associate with the AQ-10 in the current study. It is important to note that systemizing, empathizing and autism measures here are not based on an individual's measurable response to a stimulus in an experimental task. They instead reflect a parent or carer's perception of their child's traits, which may have an impact on the accuracy of those variables, and also their association with other whole-part measures.

No further cross-construct associations were identified after controlling for age and gender effects. This supports previous studies which also did not identify common whole-part responses over a number of tasks (Chamberlain et al., 2017; Milne & Szczerbinski, 2009), and indicates that global and local processing, FI, and systemizing are measuring largely distinct processes.

2.4.5 Conclusion

In summary, the age at which a local-to-global change is observed in children aged 5 to 10 years varied across Navon task. This supports the notion that attentional demands of the task have an influence on responses, and that these task demands may mask the effects of underlying global or local advantage. There were changes in global and local responses which consistently occurred between Years 1 and 3 (ages 6-7 years) across tasks, which indicates that this may be a pivotal age in children's development with respect to visual processing. However, it is not clear from these data whether the changing responses reflect attentional development or perceptual development. Despite the lack of a common global and local processing development pattern in this age group, these findings support the predicted diminishing local processing advantage with age.

FI increased with age, and accuracy positively associated between tasks measuring disembedding and segmenting. However, it is unclear whether there are additional domain-general factors which may explain both the developmental trajectory of FI and its cross-task association. In contrast, systemizing and empathizing did not differ by age, suggesting that they are less influenced by domain-general development.

It is clear that there is a lack of clarity regarding the relative influence of whole-part processing and domain-general attentional processes on observed behavioural responses. This will be investigated further in **Chapter 3**, which focusses on the extent to which FI tasks can be explained by IQ and EF. Additionally, performance on the CEFT will be assessed using eye-tracking technology to reveal whether individual differences in working memory and inhibitory control relate to differences in fixation patterns on the task, which in turn may relate to response accuracy on the task.

Chapter 3: Field independence and domain-general processes

3.1 Introduction

As noted in **Chapter 1**, responses on visual field independence (FI) tasks are dependent on a combination of perceptual and cognitive factors. Perceptual factors, such as targets blending into their complex background due to automatic Gestalt processes, may increase the difficulty of disembedding or segmenting. Cognitive processes are required to analyse a stimulus, remember the features of the target, and overcome distractions which can also impact speed and accuracy of response. This chapter will examine associations between FI tasks and domain-general factors, to understand the extent to which FI responses are related to individual differences in general intelligence and executive functions (EF). Further, eye-tracking technology will be used to measure eye movements on an FI task to gain an understanding of how these domain-general factors may associate with differences in gaze patterns and task accuracy.

3.1.1 Domain-general associations with field independence

Studies discussed in **Chapter 1** revealed associations between FI tasks and domain-general factors such as general intelligence (IQ) and EF. For example, a study with 11- to 13-year-olds identified positive correlations between the Group Embedded Figures Test (EFT) and a general IQ measure (Flexer & Roberge, 1980). Similarly, positive correlations were identified in 7- to 9-year-olds between the Children's Embedded Figures Test (CEFT) and visuospatial IQ (Swyter & Michael, 1982). However, a study with 5- to 10-year-olds revealed no significant verbal IQ difference in CEFT scores between those in the high and low FI groups (Bigelow, 1971), indicating that FI may more closely associate with visuospatial rather than verbal intelligence measures.

Working memory (WM) measures the ability to manipulate information whilst it is held in the memory systems (Alloway, Gathercole, Kirkwood, & Elliott, 2008; Raghubar, Barnes, & Hecht, 2010). When completing FI tasks, individuals need to have a clear representation of the target shape, and remember which areas of the complex figure have already been examined. FI has been shown to involve visuospatial and executive elements of WM in undergraduate students (Huygelier et al., 2018; Miyake et al., 2001), and a study with 8- to 11-year-olds found an association between verbal WM and FI (Guisande, Páramo, Tinajero, & Almeida, 2007; Guisande et al., 2008), indicating that all aspects of WM have revealed associations with FI. Fewer studies have examined associations between FI and inhibitory control (IC), but differences have been observed in inhibitory control tasks between high and low FI undergraduates (Imanaka et al., 2017; Jia et al., 2014). This likely reflects the FI task demands where IC could appropriately guide attentional processes towards the embedded target, and stop a premature response in distractor areas.

Although the strength of evidence suggests that domain-general factors associate with achievement on FI tasks, there has been little analysis of the mechanism of that association. One suggestion is that domain-general processes determine how efficiently a stimulus is examined and analysed (Evans et al., 2013). Better WM allows shapes and patterns to be matched more effectively, and better IC allows irrelevant areas of the stimulus to be ignored. Together, these abilities could determine how long an individual looks at task-relevant parts of a stimulus, which in turn could explain individual differences in responses to FI tasks.

3.1.2 Field independence and eye tracking

3.1.2.1 Eye movement

Humans have a limited capacity for receiving and processing information, so selective attention is deployed to isolate and focus on task-relevant information within the cluttered environment (Chen, Meier, Blair, Watson, & Wood, 2013; Ludwig, Davies, & Eckstein, 2014). Visual information is perceived by light entering the eye and stimulating the cells in the retina at the back of the eye. The type and density of retinal cell determines the level of detail which can be perceived, so there is variation in the amount of detailed information perceived across the visual field (the area of visible space when gazing in a single direction). The level of clarity in a person's vision, known as visual acuity, is highest when the image falls on the fovea. This is an area of the retina approximately 0.5 mm in diameter, where the high density of photosensitive cells allows detailed information to be collected (Tobii, 2019b). Visual acuity is weaker in the parafoveal region which extends around the fovea, and is poor in the peripheral regions which extend beyond the parafoveal region to the edge of the visual field (Lai et al., 2013; Rayner, 2009). Therefore, once parts of a stimulus are selected for attention, we need to orient our eyes so that the area of interest falls in the fovea to best capture the information of interest (Helo, Pannasch, Sirri, & Rämä, 2014). This can be achieved by moving our head and body to gather information across a wide area, and moving our eyes to collect information from across the visual field (Solman, Foulsham, & Kingstone, 2017; Spector, 1990). This means our eyes move to locations predicted to contain features of interest so that information can be collected and processed (Chen & Choi, 2008; Henderson, 2017).

Eye tracking technologies can be used to record eye movement patterns. These patterns give an insight into the features of the environment or stimulus which are capturing an individual's visual attention. Therefore, eye movement analysis is an indirect measure of brain functions involved in visual processing, and a direct measure of the relationship between the stimulus and an individual's response (Eckstein, Guerra-carrillo, Singley, & Bunge, 2016; Luna et al., 2008). Using this information, inferences can be made about an individual's allocation of attention during a goal-directed task and when they are viewing their surroundings freely with no task goals (Helo et al., 2014; Hessels, Niehorster, Andersson, & Hooge, 2018).

There are two types of eye activity which are typically measured: fixations and saccades. Fixations refer to periods of time when the eye is stationary, although not completely still, and information is acquired from that space in the visual field (Eckstein et al., 2016; Holmqvist, Nyström, Andersson, Dewhurst & ven der Weijer, 2011; Rayner, 2009). The purpose of fixations is to allow fine details to be collected and processed when objects, stimuli, or locations in the environment fall onto the fovea (Ludwig et al., 2014). Although fixations are periods when eyes are relatively stationary, each fixation also includes a series of minute movements, known as microsaccades, tremors, and drift, which help to keep the eye aligned with the object of interest (Holmqvist et al., 2011; Tobii, 2019). The definition of a fixation with respect to its duration varies widely in the literature (Hessels et al., 2018) but typically ranges from 80-600 ms, although longer fixations up to several seconds have been observed (Holmqvist et al., 2011; Tobii, 2019).

Fixation patterns vary according to the type of activity being undertaken. For example, a typical fixation length in reading is 225-250 ms, but for scene perception, it is 300 ms (Rayner, 2009). Viewing patterns of the same stimulus differ according to task goals. This was famously demonstrated in a study by Yarbus (1967). Widely differing patterns of fixations were observed on the same static stimulus (a painting by Ilya Repin called *The Unexpected Visitor*) depending on the question that was asked of the participants, such as estimating the ages of the people in the picture or suggesting what the family might have been doing before the visitor arrived (Yarbus, 1967). Goal-directed variations in gaze across a stimulus were also observed in a category-learning task where undergraduates had to sort fictitious microorganisms into one of four categories based on a set of features. Fixations towards irrelevant parts of the stimulus were fewer and shorter than fixations towards relevant features (Chen et al., 2013). Distinct gaze patterns can also be observed where different parts of the stimulus may be more or less interesting to the observer. For example, a study with undergraduate students who were given no specific task-goal, revealed a greater preference for attending to pictures with a person present than to pictures without a person (Fletcher-Watson, Findlay, Leekam, & Benson, 2008).

Saccades are very quick eye movements between fixations that allow visual attention to be oriented and centred on a new area of the visual field (Chen & Choi, 2008; Holmqvist et al., 2011; Luna et al., 2008). Visual information is not collected during saccades under normal circumstances as the image would be blurry due to the speed of movement, with an average saccade lasting only 30-80 ms (Eckstein et al., 2016; Holmqvist et al., 2011; Rayner, 2009). Saccades can reflect either planned motor movements, for example when reading, or automatic stimulus-driven movements, such as looking at a flashing bicycle light (Eckstein et al., 2016; Luna et al., 2008; Rayner, 2009). Studies suggest that saccades are guided by the analysis of information in peripheral vision 80-100 ms before the saccade is actioned (Ludwig et al., 2014). This is important for moving around an environment, hazard-awareness, and navigation, and enables the rapid orienting of the eyes so the most important information falls on the fovea. We shift our gaze on average three times every second (Holmqvist et al., 2011; Ludwig et al., 2014).

Patterns of fixations and saccades vary by time spent on an activity. For example, in the first two seconds of freely surveying a new scene, when attention is more likely to be directed by bottom-up processes, fixations are shorter and saccades longer (distance) in order to gain information about the scene and to detect movement. After the first two seconds, when object identification is more important, fixations are longer and saccades shorter, associated with the involvement of top-down processes (Helo et al., 2014). In a study involving a category-learning task, individuals fixated across all parts of the stimulus early in the task, but restricted fixations to relevant parts of the stimulus over time (Rehder & Hoffman, 2005). This change in viewing patterns with familiarity and expertise was also observed in a small-scale multiple choice study comparing self-categorised novices and experts, where fixations and saccades were fewer in number in the expert group compared with the novice group (Tai, Loehr, & Brigham, 2006). These examples demonstrate that more effective and efficient gaze patterns develop over time to meet task demands; fixations are likely to be longer and fewer in number as task-experience increases, and the most relevant features of a stimulus are viewed relatively longer than irrelevant features.

3.1.2.2 Field independence and eye movement

There are very few studies that have related performance on FI tasks to the patterns of fixations and saccades observed whilst participants complete the task. Nisiforou & Laghos (2016) divided adults into three groups according to their performance on a pen-and-paper hidden figures task, where one of a choice of four simple shapes were embedded within a complex figure. The participants then completed a different hidden figures test while their eye movements were recorded. Group-level analyses revealed a greater number of fixations in each of the lower performing groups compared with the higher performing group (Nisiforou & Laghos, 2016). A more efficient pattern of fixations and saccades was also observed in the high FI group of a small-scale study measuring eye movements during a hidden face search task. Here, the high FI group (as determined by response time (RT) on the task) spent relatively less time on distractor areas of the stimulus compared with the low FI group (Nisiforou & Laghos, 2013). A further study with adults revealed no differences in fixation length between high and low FI participants, however there was a significant difference in the spatial pattern of fixations, with high FI individuals fixating more on information-rich areas than low FI individuals (Conklin, Muir, & Boersma, 1968). In sum, there is limited information about how FI relates to eye movements. However, there is some evidence that a more efficient viewing strategy, i.e. longer fixations, and fewer fixations on distractor areas, is observed in individuals with higher accuracy or quicker RT.

Although eye movements recorded during FI tasks have not previously been measured in children, there are eye-tracking studies which have examined how eye movement changes over development as fixation control, and saccadic speed, length, and accuracy mature through childhood and into adolescence (Helo et al., 2014; Luna et al., 2008). This developmental change in eye movement control results from a combination of physiological and cognitive development, including WM, IC, and processing speed, which continue to develop into adolescence (Luna et al., 2008). For example, at the age of 8- to 10-years-old, fixation control has improved compared with younger children such that fewer saccades are made to non-task related distractors. This reflects the growing maturity of higher-order cognitive processes which inhibit eye movement to distracting stimuli (Helo et al., 2014; Luna et al., 2008). The accuracy of saccades continues to develop until the age of 8 years, while RT to initiating saccades does not reach adult-levels until adolescence (Alahyane et al., 2016).

3.1.3 Current studies

FI is conceptualised as an ability to discriminate a target from its context (Huygelier et al., 2018; White & Saldaña, 2011; Witkin & Asch, 1948). However, little is known about the extent to which this FI ability is simply a measurable behavioural response to a particular type of stimulus which requires a combination of IQ and EF to process and respond effectively. Research indicates that there is an association between FI and domain-general processes (Guisande et al., 2008; Imanaka et al., 2017; Miyake et al., 2001; Swyter & Michael, 1982). Although much of this research has been undertaken with adults, the limited evidence suggests that this association is also apparent in children (Flexer & Roberge, 1980; D. Goodenough & Karp, 1961; Guisande et al., 2007). The first aim of this chapter was to evaluate associations between behavioural responses to FI tasks and domain-general abilities. This was undertaken using two distinct samples of children. The first was the group of 5- to 10-year-olds presented in **Chapter 2**. Associations were examined between two measures of FI and IQ, visuospatial WM, verbal WM, and semantic IC. The second was a group of 9- to 10-year-olds. In this group FI accuracy was correlated with four measures of EF; visuospatial WM, verbal WM, semantic IC, and response IC. In both samples, it was hypothesised that higher scores on the FI tasks would associate with higher scores on the IQ and EF measures, based on previous studies with adults and children.

The second aim of this chapter was to examine whether individual differences in EF associate with different fixation patterns on the CEFT, and whether these in turn associate with accuracy on the task. There is some indication that high FI adults view stimuli in a more efficient manner than low FI adults (Nisiforou & Laghos, 2016). This has not previously been examined in children. It is possible that these individual differences in eye movement patterns, and associated differences in performance, can be explained by differences in WM and IC. The goal of the CEFT is to measure how accurately participants disembed the target shape from a complex figure. The length and direction of fixation during the task may depend on an ability to accurately remember the target and an ability to overcome the influence of distracting shapes. This was examined here in 9- and 10-year-olds using the ‘tent’ stimuli from the CEFT, and tests of verbal and visuospatial WM, and semantic and response IC. At the age of 9 to 10 years, fixation control and saccade accuracy are better developed than in younger age groups although they have not yet reached full maturation (Aring & Grönlund, 2007; Ygge, Aring, Han, Bolzani, & Hellström, 2005). With a narrower age range, differences in performance were more likely to be task-related rather than due to contrasting levels of maturity of the visual system. The predicted association between the behavioural FI measure and fixation patterns was that those scoring more highly on the CEFT would have fewer, longer fixations, in line with findings with adult participants where more efficient gaze patterns were observed in more experienced and more expert participants. It is also likely that higher FI children would have more and longer fixations on the embedded target shape and shorter and fewer fixations on distractor areas. The expected association between behavioural EF measures and gaze patterns was that those with higher WM scores would need to look at the non-embedded target displayed to the side of the main figure less often, while those with higher IC scores would have fewer and shorter fixations on distractor shapes.

3.2 Method

3.2.1 Participants

The data examined in this chapter were collected as two separate studies; study 2 involved participants from sample 1 and study 3 involved participants from sample 2. Both studies were approved by the local ethics committee.

Study 2 participants, as introduced in **Section 2.2.1**, were from a single state primary school. Participants were drawn from Year 1 (5-6 years), Year 3 (7-8 years), and Year 5 (9-10 years), and analyses here used data from Time 1 (T1) only. Parents and carers had the opportunity to opt-out of the study, and all children gave verbal consent before any testing began.

Study 3 participants were recruited from four state primary schools. Parents and carers gave their permission for their child to be invited to participate via an opt-in reply slip, and all children completed a simple age-appropriate form giving their consent to take part. The study involved 71 children from Year 5 (aged 9 to 10 years). There were no gender differences between schools, $\chi^2(3) = 1.43$, $p = .698$. The inclusion criterion was that participants did not have a statement of special educational needs and disabilities (SEND).

Table 3.1: *Participant demographics of study 3*

	School 1	School 2	School 3	School 4	**Total**
Number of participants	6	13	29	23	**71**
Males : females	2 : 4	7 : 6	16 : 13	10 : 13	**35 : 36**
Mean age in months (SD)	124.5 (3.7)	123.9 (2.6)	124.6 (3.4)	123.7 (2.8)	**124.1 (3.1)**
Mean age in years, months	10, 4	10, 3	10, 4	10, 3	**10, 4**

3.2.2 Summary of tasks

A summary of the tasks completed by participants in each study can be found in **Table 3.2.**

Table 3.2: *Summary of tasks and measures for each study.*

Study	Measure	Task
Study 2	Disembedding	Children's Embedded Figures Test (CEFT)
	Segmenting	Design Organisation Test (DOT)
	General intelligence	Visuospatial: Raven's Coloured Progressive Matrices
		Verbal: British Picture Vocabulary Scale 3 (BPVS-3)
	Visuospatial WM	Reverse spatial span
	Verbal WM	Reverse digit span
	Semantic IC	Animal size Stroop
Study 3	Disembedding	CEFT: behavioural and eye-tracking measures
	Visuospatial WM	Reverse spatial span
	Verbal WM	Reverse letter span
	Semantic IC	Animal size Stroop
	Response IC	Whack-a-mole

3.2.3 Field independence tasks

3.2.3.1 Behavioural tasks (study 2 and 3)

Disembedding was measured with the CEFT in study 2 as described in **Section 2.2.4.1**. The same subset of 13 images was presented to all participants on paper, enabling comparisons in percentage accuracy to be made between individuals and across Year groups.

Segmenting was measured with the Design Organisation Test (DOT) as described in **Section 2.2.4.2**. This was a time-limited task, with the total number of correctly identified squares as the measure of success on the task.

The CEFT was also used in study 3, but with a different selection of stimuli presented on a screen, as described below in **Section 3.2.3.2.2.**

3.2.3.2 Eye-tracking task (study 3)

3.2.3.2.1 Eye-tracking equipment

Remote eye trackers measure the reflection of light as it hits the curved surface of the cornea (Hansen & Ji, 2010). This creates a 'corneal reflection', a small glint next to the pupil (Hooge, Holmqvist, & Nyström, 2016). Eye trackers then calculate the distance of the corneal reflection from the centre of the pupil, as this varies when the eye changes position, creating a record of the time and spatial location of where the participant is looking, or their ‘gaze point’ (Holmqvist et al., 2011; Tobii, 2014). The number of gaze points that can be recorded by an eye-tracker varies according to its specification. The eye-tracker used here recorded binocular eye movements with a sampling rate of 60 Hz, which means it recorded a gaze sample 60 times per second (Holmqvist et al., 2011; Tobii, 2014).

An eye-tracking tent was set up to create a space with equalised levels of light. This acts to reduce interference from additional light sources such as windows and thus enables more accurate measurement of fixations and saccades. A Tobii X2-60 eye-tracker (Tobii Technology, Stockholm, Sweden) was set up in the eye-tracking tent, attached to an 18-inch computer monitor with a 1920 x 1080 resolution. This was connected to the researcher’s laptop outside the tent, from where the participants’ eye movements and progress through the task could be monitored (**Figure 3.1A**). Participants sat on a stool 50cm away from the monitor, such that the area within which the picture stimuli were presented subtended a visual angle of 10.6° x 15.9° portrait and 15.9° x 10.6° landscape. As the stimuli were taken from the paper-based CEFT, they included a mix of portrait and landscape designs. The total screen area, which included the target shape for reference, subtended a visual angle of 28.0° x 15.9°, and the target shape itself subtended a visual angle of 2.5° x 2.5° (**Figure 3.1B**).

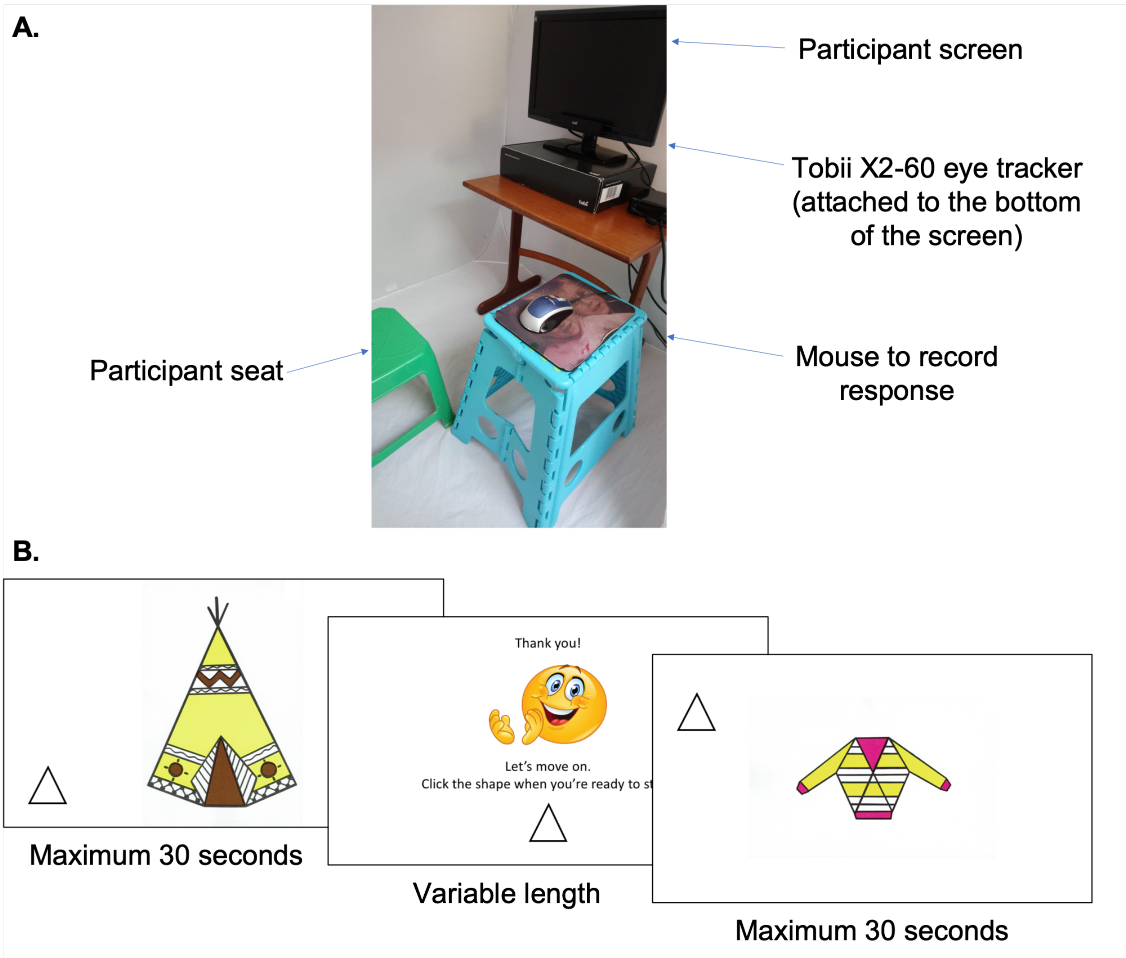


Figure 3.1: *Eye tracking study of Children's Embedded Figures Test (CEFT).* ***A.*** *A photo of the equipment, with the eye-tracking tent, participant seat, screen, and mouse, and the eye-tracker. The researcher sat outside the tent (position not shown) monitoring the participant's progress.* ***B.*** *Example trials from the task.*

The radius of an eyeball can vary between participants, therefore it was necessary to calibrate the eye tracker for each participant so its measurements of eye movements were accurate (Holmqvist et al., 2011; Tobii, 2019c). Calibration needed to be completed in the same conditions as the test so that the levels of infra-red light remain constant, and calibration points need to cover the whole area where the participant could look (Holmqvist et al., 2011). Each participant completed a five-point calibration procedure, which was the most time-efficient calibration method of covering all areas of the screen where stimuli were located. If the calibration was not successful, as recognised by the qualitative visual assessment provided by Tobii Studio, positioning of the participant or the equipment was adjusted and the calibration re-run. All participants had to achieve a good overlap of their corneal reflection relative to their gaze point (Dalrymple, Manner, Harmelink, Teska, & Elison, 2018) before taking part in the task. This was achieved for all participants.

The task was presented using Tobii Studio version 1.7.3. The stimuli were taken from the CEFT (Karp & Konstadt, 1963) which had been scanned to create a computerised task. The task began with a set of instructions followed by a set of two practice screens to ensure the participants could match the target shape from a choice of four triangles. Once completed, there were two practice trials, in which participants were asked to find the target triangle shape in the picture. Participants responded using a mouse to click on the embedded target shape. There was no time limit during the practice trials, and the researcher was able to guide participants when necessary using a second mouse operated from the researcher laptop outside the tent.

3.2.3.2.2 Eye-tracking CEFT

The main task was composed of 11 trials, consisting of all the 'tent' stimuli from the CEFT. The 'house' stimuli from the CEFT were excluded from this task, as it would be less obvious that participants had correctly located the larger house shape from a single mouse click. Each stimulus contained the target shape in one of the four corners, with the picture in which the target shape was embedded presented centrally (**Figure 3.1B**). By presenting the target shape and the picture together, it was possible to measure how often the target shape was referred to while the participant looked for the embedded shape in the picture. The position of the target shape changed for each trial in a pseudo-random order, ensuring that a single location of the target shape did not impact performance. The target shape was not placed in the extreme corner as accurate measurement in the extremes of the screen, i.e. the top and bottom corners, is a challenge due to interference from eye lashes and the position of the corneal reflection (Holmqvist et al., 2011). As with the CEFT task in study 2, participants had a maximum of 30 seconds to find the shape. When they found the embedded target shape, they clicked on the location using the mouse, and the task moved on to a 'rest' screen. The researcher timed each trial from outside the eye-tracking tent, and if no response was made within that time, the researcher progressed the task onto the rest screen using a mouse click from the researcher laptop outside the tent. Tobii Studio did not have the capacity to automatically move the test forwards if no response was made within a maximum RT when it was also programmed to move forward to the next image on a mouse click. Therefore, the researcher moved the test forward only when no response was made within 30 seconds. No feedback was given regarding the accuracy of response, and the participant clicked to start the next test trial when they were ready. Both behavioural and eye-tracking data were collected.

3.2.3.3 Eye-tracking analysis

Once the data were collected, areas of interest (AOIs) were set up to group the data. This enabled gaze behaviour to be analysed across different areas or features of the stimulus. There are a number of different methods for determining AOIs, including researcher-defined hand-drawn areas, computer-applied, or through the use of a grid (Hessels, Kemner, van den Boomen, & Hooge, 2016). Here, the AOIs were created to analyse the fixations in pre-determined areas of the stimulus which related to the task demands of the CEFT. There were four AOIs for each trial (**Figure 3.2**): AOI-1 was the non-embedded target shape located in one of the four corners of the screen; AOI-2 was the embedded target shape; AOI-3 were distractor triangle shapes in the picture stimulus; and AOI-4 was everywhere else on the screen.

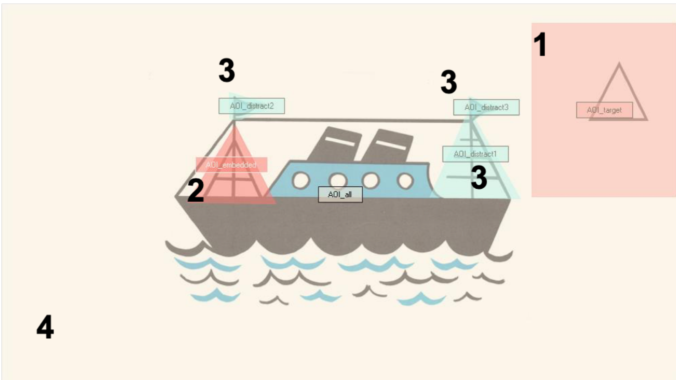


Figure 3.2: *Example trial with four areas of interest (AOI).*

AOI-1 covered the largest area to capture any eye movements made away from the picture and towards the target shape. This was designed to include eye movements enabling the participant to capture the target shape in their peripheral vision which is likely sufficient for reminding themselves of the size and shape of the target. The shape and size of AOI-1 and AOI-2 was kept constant across trials, although the location of AOI-1 varied depending on which corner the target shape was located. The number and shape of AOI-3s varied between trials as a result of the design of the stimulus pictures. When designing novel tasks for use with eye-trackers, stimuli should have sufficient space between AOIs, so that information collected in peripheral vision does not interfere with looking patterns (Holmqvist et al., 2011). Due to the nature of the stimuli here, it was not always possible to space out further the AOIs, but there were never any overlapping AOIs.

The data were analysed using Tobii Studio version 3.4.5. The gaze points were combined into fixations and saccades with a fixation filter which uses specified dimensions in space and time to group measurements (Hessels et al., 2018; Tobii, 2015). The velocity-threshold identification (I-VT) fixation filter, which determines how gaze-points are grouped into fixations and saccades based on the velocity of movement, was applied to the dataset (Salvucci & Goldberg, 2000). The default setting was used whereby velocities less than 30°/second were classed as fixations, while velocities above this threshold were classed as saccades (Tobii, 2012). The minimum length of fixation was set at 100 ms, in line with other studies (Bott et al., 2017; Rayner, 2009; Salvucci & Goldberg, 2000; Wass, Smith, & Johnson, 2013). Adjacent fixations were merged when they were less than 0.5° apart, and a 75 ms limit was set as the maximum saccade time, in line with the default fixation filter settings (Tobii, 2012).

Eye tracking data can be analysed both quantitatively and qualitatively. The number and length of fixations can be used as numerical variables. Heat maps and scan paths can be used to provide a visual representation illustrating how single participants or groups view the stimulus. Heat maps show locational information of fixations with green areas representing shorter or fewer fixations and red areas representing longer or more numerous fixations. These can either represent the raw total number of fixations in each area, or the proportion of time that gaze rests in an area as a percentage of total fixation length. Scan paths reveal the directions of saccades between fixations.

In this study, the number and length of fixations were converted into percentage data so that comparisons of gaze patterns could be made without the confound of varying RT of that trial. The quantitative eye-tracking variables for each AOI were the proportion of fixations of the total number of fixations in each trial, and the proportion of the total length of duration for that trial. Additionally, the proportion of the number of visits to each AOI as a percentage of the total number of visits in that trial was calculated. The 'number of visits' measure combines fixations within an AOI for each visit, so it represents the number of times a participant looks at an AOI and then looks away by collapsing the fixations within a single AOI into a single record. In addition to the eye tracking metrics, performance data were collected in the form of CEFT accuracy, as measured by percentage of mouse clicks in AOI-2 out of a total of 11. A % of distractor responses was also calculated, representing mouse clicks on AOI-3, enabling an assessment of whether error responses related to difficulties in ignoring non-target triangles. Fixations on AOI-4 were not analysed, as the areas of interest relating to the research questions were areas of the complex picture and the target shape. Saccade data were not examined as part of this study, as Tobii Studio does not calculate a quantitative saccade measure.

3.2.4 General intelligence tasks (study 2)

Non-verbal IQ was measured using Raven's Coloured Progressive Matrices (RCPM; Raven, Raven and Court, 1998), which has a split-half reliability of r = .82 in children aged 5- to 8-years-old (Carlson & Jensen, 1981). Receptive vocabulary, representing general verbal ability, was measured using the British Picture Vocabulary Scale 3 (BPVS-3; Dunn, Dunn, Styles and Sewell, 2009). BPVS-3 norms for individuals aged 3-16 years indicate excellent reliability of .91 (Dockrell & Marshall, 2015). Both were administered following the standard guidelines, and raw scores were recorded for each.

3.2.5 Executive functions tasks

3.2.5.1 Verbal and visuospatial working memory (study 2 and 3)

Verbal WM was measured using a backwards digit or letter span task. In study 2, sequences of digits from 1 to 9 were read to participants. In study 3, nine consonants were selected, each with a distinctive-sounding names: B, F, K, M, Q, R, T, X, Y. Participants were read a sequence of letters to remember. They responded by verbally repeating the sequence of digits or letters in reverse order. There was no time limit for responding.

Visuospatial WM was measured using an age-appropriate computerised backwards spatial span. Participants watched a frog complete a sequence of jumps on a 3x3 grid of nine lily-pads on a laptop screen (**Figure 3.3A**). They then clicked on the lily-pads in reverse order using the mouse, to show where the frog had jumped.

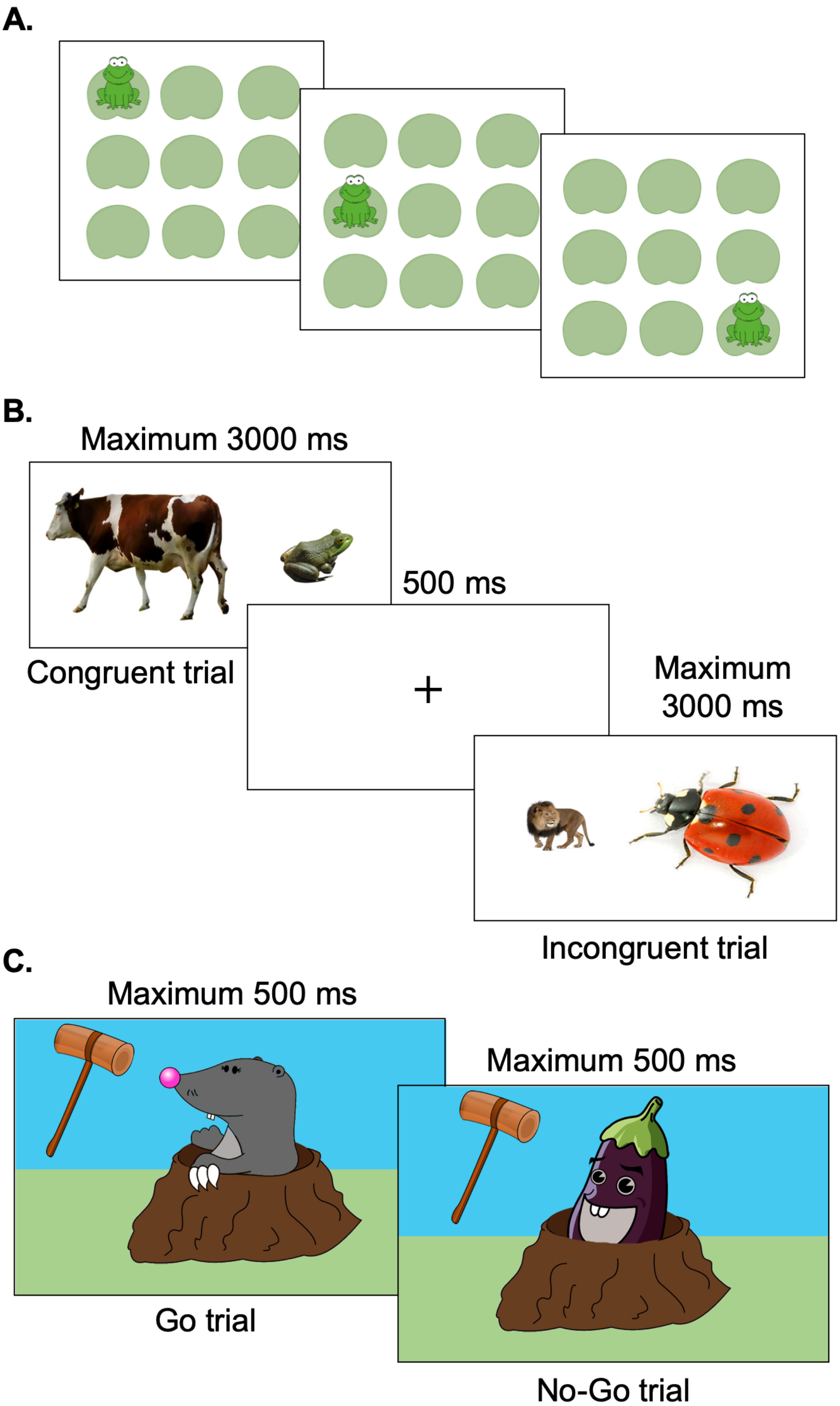


Figure 3.3: *Example trials of the executive functions tasks.* ***A****. Example sequence from the reverse spatial span task (visuospatial working memory).* ***B****. Example trials from the animal size Stroop task (semantic inhibitory control).* ***C****. Example trials from the whack-a-mole task (response inhibitory control task).*

In both tasks, participants completed practice trials with a sequence length of two, and could repeat these until they understood the task up to a maximum of five repeats. Participants then began the experimental trials with a sequence length of two. Each sequence length was repeated four times, and each new block increased in sequence length by one. Participants progressed to the next sequence length if they performed at least three trials correctly in that block. If they made two or more errors in a block, the task did not continue. The total number of correctly recalled sequences was used as a measure of performance on each of the WM tasks. Test/retest correlations partialling out Year were r_p = .536 for the verbal WM task and r_p = .357 for the visuospatial WM task, which were calculated from a separate study where more than 300 Years 3 and 5 children completed the tasks at a 3 to 5.5 month interval.

3.2.5.2 Semantic inhibitory control (study 2 and 3)

Semantic inhibitory control (IC) was measured using a computerised animal size Stroop task presented on a laptop (Merkley et al., 2016). Participants were asked to identify which of a pair of animals was the largest in real life by pressing the C- or M-key on a keyboard using their left or right hand respectively. Trials could either be congruent, where the relative size of the animal images matched the relative size in real life, or they could be incongruent and required children to inhibit the size of the image (**Figure 3.3B**). The stimuli consisted of animal photos with white backgrounds: a lion, elephant, horse and cow as large animals, and a ladybird, mouse, rabbit, and frog as small animals. The size of each large image was 72 mm x 54 mm which subtended a visual angle of 8.2° x 6.2° based on a viewing distance of 50 cm. The size of each small image was 29 mm x 21 mm, which subtended a visual angle of 3.3° x 2.4°. Each trial was presented for a maximum of 3 seconds, and if no response was made within that time, the task would proceed to the next trial and an error would be recorded. An inter-trial fixation cross was presented for 500 ms. Participants completed six practice trials and could only progress onto the main task if they responded correctly to at least four of these. The practice block could be repeated until the participant understood the task, up to a maximum of five times. Each participant then completed two blocks of 36 trials, with an equal number of congruent and incongruent trials. Children were asked to respond as quickly, but as accurately, as they could. A score was calculated as the percentage difference between congruent and incongruent RT trials ((incongruent – congruent) / incongruent), with positive scores denoting a RT cost of incongruent trials. Split-half analyses were carried out on the incongruent RT cost percentage for each study, which revealed correlations of .882 for study 2 and .764 for study 3.

3.2.5.3 Response inhibitory control (study 3)

Response IC was measured using a computerised ‘whack-a-mole’ task based on a task used by Shapiro, Wong, & Simon (2013), with a screen resolution of 1280 x 1024. Participants had to press the space bar when they saw a mole rising from its hole, but had to resist pressing the space bar when an aubergine rose from the hole (**Figure 3.3C**). The moles sometimes wore a wig or a hat, to help maintain interest on the task. There were 100 trials with 75 Go trials (mole) and 25 No-Go trials (aubergine). A mallet ‘whacked’ the mole on correct go trials, and an ‘oops’ speech bubble appeared for any commission errors when participants ‘whacked’ the aubergine. The maximum RT was 500 ms, and if no response was made on Go trials within that time, the trial moved on and an error was recorded. Accuracy on two types of errors were recorded: omission errors when no response was made on Go trials, and commission errors when a response was made on No-Go trials. In order to take account of both error types, performance was measured using a D-Prime calculation (Stanislaw & Todorov, 1999). Z-scores of the hit rate (correct Go trials) and false hit rate (incorrect No-Go trials) were calculated in Excel using the NORMSINV function. As this is unable to cope with values of 0 or 1, loglinear scores were calculated where 0.5 was added to the hit rate and false hit scores, then 1 was added to the total possible scores on each (Hautus, 1995). A D-Prime score was then calculated as follows, such that higher figures denote better performance on the task and therefore better response IC:
NORMSINV (loglinear hit rate / 100) – NORMSINV (loglinear false hit rate / 100). A split-half analysis was carried out on the D-Prime calculation, revealing a correlation of .987.

3.2.6 Procedure

Testing took place at the participants’ primary school in a quiet space during lesson times. Study 2 data were collected with the other whole-part tasks described in **Chapter 2**, and the mathematics and science tasks described in **Chapter 4**. The whole battery of tasks were completed over several sessions, with the CEFT and DOT tasks being presented in separate sessions from each other.

The study 3 data were collected along with mathematics questions (see **Chapter 4**). Children completed the tasks over three sessions of variable lengths. The four EF tasks were completed in the first session, the mathematics task in the second session, and the eye tracking CEFT in the final session. There were two orders of the EF tasks which were counterbalanced across the participants. The two WM and two IC were separated in the task order.

3.2.7 Statistical analyses

Behavioural accuracy data were analysed with correlations and regressions, to determine associations between FI responses and domain-general abilities.

The associations between CEFT performance and eye movements were analysed both quantitatively and qualitatively. Associations with the quantitative fixation data were analysed using a series of correlations and partial correlations. Fixations were also assessed using heat maps. Scan paths were not used, as the main eye tracking variables of interest were percentages of fixations in different AOIs rather than saccadic patterns between AOIs.

In order to view group-level heat maps, participants were divided into three groups using their CEFT accuracy score, in line with previous studies: low FI, mid FI, and high FI. There are several methods for splitting participants into FI groups, such as plus or minus a quarter of a standard deviation (SD) to create the group boundaries (Alamolhodaei, 2011), or using pre-set scores (Angeli, Valanides, Polemitou, & Fraggoulidou, 2016; Nisiforou & Laghos, 2016), or dividing the participants' scores into three equal groups (Guisande et al., 2008). In order to create three fairly equally-sized groups, this study created group boundaries based on the Mean plus or minus half a SD (Daniels & Moore, 2000). A subset of trials were selected for the qualitative heat map analysis by sorting overall accuracy by trial. The most and least accurate trials were selected along with the median trial. This created a group of three trials for heatmap analysis with a spread of difficulty.

Qualitative heat maps were created using relative duration, which is the amount of time an individual looks at a particular area as a proportion of their total viewing time of the stimulus. For example, if one participant looked at AOI-2 for 500 ms which was 25% of their total viewing time of that stimulus, this would carry the same weight as a participant who viewed AOI-2 for 1000 ms which was also 25% of their total viewing time. This is similar to the quantitative measure of duration length % in each AOI described in **Section 3.2.3.3**, but relative duration in a heat map is calculated across the stimulus, rather than grouped by AOIs. Relative duration is a more informative measure than absolute fixation length when participants have different total viewing times.

Any outliers on FI tasks or the domain-general measures, were identified as scores $\pm$ 3.29 SD from the mean (Field, 2013). As outliers should not be removed from datasets unless they represent a misunderstanding of the task or errors in recording (Goodwin & Leech, 2006), the analyses were carried out with and without outliers to assess their influence (see **Section 2.2.7.2** for further detail). Significant results were interpreted as having alphas less than .05, except where Bonferroni corrections were made for multiple analyses.

3.3 Results

All children from study 2 (N = 135) and study 3 (N = 71) completed all tasks. A series of Analysis of Variance (ANOVAs) were carried out with Year and gender as independent variables (IV) and each of the FI, domain-general, and-eye tracking variables as the dependent variables (DV). All Year and gender main effects are reported in **Table 3.3A** for study 2, and **Table 3.3B** for study 3.

Table 3.3A: *Study 2 (time 1): Main effects of Year and gender on Children's Embedded Figures Test (CEFT), Design Organisation Task (DOT), visuospatial IQ, verbal IQ, visuospatial working memory (WM), verbal WM, and semantic inhibitory control (IC), with estimated margin means (M) and standard error (SE). Significant effects in bold.*

DV	**Year *F*(2,129)**	**Y1 *M* (*SE*)**		**Y3 *M* (*SE*)**		**Y5 *M* (*SE*)**	**Gender *F*(1,129)**
CEFT %	**35.74, $p < .001$, $\eta_p^2 = .357$**	33.3% (2.3)	<	49.3% (2.3)	<	60.2% (2.2)	$F < 1$
DOT	**80.03, $p < .001$, $\eta_p^2 = .554$**	10.7 (0.9)	<	16.8 (0.9)	<	26.6 (0.9)	2.20, $p = .141$
Visuospatial IQ	**73.49, $p < .001$, $\eta_p^2 = .533$**	17.6 (0.7)	<	26.0 (0.7)	<	28.8 (0.7)	$F < 1$
Verbal IQ	**104.38, $p < .001$, $\eta_p^2 = .618$**	84.4 (2.0)	<	105.3 (2.0)	<	124.6 (1.9)	$F < 1$
Visuospatial WM	**60.78, $p < .001$, $\eta_p^2 = .485$**	4.2 (0.4)	<	7.8 (0.4)	<	10.9 (0.4)	$F < 1$
Verbal WM	**41.79, $p < .001$, $\eta_p^2 = .393$**	5.4 (0.3)	<	7.9 (0.3)	<	9.6 (0.3)	1.47, $p = .227$
Semantic IC ms cost %	3.04, $p = .051$	11.3% (.013)	=	7.5% (.013)	=	7.3% (.013)	$F < 1$

Table 3.3B: *Study 3: Main effects of age in months and gender on Children's Embedded Figures Test (CEFT), quantitative eye tracking measures, and executive functions with estimated margin means (SE). Significant effects in bold.*

Task	**Dependent variable**	**Mean (SE)**	**Main effect of age (months)**	**Main effect of gender**
CEFT	Accuracy (%)	74.0% (3.1)	$F(11,48) = 1.15$, $p = .346$	$F < 1$
	Distractor responses (%)	20.5% (2.6)	$F(11,48) = 1.03$, $p = .440$	$F < 1$
Eye tracking	Total number of fixations	215 (10)	$F < 1$	$F < 1$
	Total length of fixation (ms)	62759 (3416)	$F < 1$	$F < 1$
	Average length of fixation (ms)	288.6 (5.4)	$F(11,48) = 1.74$, $p = .093$	$F < 1$
	Fixations in AOI-1 (%)	3.6% (0.3)	$F(11,48) = 1.36$, $p = .222$	$F < 1$
	Fixations in AOI-2 (%)	14.2% (0.7)	**$F(11,48) = 3.00$, $p = .004$, $\eta^2_p = .407$**	$F < 1$
	Fixations in AOI-3 (%)	23.5% (0.5)	$F(11,48) = 1.35$, $p = .228$	$F(1,48) = 2.79$, $p = .101$
	Fixation length in AOI-1 (%)	2.4% (0.3)	$F < 1$	$F < 1$
	Fixation length in AOI-2 (%)	23.4% (0.9)	$F < 1$	$F < 1$
	Fixation length in AOI-3 (%)	24.6% (0.7)	$F < 1$	$F(1,48) = 2.37$, $p = .130$
	Fixation visit in AOI-1 (%)	9.0% (0.8)	$F < 1$	$F < 1$
	Fixation visit in AOI-2 (%)	38.6% (1.0)	$F < 1$	$F < 1$
	Fixation visit in AOI-3 (%)	52.4% (0.9)	$F(11,48) = 1.50$, $p = .161$	$F(1,48) = 2.27$, $p = .138$
Verbal WM	Total correct	9.2 (0.4)	$F(11,48) = 1.20$, $p = .316$	$F(1,48) = 1.33$, $p = .255$
Visuospatial WM	Total correct	11.7 (0.6)	$F(11,48) = 1.18$, $p = .329$	$F < 1$
Response IC	D-Prime	3.0 (0.1)	$F < 1$	$F < 1$
Semantic IC	% cost of incongruent trials	6.7% (0.1)	**$F(11,48) = 2.04$, $p = .045$, $\eta^2_p = .318$**	$F < 1$

Notes: AOI-1: target shape; AOI-2: embedded target shape; AOI-3: distractor triangles; WM: working memory; IC: inhibitory control.

There was no significant effect of task order on responses to almost all measures (*p*'s > 209). The only exception was the response IC task in study 3, where there was a small effect of order, $F(1,69) = 5.12$, $p = .027$, $\eta^2_p = .069$ (order 1: $M = 3.2$, $SE = 0.1$; order 2: $M = 2.9$, $SE = 0.1$). The task was completed second in order 1 and third in order 2, so it is possible that mental fatigue played a role as the task was fairly long and repetitive so likely required more sustained attention than the other cognitive tasks. It may also have been impacted by lingering effects of the task undertaken immediately prior to the task; verbal WM for order 1 and visuospatial WM for order 2. Order of task was therefore included as a covariate in analyses involving response IC. There were no main effects of gender in any of the FI or domain-general measures, therefore gender was not included in the following analyses. In study 2, almost all DVs revealed Year effects, and there were age effects in study 3 on the semantic IC task and the percentage of fixations in AOI-2. Therefore, age was included as a co-variate in the following analyses.

3.3.1 Field independence and domain-general factors

3.3.1.1 Study 2

In order to evaluate associations between FI tasks and domain-general factors, Pearson's correlations were carried out with study 2 data. Partial correlations, controlling for age in months, were also examined. There was one outlier identified in the semantic IC data, where one boy in Year 1 was much slower (and less accurate) on congruent trials than incongruent trials, and therefore the RT cost calculation was greater than 3.29 SD from the Mean.

Table 3.4: *Study 2: Pearson's correlations between responses to the Children's Embedded Figures Test (CEFT), Design Organisation Task (DOT), visuospatial IQ, verbal IQ, working memory (WM), verbal WM, and semantic inhibitory control (IC). Partial correlations controlling for age in months are shown below the diagonal. *p < .05; **p < .01; ***p < .001. Significant correlations are highlighted in bold.*

	CEFT	**DOT**	**Visuospatial IQ**	**Verbal IQ**	**Visuospatial WM**	**Verbal WM**	**Semantic IC**
CEFT accuracy %		**.624 *** **	**.627 *** **	**.604 *** **	**.608 *** **	**.587 *** **	-.081
DOT correct	**.323 *** **		**.681 *** **	**.699 *** **	**.691 *** **	**.568 *** **	-.114
Visuospatial IQ	**.346 *** **	**.320 *** **		**.702 *** **	**.639 *** **	**.643 *** **	**-.208 * **
Verbal IQ	**.248 ** **	**.267 ** **	**.318 *** **		**.605 *** **	**.640 *** **	-.161
Visuospatial WM	**.320 *** **	**.358 *** **	**.280 ** **	.116		**.632 *** **	-.128
Verbal WM	**.336 *** **	**.202 * **	**.364 *** **	**.309 *** **	**.353 *** **		**-.237 ** **
Semantic IC % cost	.050	.049	-.098	-.009	-.013	-.149	

N = 135, df = 132

After controlling for age, each of the FI tasks significantly associated with all domain-general measures except for semantic IC, such that those achieving higher scores on the CEFT and DOT also recorded higher scores on the domain-general tasks. There were also significant correlations between the IQ variables and the WM variables. When the analyses were re-run without the outlier, the only change was that semantic IC significantly associated with verbal WM (r_p = -.172, p = .048).

In recognition of the potential non-linear associations between FI and EF, a series of regressions were carried out with either CEFT or DOT as the DV, with dummy Year variables and all EF measures in block 1. Interaction terms between the dummy Year variables and EF measures were entered stepwise in block 2. The dataset had outliers excluded. In the CEFT model, there were no significant differences in association between the Years (p's > .054). In the DOT model, the association with visuospatial IQ differed in Year 1 compared with the other Years, β = -.504, p = .017. **Figure 3.4** reveals a weaker association between DOT scores and visuospatial IQ in Year 1 compared with Years 3 and 5.

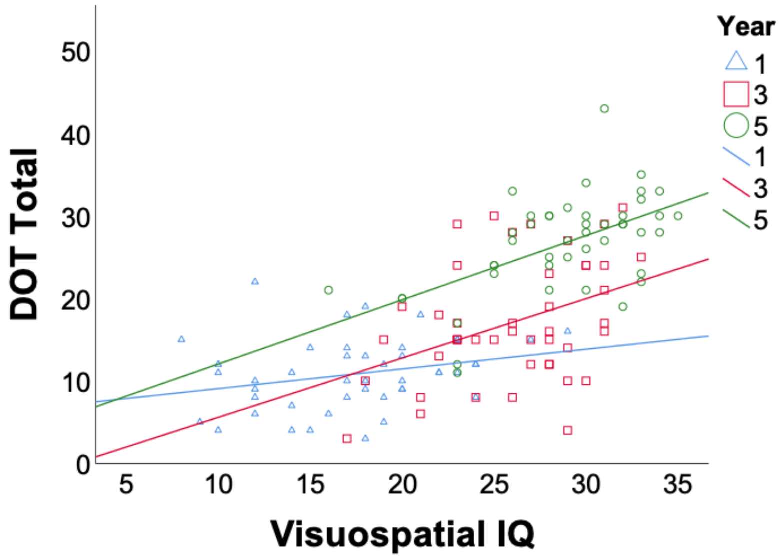


***Figure 3.4**: Comparison of associations between Design Organisation Test (DOT) and visuospatial IQ as a function of Year.*

To determine how much variance in the FI measures could be explained by age, IQ, WM and IC, multiple regressions were carried out with CEFT accuracy and DOT scores as DV, and age in months, IQ, visuospatial WM, verbal WM and IC as IV (**Table 3.5**). The models explained a significant amount of variance in each task; 51% of the observed individual differences in CEFT accuracy and 65% of DOT correct responses. Significant unique contributors to the CEFT model were visuospatial IQ and visuospatial WM, while significant unique contributors to the DOT model were IQ, visuospatial WM, and age. The fact that some of the significant correlations were non-significant in the regression indicates that there was overlapping variance between the domain-general tasks. There was no difference in the pattern of significance after the semantic IC outlier was removed.

Table 3.5: *Study 2: Hierarchical regressions with DV of Children's Embedded Figures Test (CEFT) and Design Organisation Test (DOT) and IV of age, IQ, working memory (WM) and inhibitory control (IC). Significant effects are highlighted in bold.*

	CEFT Accuracy		DOT score	
	β	*p*	β	*p*
Age in months	.091	.438	**.285**	**.005**
Visuospatial IQ	**.232**	**.021**	**.186**	**.029**
Verbal IQ	.141	.204	**.195**	**.039**
Visuospatial WM	**.207**	**.029**	**.269**	**.001**
Verbal WM	.178	.053	-.013	.863
IC	.077	.235	.044	.863
R^2	**.511**	**< .001**	**.649**	**< .001**

3.3.1.2 Study 3

To evaluate associations between CEFT and WM and IC, Pearson's correlations were carried out with study 3 data (**Table 3.6**). Both CEFT accuracy and proportion of distractor triangle responses were included in the analyses, as it was possible that the EF measures would reveal different associations. As there was a significant order effect identified in the response IC data, partial correlations were also examined, controlling for age and order of test. There were no outliers identified in the data.

Table 3.6: *Study 3: Pearson's correlations between responses to the Children's Embedded Figures Test (CEFT) accuracy, CEFT distractor responses, visuospatial IQ, verbal IQ, working memory (WM), verbal WM, semantic inhibitory control (IC), and response IC. Partial correlations controlling for age in months and task order are shown below the diagonal. *p < .05; **p < .01; ***p < .001. Significant correlations are highlighted in bold.*

	CEFT accuracy	**CEFT distractor**	**Visuospatial WM**	**Verbal WM**	**Semantic IC**	**Response IC**
CEFT accuracy %		**-.929 *****	**.257 ***	.061	-.008	**.272 ***
CEFT distractor %	**-.928 *****		**-.301 ***	-.159	.016	**-.267 ***
Visuospatial WM	**.258 ***	**-.300 ***		**.413 *****	-.054	**.286 ***
Verbal WM	.059	-.157	**.401 ****		-.057	-.046
Semantic IC % cost	-.009	.001	-.032	-.042		-.158
Response IC	**.244 ***	**-.245 ***	**.275 ***	-.065	-.124	

N = 71, df = 67

There was a strong negative correlation between CEFT accuracy and proportion of CEFT distractor responses, indicating that the vast majority of incorrect responses involved falsely selecting distractor triangles, rather than running out of time. After controlling for age in months and task order, visuospatial WM significantly correlated with both CEFT measures revealing that those with better visuospatial WM achieved higher accuracy and selected distractor triangles less often. The association between response IC and CEFT revealed that those with poorer response IC were more likely to select a distractor triangle as the embedded target and less likely to select the embedded target. Both of these were weak associations (Cohen, 1992) as revealed in **Figure 3.5**. There were no other significant associations between FI and EF.

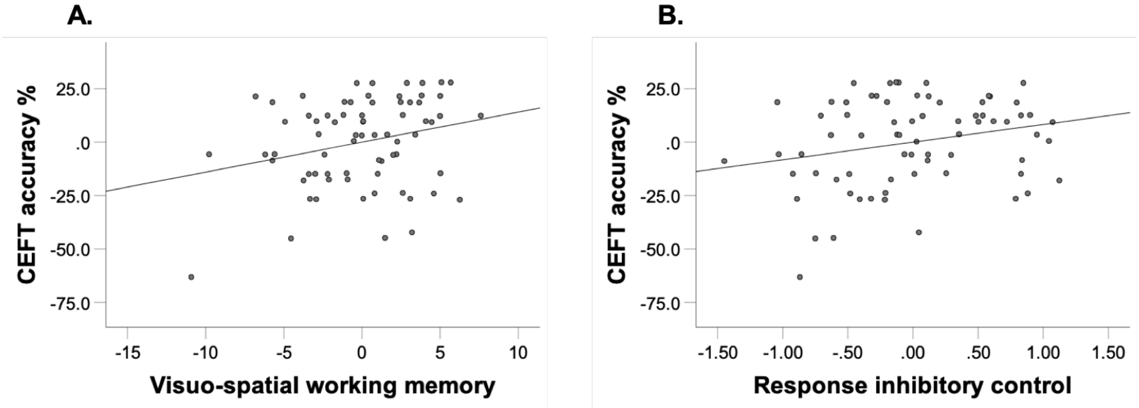


Figure 3.5: *Plots of partial correlations, controlling for age in months and order of task, between the Children's Embedded Figures Test (CEFT) accuracy and executive functions.* ***A****. CEFT accuracy % and visuospatial working memory score.* ***B****. CEFT accuracy % and response inhibitory control D-Prime.*

A significant, positive correlation was revealed between verbal WM and visuospatial WM but there was no association between the two measures of IC. This indicates that semantic IC and response IC represent two distinct processes, further supported by the fact that response IC positively related to visuospatial WM but semantic IC did not.

To determine how much variance in this CEFT test could be explained by age, WM and IC, a multiple regression was carried out with CEFT accuracy as the DV, and age in months, visuospatial WM, verbal WM, semantic IC, and response IC as IV (**Table 3.7**).

Table 3.7: *Study 3: Hierarchical regressions with DV of Children's Embedded Figures Test (CEFT) and IV of age, working memory (WM) and inhibitory control (IC). Significant effects are highlighted in bold.*

	CEFT Accuracy	
	β	*p*
Age in months	-.070	.559
Visuospatial WM	.197	.155
Verbal WM	.005	.970
Semantic IC	.028	.812
Response IC	.225	.079
R^2	.111	.166

Contrary to study 2, the regression revealed that age and EF did not significantly explain CEFT accuracy. Further, none of the individual variables uniquely contributed to CEFT accuracy. It is worth noting that there were several participants achieving 100% on this CEFT task (N = 12). These ceiling effects may explain the limited associations identified in this study between FI and EF. A further difference between the studies was that the task in study 2 included both tent and house stimuli. The accuracy on Year 5 responses to the tent stimuli in study 2 was similar to study 3 (study 2: 71.9%; study 3: 75.4%) but they revealed a lower accuracy percentage in the house trials (study 2: 60.3%).

3.3.2 Field independence and fixation patterns

3.3.2.1 Individual quantitative fixation analyses

This set of analyses examined associations between fixations and behavioural responses to the CEFT (**Table 3.8**). Pearson's correlations and partial correlations were carried out controlling for age in months only. Overall eye-tracking patterns across the task of total number of fixations and mean length of fixation (total length of fixation / total number of fixations) were examined, as well as percentage of fixations, fixation length, and visits in each AOI.

Table 3.8: *Study 3: Correlations between the Children's Embedded Figures Test (CEFT) and eye tracking variables. *p < .05; **p < .01; ***p < .001. Significant correlations are highlighted in bold.*

	Pearson's correlations		**Partial correlations**	
	CEFT accuracy	CEFT distractor responses	CEFT accuracy	CEFT distractor responses
Total number	-.201	.095	-.201	-.094
Mean length (ms)	-.115	.149	-.115	.148
Number of fixations AOI-1 (%)	-.182	.196	-.184	.201
Number of fixations AOI-2 (%)	-.040	.060	-.040	.060
Number of fixations AOI-3 (%)	**-.372 ****	**.418 *****	**-.372 ****	**.418 *****
Length of fixations AOI-1 (%)	-.176	.165	-.179	.170
Length of fixations AOI-2 (%)	**.657 *****	**-.597 *****	**.660 *****	**-.598 *****
Length of fixations AOI-3 (%)	**-.398 ****	**.477 *****	**-.399 ****	**.480 *****
Visit AOI-1 (%)	-.165	.182	-.168	.188
Visit AOI-2 (%)	**.577 *****	**-.572 *****	**.577 *****	**-.572 *****
Visit AOI-3 (%)	**-.479 *****	**.459 *****	**-.485 *****	**.462 *****

N = 71, df = 68
Notes: AOI-1: target shape; AOI-2: embedded target shape; AOI-3: distractor triangles.

Contrary to expectation, CEFT accuracy did not associate with either the total number of fixations or the average length of fixation on the task. However, there were some significant associations between the spatial distribution of fixations and both CEFT accuracy and percentage of CEFT distractor responses. After controlling for age effects, those achieving higher accuracy made fewer (**Figure 3.6A**) and shorter (**Figure 3.6B**) fixations and fewer visits (**Figure 3.6C**) to areas containing distractor triangles, whereas the opposite was the case for those selecting a distractor response more often. Similarly, longer fixations (**Figure 3.6D**) and more visits (**Figure 3.6E**) to the embedded target shape were observed in those with higher accuracy, while the opposite was the case for those selecting a distractor response.

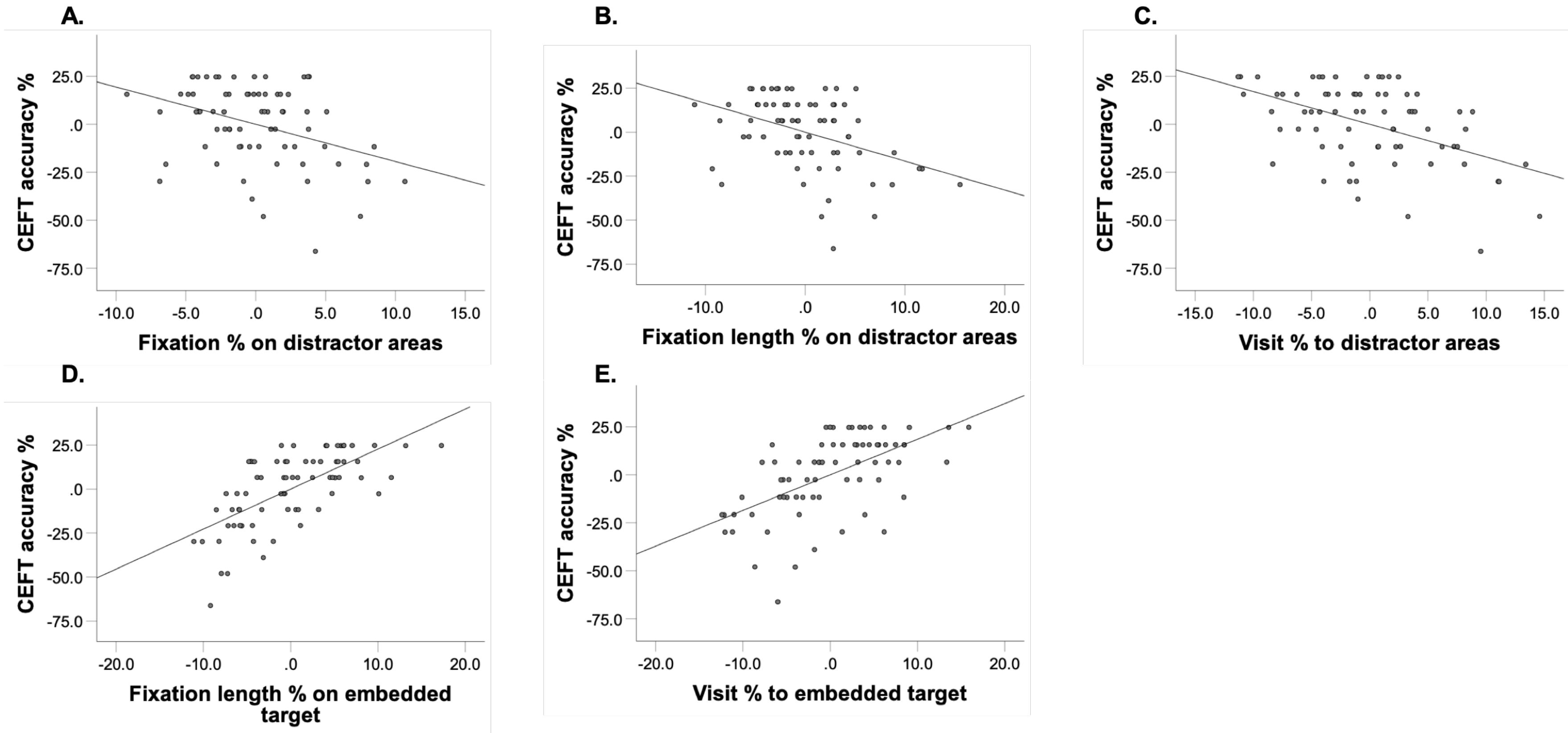


Figure 3.6: *Plots of partial correlations controlling for age in months between Children's Embedded Figures Test (CEFT) accuracy and spatial pattern of fixations.* ***A.*** *CEFT accuracy and percentage of fixations on distractor triangles (AOI-3).* ***B.*** *CEFT accuracy and percentage of fixation length on AOI-3.* ***C.*** *CEFT accuracy and percentage of visits to AOI-3.* ***D.*** *CEFT accuracy and percentage of fixation length on the embedded target shape (AO1-2).* ***E.*** *CEFT accuracy and percentage of visits to AOI-2.*

3.3.2.2 Quantitative analysis of group differences in fixation

Participants were split into three groups according to their FI accuracy scores (**Table 3.9**).

Table 3.9: *Study 3: Field independence (FI) group-level demographics determined by accuracy on the Children's Embedded Figures Test (CEFT).*

	Low FI	Mid FI	High FI
Number of participants	24	21	26
Male : Female	12 : 12	11 : 10	12 : 14
Mean age months (SD)	123.8 (3.3)	124.2 (3.4)	124.4 (2.7)
Mean age years, months	10, 3	10, 4	10, 4
CEFT accuracy (%)	51.1 (14.4)	78.8 (5.3)	95.1 (4.6)

A series of ANOVAs were carried out with each of the fixation variables as DVs and FI group as the IV. Group effects are summarised in **Table 3.10**. Post-hoc Tukey comparisons were run to compare fixation patterns between the three FI groups. In addition to the differences reported in **Table 3.10** there was a significant difference between the low and high FI groups for each significant group-level analysis (p's < .008).

Table 3.10: *Study 3: Group-level field independence (FI) effects on fixation % on the target shape (AOI-1), embedded shape (AOI-2), and distractor triangles (AOI-3), with estimated marginal means (M) and standard error (SE). Significant effects are highlighted in bold.*

Dependent variable	FI *F*(2,68)	Low FI *M* (*SE*)		Mid FI *M* (*SE*)		High FI *M* (*SE*)
Number of fixations AOI-1 (%)	7.21, *p* = .290					
Number of fixations AOI-2 (%)	*F* < 1					
Number of fixations AOI-3 (%)	**5.45, $p = .006$, $\eta^2_p = .138$**	25.5% (0.8)	=	22.9% (0.8)	=	22.2% (0.7)
Length of fixations AOI-1 (%)	1.10, *p* = .339					
Length of fixations AOI-2 (%)	**19.85, $p < .001$, $\eta^2_p = .407$**	18.7% (1.0)	<	25.0% (1.1)	=	27.1% (1.0)
Length of fixations AOI-3 (%)	**6.85, $p = .002$, $\eta^2_p = .168$**	27.4% (1.0)	<	23.7% (1.0)	=	22.7% (0.9)
Visit AOI-1 (%)	1.44, *p* = .245					
Visit AOI-2 (%)	**18.77, $p < .001$, $\eta^2_p = .356$**	34.5% (1.1)	<	39.5% (1.1)	<	43.6% (1.0)
Visit AOI-3 (%)	**9.50, $p < .001$, $\eta^2_p = .218$**	55.2% (1.1)	=	51.6% (1.2)	=	48.6% (1.1)

High FI participants had a lower percentage of fixations in distractor AOIs and a lower percentage of visits to distractor triangles than low FI participants. They also had a higher percentage of visits to the embedded shape than mid FI participants who in turn had a higher percentage of visits than low FI participants. The length of fixation on the embedded shape was lower for low FI participants than mid and high FI participants, with a longer length of fixation on the distractor triangles for low FI than both mid and high FI participants. There was no difference in fixation length between mid and high FI participants. There were no significant differences in fixation patterns to the non-embedded target shape between the three groups.

3.3.2.3 Group-level qualitative fixation analysis

Three trials were selected for the heat-map analysis based on overall accuracy (**Table 3.11**). The order by accuracy differed from the order in which the stimuli were presented; the stimuli were presented trial 1 to 11 in ascending order, in line with the standard CEFT.

Table 3.11: *Mean accuracy by trial.*

Trial number	1	2	5	3	7	9	8	10	6	4	11
Accuracy %	98.6	97.2	93.0	93.0	81.7	77.5	66.2	84.8	63.4	49.3	43.7
Presented?	Y					Y					Y

Heat maps were then created for each FI group using relative duration (**Figure 3.7**). This visually demonstrates the spread of fixations across the stimulus, revealing where participants in each of the three groups had the longest fixations relative to the total length of time on each trial.

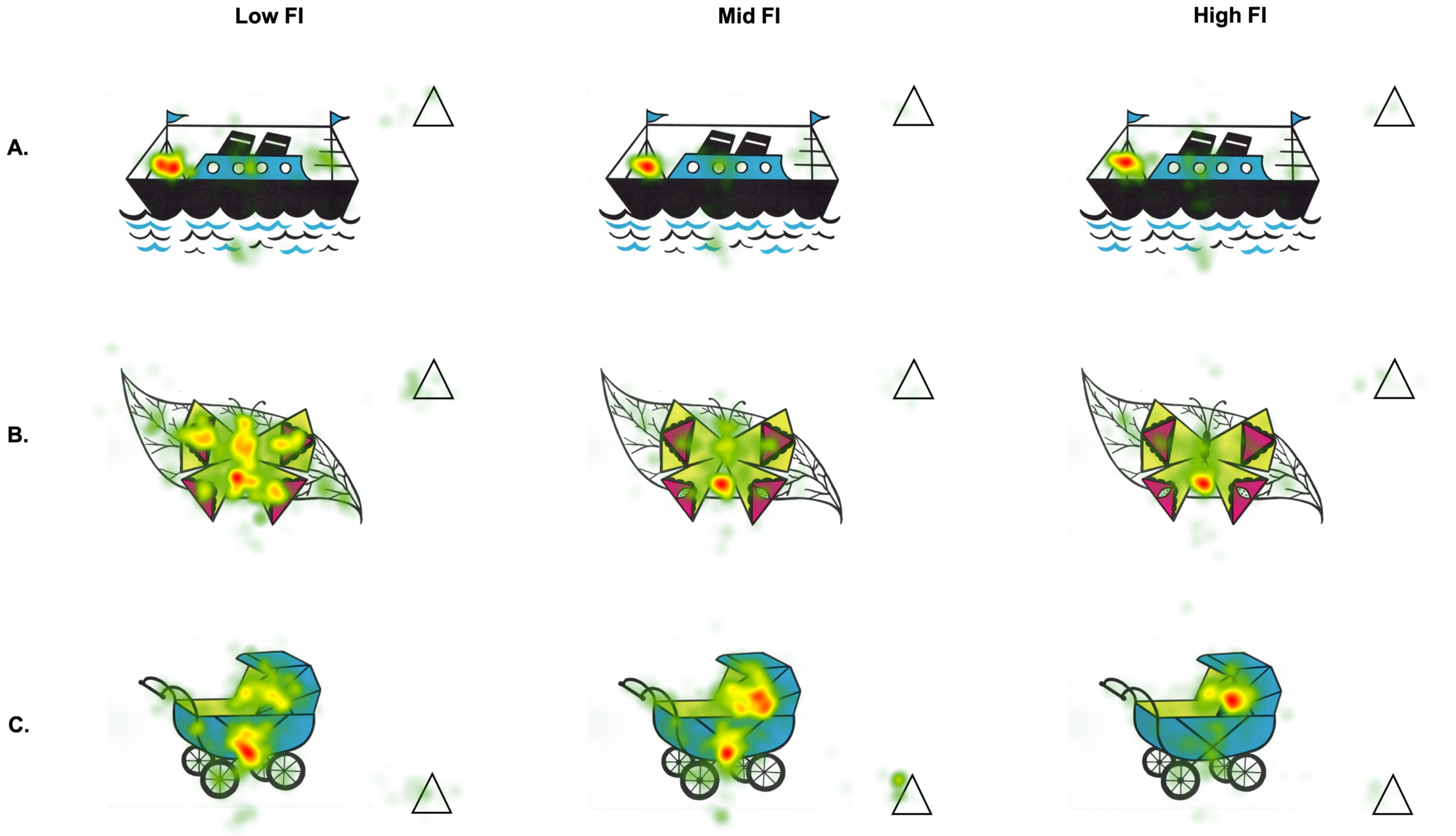


Figure 3.7: *Heat maps of relative fixation length for low, mid, and high FI groups on trials with high (**A**.), mid (**B**.), and low (**C**.) accuracy. Low to high percentage of fixation length.*

Figure 3.7 reveals that there were some group-level differences in fixation pattern between trial. The heat maps indicate that the highest proportion of fixations were concentrated on the embedded target shape and the distractor triangles, but that there was some variation between the three groups. In the easiest trial (based on highest accuracy), there was very little difference in fixation pattern between the groups. The trial with middle-difficulty revealed similar gaze patterns between the high and mid FI groups, and a more diffuse fixation pattern in the low FI group. The most difficult trial revealed a greater focus on the main distractor triangle than the embedded target in the low FI group, a high proportion of fixation length in both areas in the mid FI group, and a greater focus on the embedded target shape than the distractor triangle in the high FI group. This suggests that participants with lower overall accuracy, who are therefore described as low FI, may find it increasingly difficult to overcome distractor triangles in trials which are overall more challenging. Those described as high FI spend relatively longer looking at the embedded target shape, even in the more challenging trials. Those described as mid FI sometimes have similar gaze patterns as the high FI group, but in the more challenging trials, find it harder to ignore distractor triangles. However, these patterns would need confirming across other trials with similar overall accuracy scores, as these group-level patterns may be specific to these stimuli. The next section explores whether these gaze patterns can be explained by individual differences in EF.

3.3.3 Executive functions and fixation patterns

3.3.3.1 Individual quantitative associations

Pearson's correlations and partial correlations controlling for age and task order were carried out to identify significant associations between the EF variables and the fixation variables. The only significant associations after controlling for age and task order were between semantic IC and fixation % on distractor areas, r_p = -.340, p = .004 (**Figure 3.8A**), and percentage of fixation length on distractor areas, r_p = -.332, p = .005 (**Figure 3.8B**). This indicated that those with better semantic IC fixated less on distractor areas of the stimulus.

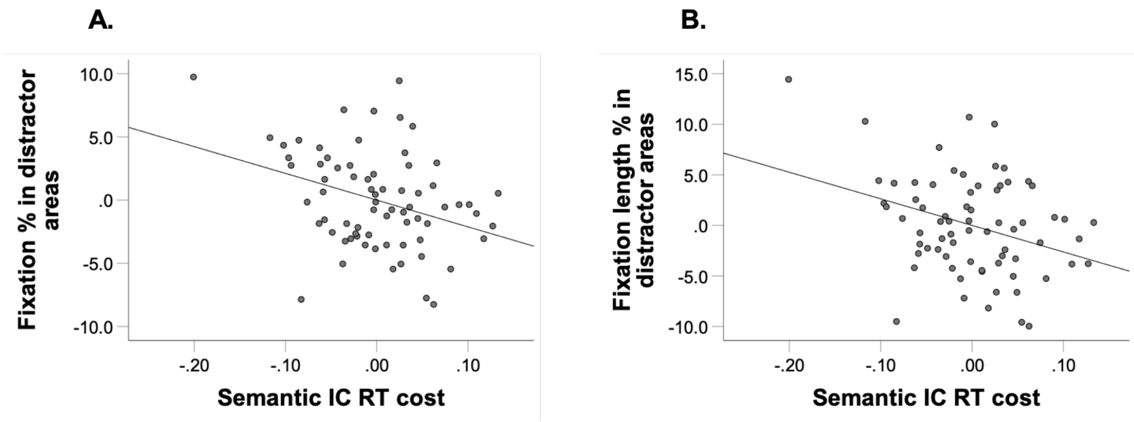


Figure 3.8: *Plots of partial correlations between semantic inhibitory control (IC) and fixations on Children's Embedded Figures Test (CEFT).* ***A****. Semantic IC and percentage of fixations on distractor triangles (AOI-3).* ***B****. Semantic IC and percentage of fixation length on AOI-3.*

The association between response IC and percentage of fixation length on the embedded target was significant in the direct correlations, r = .258, p = 030, but became non-significant in the partial correlations (p = .064). All other associations were non-significant, p's > .061. Contrary to expectation, participants with poorer WM did not refer to the target shape more often than those with better WM.

3.3.3.2 Differences between groups

As the qualitative heat maps revealed differences in fixation depending on FI group, regressions were carried out with dummy FI group variables to determine whether associations between fixation and EF differed by FI group. For each fixation DV, dummy FI group variables, visuospatial WM, verbal VW, semantic IC, response IC were entered in block 1. Interaction terms between each of the dummy terms and each of the EF variables were added stepwise in block 2.

There were two regressions where low FI participants had a significantly different association between fixations and EF compared with the mid and high FI groups; % of fixations on distractor areas, Low FI x verbal WM, β = -.856, p = .027 (**Figure 3.9B**); % of fixation length on distractor areas, Low FI x verbal WM, β = -1.152, p = .002 (**Figure 3.9C**). The negative association between verbal WM and more frequent and longer fixations on distractor areas was greater in the low FI group than the mid and high FI groups.

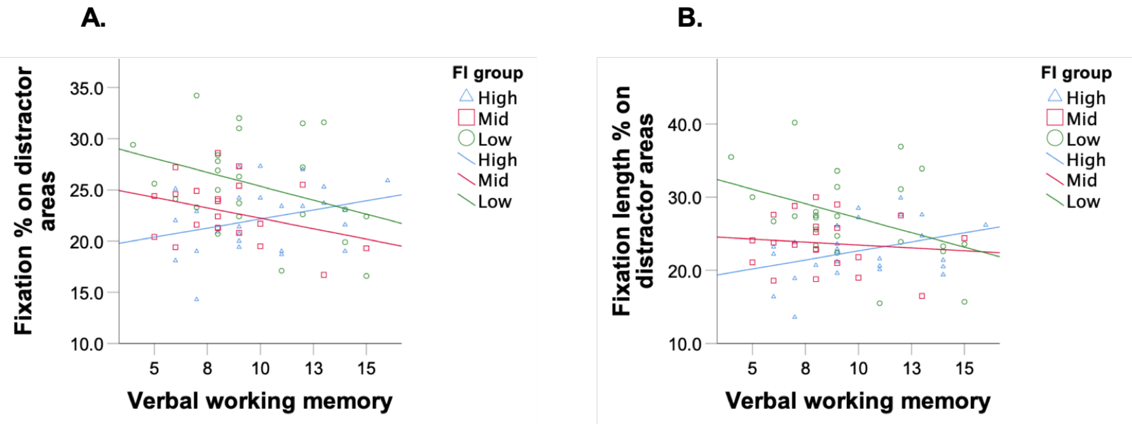


Figure 3.9: *Comparison of fixation patterns and verbal working memory by field independence (FI) group.*

3.4 Discussion

The purpose of this chapter was to further understand the relationship between FI and domain-general abilities. This was achieved through two studies; one behavioural with children in three age groups between the ages of 5 and 10 years, and one which tracked task-related fixation patterns, relating them to behavioural measures of FI and EF in children aged 9 to 10 years. The eye-tracking technology allowed for the identification of areas of the stimulus attracting most attention during the task. This is the first study to examine eye movement activity in children during an FI task.

3.4.1 Field independence and domain-general abilities

Previous studies have identified an association between FI and IQ, WM, and IC (Flexer & Roberge, 1980; Guisande et al., 2007, 2008; Huygelier et al., 2018; Imanaka et al., 2017; Jia et al., 2014; Miyake et al., 2001; Swyter & Michael, 1982). The two studies described here examined these associations in greater depth using a number of domain-general tasks and FI tasks.

Study 2 aimed to gain a clearer picture of the predictors of individual differences in two tasks measuring FI in children aged 5-10 years, as well as of the correlations observed between these two tasks. After controlling for age, there was only a moderate correlation between the two FI tasks, which is suggestive of the impact of the contrasting FI processes (disembedding and segmenting) involved in the separation of the target from its context in these two tasks. Age, IQ, WM and IC explained 50% of variance in CEFT accuracy and 65% of variance in DOT score. Higher visuospatial IQ and higher visuospatial WM were the only significant predictors of CEFT accuracy. In contrast, individual differences in DOT scores were explained by age, visuospatial IQ, verbal IQ, and visuospatial WM. In the CEFT, participants have to hold a single target image in mind while they search for a match in the complex figure, a process that likely recruits visuospatial WM. Research with adults showed that achievement on the EFT was diminished when participants performed a secondary visuospatial WM task concurrently (Miyake et al., 2001). The DOT has a more complicated series of steps required to complete the task, including segmenting the elemental parts, identifying and matching the elemental pattern, and then writing the correct number in the appropriate space on the grid, which may explain why age and verbal IQ contributed to individual differences in performance. Although verbal WM correlated with both the CEFT and DOT, in line with previous research in children (Alloway & Alloway, 2010; Guisande et al., 2008), it was not a significant unique predictor of either FI measure when the additional general cognitive measures were accounted for.

The purpose of study 3 was to examine behavioural associations between the CEFT task and EF by identifying associations with fixation patterns during the completion of the task. In line with study 2, CEFT task accuracy associated with visuospatial WM. Higher CEFT accuracy also weakly associated with higher response IC, which was an additional task included in study 3. This supports previous studies where associations were revealed between FI measures (EFT accuracy and Hidden Figures Test accuracy) and visuospatial memory (Huygelier et al., 2018; Miyake et al., 2001) and response IC (Imanaka et al., 2017).

The more limited associations between FI and EF revealed in study 3 compared with study 2 may be due to a number of differences in the task design and administration. Study 2 used the usual method of presentation (i.e. paper-based stimuli) whereas the stimuli in study 3 were presented on a screen and responses were made with a mouse. There is some evidence that the mode of task presentation can lead to differences in performance (Carpenter & Alloway, 2018; Noyes & Garland, 2008). Additionally, studies 2 and 3 used different sets of CEFT stimuli, with both 'tent' and 'house' trials presented in study 2 but only 'tent' trials in study 3. It is possible that associations with domain-general abilities may be more closely linked to the more complex 'house' target shape, and that some participants in study 3 found the task sufficiently straightforward not to tax other systems, as 12 participants achieved 100% accuracy. Finally, there was a wider range of participant age in study 2 compared with study 3. This is unlikely to be the main source of the differences between the two studies however, as the dummy Year analysis undertaken in **Section 3.3.1.1** indicated that there were no non-linear associations between the CEFT and the EF measures in study 2. Therefore, there is no evidence of distinct patterns of association in Year 5 compared with Years 3 and 1, or compared with the overall level.

In sum, this chapter has revealed significant associations between FI and domain-general abilities, with some contrasts between FI task. Although verbal WM significantly correlated with CEFT accuracy in study 2 but not study 3, both studies were consistent in the significant association between CEFT accuracy and visuospatial WM. Differences between the two studies are likely to be due to the differing designs of the task; whilst this enabled eye movements to be measured, it may also have impacted on the level of EF being engaged to complete the task.

3.4.2 Field independence and fixation patterns

Previous studies have identified fewer fixations and fewer saccades in adults achieving higher scores on the FI tasks (Nisiforou & Laghos, 2013, 2016), so the prediction for study 3 was that children with higher scores on the CEFT would have fewer and longer fixations. However, evidence here runs counter to this, with no association between CEFT accuracy and the overall number of fixations or mean length of fixations. This may be due to the ceiling effects in the behavioural element of the task. Nonetheless, there were differences in the spatial distribution of fixations during the task which related to CEFT accuracy. Those with longer and more fixations in the area of the embedded target shape achieved a higher score, as did those who spent less time, had shorter fixations, and fewer visits on distractor areas. When analysing these associations at the group level, it was generally revealed that the low FI group had different patterns of fixation compared with the mid and/or high FI groups. This was somewhat supported by the qualitative analysis which showed that attention appears to be more focussed on the embedded target and less on distractors in the high FI group, while the attention of the low FI group was directed to a greater extent towards distractor triangles. There was some inter-trial variation in the groups' fixation patterns suggesting that these patterns may depend on the individual image or the difficulty of the trial. In the trial with the highest accuracy, there was very little difference between the three groups, while the trial with the lowest accuracy revealed a difference in attention on the target and distractor triangles between the three groups. This suggests that there may be differences in fixation pattern relating to the ease with which the embedded target can be disembedded, such as where it becomes harder to distinguish from the complex background due to the Gestalt principle of good continuation. This supports a previous study which found variation in response accuracy depending on the extent to which a target was embedded into the complex figure using Gestalt principles (De-Wit et al., 2017).

Distractor triangles received a lot of attention, even when they did not match the target shape either in size or configuration. This suggests that participants were looking for triangular features during their search or any three-sided enclosed shape, rather than the specific equilateral triangle. However, attention was also distributed across other parts of the stimuli, away from the target shape, the embedded shape, and distractor triangles. This likely reflects the nature of the task, requiring all parts of the stimulus to be examined in order to locate the target.

In sum, the overall number or mean length of fixation did not vary with CEFT accuracy performance, contrary to expectation. However, there were associations between CEFT accuracy and fixation patterns across the stimuli. Those who focussed their attention more on the embedded target and less on distractor areas achieved a higher level of accuracy. Triangular shapes received a high level of attention even when they were substantially different in size and configuration from the target shape. The next section will examine the extent to which these associations between CEFT accuracy and fixation patterns can be explained by individual differences in EF abilities.

3.4.3 Executive functions and fixation patterns

As discussed in **Section 3.4.1**, accuracy on the study 3 CEFT task revealed weak associations with visuospatial WM and response IC. Despite this limited association at an overall level, it was predicted that fixation patterns would vary according to an individual's WM and IC scores. Contrary to expectation, those with lower WM scores did not look at the target shape more frequently. It is possible that this had an impact on individuals' performance, in that lower visuospatial WM weakly associated with lower CEFT accuracy. If individuals who are less able to hold an accurate representation of the target shape in mind referred to the given target shape more frequently, they may have been able to achieve a higher accuracy score on the CEFT. There was only one association between EF and fixation patterns. After controlling for age and task order, better semantic IC related to a lower percentage of fixations and fixation length in areas of distractor triangles, in line with expectation. This indicates that an ability to overcome a distracting meaning enabled attention to be directed away from distractor areas containing non-matching target triangles.

In sum, the evidence for associations between EF and fixation patterns was fairly limited in this study. Differing group-level fixation patterns were observed in those with higher and lower FI, and a significant association was observed between semantic IC and fixations in distractor areas. Other WM and IC measures did not associate with fixation patterns in this study. It is possible that the CEFT task in study 3 was not challenging enough for all participants, which limited the extent to which WM and IC were engaged to find the embedded target. This is supported by the fact that WM in particular had a closer association with the CEFT task in study 2, where participants encountered mixture of the tent and house stimuli. It would be interesting to use eye tracking technology while participants try to disembed house stimuli from the complex stimulus to find out whether more challenging stimuli reveal stronger associations with EF.

3.4.4 Considerations of using eye-tracking technology

Although eye tracking technology records where an individual is looking at a particular point in time, it cannot directly measure what an individual is thinking (Holmqvist et al., 2011). Eye movements infer an individual's selective attention towards a spatial location, but cannot measure attention towards specific dimensions of a stimulus (Rehder & Hoffman, 2005). It is possible to be looking at an object whilst thinking about a completely different topic. By measuring task behaviour as well as eye movements, it is hoped that the eye tracking data here should reflect task-related eye movements, and they have been analysed as such. A longer gaze time at a particular point may indicate that the individual is consciously or unconsciously aware that the target shape is embedded at that location in the complex picture, or the individual's attention may be drawn to that particular location due to a salient colour or feature. Importantly, objects which are fixated upon may be processed but not remembered, while objects falling into peripheral vision and not processed in detail may be remembered. In one study looking at eye movements during a driving task, 20% of objects which were not fixated on, were recalled by the participants (Underwood, Chapman, Berger, & Crundall, 2003). As the design of the stimuli did not always allow for space between the AOIs, it is possible that the content of a neighbouring AOI was perceived without the eye tracking technology recording a fixation within that AOI. This illustrates the challenge of adapting previously-existing stimuli for use with eye-tracking technology, and the subsequent difficulty of comparing findings with previous studies, if the task has been adapted for the purposes of eye-tracking.

Nonetheless, eye-tracking technology can reveal additional information about the allocation of visual attention which cannot be collected by behavioural responses alone. For example, CEFT accuracy associated here with response IC but not semantic IC, while percentage of fixations on distractor areas revealed an association with semantic IC. Either measure taken in isolation may lead to incomplete understanding of the processes involved in the CEFT.

3.4.5 Conclusion

The associations between FI and domain-general abilities differed by task. Visuospatial WM consistently associated with accuracy on the FI tasks, although the strength of correlation varied by task. Through using eye-tracking technology, it was revealed that those achieving higher accuracy on the CEFT task had a higher percentage of fixations and length of fixations on the embedded target and a lower percentage of fixations and shorter fixations on distractor areas. CEFT accuracy had a weak association with response IC and no significant association with semantic IC, however, those with better semantic IC had a lower proportion of fixations on distractor areas. Based on study 3, there is some evidence that associations between FI and EF can be explained by differing patterns of fixation across the stimulus, however this was more limited than expected which may be due to the relatively low level of challenge in the eye-tracking CEFT task.

Studies discussed in **Chapters 2** and **3** have presented insight into the development and explanatory factors involved in whole-part tasks. The focus of **Chapter 4** will be to examine the development and explanatory factors associated with mathematics and science tasks.

Chapter 4: Developmental changes in mathematics and science

4.1 Introduction

The literature reviewed in **Chapter 1** indicated that mathematics and science achievement associate with individual differences in general cognitive processes including IQ and EF, while the studies in **Chapter 3** revealed positive associations in the scores of field independence (FI) tasks and executive functions (EF). Before studying possible associations between whole-part measures and mathematics and science achievement (see **Chapter 5**), the development of mathematics and science achievement, as well as differential associations with EF and IQ over development, need to be understood.

4.2 Development of mathematics and science

Variation in mathematics achievement has been shown to relate to both domain-specific and domain-general factors (Cragg, Keeble, Richardson, Roome, & Gilmore, 2017; Geary, 2011; Gilmore, Keeble, Richardson, & Cragg, 2017). Domain-specific factors include procedural skills, which determine an individual's knowledge of how to solve mathematical problems, and conceptual understanding, which refers to individuals' understanding of mathematical principles underlying the arithmetic (Cragg, Keeble, et al., 2017; Gilmore et al., 2018, 2017; Rittle-Johnson, Siegler, & Alibali, 2001). Domain-general factors include general intelligence (Geary, 2011), EF (Bull & Lee, 2014; Cragg, Richardson, Hubber, Keeble, & Gilmore, 2017; Van der Ven, Kroesbergen, Boom, & Leseman, 2012), and spatial abilities (Gilligan et al., 2018). Together, these factors explain a large amount of variance in mathematical achievement, for example Gilmore et al (2017) reported that 71% of variance in mathematics achievement was explained by procedural skills, conceptual understanding, working memory (WM), and interactions between these factors, in children aged 5-6 years. This highlights the importance of understanding associations between mathematical achievement and these domain-general and domain-specific factors when examining other possible predictors of mathematics achievement.

WM may be important for solving mathematics problems through representing the problem, temporarily storing part-solutions, and retrieving information from long-term memory (Gilmore et al., 2018; Raghubar et al., 2010; Van der Ven et al., 2012). WM has a large effect on mathematics achievement, although the strength of this relationship varies according to the mathematical elements being assessed (Lee &

Bull, 2016), whether rate of growth or current performance level is being measured (Wei, Guo, Georgiou, Tavouktsoglou, & Jordan, 2018), and whether visual or visuospatial WM tasks are being used (Cragg, Richardson, et al., 2017). The strength of association between mathematics and WM also varies at different timepoints across development as children encounter mathematical problems requiring different resources (Geary, 2011). For example, a study found that the association between WM and mathematics in 5- to 6-year-olds was weaker than in children aged 6-8 years, likely due to its lesser involvement in early activities of number recognition, writing, counting, and sequencing, compared with early arithmetic (Lee & Bull, 2016). In parallel, WM is less required when calculations become more automatic (i.e. when individuals become more expert), although it is still tapped for higher difficulty and multi-step problems (Geary, 2011; Lee & Bull, 2016; Wei et al., 2018). In contrast to the variation in strength of association between WM and mathematics at different ages, the association between WM and the rate of growth in mathematics achievement across development, after controlling for prior achievement, remains constant (Lee & Bull, 2016; Wei et al., 2018). Although this appears contradictory, it may reflect a difference in WM contribution between learning novel concepts or procedures, and carrying out previously-learnt calculations and procedures (Cragg, Keeble, et al., 2017; Lee & Bull, 2016). This is consistent with another study which found varying contributions of WM depending on whether the mathematical skill was being learnt, being consolidated, or had been mastered (Raghubar et al., 2010).

To date, there has been a greater emphasis on using verbal WM tasks than visuospatial WM tasks (Gilmore et al., 2018; Raghubar et al., 2010). Nonetheless, there are some inconsistencies in the literature as to whether verbal WM or visuospatial WM has the greater explanatory power on mathematics performance. These conflicting findings may result from differing WM systems being required to complete different aspects of mathematics. Arithmetic tasks may draw on verbal WM for the retrieval of number facts from long-term memory which are stored in verbal code (Cragg, Keeble, et al., 2017; De Smedt et al., 2009), and geometry requires visuospatial WM (Kyttala & Lehto, 2008). Even within these broad categories, different patterns have been identified. For example, in arithmetic questions completed by children aged 7 to 9 years and adults, addition and subtraction had a stronger association with visuospatial WM than multiplication and division questions (Lee & Kang, 2002; van der Ven, van der Maas, Straatemeier, & Jansen, 2013). Differences in relative importance of verbal and visuospatial WM have also been revealed over development. Visuospatial WM is particularly important in early mathematics development, perhaps as children rely on concrete representations of number and arithmetic (De Smedt et al., 2009; van der Ven et al., 2013) or as younger or less experienced children store information in a visual way rather than converting it to a verbal code (Holmes & Adams, 2006; Meyer, Salimpoor, Wu, Geary, & Menon, 2010). Other findings indicate that verbal WM may have a greater role in initial learning while visuospatial WM may be more important for developing a deeper understanding of arithmetic and for more complex mathematics (Cragg, Richardson, et al., 2017; Meyer et al., 2010; Raghubar et al., 2010; Soltanlou, Pixner, & Nuerk, 2015). Together, this suggests that visuospatial WM may play a greater role in numerical skills in younger children, but as children develop and encounter more challenging arithmetic, verbal WM becomes more important, particularly for retrieval. As mathematics tasks start requiring reasoning abilities, then visuospatial WM becomes more important once more.

The role that inhibitory control (IC) plays in mathematics is in identifying the most relevant pieces of information and ensuring that these receive more attention than distractor information (Szucs, Devine, Soltesz, Nobes, & Gabriel, 2014), as well as ensuring previously-used but inefficient strategies are inhibited when solving mathematical problems (Lubin, Vidal, Lanoë, Houdé, & Borst, 2013; Van der Ven et al., 2012). IC is also involved during the retrieval of arithmetic facts where competing responses need to be suppressed. For example, to correctly retrieve the answer to 2+5, the answer to 2x5 needs to be inhibited (Hubber, Gilmore, & Cragg, 2014; Wei et al., 2018). Overall, studies examining the relationship between IC and mathematics have had less consistent outcomes than those studying WM. This may be due to IC making a smaller contribution to performance which only sometimes reaches significance, or it may only contribute unique variance when other EF are not included in the analysis, particularly WM (Cragg, Keeble, et al., 2017).

Although the associations between EF and science have received much less attention, there are studies which reveal that WM and IC relate to achievement. The results are quite variable, however. One study with 9- to 11-year-olds only identified a relationship between IC and science when a prior concept needed to be inhibited (Vosniadou et al., 2018), and a study with 10- to 11-year-olds using a pen-and-paper IC task and science task did not reveal an association between the measures (Mayer, Sodian, Koerber, & Schwippert, 2014). However this latter finding may simply indicate that the content of the science task did not require IC, or that the IC task itself was not sufficiently sensitive to be used as the sole IC measure (Mayer et al., 2014). In contrast, a study with 11- to 12-year-olds found that inhibitory control predicted science achievement scores when WM was partialled out, and that visuospatial WM significantly predicted science after controlling for verbal WM, although the reverse was not significant (St Clair-Thompson & Gathercole, 2006). Similarly, visuospatial WM, but not verbal WM, significantly associated with science in 12- to 13-year-olds (Rhodes et al., 2016). Together, these findings indicate that visuospatial WM may have a stronger association with science achievement than verbal WM, and that IC is a significant explanatory variable particularly when the science task requires prior concepts to be ignored.

4.2.1 Misconceptions

Naïve conceptions, counterintuitive concepts, or previous misconceptions do not get replaced by new learning, but rather need to be suppressed using IC in order for a more accurate response to be reached (Brookman-Byrne et al., 2018; Mareschal, 2016; Vosniadou, 2013). For example, in mathematics, the application of the size of positive numbers such as 1, 2, and 3, to either negative numbers (-1, -2, or -3) or fractions (1/1, 1/2, or 1/3), will result in an error as the magnitude of negative numbers and fractions get smaller as the digit increases (Hansen, Drews, Dudgeon, Lawton & Surtees, 2017). In science, the concept of insulation relating to maintaining warmth may be understood by children who have practical experience of wearing a coat, but this experience may be wrongly applied to the question of whether they should put a coat around a snowman to keep him frozen (Naylor, Naylor & Mitchell, 2000). In both examples, the prepotent response needs to be inhibited in order for the correct response to be attained. Additionally, perceptual cues can lead to the formation of misconceptions, as well as interfering when working out the solution to a problem (Mareschal, 2016). For example, an angle may be interpreted as being larger if the angle arc is drawn further away from the angle point and is therefore longer (**Figure 4.1A**), or a shape may not be perceived as a square if its lower side is not horizontal (**Figure 4.1B**). Similarly, in science, the sun appears to move while the earth feels like it is stationary, creating a misconception about the celestial mechanics of the solar system.

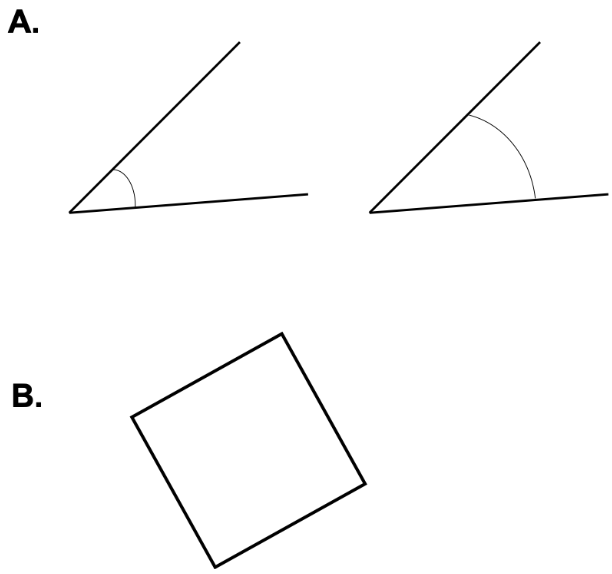


Figure 4.1: *Examples of common mathematics misconceptions.* ***A.*** *The angles are identical but the angle to the right may be perceived as larger as the angle arc is longer.* ***B.*** *The shape may not be identified as a square as it does not have a horizontal side.*

The results of a recent large-scale intervention study revealed that children aged 7-10 years with better semantic IC achieved a higher proportion of correct answers on a misconceptions task, with a significant association between IC and science in the younger children, and both science and mathematics in the older children (Wilkinson et al., 2019). Similarly, in a study with young adolescents aged 11-15 years, both semantic and response IC measures positively related to outcomes on mathematics and science misconception tests (Brookman-Byrne et al., 2018). Both studies indicate that children and teenagers who are more proficient in overcoming the distractor in the IC tasks, are also able to overcome incorrect, competing responses to common misconception questions.

4.2.2 Current studies

Although there is some variation between studies, the literature overall has identified an association between domain-general abilities and achievement in both mathematics and science in children. The variation between studies is largely due to differing demands of the mathematics and science tests, the age of the children participating, and the type of EF and other domain-general tasks involved in the studies. The studies discussed below aim to give a foundation for further analysis with whole-part measures, but also have a valuable contribution to the literature in their own right.

In this chapter, two distinct studies are reported which examined associations between mathematics and science, and domain-general abilities of either general intelligence and EF, or just EF. In **Chapters 2** and **3**, intelligence measures were separated into distinct verbal and visuospatial abilities in order to understand differing levels of contribution. In **Chapters 4** and **5**, a single measure will be used to represent general intelligence, so general domain-general abilities can be covaried. Typically, EF are grouped into three broad abilities, namely WM, IC, and shifting. Only WM and IC will be considered here, as these are most likely to be tapped during whole-part tasks, and some studies have found the contribution of shifting to academic achievement to be non-significant or inseparable from IC in children (St Clair-Thompson & Gathercole, 2006; Van der Ven et al., 2012). The EF measures will remain separated in **Chapters 4** and **5** in order to examine specificity of association.

Study 4 was designed to examine the cross-sectional changes in mathematics and science scores across Years 1, 3, and 5 (children aged 5 to 10 years). The tasks comprised reasoning and written arithmetic mathematical questions drawing on a combination of procedural and conceptual mathematical skills, reasoning and scientific inquiry science questions, and a novel set of common misconception questions which are frequently encountered in mathematics and science. The aim was to identify associations between the measures and reveal the relative contribution of EF and general intelligence in explaining mathematics and science performance, as well as identifying how these relationships vary over time, if at all. It was predicted that general intelligence and WM would significantly explain performance on mathematics and science measures, but semantic IC may have a weaker association, in line with previous literature with children of a similar age (Van der Ven et al., 2012). It was also predicted that associations with verbal and visuospatial WM would vary by mathematical task; a stronger association was expected between visuospatial WM and the procedure-focussed task, while verbal WM would likely have a stronger association with the word problem task. Additionally, as the literature has identified stronger associations between mathematics and verbal WM in 7-year-old children and visuospatial WM in 8-year-olds (Alloway & Passolunghi, 2011; Meyer et al., 2010), it was expected that there would be some differences in association between Year 1, Year 3, and Year 5. As prior misconceptions may need to be inhibited in order to give a correct response, a positive association with IC was expected to be observed.

Study 5 was designed to examine relationships between EF and a novel task comprising four sub-sets of mathematical questions; ‘Number’, ‘Word’, ‘Shape’, and ‘Graphs’. This was designed to give a more nuanced snapshot of how the explanatory power of EF varies depending on the type of mathematics questions being completed. This study also included both a semantic IC task and a response IC task, to gain insight into how these differ in children. The prediction was that a stronger association would be found between the WM measures and mathematics than IC, reflecting the literature. In addition, it was thought probable that there would be a stronger relationship between visuospatial WM and the shape and graph questions than between verbal WM and those subsets, due to the visuospatial nature of those questions. Finally, verbal WM was expected to associate more strongly with the word problems than the number questions, due to the process of separating the verbal 'story' information from the mathematical procedure embedded within the questions.

It should be noted that the studies in this chapter were not specifically designed to evaluate the predictive power of EF and general intelligence on mathematics and science achievement across primary school. To do this effectively, a larger battery of EF tasks would need to be included for each measure of EF. This chapter should instead be considered an important first step in understanding associations between part-whole measures and academic performance. The identification of factors which associate with academic performance, and the understanding how these relationships change over time, gives a clearer context for evaluating and understanding part-whole relationships with mathematics and science.

4.3 Method

4.3.1 Participants

Both studies were approved by the local ethics committee. Study 4 included cross-sectional data for the mathematical and science tests, as well as longitudinal data for the misconceptions test with time-points two years apart. The participants for the study 4 were the same as the group introduced in **Chapter 2** (sample 1), and the participants for study 5 were introduced in **Chapter 3** (sample 2). Study 5 involved a single Year group, and therefore does not examine the associations from a developmental perspective. A summary of participants can be found in **Table 4.1**.

Table 4.1: *Summary of participant demographics from study 4 and study 5.*

	Study 4					**Study 5**
	Year 1	**Year 3**	**Year 3**	**Year 5**	**Year 5**	**Year 5**
	T1___	**___T2**	**T1___**	**___T2**	**T1**	
Number	45 [(a)]	28 [(a)]	45 [(b)]	32 [(b)]	45	71
Male : Female	16 : 29	10 : 18	27 : 18	20 : 12	24 : 21	35 : 36
Mean age in months (SD)	68.6 (3.3)	92.2 (3.6)	92.8 (3.5)	114.2 (3.6)	115.4 (3.6)	124.1 (3.1)
Mean age in years, months	5, 8	7, 8	7, 8	9, 6	9, 7	10, 4

[(a)] Longitudinal cohort in Year 1 at T1 and Year 3 at T2
[(b)] Longitudinal cohort in Year 3 at T1 and Year 5 at T2
Notes: T1 = Time 1; T2 = Time 2

4.3.2 Summary of tasks

A summary of the tasks completed by participants in each study can be found in **Table 4.2.**

Table 4.2: *Summary of tasks and measures for each study.*

Study	Measure	Task
Study 4	Mathematics	WIAT II Mathematical Reasoning
		WIAT II Numerical Operations
	Science	Novel science test
	Misconceptions	Novel misconceptions test (Time 1 and Time 2)
	General intelligence	Visuospatial: Raven's Coloured Progressive Matrices
		Verbal: British Picture Vocabulary Scale 3 (BPVS-3)
	Visuospatial WM	Reverse spatial span
	Verbal WM	Reverse digit span
	Semantic IC	Animal size Stroop
Study 5	Mathematics	Novel task with four subsets (Number, Word, Shape, Graph)
	Visuospatial WM	Reverse spatial span
	Verbal WM	Reverse letter span
	Semantic IC	Animal size Stroop
	Response IC	Whack-a-mole

Notes: WIAT: Wechsler Individual Achievement Test; WM = working memory; IC = inhibitory control

4.3.3 Mathematics tests (study 4)

Mathematics ability was assessed using two subtests of the Wechsler Individual Achievement Test-II (WIAT-II; Wechsler, 2002). Mathematical Reasoning is verbally administered and tests reasoning through word problems and questions using visual stimuli; Numerical Operations is a pen-and-paper test of counting and computation (Meyer et al., 2010). As they assess different mathematical skills, and therefore may differentially relate to general intelligence and EF as well as whole-part variables, the subtest scores were kept as separate variables in the analyses. The standard testing and scoring procedures of the WIAT-II were used. Test/retest reliability ranges from .85 to .98 for both subtests (Wechsler, 2002).

4.3.4 Mathematics test (study 5)

This was a novel pen-and-paper task (which can be viewed here: https://osf.io/s9f7u/) based on the content of the Year 5 UK National Curriculum and designed using a layout with which participants would be familiar (**Figure 4.2**). During the design stage, the questions were reviewed by a senior teacher from a local primary school to ensure the questions were clear and of an appropriate standard. There were 40 questions in the task, with 10 questions on each subset of ‘Number’, ‘Word’ problems, ‘Shape’ (‘geometry’ in the National Curriculum), and ‘Graphs’ (‘’statistics’ in the National Curriculum). The questions from these subsets were interleaved throughout the task, requiring participants to switch between subsets throughout the task. The arithmetic content of the number and word problems were equivalent in terms of operation used, magnitude of number, and steps required to complete the task. This enabled a direct comparison between children’s performance when the arithmetic problem was embedded within a ‘real-life’ context and when it was presented without a context. The content of both these subsets tested all four operations, fractions, sequencing and ordering numbers, and rounding numbers. The shape questions tested understanding about 2D and 3D shapes, angles, and transformations. The graph questions included line graphs, pictograms, and bar charts, as well as interpreting information presented in a table. There was a mix of difficulty levels across the questions in each maths subset, designed to elicit a range of accuracy responses across the sample.

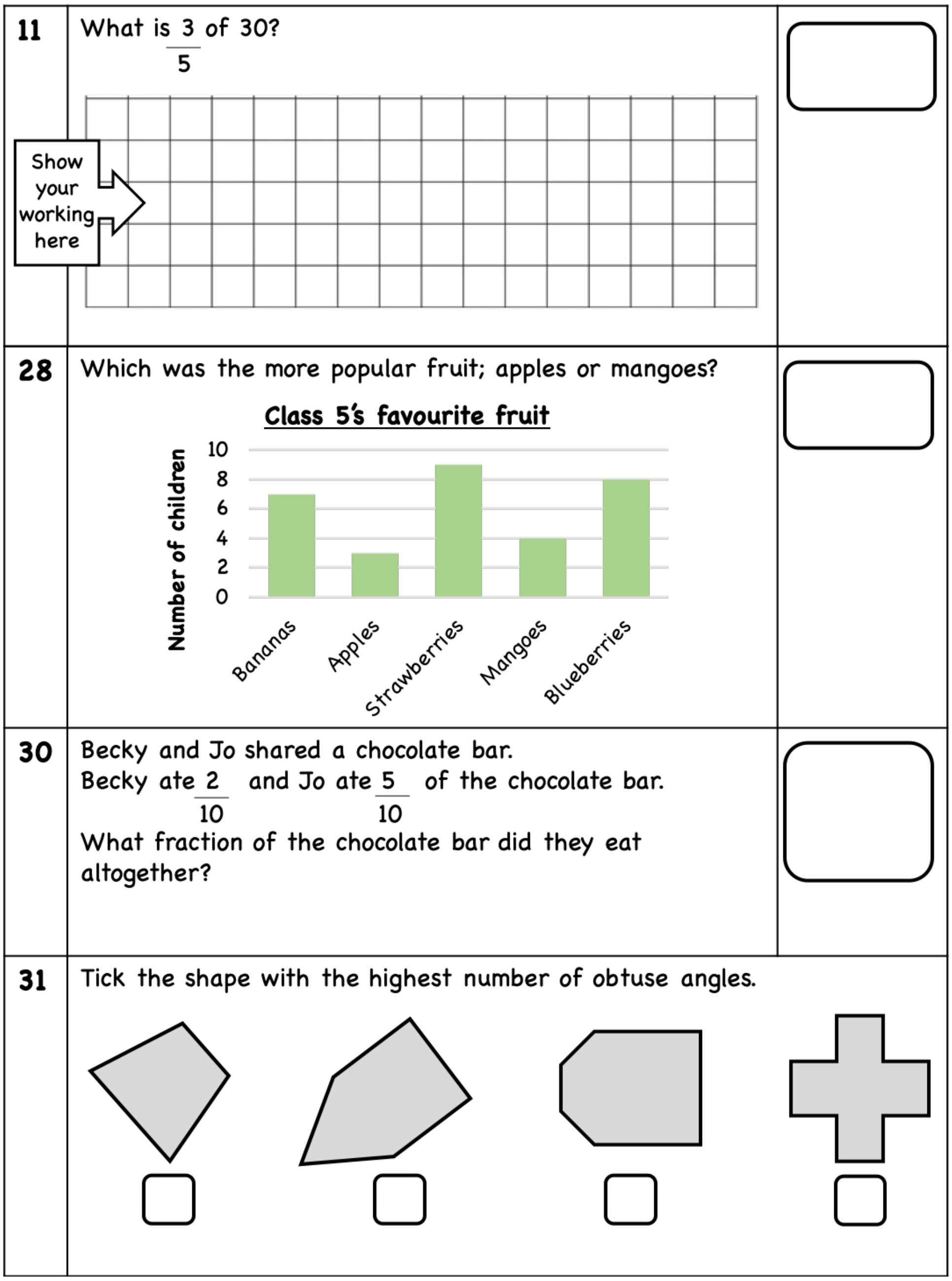


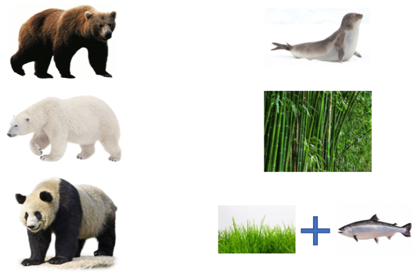
Figure 4.2: *An example from each subset of the mathematics task. Question 11: Number; Question 28: Graph; Question 30: Word; Question 31: Shape.*

There was no time limit, but the task was designed to take approximately 30 minutes for each child to complete. The questions were read to each participant, to ensure there were no effects of reading ability or of children misreading the question. This task gave an overview of the participants' performance over much of the Year 5 mathematics curriculum and therefore gave a fair reflection of their current subject ability. A mean percentage accuracy score was calculated for each subset and for the task as a whole, as well as an overall arithmetic score which represented the mean accuracy from the Number and Word subsets combined. As this was a novel task, there was no previous reliability analysis. Therefore, a split-half analysis was run using the overall mathematics score, with participants being sorted alternately into each group to ensure a randomly-sorted mix in each group. The Pearson's correlation between the scores was .961, demonstrating good reliability of the overall task.

4.3.5 Science test (study 4)

In the absence of a suitable standardised test, science ability was assessed using a test designed for this study which can be viewed here: https://osf.io/d58ht/. Questions were based on content from the current UK National Curriculum and were designed to test scientific reasoning and scientific inquiry (**Figure 4.3**).

Year 1



Group 2 Question 1

Here are 3 bears and the food that they eat.
Match the bear to its food.
Give the child the pictures.

We can use special words to describe these animals: carnivore, herbivore and omnivore.
Point to the carnivore's food.
Now point to the herbivore's food.

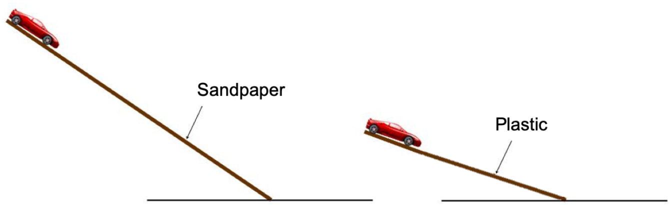
Year 3

Group 3 Question 3

Some children were making shadow puppets.
Which of these shadows would be made by this shadow puppet?

The children wanted their shadow to have an eye.
What could the children do to their shadow puppet so that the shadow has an eye?

Year 5



Group 4 Question 3

Some children wanted to explore how the ramp surface affects the speed of a car.
This one has sandpaper on the ramp, and this one has smooth plastic on the ramp.

What 2 things should the children change to make it a fair test?

Figure 4.3: Example questions from each Year of the science test.

To ensure the task was representative of science learning in Years 1, 3, and 5, a question was included from every topic in the curriculum for those Years, with the exception of human development in Year 5. Similar to the WIAT II, participants within the same Year started at the same level, but would then attempt more challenging or easier questions depending on their accuracy scores. A score of eight or more out of 10 on a level enabled progression to the next level; a lower score led children to complete the lower difficulty level until they scored at least eight. Lower levels which were not attempted were credited with full marks. The reliability of this novel task was measured by running a split-half analysis in each Year group separately. The Pearson's correlations between the scores were: Year 1: .760; Year 3: .870; and Year 5: .798.

4.3.6 Misconceptions (study 4)

A novel misconceptions test was designed for this study and can be viewed here: https://osf.io/47r2u/. The test comprised 10 multiple choice questions, five of which had mathematical content and five were related to science (**Figure 4.4**).

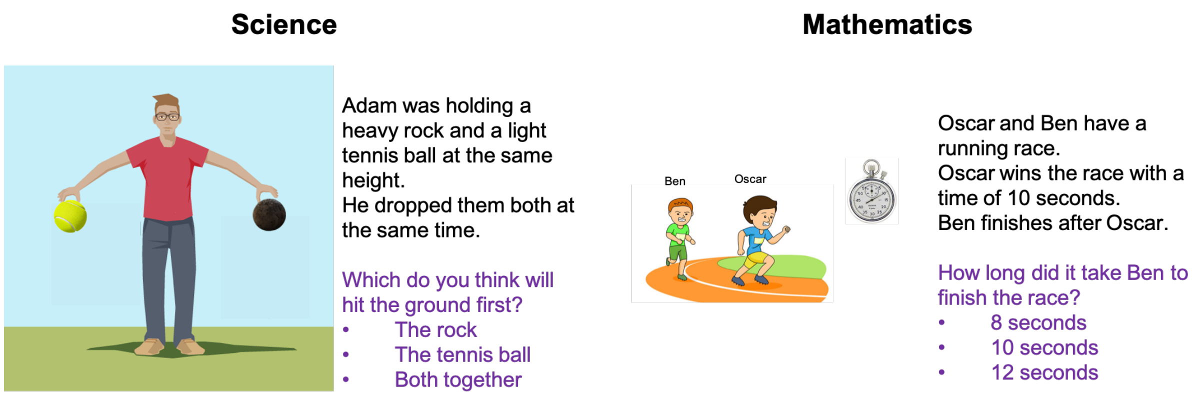


Figure 4.4: *An example of a science misconception question and a mathematics misconception question.*

The subject areas were all identified as being common misconceptions held by primary school children (Hanson, 2014; Allen, 2014), and children in all age groups were presented with the same set of questions. The questions largely related to content not specifically covered by the mathematics and science curricula in order to ensure the younger children were not at an immediate disadvantage. The multiple choice design of the questions was a further step taken to ensure younger children had the potential to answer the questions as easily as the older children.

Children were presented with an image and were then asked the question associated with that image and provided with the multiple choice responses. Each question contained three response options: the common misconception, an additional incorrect response (an uncommon error), and the correct answer. It should be noted that children choosing the uncommon error may either have limited domain-specific knowledge of this topic, or may hold a misconception which is different from the common misconception associated with that question. The proportion of each response type was recorded. The variables of interest were the proportion of misconception responses and the proportion of uncommon errors, as this task was designed to assess children's misconceptions and how they interfere with correct responses. The reliability for this novel task was calculated using split-half analysis on Time 1 data for each Year group, with participants sorted into the two groups by odd and even participant number. The Pearson's correlations between the groups were: Year 1: .900; Year 3: .890; and Year 5: .910.

4.3.7 General intelligence and executive functions (study 4)

The general intelligence tasks are described in **Section 3.2.4**. The measure used here is a general intelligence measure calculated from the mean of the Z-scores of the verbal and non-verbal IQ tasks. The WM and semantic IC tasks are described in **Section 3.2.5**. The verbal WM task used here is the reverse number task. Numbers were used in this study as it was considered they would be very familiar to all participants, from Year 1 children to Year 5 children, and children would be used to hearing different numbers being grouped together.

4.3.8 Executive functions (study 5)

There were four tasks completed by the participants: Visuospatial WM, measured using an age-appropriate reverse spatial span (**Section 3.2.5.1**); verbal WM, measured using a reverse letter span (**Section 3.2.5.1**); semantic IC, measured using the animal size Stroop task (**Section 3.2.5.2**); and response IC, measured using the whack-a-mole task (**Section 3.2.5.3**).

4.3.9 Design and procedure

The testing for both studies took place at the participants' primary school during lesson times. For study 4, the data were collected as part of a larger test battery completed over three or four sessions, which included the global and local processing tasks discussed in **Chapter 2**. The Time 2 misconceptions data were collected two years later from consenting Time 1 (T1) participants in the lower two Year groups. For study 5, the mathematics and EF data were collected as part of the eye tracking study described in **Chapter 3**.

Each task was explained to the participants immediately prior to them attempting the activity, and practice trials were completed for the EF tasks and IQ tasks. Accuracy was collected, in the form of raw scores for Mathematical Reasoning, Numerical Operations, and science. The percentage of each response choice was calculated for the misconceptions test. Mean RTs were calculated from correct responses only for the animal size Stroop task. Computerised stimuli were presented via MATLAB (R2010b), using the Cogent toolbox (http://www.vislab.ucl.ac.uk/cogent_2000.php), on a 12-inch Dell laptop. The task order was counterbalanced across participants in each Year in study 4. The EF tasks in study 5 were counterbalanced, but all participants completed the groups of tasks in the same order over three sessions: Session 1 for the EF tasks, session 2 for the mathematics test, and session 3 for the eye-tracking task.

4.3.10 Statistical analyses

Response data from the mathematics and science tests, EF tasks, and general intelligence tasks, were analysed using correlations, analysis of variance (ANOVAs), and regressions. In order to determine whether the associations between the general intelligence and EF variables and mathematics, science and misconceptions remained constant over the three Year groups in study 4, analyses were run with dummy Year variables. For this analysis, all EF and intelligence variables were transformed into Z-scores to enable comparison across the measures. Any significant interactions with the dummy Year variables indicated that the relationship varied between Years. Linear mixed models (LMMs) were used to analyse the cross-sectional and longitudinal misconception data from timepoints two years apart (see **Section 2.2.7.1** for more information about LMMs).

All children attempted and completed all tasks. The data were examined to identify any outliers, identified as data points ± 3.29 SD from the Mean. In study 4, there was one Year 5 participant who exceeded the higher boundary in Numerical Operations, and one Year 1 participant who exceeded the lower boundary in Semantic IC (where they were much slower to respond to the congruent trials than the incongruent trials). There were no outliers in study 5. As these outliers were not due to recording errors or inability to complete the task, the analyses were run with the outliers included and then excluded to assess their impact on the outcomes (see **Section 2.2.7.2** for further details).

4.4 Results

4.4.1 Order effects

As revealed in **Section 3.3.1.2**, there was an order effect in the study 5 response IC task. Therefore, order was included as a covariate in analyses with the response IC measure.

4.4.2 Study 4

4.4.2.1 Mathematics and science development

A series of ANOVAs were carried out with Year and gender as independent variables (IV), and Mathematical Reasoning, Numerical Operations, or science as the dependent variable (DV). All main effects are reported in **Table 4.3**. There was a main effect of Year in each test with increasing scores with each ascending Year (paired comparisons p's < .001), but there was no significant effect of gender (p's > .232), and no significant interactions (p's > 231).

Table 4.3: *Study 4: Main effects of Year (Y) on Mathematical Reasoning (Maths), Numerical Operations (Num. Op.), and Science, with estimated marginal means (M) and standard error (SE). Significant effects in bold.*

DV	**Year *F*(2,129)**	**Y1 *M* (*SE*)**	**Y3 M (*SE*)**	**Y5 M (*SE*)**	**Gender *F*(1,129)**
Maths	**175.56, *p* < .001, η_p^2 = .728**	17.1 (0.9)	31.2 (0.9)	39.5 (0.8)	1.44, *p* = .233
Num. Op.	**123.30, *p* < .001, η_p^2 = .657**	8.3 (0.6)	12.8 (0.6)	21.2 (0.6)	*F* < 1
Science	**196.44, *p* < .001, η_p^2 = .753**	11.0 (1.2)	26.8 (1.2)	43.9 (1.2)	*F* < 1

4.4.2.2 Misconceptions development

An LMM was carried out with error type (misconception, uncommon error) as the within-subject factor, with Year and gender as between-subject factors. The significant main effects are reported in **Table 4.4**.

Table 4.4: *Study 4: Linear mixed model main effects of Year and error type on responses to the misconceptions test at Times 1 and 2.*

Main effect	**Statistics**	**Estimated means (*SE*)**
Year	*F*(2,312.4) = 25.22, *p* < .001	Year 1: 34.0% (1.3)
		Year 3: 28.3% (0.9)
		Year 5: 23.1% (0.9)
Error type	*F*(1,284.2) = 456.51, *p* < .001	Misconception: 43.3% (0.9)
		Uncommon error: 13.7% (0.9)

There were no main effects or interactions with gender (*p*'s > 201). There was a main effect of Year, with a falling proportion of incorrect responses made with each ascending Year (Year 1 to 3: *p* = .002; Year 3 to 5: *p* = .001). There was also a main effect of error type, with a higher proportion of misconception responses selected than other error responses (*p* < .001). These main effects were modulated by an interaction, *F*(2,284.2) = 11.22, *p* < .001. A follow-up LMM run for each error type revealed a significant Year effect in the proportion of misconception responses, *F*(2,150.4) = 22.91, *p* < .001, with no significant difference in proportion of misconception responses between Years 1 and 3 (*p* = .097), but a significant decrease in percentage of misconception responses between Years 3 and 5 (*p* < .001). Conversely, the significant Year effect in the proportion of uncommon error responses, *F*(2,284.2) = 11.22, *p* < .001, revealed a reduction in uncommon errors between Years 1 and 3 (*p* = .033) but no difference in proportion of responses between Years 3 and 5 (*p* = .116) (**Figure 4.5**).

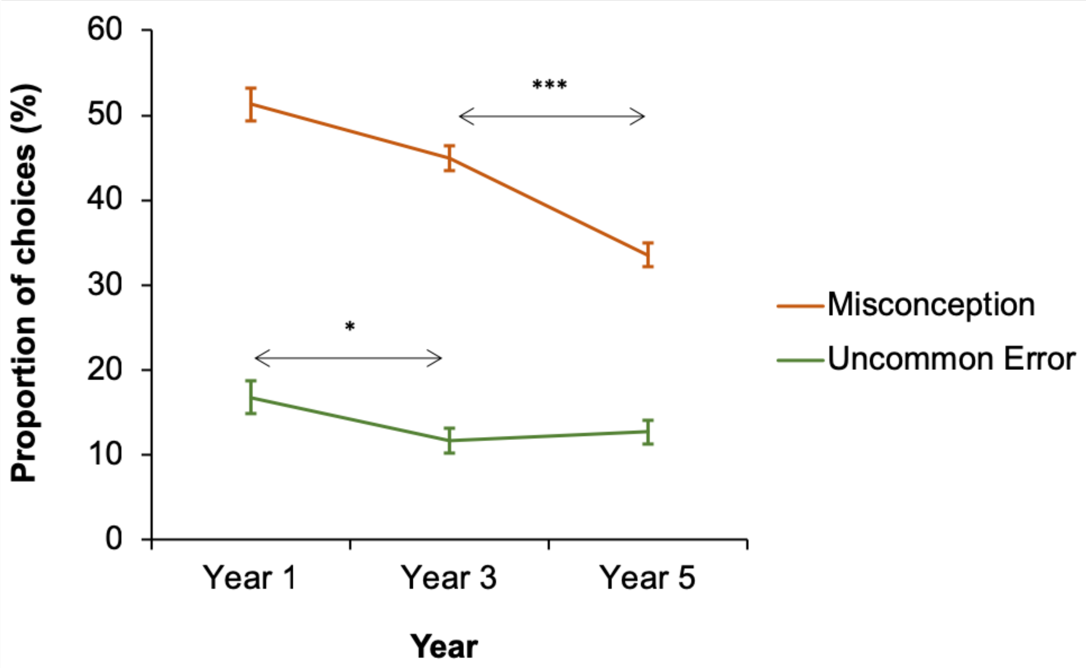


Figure 4.5: *Study 4: Incorrect misconception question responses as a function of error type and Year (estimated mean ± SE). *p < .05; **p < .01; ***p < .001.*

4.4.2.3 Executive Functions and IQ development

A further series of ANOVAs were carried out with Year and gender as IV, and each of the general intelligence and EF measures as the DV. All main effects are reported in **Table 4.5**. There was a main effect of Year in each test with increasing scores with each ascending Year for general intelligence, verbal WM, and visuospatial WM (paired comparisons *p*'s ≤ .001), and an increased effect of incongruency in Year 1 compared to Years 3 and 5 (*p*'s = .048 and .031), but no difference between Years 3 and 5 (*p* > .999). There was no significant effect of gender (*p*'s > .226), and no significant interactions (*p*'s > .239).

Table 4.5: *Study 4: Main effects of Year (Y) on general intelligence, visuospatial working memory, verbal working memory, and inhibitory control, with estimated marginal means (M) and standard error (SE). Significant effects in bold.*

DV	**Year *F*(2,129)**	**Y1 *M* (*SE*)**		**Y3 M (*SE*)**		**Y5 M (*SE*)**	**Gender *F*(1,129)**
General intelligence	**128.3, $p < .001$, $\eta_p^2 = .665$**	-1.0 (0.1)	<	0.2 (0.1)	<	0.8 (0.1)	$F < 1$
Visuospatial working memory	**60.78, $p < .001$, $\eta_p^2 = .485$**	4.2 (0.4)	<	7.8 (0.4)	<	10.9 (0.4)	$F < 1$
Verbal working memory	**41.79, $p < .001$, $\eta_p^2 = .393$**	5.4 (0.3)	<	7.9 (0.3)	<	9.6 (0.3)	1.47, $p = .227$
Semantic inhibitory control	**4.22, $p = .017$, $\eta_p^2 = .061$**	0.14 (0.02)	<	0.09 (0.02)	=	0.09 (0.02)	$F < 1$

4.4.2.4 Associations between mathematics, science, intelligence, and EF

Pearson's correlations and partial correlations controlling for age in months, were carried out to examine overall associations between all of the academic and domain-general variables (**Table 4.6**).

Table 4.6: *Pearson's correlations between scores on the mathematics and science tests, misconceptions test, general intelligence, working memory (WM), and inhibitory control (IC) measures above the diagonal, and partial correlations controlling for age in months below the diagonal. * p < .05, ** p < .01, *** p < .001. Significant effects are highlighted in bold.*

	Mathematical Reasoning	Numerical Operations	Science	Misconceptions	Intelligence	Visuospatial WM	Verbal WM	IC	Age
Math. Reas.		**.858 *** **	**.832 *** **	**-.546 *** **	**.875 *** **	**.709 *** **	**.716 *** **	**-.189 ***	**.842 *** **
Num. Op.	**.581 *** **		**.773 *** **	**-.501 *** **	**.778 *** **	**.678 *** **	**.658 *** **	-.167	**.793 *** **
Science	**.377 *** **	**.280 ** **		**-.488 *** **	**.828 *** **	**.637 *** **	**.629 *** **	**-.172 ***	**.869 *** **
Misconceptions	**-.332 *** **	**-.252 ** **	**-.201 ***		**-.497 *** **	**-.415 *** **	**-.347 *** **	-.099	**-.459 *** **
Intelligence	**.600 *** **	**.371 *** **	**.413 *** **	**-.236 ** **		**.674 *** **	**.696 *** **	**-.235 *** **	**.817 *** **
Visuospatial WM	**.312 *** **	**.283 ** **	.083	-.148	**.250 ** **		**.632 *** **	-.150	**.699 *** **
Verbal WM	**.457 *** **	**.346 *** **	**.230 ** **	-.088	**.417 *** **	**.353 *** **		**-.277 ** **	**.621 *** **
IC	.039	.030	.073	-.004	-.070	.013	-.149		**-.232 ** **

N = 135, df = 132
Notes: Math. Reas. = Mathematical Reasoning; Num. Op. = Numerical Operations

These analyses revealed significant positive associations after controlling for age between each mathematics and science measure, and the general intelligence measure. The associations with EF were more specific. Both mathematics tests positively correlated with both WM measures. However, science scores were only significantly correlated with verbal WM, and the proportion of misconception errors did not relate to either WM measure. IC did not associate with any mathematics or science measure after controlling for age.

As there was a participant identified as an outlier in the Numerical Operations data and another participant identified as an outlier in the IC data, the correlations were re-run with these participants removed from the sample to reveal their influence on the outcomes. After controlling for age, there were no differences in the pattern of results with the mathematics and science measures, indicating that the outliers are not unduly influencing the analyses.

A series of hierarchical regressions were run to identify significant independent predictors of variance in each mathematics and science measure. All general intelligence and EF measures were entered into the same model along with age in months as IVs (**Table 4.7**). DVs were Mathematical Reasoning, Numerical Operations, science, and percentage of misconception choices from Time 1.

Table 4.7: *Hierarchical regressions with DV of Mathematical Reasoning (Maths), Numerical Operations (Num. Op.), science, and misconceptions (Miscon.), and IV of age, general intelligence (IQ), visuospatial working memory (VSWM), verbal working memory (VWM), and inhibitory control (IC). Significant effects are highlighted in bold.*

	Maths		**Num. Op.**		**Science**		**Miscon.**	
	β	*p*	β	*p*	β	*p*	β	*p*
Age (months)	**.327**	**<.001**	**.396**	**<.001**	**.591**	**<.001**	-.112	.428
IQ	**.453**	**<.001**	**.272**	**.004**	**.337**	**<.001**	**-.353**	**.017**
VSWM	.081	.136	.126	.082	-.037	.539	-.132	.251
VWM	**.160**	**.003**	**.154**	**.031**	.065	.266	.048	.665
IC	.058	.116	.047	.342	.059	.151	-.011	.887
R^2	**.836**	**<.001**	**.706**	**<.001**	**.801**	**<.001**	**.263**	**<.001**

The regressions revealed that a large amount of variation was explained by age, general intelligence and EF factors in Mathematical Reasoning, Numerical Operations, and science, and a significant but lower amount of variation was explained in the misconception response choices. General intelligence was a significant unique predictor in all the models, while verbal WM significantly uniquely contributed to performance on the mathematics tasks only. Age was also a significant contributor to all the models except the misconception errors. When re-running the regressions without the outliers, there was very little difference in results in all models except Numerical Operations. Here, visuospatial WM explained unique variance (β = .204, *p* = .006) while verbal WM was non-significant (β = .120, *p* = .086).

As misconception data were collected at two time-points, a regression was carried out to determine whether the domain-general measures predicted the percentage of misconception choices at Time 2 (T2) after accounting for percentage of misconception choices at T1. T2 misconceptions was the DV, with T1 misconceptions, general intelligence, visuospatial WM, verbal WM, and semantic IC as the IVs. This model was not significant (*p* = .164) indicating that domain-general abilities at T1 did not predict choices on the misconception task at T2 after accounting for previous misconception choices.

Regressions were run to establish whether the relationships revealed above remained constant across Year groups. General intelligence and EF measures, and the dummy Year variables were entered in regression analyses in block 1, with each of the mathematics and science variables as DVs. Only misconception data from T1 were entered into this analysis, as these were collected at the same time as the general intelligence and EF data. Interaction terms between the dummy Year variables and the EF and intelligence IVs were entered stepwise in block 2. In three of the regressions, general intelligence was revealed as having a different association with the DVs in Year 1 than the other Years; Mathematical Reasoning, Year 1 x IQ, β = -.242, *p* = .002 (**Figure 4.6A**); Science, Year 1 x IQ, β = -.233, *p* = .007 (**Figure 4.6B**); Misconceptions, Year 1 x IQ, β = .416, *p* = .012 (**Figure 4.6C**). In the Numerical Operations analysis, Verbal WM had a different association in Year compared with the other Years; Year 1 x VWM, β = -.203, *p* = .004. In all cases, the association between the variables was weaker in Year 1 than in Years 3 and 5 (**Figure 4.6D**).

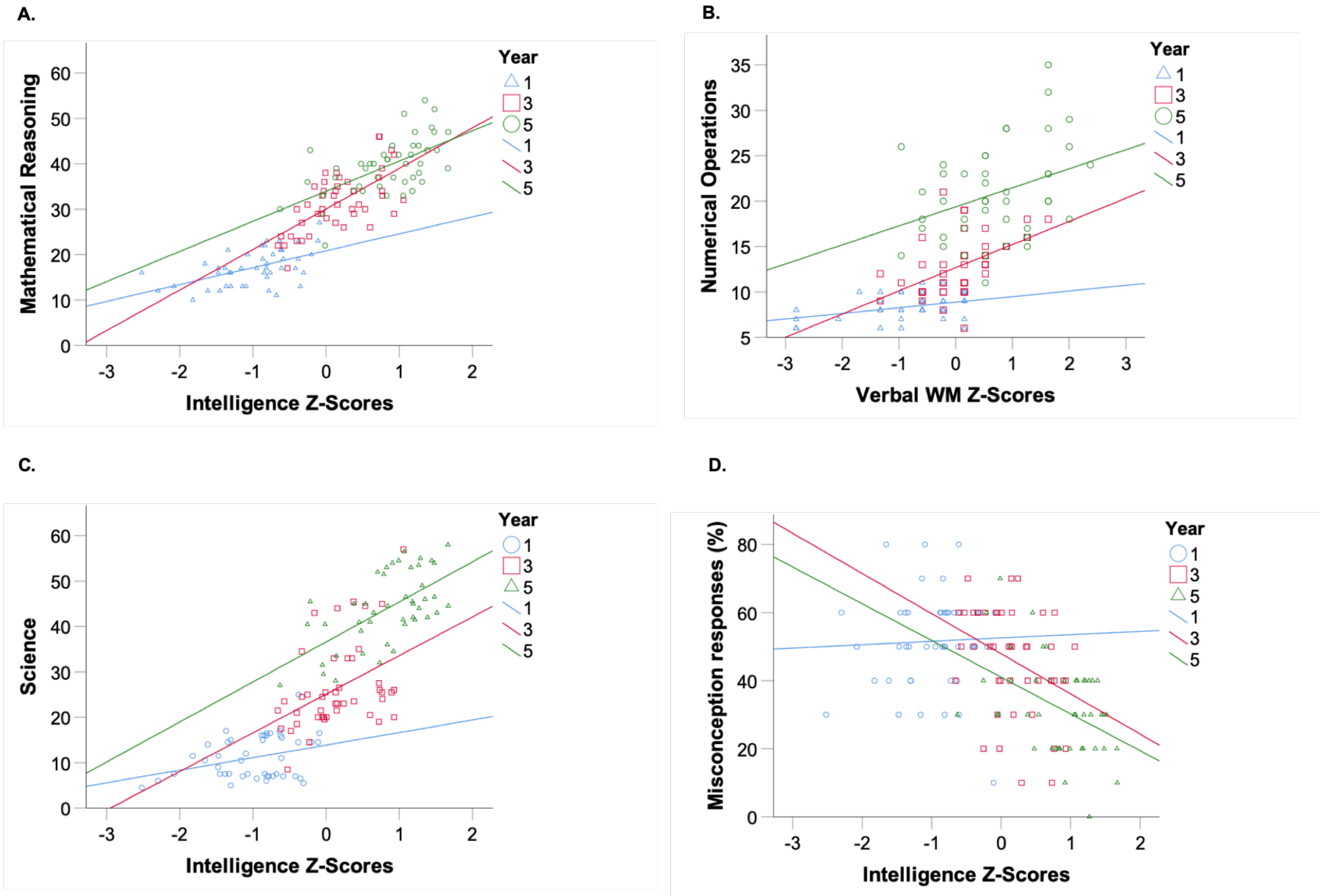


Figure 4.6: *Comparison of general intelligence z-scores and verbal working memory (WM) z-scores by Year with* ***A.*** *Mathematical Reasoning,* ***B.*** *Numerical Operations,* ***C.*** *Science, and* ***D.*** *Misconceptions.*

4.4.3 Study 5

4.4.3.1 Mathematics and EF main effects

To examine whether there were main effects of gender, a series of ANOVAs were carried out with each of the mathematics measures as the DV and gender as the IV. There were main effects of gender in every mathematics subset (**Table 4.8**), with higher scores achieved by males compared with females. However, the effect in the Number dataset would not survive Bonferroni corrections for multiple analyses; Number data were analysed as part of the arithmetic and total measures, therefore the corrected alpha is .017.

Table 4.8: *Main effects of gender with male and female estimated marginal means (M) and standard error (SE).*

DV	Gender F(1,48)	Male % M (SE)	Female % M (SE)
Number	4.52, p = .039, η_p^2 = .086	92.06 (3.10)	81.77 (3.03)
Word	**8.23, p = .006, η_p^2 = .146**	83.32 (3.57)	68.27 (3.49)
Arithmetic	**7.48, p = .009, η_p^2 = .135**	87.69 (3.08)	75.02 (3.01)
Shape	**5.95, p = .018, η_p^2 = .110**	72.62 (4.39)	56.19 (4.29)
Graph	**7.76, p = .008, η_p^2 = .139**	82.12 (3.10)	69.24 (3.03)
Total	**9.16, p = .004, η_p^2 = .160**	82.53 (2.99)	68.87 (2.92)

Further ANOVAs were run with individual EF measures as the DV. There were no main effects of gender (*p*'s > .256) in any of the EF measures.

4.4.3.2 Associations between mathematics subsets and EF

Pearson's correlations and partial correlations, controlling for age in months, gender and task order, were run to identify associations between the mathematics subsets and EF (**Table 4.9**).

Table 4.9: *Pearson's correlations between performance on the mathematics subtests, working memory (WM), and inhibitory control (IC) above the diagonal, and partial correlations controlling for age in months, gender, and task order, below the diagonal. * p < .05, ** p < .01, *** p < .001. Significant effects are highlighted in bold.*

	Number	Word	Arithmetic	Shape	Graph	Total	VSWM	VWM	SIC	RIC
Number		**.698 *****	**.904 *****	**.687 *****	**.599 *****	**.861 *****	**.334 ****	**.358 ****	-.087	-.048
Word	**.617 *****		**.937 *****	**.623 *****	**.618 *****	**.859 *****	**.294 ***	**.435 *****	-.126	.105
Arithmetic	**.882 *****	**.915 *****		**.707 *****	**.661 *****	**.932 *****	**.338 ****	**.434 *****	-.118	.040
Shape	**.627 *****	**.544 *****	**.647 *****		**.602 *****	**.877 *****	**.290 ***	**.379 ****	**-.282 ***	.008
Graph	**.530 *****	**.544 *****	**.597 *****	**.550 *****		**.813 *****	.197	**.409 *****	-.223	-.144
Total	**.825 *****	**.814 *****	**.911 *****	**.857 *****	**.788 *****		**.327 ****	**.463 *****	-.220	-.015
Visuospatial WM	**.341 ****	**.309 ***	**.360 ****	**.274 ***	.194	**.338 ****		**.413 *****	-.054	**.288 ***
Verbal WM	**.352 ****	**.457 *****	**.454 *****	**.364 ****	**.409 ****	**.479 *****	**.403 ****		-.057	-.051
Semantic IC	-.036	-.072	-.062	**-.249 ***	-.198	-.181	-.032	-.039		-.136
Response IC	-.054	.126	.048	.005	-.150	-.020	**.275 ***	-.061	-.126	

N = 71; df = 66

Notes: VSWM = visuospatial working memory; VWM = verbal working memory; SIC = semantic inhibitory control; RIC = response inhibitory control.

As expected, there were strong positive correlations between all the mathematical subsets after controlling for age, gender, and order. There was a strong correlation was between the Number and Word subsets, which is not surprising as the subsets included the same arithmetic content, however the strength of this relationship did not differ a huge extent from the strength of the associations between other subsets.

After controlling for age, gender, and order, verbal WM had a significant, positive association with all subsets, while visuospatial WM had no significant relationship with the Graph subset and a slightly weaker association than verbal WM with all other subsets. Semantic IC only significantly correlated with the Shape subset, while response IC revealed no significant association. Interestingly, the only significant associations between the EF measures were within WM, and a significant but weaker association between visuospatial WM and response IC. This suggests that the tasks were tapping unique aspects of the IC system but that the WM tasks measured common central elements as well as unique verbal and visuospatial elements of the WM system. However, it should be noted that the tasks are not pure measures of EF and that there are general task-specific processes which also contribute to individual differences in scores.

A series of hierarchical regressions were run to identify significant predictors of each mathematics subset. All EF measures, age, and gender were entered as IV, with each mathematical subset as the DV (**Table 4.10**). Task order was not entered here as the mathematical subsets were presented in the same order to all participants.

Table 4.10: *Hierarchical regressions with DV of Number, Word, Shape, and Graph, and IV of age in months, gender, visuospatial working memory (VSWM), verbal working memory (VWM), semantic inhibitory control (SIC), and response inhibitory control (RIC). Significant effects are highlighted in bold.*

	Number		**Word**		**Shape**		**Graph**	
	β	*p*	β	*p*	β	*p*	β	*p*
Age	.055	.623	.035	.743	.138	.201	-.013	.905
Gender	**-.264**	**.018**	**-.282**	**.008**	**-.235**	**.029**	-.190	.085
VSWM	.235	.069	.068	.573	.143	.245	.071	.577
VWM	.230	.065	**.386**	**.001**	**.272**	**.024**	**.349**	**.005**
SIC	-.052	.643	-.056	.594	**-.236**	**.030**	-.211	.060
RIC	-.026	.825	.204	.070	-.002	.989	-.097	.408
R^2	**.247**	**.005**	**.325**	**<.001**	**.304**	**<.001**	**.254**	**.004**

These regressions revealed that each model was significant, with the DVs explaining 24-33% of the variance in scores. Gender was a significant predictor in the Number, Word and Shape subsets, but not in the Graph subset. Verbal WM explained significant unique variance in the Word, Shape, and Graph subsets, but not the Number subset, while visuospatial WM was not a significant contributor in any model. Semantic IC explained significant unique variance in the Shape subset only, and response IC did not make a significant unique contribution to any model.

The arithmetic demands of the Number and Word subsets were equivalent in terms of operation, number of steps, and difficulty. This allowed the opportunity to separate the demands of the arithmetic calculation from the process of interpreting and responding to the word problem. To control for the arithmetic element of the word problem, a two-step regression was run with the Word subset score as the DV, and the Number subset score as the IV in block 1, and age, gender, and the EF measures in block 2 (**Table 4.11**).

Table 4.11: *A two-step hierarchical regression with DV of Word subset. In block 1, IV of Number subset was added, and in block 2, IVs of gender, age in months, visuospatial working memory (VSWM), verbal working memory (VWM), semantic inhibitory control (SIC), and response inhibitory control (RIC) were added. Significant effects are highlighted in bold.*

	IV	R^2	*P*	$R^2 \Delta$	*p*	β	*p*
Block 1		**.487**	**<.001**				
	Number					**.698**	**<.001**
Block 2		**.584**	**<.001**	**.097**	**.035**		
	Number					**.586**	**<.001**
	Age					.003	.976
	Gender					-.128	.140
	VSWM					-.069	.480
	VWM					**.251**	**.010**
	SIC					-.026	.756
	RIC					**.220**	**.015**

The score on the Number subset explained nearly half of the variance on the Word subset. The combination of EF measures, age, and gender explained a further significant portion of variance on the Word subset. After controlling for the Number subset, age, gender, and other EFs, both verbal WM and response IC were significant unique predictors of the Word subset.

4.5 Discussion

The studies in this chapter sought to examine the development of mathematics, science, and domain-general abilities. This is important for clarifying the changing relationships between mathematics or science, and domain-general abilities over time, and has made a valuable contribution to the literature as it compared different mathematical tasks and different types of WM and IC. Additionally, this chapter is an important first step for investigating and explaining the relationship between whole-part measures and mathematics and science (discussed in **Chapter 5**).

4.5.1 Developmental changes in mathematics and science

In study 4, mathematics and science scores improved with age, supporting the expected outcome that scores on these tasks will improve as children develop. This is likely to be partly due to improvements in domain-specific understanding, where older children have built on previous conceptual ideas and procedural knowledge (Aubrey, Dahl, & Godfrey, 2006; Canobi, 2004) and partly related to the development of IQ and EF (Cragg & Gilmore, 2014).

Children in all Years selected a greater proportion of misconception responses than uncommon errors, which replicates findings in a large-scale study with children aged 7 to 10 years completing multiple choice misconception questions (Wilkinson et al., 2019). In the current study, the changes in responses to the misconceptions questions were non-linear, with the proportion of uncommon errors reducing between Year 1 and Year 3, and the proportion of misconceptions reducing between Year 3 and Year 5. This may indicate that the uncommon errors were generally easier to dismiss or overcome in order to reach the correct response than common misconceptions. Younger children base their conceptual understanding on their first-hand experiences with the world, and the learning of new concepts can take time to assimilate with this early scientific framework (Vosniadou, 2019; Vosniadou, 2013). One small-scale study found that children aged 6-7 years responded to misconceptions based on their own experiences and observations but children aged 14-15 years responded with a mix of their intuitive knowledge and their learnt concepts (Thompson & Logue, 2006). It is important to note that although the content of this task was designed to test common misconceptions held by children, it is impossible to determine from these data whether individuals had competing conceptions.

4.5.2 Developmental changes and associations between domain-general tasks

4.5.2.1 Development of domain-general tasks

In study 4, general intelligence and WM improved with each Year group, but IC only improved between Years 1 and 3, and then remained stable between Years 3 and 5. This is supportive of studies revealing a protracted development of WM into adolescence (Isbell, Fukuda, Neville, Vogel, & Conklin, 2015; Thomason et al., 2009). Studies have also consistently identified improvement in IC in preschool and early schooling, whereas there has been mixed evidence for IC development after the age of 5 or 6 years (Best & Miller, 2010).

4.5.2.2 Associations between domain-general tasks

The two tasks measuring WM were designed to tap primarily the phonological loop (verbal WM) and the visuospatial sketchpad (visuospatial WM) although both tasks likely draw on the central executive resources as well (Cragg, Keeble, et al., 2017; Miyake et al., 2000). The correlations between the WM tasks in both studies support this: The tasks positively correlated with each other after controlling for any age or gender effects, indicating that they were measuring similar central resources. However, this correlation was only moderately strong which is suggestive of the fact that they were also measuring separable constructs. In contrast, the IC tasks in study 5 did not significantly correlate with each other, revealing that they are independent cognitive constructs (Tiego et al., 2018) or that they follow different developmental trajectories (Huizinga, Dolan, & van der Molen, 2006), impacting on their overall correlation.

There was very little association between WM and IC variables after controlling for covariates. Analyses from study 5 revealed a weak but positive association between response IC and visuospatial WM after controlling for age in months, gender, and task order, whereby participants who responded more accurately to the whack-a-mole task were more likely to have better visuospatial WM. Results from other studies looking at associations between WM and IC suggest that IC tasks require WM skills (Tiego et al., 2018; Van der Ven et al., 2012). Here, children may create a short-term representation of the mole and aubergine in WM which they then match to the stimulus and respond accordingly.

The only task with order effects was response IC. The task represents an individual's response IC, however it also draws on sustained attention to maintain focus on what is a straightforward but repetitive game. The group who completed this task third (out of 4 tasks) achieved lower scores than those who completed the task second. This may indicate that participants were more tired after completing two other tasks, leading to mistakes on the whack-a-mole task. A further possibility is that there was a lingering impact of the task completed prior to the response IC task. The lower-scoring group completed the visuospatial WM task before the response IC task. This was the only measure that significantly correlated with response IC scores, so may have contributed to the poorer performance as they draw on similar resources.

In sum, all general intelligence and EF tasks revealed improvements with development except semantic IC which remained stable between Years 3 and 5. The patterns of correlation suggest that the EF tasks were measuring distinct domain-general abilities. The overlap between WM tasks is likely to reflect the involvement of central executive WM resources, while the fact that the IC responses had no significant overlap suggests that semantic and response IC are two independent constructs. However, it should be noted that the task design was more similar between the two WM tasks than the two IC tasks, which may also have impacted on the strength of associations.

4.5.3 Associations between academic tasks and domain-general abilities

4.5.3.1 Mathematics

Previous research has consistently identified associations between WM and mathematics (Cragg, Keeble, et al., 2017; Geary, 2011; Lee & Bull, 2016), while associations between IC and mathematics has been more variable (Cragg, Keeble, et al., 2017; Hubber et al., 2014; Lubin et al., 2013). The prediction here was that verbal WM would associate more closely with Mathematical Reasoning than Numerical Operations, while for visuospatial WM, the reverse would be found. However, it was suggested that associations would change with development.

In line with expectation, general intelligence and WM revealed a positive association with Mathematical Reasoning and Numerical Operations after controlling for age and task order. The associations were stronger for Mathematical Reasoning than Numerical Operations which likely reflects the differing cognitive demands required for each task. Mathematical Reasoning questions were presented verbally and supported with visual stimuli, which required initial processing to assimilate and make sense of the information before solving the mathematical element. In contrast, Numerical Operations assesses counting and computation abilities which requires participants to follow a procedure and then retrieve and use mathematical facts. There was no significant mathematics association with IC in study 4, however semantic IC did associate with the Shape subset in study 5. This is consistent with the literature which has found a reliable association between WM and mathematics, and has only found an intermittent association with IC (Van der Ven et al., 2012).

As predicted, there were differences in association between verbal and visuospatial WM depending on the type of mathematics task. Significant positive associations were identified with each mathematical task in both study 4 and 5, after controlling for covariates. In all cases, verbal WM associated more closely with mathematical achievement than visuospatial WM, although the difference was often small.

In study 4, verbal WM and general intelligence were independent predictors of Mathematical Reasoning after accounting for age and the other EF measures. The data from study 5 supported this finding, with verbal WM uniquely predicting the Word subset scores after controlling for age, gender, and other EF. The Mathematical Reasoning task is a verbally-administered test with visual stimuli to support the question. Similarly, the Word subset of the mathematics test in study 5 embedded the arithmetic content within a 'real-life' situation, requiring participants to first make sense of the story in order to identify the arithmetic required to answer the question. In fact, only 49% of the variance on the Word subset was explained by the Number subset, indicating that half of the variance related to other factors. The positive association with verbal WM in both these tasks may reflect the resources required to remember and interpret the question, and may also be due to the retrieval of previously-learnt number facts from long-term memory (Cragg, Keeble, et al., 2017; De Smedt et al., 2009; Holmes & Adams, 2006).

When the outliers were excluded, visuospatial WM and general intelligence were independent predictors of Numerical Operations in study 4. This supports previously-identified associations between visuospatial WM and arithmetic tasks in children aged 6 to 10 years (Holmes & Adams, 2006), and 8 to 9 years (Alloway & Passolunghi, 2011; Meyer et al., 2010), although note that in these latter studies, the association with visuospatial WM was not significant in 7- to 8-year-olds. A study with 8- to 25-year-olds found an association between both verbal and visuospatial WM and procedural questions (Cragg, Keeble, et al., 2017). However, an important distinction in this study is that the task was a mental calculation task rather than a task using written methods, which may explain the additional involvement of verbal WM. The positive association with visuospatial WM may result from participants using visuospatial strategies such as place value columns or mental number-lines during the completion of the task. In study 5, the Number subset also required participants to complete written calculation methods, however the only significant unique contribution to the

Number subset scores was gender. Although visuospatial WM significantly associated with all subsets, it did not explain unique variance on any of the subsets. It is unclear why the associations differed between the two written calculation tasks. It is possible that the variance attributed to gender may have overlapped with variance associated with visuospatial WM, and that this was more apparent in the Number subset than in the Numerical Operations task. Interestingly, gender did not make a unique contribution to Word subset scores after controlling for the Number subset, which is suggestive of the arithmetic element driving the gender effect on the Word subset.

In study 5, verbal WM significantly associated with both the Shape and Graph subsets, but visuospatial WM only weakly associated with the Shape subset and the relationship was non-significant with the Graph subset. When accounting for variation associated with age, gender, and other EFs, only verbal WM explained unique variance on both the Shape and Graph subsets. It is perhaps surprising that visuospatial WM was not an independent contributor, given that the stimuli required participants to read the spatial elements of the shapes, graphs, and tables. Previous research has revealed an association between visuospatial WM but not verbal WM and ‘shape, space and measures’ questions (Holmes & Adams, 2006). It is possible that the association with verbal WM in the current study relates to the interpretation of the mathematical language used in the shape questions such as acute angle, quadrilaterals, and net. If these concepts are still at an early stage of familiarity, it may be that verbal WM is required to temporarily store the meaning of these words whilst working out the answer to the question. In the Graph subset, verbal WM may have been tapped to interpret the question as well as to identify and remember key elements, before the stimulus could be used to answer the question.

The only significant correlation with IC in both studies was between the Shape subset in study 5 and semantic IC, whereby those with a greater RT cost of incongruent stimuli (and therefore a lower IC score) achieved lower scores on the Shape subset. Semantic IC continued to explain unique variance on the Shape subset after accounting for age, gender, and other EFs. This likely reflects the importance of overcoming distracting visual information, such as questions about acute or obtuse angles that require all other angles to be ignored. It may also reflect common processes involved with the understanding of visual size, as the animal size Stroop required participants to select the largest animal, and the Shape subset included several questions relating to angle size. Response IC, but not semantic IC, explained variance on the Word subset after controlling for accuracy on the Number subset, age, gender, and other EFs. This may reflect processes involved in interpreting the question, whereby a more delayed response could allow the question to be fully understood and the calculation to be planned.

It is interesting to note that verbal WM associated with the mathematics tasks in both studies despite the differing task design; reverse digit span in study 4 and reverse letter span in study 5. This suggests that the significant association in study 4 was not confounded by the use of numeric data in the verbal WM task. This contrasts with a study where an association was found between WM tasks using numeric content but not word content (Passolunghi & Cornoldi, 2008). However, the participants in this study had mathematical difficulties which may have impacted on WM associations using domain-specific content.

4.5.3.2 Science

The limited research examining domain-general associations with science has identified a possible closer association with visuospatial WM than verbal WM (Rhodes et al., 2016; St Clair-Thompson & Gathercole, 2006), and an association with IC when a prior concept is needed to be suppressed (Vosniadou et al., 2018). In contrast to these findings, the science test here did not significantly associate with either visuospatial WM or IC after controlling for age and task order. This may be explained by the design of this science task.

Significant correlations were identified with general intelligence and verbal WM, however, after controlling for other EFs only general intelligence and age in months made significant unique contributions to science scores. Although verbal WM may play a role in science achievement, the results here suggest that this represents overlapping variance with other factors. The science test was administered towards the beginning of the academic year, therefore the extent to which students were using previously-learnt knowledge may be variable. This would explain the lack of association with WM and IC, as participants may not have been recalling, manipulating, and selecting concepts.

A second possibility is that WM and IC are more relevant for some types of science questions than others, such that their individual contributions are not sufficient to elicit a significant result when examining achievement at an overall level. The questions in this science test required some knowledge recall, as well as the use of reasoning skills to apply knowledge to the context of the question. For example, in order to group pictures of animals into birds, mammals, and fish, participants had to recall features of each group, and then apply these to the animals presented to them. However, a study with adolescents aged 13-14 years which used a general measure of science, did identify associations with both visuospatial and verbal WM (Alloway, Banner, & Smith, 2010), which could indicate that these associations are more likely to be revealed only in older children and adolescents.

4.5.3.3 Misconceptions

Recent studies have identified associations between misconceptions and IC in children (Wilkinson et al., 2019) and adolescents (Brookman-Byrne et al., 2018). In line with expectation, the current study revealed a negative association between the percentage of misconception responses and achievement on the mathematics and science tests. Participants scoring higher on the mathematics and science tests were therefore less likely to select the misconception response. However, contrary to expectation, IC did not significantly associate with the proportion of misconception responses. In fact, the only domain-general measure which significantly associated with misconceptions was general intelligence, which was also the only unique predictor of misconception choices. The lack of association with IC may indicate that the short misconceptions test did not include a sufficient number of questions where participants had competing concepts which needed to be overcome. It was perhaps surprising that the correlation between misconceptions and age was much lower than between age and the other academic measures, given that all Year groups completed the same task. This is likely to be due to the non-linear changes in misconception responses described in **Section 4.4.2.2**.

4.5.3.4 Changes in association with development

The associations between domain-general and mathematics and science scores remained constant between the ages of 5 and 10 years with the exception of general intelligence associations with Mathematical Reasoning, science and misconceptions, and between verbal WM and Numerical Operations. In all these cases, the strength of the relationship was weaker in Year 1 than in Years 3 and 5. The weaker relationship with general intelligence may reflect differing cognitive resources required between the Year groups to complete the task, or it may be due to a lower variance in the academic scores in children 5- to 6-years-old. This highlights the importance of designing tasks which elicit a variety of responses across all age groups.

The weaker association between Numerical Operations and verbal WM in Year 1 may be related to the type of questions that the younger Year group would have completed on the task. The early questions involved number recognition, writing number, and very early arithmetic. It is possible that these early mathematical skills do not tap WM in the same way as more complex arithmetic tasks which require the participant to remember the operand and access number facts from long-term memory in order to find a solution (Lee & Bull, 2016). Although younger children likely use WM resources to help support early mathematical tasks such as counting, there may be less of a need to access verbal WM in order to succeed on these early questions, than the more challenging questions for older children.

4.5.4 Limitations

The academic tests completed in both studies have been used as an indicator of the participants' ability in those subjects. However, this only gives a snapshot of their abilities, and it may not reveal their true performance on specific concepts. In order to achieve a more complete view, the participants would need a greater number of repetitions of the conceptual content and it should be repeated a number of times over a period of time. This would additionally allow an analysis of how the general intelligence and EF measures predict conceptual development over time, rather than concurrent conceptual understanding and procedural skill.

The content of the two studies were quite distinct, with study 4 providing an overall picture, while study 5 examined specific EF differences between subsets of mathematics. However, a key difference was that general intelligence was accounted for in study 4, but not in study 5. A more complete comparison could have been made with more consistent task selection in both studies. It would also be interesting to repeat the tasks in study 5 with children of different ages, to assess whether these associations changed with development.

4.5.5 Conclusion

Domain-general abilities are significantly associated with achievement in mathematics and science, with WM revealing a closer association with the academic measures than IC. There were distinct patterns of association in each mathematical test, with reasoning and word problems associating more closely with verbal WM, and the Numerical Operations task associating more closely with visuospatial WM. The only mathematics task which associated with IC was the Shape subset. Contrary to expectation, there was no association between IC and misconception responses. General intelligence and age were the only unique predictors of the science scores.

Having examined the associations between domain-general variables and the mathematics and science scores, whole-part task associations with mathematics and science will be examined in the next chapter. The analyses in **Chapter 5** will evaluate the direct associations, as well as examining overlapping variance with domain-general factors which may explain the direct associations.

Chapter 5: Associations between whole-part measures and mathematics and science achievement

5.1 Introduction

Chapter 1 highlighted the incomplete literature on associations between whole-part visual processing and achievement in mathematics and science. Through the course of this thesis, the developmental changes in whole-part visual processing tasks (**Chapter 2**) and mathematics and science tests (**Chapter 4**) have been examined in depth. Alongside this, studies in **Chapters 3 and 4** have revealed associations with domain-general factors which may help to explain any direct relationships between whole-part tasks and mathematics and science. This chapter will now draw these analyses together, looking first at direct associations between whole-part measures and mathematics and science achievement, and then at associations after accounting for domain-general factors.

5.2 Current studies

5.2.1 Overview

Analyses in **Chapter 2** revealed that there were few associations between the whole-part tasks, indicating that the tests and questionnaires were generally measuring distinct constructs. This supports previous studies comparing performance on a number of whole-part measures (Chamberlain et al., 2017; Milne & Szczerbinski, 2009). Therefore, the relationships between whole-part tasks and mathematics and science will be examined separately in this chapter, as it is not possible to calculate a valid composite which reflects an individual's overall advantage towards either the whole or the parts.

5.2.1.1 Global / local processing

There has been no research examining the direct association between global and local processing, and mathematics and science. However, the indirect associations between atypical global and local processing in autism (Almeida et al., 2014; Booth & Happé, 2010; Happé & Frith, 2006), and the higher proportion of people with autism involved in STEM courses and careers (Baron-Cohen et al., 2007; Wei et al., 2013), could indicate that an association exists. Therefore, the first aim was to clarify whether global and local measures associate with mathematics and science achievement in neurotypical children aged 5 to 10 years. The second aim was to consider the role of domain-general factors in the associations. We predicted that local processing may have a closer association with the academic tasks, on the basis that the literature broadly identifies a relative advantage in local processing in autistic populations. It is possible that relationships between the global and local processing measures and mathematics and science performance could be partly explained by inhibitory control (IC), as responses on the Navon task might be affected by conflicting information at the non-target level. We therefore anticipated that any association between incongruent trials on the Navon selective attention task and the mathematics and science tests, might be driven by individuals differences in IC. **Chapter 2** revealed that there were differences in response depending on the attentional demands of the Navon task, and that changes in global and local processing over development were non-linear. Therefore, associations were examined for each Navon task separately, and for each Year separately, to establish how they may change over development.

5.2.1.2 Field independence

Research has identified a positive association between field independence (FI) and academic measures in children and adolescents (Alamolhodaei, 2002; Buriel, 1978; Tinajero & Paramo, 1997). However, there is limited understanding about the underlying processes which explain this association, and whether these change with development. The aim here was to establish whether FI associated with achievement on mathematics and science tests, and whether these associations were explained by individual differences in age, general intelligence, working memory (WM) and IC. The prediction was that higher FI scores would associate with higher mathematics and science scores, and that much of this relationship would be explained by overlapping variance with general intelligence and executive functions (EF) based on the associations revealed in **Chapters 3 and 4**. Further, there was an expectation that Mathematical Reasoning and science scores would more closely associate with the FI measure of disembedding (where an embedded or hidden element is located within a more complex background) than the FI measure of segmenting (where elements are separated at predetermined locations), due to the need to separate concepts from the context of the question. Additionally, it was predicted that associations may strengthen with age, due in part to EF development, and therefore differing patterns may be observed between school Years.

5.2.1.3 Systemizing, Empathizing, and autistic traits

Higher systemizing (SQ) scores have been found to associate with higher mathematics scores in adults (Bressan, 2018), and those enrolled on science courses at university have been found to have higher SQ scores than those on humanities courses (Billington et al., 2007; Focquaert et al., 2007; Svedholm-Häkkinen & Lindeman, 2016). There has been limited research on associations between systemizing and academic achievement in children, but in one study with 7- to 12-year-olds, SQ scores had a positive association with problem-solving before accounting for IQ and reading ability, while empathizing quotient (EQ) had a negative association with calculation scores (Escovar et al., 2016). We therefore predicted that higher systemizing scores using the children's scale (SQ-C) would associate with higher Mathematical Reasoning achievement, and higher scores on the questionnaire designed to measure empathizing scores in children (EQ-C) would associate with lower Numerical Reasoning achievement. Additionally, we expected to find a positive association between scores on the short autism trait questionnaire (AQ-10) and mathematics achievement. We did not expect these associations to change over development, as children's SQ and EQ scores have been measured as remaining stable over time in previous research (Wakabayashi, 2013), and there were no Year effects in the SQ-C, EQ-C, or AQ-10 data identified in **Chapter 2**.

5.3 Method

5.3.1 Participants

The data here were drawn from two samples. Both were approved by the local ethics committee. Study 6 included cross-sectional and longitudinal data with the same participant group introduced in **Chapter 2** (sample 1). The IQ, EF, mathematics and science measures were collected at Time 1 (T1) only, while Navon, FI, and misconceptions data were collected at T1 and Time 2 (T2). At T1, an opt-out procedure was used, and at T2 an opt-in procedure was used. Study 7 included data from Year 5 children who were introduced in **Chapter 3** (sample 2), and used an opt-in method. The children in both studies gave verbal consent before any testing. A summary of participants can be found in **Table 5.1**.

Table 5.1: *Summary of participant demographics from study 6 and study 7. Participant demographics for the study 6 subset of participants with systemizing, empathizing, and autistic traits data are also shown.*

	Study 6					Study 6 subset			Study 7
	Year 1	**Year 3**	**Year 3**	**Year 5**	**Year 5**	**Year 1**	**Year 3**	**Year 5**	**Year 5**
	T1___	**___T2**	**T1___**	**___T2**	**T1**	**T1**	**T1**	**T1**	
Number	45 [(a)]	28 [(a)]	45 [(b)]	32 [(b)]	45	22	20	23	71
Male : Female	16 : 29	10 : 18	27 : 18	20 : 12	24 : 21	8 : 14	11 : 9	11 : 12	35 : 36
Mean age in months (SD)	68.6 (3.3)	92.2 (3.6)	92.8 (3.5)	114.2 (3.6)	115.4 (3.6)	68.5 (3.3)	91.7 (3.3)	116.9 (3.6)	124.1 (3.1)
Mean age in years; months	5; 8	7; 6	7; 6	9; 6	9; 7	5; 8	7; 5	9; 8	10; 4

[(a)] Longitudinal cohort in Year 1 at T1 and Year 3 at T2
[(b)] Longitudinal cohort in Year 3 at T1 and Year 5 at T2
Notes: T1 = Time 1; T2 = Time 2

5.3.2 Summary of tasks

A summary of the tasks can be viewed in **Tables 5.2A and B**. For further details about the content and administration of the tasks, see **Chapter 2** regarding the whole-part tasks, **Chapter 3** for the general intelligence and EF tasks, and **Chapter 4** for the mathematics and science tasks.

Table 5.2A: *Summary of tasks and measures from study 6.*

Measure	**Task**
Global / local responses	Free choice Navon task
	Selective attention Navon task
	Divided attention Navon task (big-small)
Disembedding	Children's Embedded Figures Test (CEFT)
Segmenting	Design Organisation Test (DOT)
Systemizing	Systemizing quotient for children (SQ-C)
Empathizing	Empathizing quotient for children (EQ-C)
Autism traits	Autism trait questionnaire (AQ-10)
Mathematics	WIAT Mathematical Reasoning
	WIAT Numerical Operations
Science	Novel science test
Misconceptions	Novel misconceptions test (Time 1 and Time 2)
General intelligence	Visuospatial: Raven's Coloured Progressive Matrices
	Verbal: British Picture Vocabulary Scale 3 (BPVS-3)
Visuospatial WM	Reverse spatial span
Verbal WM	Reverse digit span
Semantic IC	Animal size Stroop

Notes: WIAT: Wechsler Individual Achievement Test; WM: working memory; IC: inhibitory control.

Table 5.2B: *Summary of tasks and measures from study 7.*

Measure	Task
Disembedding	Children’s Embedded Figures Test (CEFT)
Mathematics	Novel task with four subsets (Number, Word, Shape, Graph)
Visuospatial WM	Reverse spatial span
Verbal WM	Reverse letter span
Semantic IC	Animal size Stroop
Response IC	Whack-a-mole

Notes: WM: working memory; IC: inhibitory control.

5.3.3 Global and local processing

As described in **Section 2.2.3**, there were three Navon tasks completed by the participants, each with differing attentional demands. This chapter will examine whether scores on each of the global or local processing tasks relate to achievement on the mathematics and science academic tasks.

The free choice Navon task score was represented by the proportion of global matches. Linear integrated speed-accuracy scores (LISAS; see **Section 2.2.7.1**) were calculated for the selective attention and divided attention Navon tasks to represent global and local responses on each of those tasks. Composite variables using LISAS scores were calculated to represent each individual’s advantage towards either global or local processing on each of the selective and divided attention tasks. These were calculated using the following formula: (global LISAS – local LISAS) / global LISAS. The resultant LISAS measure denoted a relatively more global advantage with negative figures and a relatively more local advantage with positive figures. The further the composite calculation is from 0, the greater the percentage difference between the global and local responses. LISAS scores were also calculated for responses to different congruency trials on the selective attention Navon task. Responses to the congruent and neutral trials using LISAS data did not significantly differ from each other (see **Section 2.3.1.2**), therefore a mean LISAS score was calculated to represent an amalgamated response to congruent and neutral trials. Responses to incongruent trials enabled an analysis of how performance in a Navon task condition with interfering information on the Navon task may differentially relate to mathematics and science.

5.3.4 Field independence

The study 6 tasks are described in greater detail in **Sections 2.2.4** and study 7 tasks in **Section 3.2.3.2.** The disembedding task in study 6 was a reduced form of the standard paper version of the Children's Embedded Figures Test (CEFT), comprising a mix of tent and house trials. The disembedding task in study 7 was a computerised version of the CEFT comprising only tent trials, as these were more suitable for being adapted for use with the eye-tracking technology. In both studies, the measure of interest was percentage accuracy on the task. Segmenting was measured in study 6 using the Design Organisation Test (DOT), which was a timed pen-and-paper task. The number of correct responses within the time limit was recorded.

5.3.5 Systemizing, empathizing, and autism traits

These questionnaires are described in greater detail in **Section 2.2.5**. There are four variables used in this chapter: SQ-C, EQ-C, D-Score, and AQ-10. The systemizing and empathizing scores represented the number of responses on the questionnaire relating to those traits, while the D-Score is a composite measure giving an indication of each individual's relative strength on either trait, such that positive integers denote a relative strength in systemizing over empathizing. The autism traits score is determined by each individual's score on the questionnaire.

5.3.6 Design and procedure

The testing for both studies took place at the participants' primary school during lesson times. Study 6 data were collected in the first of three terms of the school year. The Study 6 T2 global and local processing, CEFT, DOT, and misconceptions data were collected two years later from consenting T1 participants in the lower two Year groups. Study 7 data were collected during term three of the school year, as part of the eye tracking study described in **Chapter 3**.

Each task was explained to the participants immediately prior to them attempting the activity, and practice trials were completed for all tasks. Accuracy percentage and response time (RT) data were recorded for the Navon tasks, CEFT, misconceptions test, and IC tasks. Scores reflecting the number of correct responses were recorded for Mathematical Reasoning, Numerical Operations, science test, DOT, general intelligence and WM tasks. Mean RTs were calculated from correct responses only for the selective and divided attention Navon tasks and the animal size Stroop task. Computerised stimuli were presented via MATLAB (R2010b), using the Cogent toolbox (http://www.vislab.ucl.ac.uk/cogent_2000.php), on a 12-inch Dell laptop. The task order was counterbalanced across participants in each Year in study 6. The EF tasks in study 7 were counterbalanced, but all participants completed the groups of tasks in the same order over three sessions: Session 1 for the EF tasks, session 2 for the mathematics test, and session 3 for the eye-tracking task.

5.3.7 Statistical analyses

Response data from the whole-part tasks, mathematics and science tests, and domain-general tasks were analysed using analysis of variance (ANOVAs) and correlations. Regressions were then run to determine the extent to which direct relationships between whole-part tasks and mathematics and science tests were explained by domain-general tasks. Further regressions were run with dummy Year variables to establish whether associations differed by Year group. Any significant interactions with the dummy Year variables indicated that the association differed by Year.

All children attempted and completed all tasks, except a Year 1 boy who did not meet the minimum standard on the Navon divided attention task and was therefore excluded from this dataset. Once the analyses had been carried out, outliers were excluded and the analyses re-run, in order to assess their impact on the outcomes (further details in **Section 2.2.7.2**). This applied to any outliers identified in previous chapters as well as a further outlier that was identified in the LISAS composite % score, where one Year 3 child had a large negative (more local) score which was outside the lower boundary of selective attention.

5.4 Results

5.4.1 Global / local processing and mathematics and science

In **Sections 2.3.1.2 and 2.3.1.3**, Year and gender effects were reported using linear mixed models (LMMs) which analysed T1 and T2 data together. Here, a series of ANOVAs were carried out with T1 data only as this corresponded with the time at which mathematics and science data were collected. Year and gender were the IV and each of the global / local processing variables were the DV. All Year and gender main effects are reported in **Table 5.3**.

Table 5.3: *Study 6 (time 1 only): Main effects of Year and gender on Navon task responses. LISAS scores (ms) were used for the selective attention and divided attention tasks. Composite measures were calculated as (global LISAS – local LISAS) / global LISAS. Estimated marginal means (M) and standard error (SE) are provided for each Year group and each gender. Significant effects (Bonferroni-corrected) are highlighted in bold.*

Task	**DV**	**Year *F*(2,129)**	**Y1 *M* (*SE*)**		**Y3 *M* (*SE*)**		**Y5 *M* (*SE*)**	**Gender *F*(1,129)**	**Male *M* (*SE*)**	**Female *M* (*SE*)**
Free choice	Global %	**51.18, *p* < .001, η_p^2 = .134**	55.0 (5.6)	<	90.3 (5.5)	=	72.2 (5.4)	**4.02, *p* = .047, η_p^2 = .030**	66.1 (4.5)	78.9 (4.5)
Selective attention	Global	**51.18, *p* < .001, η_p^2 = .442**	1488 (41)	>	1193 (41)	>	907 (40)	*F* < 1	1188 (33)	1204 (33)
	Local	**34.29, *p* < .001, η_p^2 = .347**	1384 (37)	=	1278 (36)	>	975 (36)	5.20, *p* = .024, η_p^2 = .039	1164 (30)	1260 (30)
	Composite	**4.74, *p* = .010, η_p^2 = .068**	4.1 (4.1)	>	-11.5 (4.0)	=	-10.9 (4.0)	3.49, *p* = .064	-1.8 (3.3)	-10.4 (3.3)
	Global cong. & neut.	**49.75, *p* < .001, η_p^2 = .435**	1404 (41)	>	1135 (40)	>	843 (39)	*F* < 1	1125 (33)	1130 (32)
	Global Incong.	**10.54, *p* < .001, η_p^2 = .140**	1416 (71)	=	1249 (69)	>	973 (68)	2.40, *p* = .124	1151 (57)	1274 (56)
	Local cong. & neut.	**31.53, *p* < .001, η_p^2 = .328**	1310 (37)	=	1209 (36)	>	919 (36)	3.39, *p* = .068	1107 (30)	1184 (29)
	Local Incong.	**24.19, *p* < .001, η_p^2 = .273**	1441 (46)	=	1339 (45)	>	1017 (44)	6.32, *p* = .013, η_p^2 = .047	1200 (37)	1332 (37)
Divided attention	Global	**36.73, *p* < .001, η_p^2 = .366**	1496 (45)	>	1206 (43)	>	964 (42)	*F* < 1	1207 (36)	1237 (35)
	Local	**20.53, *p* < .001, η_p^2 = .243**	1382 (53)	>	1130 (51)	>	915 (50)	*F* < 1	1147 (42)	1138 (41)
	Composite	*F* < 1	7.0 (3.2)		5.3 (3.1)		4.3 (3.0)	*F* < 1	5.4 (2.6)	5.7 (2.5)

Notes: Y = year; cong. = congruent; neut. = neutral; incong. = incongruent; LISAS = linear integrated speed-accuracy scores

There was a main effect of Year in every IV except the LISAS composite score on the divided attention task (p = .830). There were gender effects on the free choice Navon global choices, the selective attention Navon local scores, and local incongruent trials (all other p's > .065), with males making fewer global choices on the free choice task, and achieving a lower LISAS score (and therefore a better performance) on the selective attention local task than females. However, the selective attention gender effects would not survive a Bonferroni correction to the alpha based on multiple analyses. As the selective attention data were analysed as global and local, composite, and as congruency measures, the data were analysed in three different ways, so the corrected alpha would be .017. There were no significant interactions (p's > .232).

One prediction was that IC might explain any differences in association due to the congruency effect on the selective attention Navon task. However, Pearson's partial correlations controlling for age in months, between LISAS scores on the selective attention task and IC were non-significant (p's > .794). Differences in score between the congruent/neutral trials and incongruent trials appear not to be explained by individual differences in semantic IC score.

Pearson's correlations were carried out to examine associations between the global and local measures and the mathematics and science measures (**Table 5.4**). As significant Year and gender effects were identified, partial correlations were also run to covary age and gender. There was a small but significant positive association between the percentage of global matches on the free choice task and Mathematical Reasoning scores, and a small but significant negative association between the global selective attention LISAS scores and Mathematical Reasoning after controlling for age in months and gender. However, the selective attention association would not survive Bonferroni corrections for multiple correlations. The analyses were re-run after the outliers had been removed, revealing no difference in the significance of the partial correlations.

Table 5.4: *Study 6: Correlations between scores on the global and local processing Navon tasks and mathematics and science tests at Time 1. * p < .05, ** p < .01, *** p < .001. Significant effects (Bonferroni-corrected) are highlighted in bold.*

Task	**Measure**	**Pearson's correlations**				**Partial correlations controlling for age and gender**			
		Mathematical Reasoning	Numerical Operations	Science	Misconceptions	Mathematical Reasoning	Numerical Operations	Science	Misconceptions
Free choice	Free Choice % [(a)]	**.226 ****	.118	.140	-.057	**.172 ***	-.035	-.026	.011
Selective attention	Global [(a)]	**-.643 *****	**-.569 *****	**-.616 *****	**.363 *****	-.192 *	-.081	-.091	.081
	Local [(a)]	**-.516 *****	**-.517 *****	**-.523 *****	**.259 ****	-.062	-.137	-.077	-.024
	Composite [(a)]	**-.206 ***	-.125	-.168	.160	-.073	.078	.043	.091
	Global cong. & neut.	**-.620 *****	**-.561 *****	**-.607 *****	**.347 *****	-.150	-.074	-.079	.065
	Global incong.	**-.401 *****	**-.332 *****	**-.367 *****	**.247 ****	-.131	-.043	-.065	.070
	Local cong. & neut.	**-.506 *****	**-.503 *****	**-.506 *****	**.282 ****	-.065	-.125	-.057	.020
	Local incong.	**-.439 *****	**-.458 *****	**-.453 *****	**.175 ***	.009	-.109	-.029	-.103
Divided attention	Global [(b)]	**-.571 *****	**-.515 *****	**-.559 *****	**.304 *****	-.128	-.068	-.078	.029
	Local [(c)]	**-.436 *****	**-.361 *****	**-.392 *****	**.283 ****	-.079	.029	.048	.085
	Composite [(b)]	-.086	-.131	-.138	.031	-.036	-.112	-.139	-.007

[(a)] N = 135, df = 131
[(b)] N = 133, df = 129
[(c)] N = 134, df = 130
Note: Cong. = congruent; neut. = neutral; incong. = incongruent; WM = working memory; IC = inhibitory control

A multiple regression was carried out on T1 data with Mathematical Reasoning as the DV to reveal whether associations with global and local processing in the Navon tasks could be explained by domain-general factors. Age in months and gender as well as the domain-general IVs of general intelligence, WM, and semantic IC were entered in the first model, and the free choice Navon task measure was entered in the second model (**Table 5.5**). When the domain-general factors were included in the model, the global free choice percentage scores did not uniquely predict the Mathematical Reasoning scores.

Table 5.5: *Study 6 (time 1 only): Multiple regression to identify unique variance in Mathematical Reasoning scores explained by the free choice Navon task variable, after covarying for domain-general factors. Age in months, gender, general intelligence (IQ), visuospatial working memory (VSWM), verbal working memory (VWM) and inhibitory control (IC) were entered as predictors in model 1, and the free choice global percentage was entered in model 2. Significant effects are highlighted in bold.*

			Navon task		**Age**		**Gender**		**IQ**		**VSWM**		**VWM**		**IC**	
DV	R^2	***p***	$R^2\ \Delta$	***p***	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*
Mathematical Reasoning																
Model 1	**.838**	**< .001**			**.324**	**<.001**	-.054	.146	**.439**	**< .001**	.079	.146	**.173**	**.002**	.059	.111
Model 2 (Free choice global %)	**.841**	**< .001**	.002	.172	**.323**	**< .001**	-.060	.106	**.431**	**< .001**	.079	.141	**.169**	**.002**	.050	.172

Notes: LISAS = linear integrated speed-accuracy scores

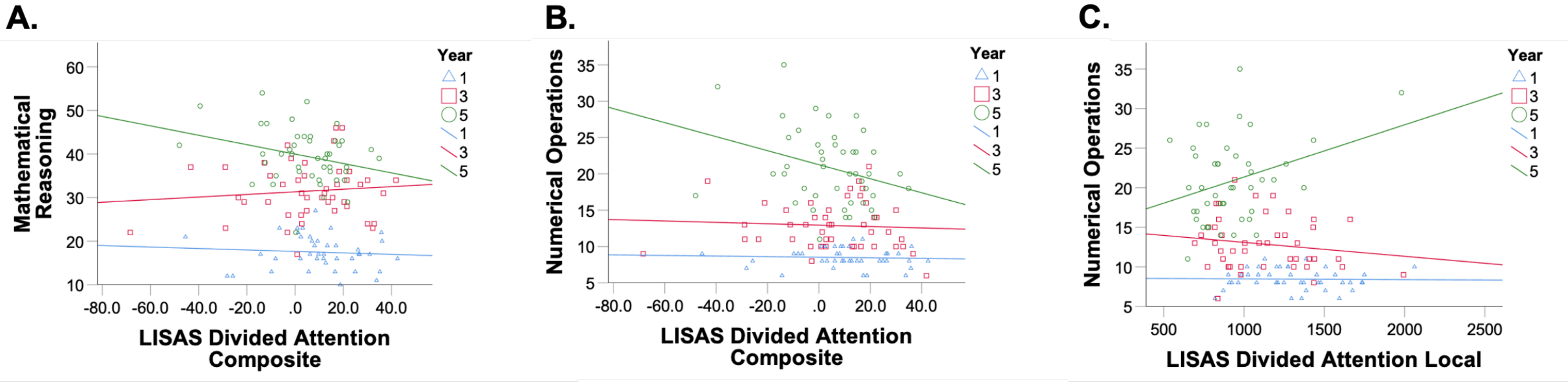
Further correlations were carried out with the T2 data, which only included two Year groups, and one measure of mathematics or science (misconceptions) (**Table 5.6**). There were no significant associations between the Navon tasks and the misconceptions test at T2. This is in line with T1 partial correlations, but contrasts with the T1 direct correlations, where misconception choices associated with almost all global and local measures.

Table 5.6: *Study 6 (time 2 only): Correlations between scores on the global and local processing tasks and misconception test. Composite measures were calculated as (global LISAS – local LISAS) / global LISAS. * p < .05, ** p < .01, *** p < .001. Significant effects are highlighted in bold.*

Task	**Measure**	**Pearson's correlations** Misconceptions	**Partial correlations controlling for age and gender** Misconceptions
Free choice	Free Choice % [a]	.016	-.011
Selective attention	Global [a]	.086	-.030
	Local [a]	-.010	-.033
	Composite [a]	.108	.026
	Global cong. & neut. [a]	.076	-.009
	Global incong. [b]	.119	.040
	Local cong. & neut. [a]	-.002	-.038
	Local incong. [b]	-.052	-.077
Divided attention	Global [a]	.028	-.065
	Local [a]	.013	-.075
	Composite [a]	-.015	-.028

[a] N = 60, df = 56
[b] N = 59, df = 55
Notes: Cong. = congruent; neut. = neutral; incong. = incongruent

As the Year effects were not linear in the Navon tasks (**Section 2.3.1**), regression analyses were carried out to determine whether the associations between global and local Navon task measures and mathematics and science achievement differed between Years. Each of the global or local processing measures and the dummy Year variables were entered in block 1 with each of the mathematics and science variables as the DV. Interaction terms between the dummy Year variables and the global or local processing IVs were entered stepwise in block 2. In four instances, Year 5 associations differed from Years 1 and 3: Mathematical Reasoning, Year 5 x divided attention composite, β = -.107, *p* = .043 (**Figure 5.1A**)(without outliers: β = -.109, *p* = .049); Numerical Operations, Year 5 x divided attention local, β = .403, *p* = .025 (**Figure 5.1B**) (without outliers: β = .556, *p* = .003); Numerical Operations, Year 5 x divided attention composite, β = -.133, *p* = .025 (**Figure 5.1C**) (without outliers: β = -.138, *p* = .022); Science, Year 5 x divided attention composite, β = -.119, *p* = .017 (**Figure 5.1D**) (without outliers: β = -.123, *p* = .018). In all cases, the interaction appeared to be driven by the fact the associations were observed in Year 5 only, and the associations were such that relative strength in global over local responses or poorer local performance was associated with better mathematics or science performance. Additionally, there was a differing association in Year 1 compared with Years 3 and 5 between Misconception responses and the selective attention composite score, β = .201, *p* = .025 (**Figure 5.1E**) (without outliers: β = .213, *p* = .021).



← Greater relative strength in global over local responses

→ Greater relative strength in local over global responses

← Greater relative strength in global over local responses

→ Greater relative strength in local over global responses

← Better performance on local responses

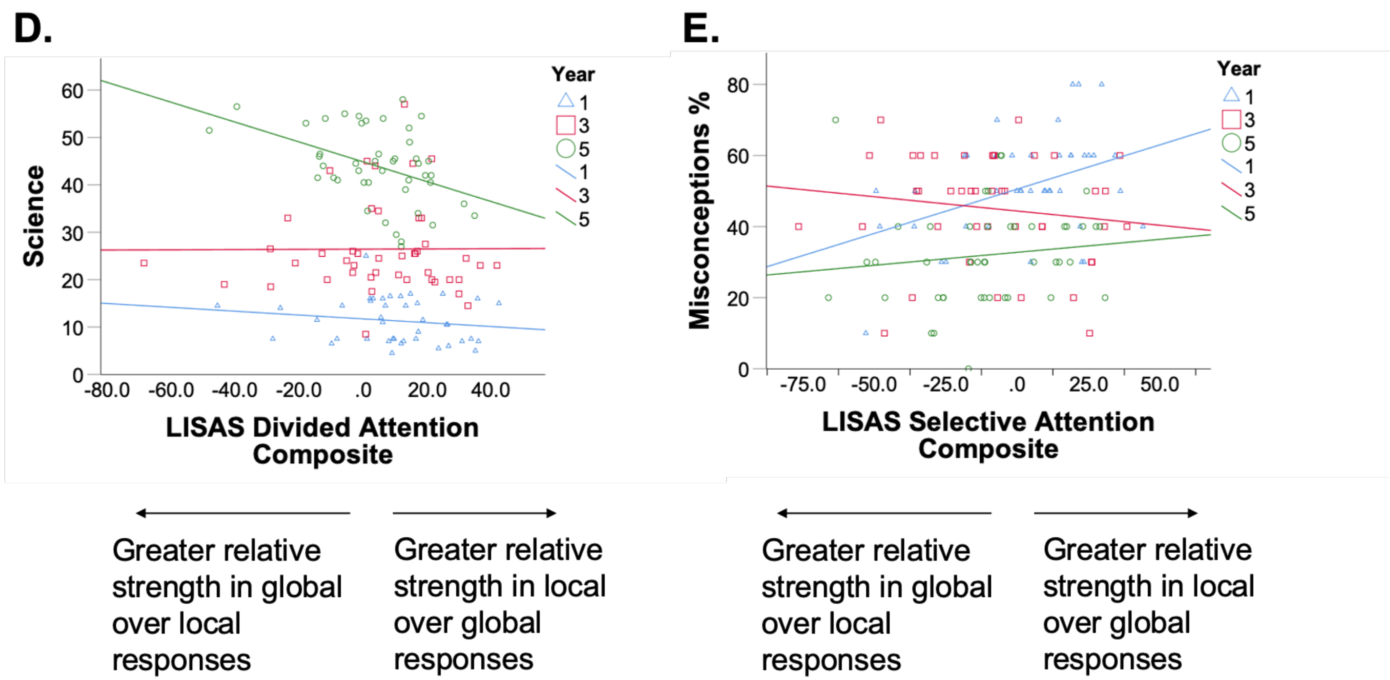


Figure 5.1: *Associations between global and local scores and mathematics and science achievement as a function of Year (outliers have been excluded).* ***A.*** *LISAS divided attention composite score and Mathematical Reasoning;* ***B.*** *LISAS divided attention composite score and Numerical Operations;* ***C.*** *LISAS divided attention local score and Numerical Operations;* ***D.*** *LISAS divided attention composite score and Science;* ***E.*** *LISAS selective attention composite score and proportion of misconception responses.*

5.4.2 Field independence and mathematics and science

5.4.2.1 Study 6

In **Section 2.3.2**, Year and gender effects were reported using linear mixed models (LMMs) which analysed T1 and T2 data together. Here, a series of ANOVAs were carried out with T1 and T2 data separately, with Year and gender as IV and each of the FI variables as the DV. All significant Year and gender effects are reported in **Table 5.7**. Scores on the tasks increased with each ascending Year (p's ≤ .002). There were no significant gender effects (p's > .124) and no interactions between Year and gender (p's > .140).

Table 5.7: *Study 6: Main effects of Year on the Children's Embedded Figures Test (CEFT) and Design Organisation Test (DOT), with estimated marginal means (M) and standard error (SE). Significant effects in bold.*

Time	DV	Year *F*(2,129)	Y1 *M* (*SE*)	Y3 M (*SE*)	Y5 M (*SE*)
1	CEFT %	**35.74, $p < .001$, $\eta_p^2 = .357$**	33.3 (2.3)	49.3 (2.3)	60.2 (2.2)
1	DOT score	**80.03, $p < .001$, $\eta_p^2 = .554$**	10.7 (0.9)	16.8 (0.9)	26.6 (0.9)

Time	DV	Year *F*(1,56)	Y1 *M* (*SE*)	Y3 M (*SE*)	Y5 M (*SE*)
2	CEFT %	**19.79, $p < .001$, $\eta_p^2 = .261$**		38.6 (3.7)	61.1 (3.4)
2	DOT score	**18.92, $p < .001$, $\eta_p^2 = .252$**		19.0 (1.3)	26.5 (1.2)

Notes: Y: Year

Pearson's correlations were carried out to examine associations between the FI measures and the mathematics and science measures. As Year was identified as a main effect in both the FI and academic measures, partial correlations were also run to control for age effects (**Table 5.8**). At T1, there were strong positive associations between FI and all mathematics and science measures, which remained significant when accounting for age in months as a covariate, except the misconceptions test which became non-significant. At T2, there was no significant association between the FI measures and misconceptions, even when age in months was not covaried. The analyses were re-run without the Numerical Operations outlier, revealing that the outlier had very little influence on the associations.

Table 5.8: *Study 6: Correlations between scores on the field independence tasks and mathematics and science tests. * p < .05, ** p < .01, *** p < .001. Significant effects are highlighted in bold.*

Time		Pearson's correlations				Partial correlations controlling for age (months)			
		Mathematical Reasoning	Numerical Operations	Science	Misconceptions	Mathematical Reasoning	Numerical Operations	Science	Misconceptions
Time 1									
	CEFT % [(a)]	**.694 *****	**.641 *****	**.650 *****	**-.367 *****	**.422 *****	**.326 *****	**.307 *****	-.124
	DOT score [(a)]	**.749 *****	**.741 *****	**.724 *****	**-.439 *****	**.340 *****	**.370 *****	**.236 ****	-.163
Time 2									
	CEFT % [(b)]				-.210				-.087
	DOT score [(b)]				-.207				-.048

[(a)] N = 135, df = 132
[(b)] N = 60, df = 57
Notes: CEFT = Children's Embedded Figures Test; DOT = Design Organisation Test

As the FI, mathematics and science achievement measures showed strong correlations with IQ and WM, and weak or non-significant correlations with IC (**Sections 3.3.1.1 and 4.4.2.4**), a series of multiple regressions were run to establish whether the observed associations between FI and mathematics and science were due to overlapping variance with the domain-general variables (**Table 5.9**). Age in months, general intelligence, visuospatial WM, verbal WM and IC were entered first (Model 1). Then, CEFT accuracy or DOT score were added to establish whether they significantly explained additional variance (Models 2A and 2B). Model 2A indicated that CEFT accuracy explained a significant amount of additional unique variance for Mathematical Reasoning and science achievement. In contrast, the DOT score in Model 2B explained additional variance in Numerical Operations only. IQ and age were also significant predictors in the models, with verbal WM predicting performance on Mathematical Reasoning as well as Numerical Operations in Model 2B only. No additional variance in the misconceptions choices was explained by CEFT or DOT after accounting for age, intelligence, and EF.

Table 5.9: *Study 6: Multiple regressions to identify unique variance in mathematics and science scores explained by field independence (FI). Age in months, IQ, verbal working memory (VWM), visuospatial working memory (VSWM) and inhibitory control (IC) were entered in model 1, and then Children's Embedded Figures Test (CEFT) in model 2A and Design Organisation Test (DOT) in model 2B. Significant effects are highlighted in bold.*

			FI		Age		IQ		VSWM		VWM		IC	
DV	R^2	*p*	R^2 Δ	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*
Mathematical Reasoning														
Model 1	**.836**	**< .001**												
Model 2A (CEFT)	**.841**	**< .001**	**.006**	**.032**	**.319**	**< .001**	**.415**	**< .001**	.057	.291	**.141**	**.009**	.050	.171
Model 2B (DOT)	**.838**	**< .001**	.002	.211	**.306**	**< .001**	**.426**	**< .001**	.061	.283	**.161**	**.003**	.055	.138
Numerical Operations														
Model 1	**.706**	**< .001**												
Model 2A (CEFT)	**.712**	**< .001**	.006	.102	**.387**	**< .001**	**.233**	**.015**	.102	.164	.134	.061	.039	.433
Model 2B (DOT)	**.721**	**< .001**	**.015**	**.009**	**.336**	**< .001**	**.199**	**.037**	.070	.341	**.157**	**.025**	.038	.435
Science														
Model 1	**.801**	**< .001**												
Model 2A (CEFT)	**.808**	**< .001**	**.006**	**.046**	**.582**	**< .001**	**.298**	**< .001**	-.060	.314	.045	.441	.050	.213
Model 2B (DOT)	**.804**	**< .001**	.002	.223	**.567**	**< .001**	**.308**	**< .001**	-.058	.347	.066	.258	.055	.177
Misconceptions														
Model 1	**.263**	**< .001**												
Model 2A (CEFT)	**.264**	**< .001**	< .001	.797	-.109	.439	**.343**	**.025**	-.126	.285	.053	.640	-.009	.909
Model 2B (DOT)	**.266**	**< .001**	.002	.511	-.087	.548	**-.323**	**.036**	-.109	.363	.047	.673	-.007	.925

In order to determine whether the Numerical Operations and IC outliers were influencing the regressions, the analyses were repeated excluding the outliers. The additional contribution of CEFT accuracy remained non-significant while that of DOT scores remained significant. However, there were some differences in the contributions of the other variables in the models. Visuospatial WM became a significant contributor in both models (p's ≤ .048), while verbal WM became non-significant in the DOT model (p = .073). There were no changes in the contribution to the models made by IC as a result of removing the outlier.

A further series of regressions were run to determine whether the associations between the FI tasks and the mathematics and science tests remained constant across the Year groups or whether there were nonlinear effects. Dummy Year variables were created and entered in regression analyses with each of the mathematics or science measures as DVs and CEFT or DOT as IVs in block 1. In block 2, interaction terms between the dummy Year variables and CEFT or DOT were entered stepwise. The only significant interaction term was observed when predicting Numerical Operations from CEFT; Year 1 x CEFT, β = -.291, p = .024. This indicated that the association between CEFT and Numerical Operations was weaker in Year 1 than in Years 3 and 5 (**Figure 5.2**). Other associations did not significantly differ across the Years.

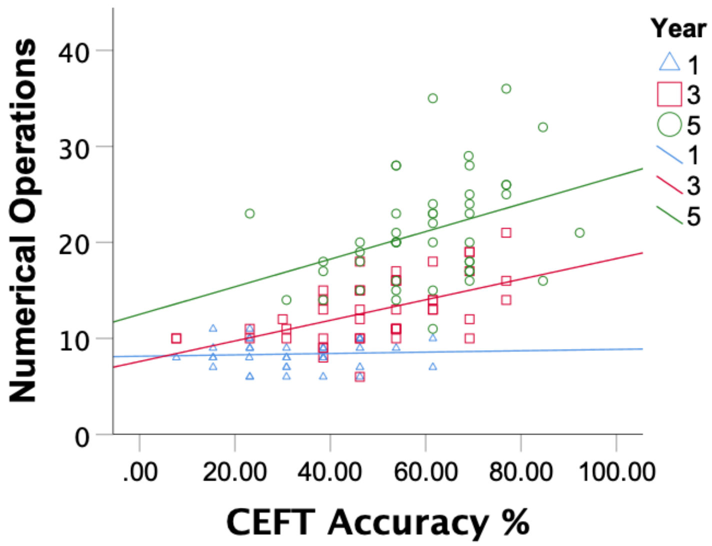


Figure 5.2: *Scatterplot of Numerical Operations score as a function of Children's Embedded Figures Test (CEFT) accuracy and Year.*

5.4.2.2 Study 7

Pearson's correlations were carried out to examine associations between CEFT accuracy and each of the mathematics subtests; Number, Word, Shape, and Graph (**Table 5.10**). Partial correlations were run with gender and age in months as covariates. Although the participants were in a single Year group, age in months was included in the partial correlations to control for difference by age within the Year group. As the data were analysed twice, once individually and once as an overall maths measure, the Bonferroni-corrected alpha is .025.

Table 5.10: *Correlations between scores on the Children's Embedded Figures Test (CEFT) and mathematics tests. * p < .05, ** p < .01, *** p < .001. Significant effects (Bonferroni-corrected) are highlighted in bold.*

	Pearson's correlations	**Partial correlations controlling for age and gender**
	CEFT	CEFT
Number	.113	.134
Word	.203	.230
Shape	**.290 ***	**.325 ****
Graph	.101	.115
Overall maths	.221	.253 *

N = 71, df = 67

After accounting for age and gender effects, there was a significant association between CEFT accuracy and the Shape subset of the mathematics task. Although the overall maths association was just outside significance based on the Bonferroni-corrected alpha, it revealed a trend correlation supportive of the findings in **Section 5.4.2.1** where both Mathematical Reasoning and Numerical Operations associated with CEFT accuracy.

In order to understand the significant association further, a regression was run to determine whether CEFT accuracy explained unique variance in the Shape subset after accounting for age, gender, WM, and IC. The mathematics variable was entered as the DV, with IVs in block 1 of age in months, gender, visuospatial WM, verbal WM, semantic IC, and response IC. CEFT accuracy was then entered in block 2 (**Table 5.11**). CEFT accuracy accounted for unique variance on the Shape subset after controlling for age, gender, and EF.

Table 5.11: *Study 7: Multiple regression to identify unique variance of Children's Embedded Figures Test (CEFT) accuracy with DV of the Shape subset; IVs of age in months, gender, verbal working memory (VWM), visuospatial working memory (VSWM), semantic inhibitory control (SIC), and response inhibitory control (RIC) were entered in the first model, and CEFT accuracy in the second model. Significant effects are highlighted in bold.*

			CEFT		**Age**		**Gender**		**VSWM**		**VWM**		**SIC**		**RIC**	
DV	R^2	***p***	R^2 Δ	***p***	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*	β	*p*
Shape																
Model 1	**.304**	**.001**			.138	.201	**-.235**	**.029**	.143	.245	**.272**	**.024**	**-.236**	**.030**	-.002	.989
Model 2 (CEFT)	**.372**	**< .001**	**.068**	**.011**	.157	.129	**-.242**	**.019**	.089	.458	**.270**	**.019**	**-.244**	**.020**	-.063	.568

The mathematical content of the Number and Word subsets were comparable in terms of operation and difficulty. Although there were no significant direct associations between CEFT and either the Number or Word subset, a final regression was run to determine whether CEFT explained variance on the Word subtest after accounting for achievement on the Number subset, age, and gender. This sought to explore whether the disembedding element of the Word subset (scores after controlling for variance associated with the arithmetic element) related to the visual disembedding task. This analysis showed that CEFT accuracy did not explain unique variance on the Word subset after accounting for variance explained by Number, age, and gender (**Table 5.12**).

Table 5.12: *Study 7: A two-step hierarchical regression with performance on the Word mathematics subset as DV. Number, age in months, and gender were added in model 1, and Children's Embedded Figures Test (CEFT) was added in model 2. Significant effects are highlighted in bold.*

			CEFT		**Number**		**Age**		**Gender**	
DV	***R*²**	***P***	***R*² Δ**	***p***	β	*p*	β	*p*	β	*p*
Word										
Model 1	**.500**	**< .001**			**.661**	**< .001**	.037	.671	-.112	.220
Model 2	**.518**	**< .001**	.018	.118	**.642**	**< .001**	.042	.629	-.137	.175

5.4.3 Systemizing / empathizing / autism traits and mathematics and science

As revealed in **Chapter 2**, there were no main effects of Year or gender on SQ-C, EQ-C, D-score, or AQ-10.

Pearson's correlations, and partial correlations controlling for age, were run between each of these part-whole measures and the mathematics and science scores. After Bonferroni-correcting the alpha for multiple analyses (where the SQ-C and EQ-C were analysed twice; once as the score and once as the D-score calculation), there were no significant correlations (*p*'s > .040). When the partial correlations were re-run without the Numerical Operations outlier, the pattern of significance did not change.

To check for nonlinear effects, regressions were run to determine whether these associations changed by Year group. Each of the mathematical and science measures were added as DV, with each SQ-C, EQ-C, and AQ-10 measure entered with dummy Year variables in block 1, and interaction terms entered in block 2. The only significant interactions were: Mathematical Reasoning, Year 1 x SQ-C, β = -.457, *p* = .031 (**Figure**

5.2A); Mathematical Reasoning, Year 5 x AQ-10, β = -.265, p = .016 (**Figure 5.2B**); Numerical Operations, Year 5 x AQ-10, β = -.364, p = .004 (**Figure 5.2C**); Science, Year 5 x AQ-10, β = -.245, p = .028 (**Figure 5.2D**). The relationship between Mathematical Reasoning and SQ-C was slightly negative in Year 1 while it was positive in Years 3 and 5 (more so in Year 3). In the AQ-10 Year comparisons, the association was larger and negative in Year 5 compared with the other Years.

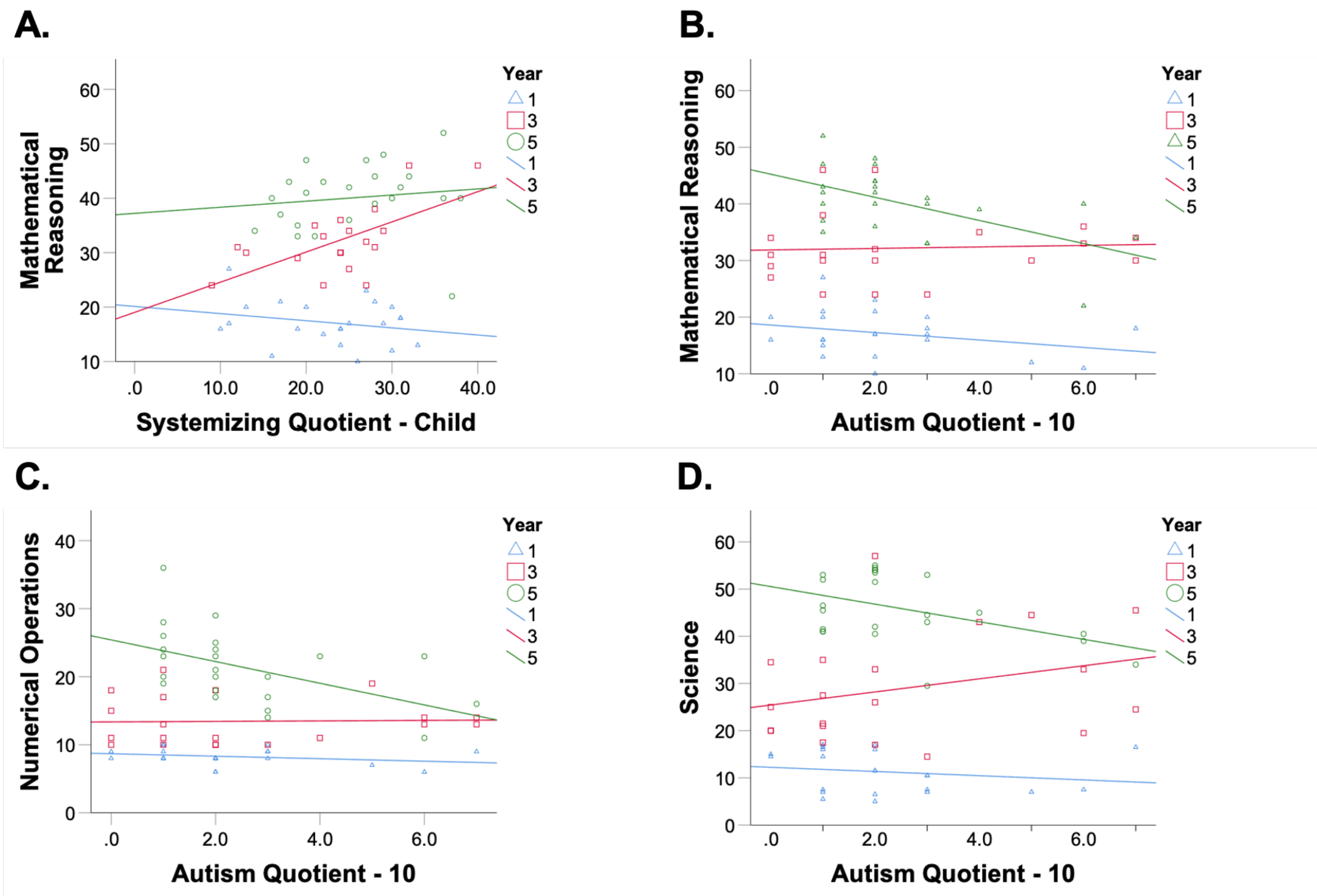


Figure 5.3: *Study 6: Comparison of systemizing and autism traits by Year.* ***A.*** *Mathematical Reasoning and systemizing,* ***B.*** *Mathematical Reasoning and autism traits,* ***C.*** *Numerical Operations and autism traits,* ***D.*** *Science and autism traits.*

5.5 Discussion

This chapter explored relationships between whole-part measures and mathematics and science. Based on findings in earlier chapters that found little overlap between global and local processing, FI, and systemizing, each whole-part construct was examined separately. The studies revealed differing associations with mathematics and science across the whole-part constructs and across tasks.

5.5.1 Global and local processing

To date, the direct association between global and local processing and academic achievement has not been evaluated. Research with autistic adults has identified a preference for selecting jobs and university courses in fields of mathematics and science (Baron-Cohen et al., 1998; Briskman et al., 2001; Wei et al., 2013). Studies with autistic children and adolescents have found mixed outcomes in mathematics achievement. This includes difficulties encountered by autistic children in specific areas of mathematics (Bae et al., 2015; Griswold et al., 2002), achievement scores which are in line with expectation based on IQ (May et al., 2013; Mayes & Calhoun, 2003), as well as examples of enhanced performance in mathematics (Jones et al., 2009). Although the mechanisms explaining these trends have not been fully investigated, it is possible that some autistic traits are advantageous for subjects such as mathematics and science. This chapter found limited evidence of an association between global or local processing and mathematics and science achievement in neurotypical children aged 5 to 10 years, but where associations were significant, the direction of association was consistent.

The first aim of study 6 was to identify whether global and local processing scores had an association with mathematics and science achievement. There were weak to moderate correlations between the selective and divided attention global and local measures and all mathematics and science scores. The selective attention composite measure and free choice global match percentage significantly correlated only with Mathematical Reasoning, while the divided attention composite measure did not significantly associate with any of the test achievement scores. After accounting for age and gender, only one association remained significant: The percentage of global matches on the free choice Navon task had a weak, positive association with Mathematical Reasoning. The LISAS scores on the selective attention global Navon task had a weak, negative association with Mathematical Reasoning but this did not remain significant after Bonferroni-correcting the alpha. Overall, this reveals little association between global and local processing and mathematics achievement, but there is some evidence to suggest that a greater global advantage associates with higher achievement on the Mathematical Reasoning task. Once domain-general abilities were accounted for, the free choice Navon measure did not make a unique contribution to the Mathematical Reasoning scores, indicating that the effect is likely to be explained by overlapping variance with general intelligence, WM and IC.

Chapter 2 revealed non-linear developmental changes in global and local processing which varied depending on the task, so associations across the Year groups were also examined here. There were four instances where associations between measures differed in the responses from the Year 5 children compared with Years 1 and 3. The LISAS divided attention composite scores in Year 5 had a larger negative association with Mathematical Reasoning and Numerical Operations than in the other Years. This indicates that higher scores on both mathematics scores were achieved in Year 5 where children had a greater relative strength in global over local responses. Additionally, higher Numerical Operations scores in Year 5 associated with poorer performance on the divided attention local trials. Together, these support the correlational findings of a small global processing advantage associating with higher mathematics scores. A similar pattern was revealed in the science scores, whereby those in Year 5 with a greater relative advantage of global over local accuracy achieved higher science scores. These specific patterns in Year 5 may reflect the developmental changes in global and local processing identified in **Chapter 2**, whereby responses on the divided attention task in Year 1 revealed a local advantage while Years 3 and 5 revealed no global or local advantage. One further association revealed a different pattern in Year 1 compared with Years 3 and 5, where those who selected fewer misconception responses also had a relatively strong performance in global over local responses on the selective attention task.

Although the strength of correlation is not convincing, there is a consistent picture emerging of higher mathematics and science achievement relating to better relative performance on global processing tasks or poorer performance on local processing tasks. This is not constant across Years or across tasks, but the direction of association is consistent. This is contrary to the prediction of an association between a local processing advantage and higher achievement in mathematics and science, and may indicate that a focus on the whole problem, calculation, or scientific concept, has a greater advantage than a detailed focus. However, the fact that the relationships became non-significant after accounting for domain-general measures indicates that these small associations can be attributed to common variance with general intelligence, WM and IC processes.

5.5.2 Field Independence

Higher FI scores associate with higher achievement in mathematics and science (e.g. Alamolhodaei, 2002; Buriel, 1978; Tinajero & Paramo, 1997). However, as studies have rarely investigated the influence of covariates such as general intelligence and EF, the cognitive processes underpinning these relationships have not been fully explored, particularly in primary school age children. In line with prediction, and supportive of previous research (e.g. Leo-Rhynie, 1985; Tinajero & Paramo, 1997), there were strong positive associations between FI and mathematics and science performance. After controlling for age, general intelligence, WM, and IC, the CEFT explained additional variance on the Mathematical Reasoning and science tests, and the DOT explained additional variance on the Numerical Operations test. This indicates that although some of the relationship between FI and mathematics and science can be explained by overlapping variance with general cognitive abilities, there were additional elements of disembedding that were important for Mathematical Reasoning and science achievement, while segmenting was important for Numerical Operations only. It is possible that disembedding explained achievement on Mathematical Reasoning and science due to the inherent need to decontextualise concepts in these tests. In contrast, the written procedural maths questions in Numerical Operations were not embedded in a context, but required the segmentation of place value columns and operation symbols as well as an understanding of abstract mathematical concepts. Contrary to expectation, we did not find a stronger relationship between FI and mathematics and science with age. Although there was a main effect of Year on all variables, the only significant interaction was revealed when predicting Numerical Operations from CEFT. The association was weaker in Year 1 children than in the older children. This may indicate that younger children do not engage similar cognitive processes when completing these two activities, or it may be driven by a lack of variance in Numerical Operation scores in that Year, highlighting the difficulty of finding a written mathematics test suitable across a wide age range. There was no significant correlation between the misconceptions test and FI tasks after controlling for age in months, either at T1 or T2.

Very few studies have examined differential associations between FI and different fields of mathematics, however one study with university students revealed that FI only related to geometry, and did not associate with other mathematics measures (Zhang, 2004). This indicates that FI may more closely associate with tasks requiring visuospatial processes. This has not previously been explored with children. The current study revealed that higher CEFT scores associated with the Shape subset. In fact, even after accounting for variance associated with gender, age, and EF, CEFT scores remained a significant unique contributor to Shape scores. Although this appears to support the Zhang (2004) study, interestingly, there was no significant association between CEFT scores and the Graph subset, despite the visuospatial nature of those questions. Graph and chart questions did not form part of the mathematics test in the study by Zhang (2004), but the finding here suggests that it was not simply the visuospatial nature of the shape questions which explained their association. The content of the Shape subset included questions relating to angles, questions which required shapes to be named or identified, and questions which required mental manipulation of the given information (creating a 3D shape from a net, drawing a reflection, and drawing the position of a translated shape). The CEFT has been identified as being an intrinsic-static spatial task; a group of tasks which rely on form perception abilities for shapes or objects to be recognised (Gilligan et al., 2018). Associations in the current study between the Shape subset and CEFT are therefore likely to result from a combination of the shared disembedding processes as well as form perception elements. For example, in the angle questions, participants may have disembedded the angle from the rest of the shape in order to attend to the relevant element and reduce distractions, while the questions about shape identification are likely to have relied on form perception. The Graph questions, although visual in nature, required participants to attend to several aspects of the stimulus at once, such as both axes and the central area. It is possible that this required a more holistic approach, resulting in a lack of correlation with the CEFT.

Studies with adolescents have identified an association between FI and word problems (Alamolhodaei, 2002; Azari et al., 2013). Contrary to expectation, CEFT scores did not associate with the Word subset, nor did CEFT uniquely explain variance in the Word subset scores after controlling for achievement on the Number subset, age, and gender. This means that the disembedding element of the Word subset scores did not significantly relate to the visuospatial disembedding task. The lack of association between the Word subset and CEFT in study 7 contrasts with a positive association between Mathematical Reasoning and CEFT in study 6. This may be partly explained by the content of the two CEFT tasks. The CEFT in study 7 included only the tent trials, which may lower the difficulty level due to the smaller number of sides in the target shape (Chamberlain et al., 2017; De-Wit et al., 2017; Huygelier et al., 2018). Therefore, the process of disembedding may not have been required to the same extent as in study 6. This requires some further investigation with the more challenging CEFT stimuli.

These studies have identified important associations between FI and mathematics and science in neurotypical children aged 5 to 10 years. The finding that an ability to identify and isolate target information from complex stimuli associates with mathematics and science achievement has potential implications for mathematics and science education. Children in this age range may benefit from information being presented in a clear, non-cluttered way, with attention being drawn to key areas of focus (Fisher, Godwin & Seltman, 2014; Harp & Mayer, 1998). This would enable children to access the target information more easily from within the context of the whole page. Performance may also be improved by encouraging children to selectively attend to important parts of a stimulus, such as a changing operation symbol in a column of number sentences or specific elements of a shape, without being distracted by the whole. Additionally, in order to encourage the decontextualisation of concepts, it may be useful to present concepts in a number of contexts so that commonalities can be identified (Ainsworth, 2008). Further research is required to investigate associations between different FI tasks and subsets of both mathematics and science across different ages. This would help clarify further the relationships identified here.

5.5.3 SQ, EQ, and AQ

Research with adults has identified an association between higher systemizing scores and studying on mathematics or science university courses (Billington et al., 2007; Svedholm-Häkkinen & Lindeman, 2016), and higher prevalence of autism or autistic traits in those working in the fields of mathematics and science (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Bressan, 2018). In children, higher systemizing related to better problem solving, while higher empathizing scores associated with lower achievement on a calculation task (Escovar et al., 2016). Together, this suggests that those scoring higher on systemizing and autism trait questionnaires are also likely to achieve higher scores on mathematics and science tasks, while those who score more highly on empathizing are likely to achieve lower scores on the mathematics and science tasks. Contrary to expectation, study 6 revealed no associations between these measures. This may be due to the small sample size in this study compared with previous studies. It is possible that the associations are small and therefore only detectable with a larger sample. Additionally, the systemizing and empathizing measures are based on questionnaire responses about the child completed by a parent or carer, rather than directly measuring the construct, which can make it an unreliable measure.

There were four associations which varied between Years. The association between systemizing and Mathematical Reasoning was weaker in Year 1 than Years 3 and 5. This may be driven by less variance in the Year 1 Mathematical Reasoning scores than in the older Year groups. Also, lower scores on Mathematical Reasoning, Numerical Operations, and Science associated with higher scores on the autism trait questionnaire to a greater extent in Year 5 than in Years 1 and 3. As there were no Year effects in the AQ-10 data (**Section 2.5.3.1**), this suggests that the content of the Year 5 mathematics and science questions involved different processes from the other Years which differentially related to the autism trait questionnaire. The direction of association in the Year 5 children contradicted the tentative prediction made here, but is supportive of studies revealing poorer achievement in mathematics in autistic children and adolescents (Aagten-Murphy et al., 2015; Bae et al., 2015; Griswold et al., 2002).

5.5.4 Conclusion

Global and local processing tasks do not have a strong association with mathematics and science measures, but there is a trend in specific tasks and specific Year groups of a global advantage relating to higher achievement on mathematics and science. There were no significant associations between systemizing and empathizing and the academic tests. A more convincing association was identified between higher FI and higher mathematics and science achievement, which is supportive of prior research (Alamolhodaei, 2002; Flexer & Roberge, 1983; Tinajero & Paramo, 1998). Although the positive relationships identified between FI and both mathematics and science were partially due to overlapping variance with age, general intelligence, and WM, there were also specific, unique contributions to achievement. This likely reflects the different levels of disembedding and segmenting involved in different mathematics and science tests.

Chapters 2 to **5** have examined the data collected as part of two studies to establish the developmental trajectories of the measures, associations between measures, and where appropriate, covariates which explain these associations. The findings of this thesis will now be summarised in **Chapter 6**, making reference to the literature review presented in **Chapter 1**. Limitations and future research will also be discussed, placing this thesis in the context of ongoing research and its potential impact.

Chapter 6: Summary and discussion

6.1 Introduction

There is a substantial body of research on the visual perception of whole-part stimuli. Much of this has focussed on the mechanics of how signals are grouped into objects and background (Brooks, 2015; Kimchi et al., 2005; Milne & Szczerbinski, 2009; Vetter & Newen, 2014; Wagemans, 2015) or has compared responses on whole-part tasks between neurotypical controls and atypical populations, particularly autistic individuals (Almeida et al., 2014; Happé, 1999; Happé & Booth, 2008; Happé & Frith, 2006; Jarrold et al., 2005; Koldewyn et al., 2013; Pellicano et al., 2006; Plaisted et al., 1998). However, there is currently a lack of clarity about how these unique whole-part constructs relate to each other. The broad themes being measured in each construct (i.e. detail-focussed, whole-focussed, and the interactions between the two) appear to overlap to some extent, but the boundaries and explanations for these overlaps are unclear. Even within a single construct, particularly global and local processing, it can be difficult to make comparisons between studies due to distinctions in task and stimuli design (Dukette & Stiles, 2001; Harrison & Stiles, 2009; Kimchi, 2015; Kimchi et al., 2005; Kinchla & Wolfe, 1979). Further, there are additional factors outside the scope of this thesis, such as those relating to culture (Caparos et al., 2012; Davidoff, Fonteneau, & Fagot, 2008; Lao et al., 2013; McKone et al., 2010; Miyamoto, Nisbett, & Masuda, 2006; Nisbett & Masuda, 2003; Oishi et al., 2014) and emotional state (Basso, Schefft, Ris, & Dember, 2009; Baumann & Kuhl, 2005; Becker et al., 2017; Fredrickson & Branigan, 2005; Gasper & Clore, 2002; Huntsinger, Clore, & Bar-Anan, 2010; Srinivasan & Hanif, 2010), that have also been shown to modulate responses. A greater understanding of mechanism is therefore important, to be able to move whole-part processing from a largely descriptive field of study, to one which is more deeply understood.

The relationships between whole-part processing and success in mathematics and science have not previously been examined in a systematic way. Individual differences in global and local processing have not been assessed as a possible factor in academic achievement. There have been some studies revealing associations between field independence (FI) and academic achievement (Alamolhodaei, 2002; Azari et al., 2013; Buriel, 1978; Flexer & Roberge, 1983; Tinajero & Paramo, 1997; Zhang, 2004), but there has been little focus on understanding the mechanisms underlying such associations. Overall, there has been far less research into associations with the academic subject of science than mathematics. Therefore, this thesis not only extends previous literature by making an important contribution to our understanding of whole-part processing, but it also adds to the literature about academic achievement in core primary school subjects.

This thesis had two broad objectives; to further understand commonalities and differences between whole-part constructs; and to examine and understand whole-part associations with mathematics and science. A key part of the first objective was to examine developmental changes in whole-part processing in 5- to 10-year-olds. This has not only contributed to our understanding of visual perception development, but also enabled the associations in the second objective to be examined from a developmental perspective. This is particularly relevant for associations which may be modulated by domain-general factors. Improvements in performance on attention and executive function (EF) tasks have consistently been revealed in children aged 5-10 years (Betts et al., 2006; Cragg, 2016; Kemps, De Rammelaere, & Desmet, 2000; Klenberg, Korkman, & Lahti-Nuuttila, 2010), therefore, it should not be assumed that associations identified at a single time-point would be constant over childhood. To that end, the studies involving sample 1 children included both cross-sectional and longitudinal data, revealing patterns across three Year groups and at two timepoints. Additionally, FI was examined in greater depth using both behavioural and eye-tracking methods with a second sample of children in a single Year group.

6.2 Objective 1: Commonalities and differences in whole-part processing

6.2.1 Whole-part development summary

There is a general consensus in the literature that responses on global and local processing tasks change from a local to a global advantage over development (Dukette & Stiles, 1996; Harrison & Stiles, 2009; Kimchi et al., 2005; Poirel et al., 2011; Scherf et al., 2009). However, the age at which this transition takes place appears to vary according to the task demands and stimuli design. Children become more field independent as they develop, reaching adult-levels at around 17 years of age (Amador-Campos & Kirchner-Nebot, 1997; Bigelow, 1971; Busch et al., 1993; Flexer & Roberge, 1983; Glynn & Stoner, 1987; Goodenough & Eagle, 1963). There has been limited research into developmental changes in children's systemizing, but one study identified no age effects (Wakabayashi, 2013). The developmental trajectories of the three constructs indicate that any commonalities in children at a single timepoint are likely to be at best minimal, and any that are identified are likely to vary over time.

There is also a lack of consensus about which measures are expected to associate with one another. For example, one study revealed higher Embedded Figures Test (EFT) accuracy associated with a local processing advantage (Poirel, Mellet, et al., 2008), while in another study, faster responses on two types of EFT associated with accuracy on global processing Navon tasks and to a lesser extent accuracy on local processing tasks (Chamberlain et al., 2017). In fact, the latter study concluded that global elements of the EFT tasks were more important for locating the embedded target than an ability to focus on detailed or local elements. A greater understanding about processes involved in the EFT would help to explain how FI is conceptualised in terms of whole-part advantage, and therefore how performance may be expected to relate to other tasks.

The research presented in study 1 was designed to evaluate developmental changes in whole-part processing in children aged between 5 and 10 years using both longitudinal and cross-sectional data. There were three types of global and local tasks – a free choice task, a selective attention task, and divided attention tasks – which each used the same stimulus design, but each task involved differing attentional demands. The two FI tasks measured disembedding, the separation of a target embedded within a context, and segmenting, the partitioning of a whole into sections. There was a single measure of systemizing, which measured the participants' focus on elements of a rule-based system through a parental questionnaire. These will each be discussed in further detail below.

6.2.2 Global and local processing development

The key findings in study 1 were that there was a change in global and local processing which consistently occurred between Years 1 and 3 (between the ages of 6 and 7 years), however the specifics of this developmental change varied by task, indicating that attentional task demands have an influence on responses. The free choice Navon task responses changed from no advantage to a global advantage; the selective attention Navon task responses changed from a local to global advantage; the big-small divided attention Navon task responses changed from a local to no advantage, while the yes-no divided attention task, which was only administered to Years 3 and 5, revealed a local advantage. The anticipated congruency effects were revealed in the selective attention task, where information at the non-target level interfered with responses to the target level, but only when the information was relevant to the task. Unexpectedly, there was no difference in accuracy or response time (RT) between stay and switch trials on the big-small divided attention task. The yes-no divided attention task was introduced to the test battery at T2, which required participants to examine both levels before responding. The differences in responses between the two divided attention tasks indicate that different strategies or resources were required for the successful completion of each task, despite the apparently similar attentional demands. Together, these results demonstrate that even comparatively small adjustments to the Navon task design can elicit different patterns of responses. This is important for understanding why findings across studies testing global and local processing may not be consistent. The fact that there were distinct patterns of responses to global and local trials within each task, as well as differences in responses between tasks with differing attentional demands, demonstrates that both perceptual and attentional factors interact to determine how children respond to Navon tasks. This study has made an important contribution to our understanding of global and local processing development in neurotypical children, by comparing individuals' responses on a number of different Navon tasks. The results have emphasised that even small differences in task design lead to variation in response. The pattern of developmental changes in this neurotypical sample may also be useful for comparisons with patterns of development in atypical populations.

6.2.3 Field independence development

Children's responses to FI tasks in study 1 became increasingly field independent with increasing age, and there was a positive association between the tasks measuring disembedding and segmenting. This was in line with expectation and reveals common processes involved in different FI tasks. As revealed in studies 2 and 3, a large portion of variation on these tasks was explained by the domain-general measures of IQ and working memory (WM), particularly visuospatial abilities (both IQ and WM in study 2, and WM in study 3). However, there were also differences between the tasks, likely reflecting the contrasting task demands. Verbal IQ explained Design Organisation Test (DOT) scores but not Children's Embedded Figures Test (CEFT) accuracy in study 2, and study 3 revealed an additional association between response inhibitory control (IC) and CEFT accuracy.

To date, there has been a lack of clarity about what FI represents; whether it reflects an ability to overcome automatic perceptual processes (De-Wit et al., 2017; Miyake et al., 2001), or whether it reflects a broader ability to analyse and decontextualise information (Pithers, 2002; Rémy & Gilles, 2014; Witkin et al., 1977). This thesis has not fully resolved this tension, however it has identified some general cognitive abilities which explain variance on two FI tasks. These domain-general abilities may represent the broader ability to analyse and decontextualise information. However, in study 7 there was no association identified between CEFT accuracy and the Word task after controlling for the arithmetic element, which may indicate that FI does not represent a general decontextualising ability. This would need further investigation to explain this apparent inconsistency. Although the studies in this thesis have primarily focussed on the cognitive abilities associated with FI performance, some of the findings indicate potential influences of Gestalt processing (automatic perceptual processes) on ease of disembedding the targets. For example, accuracy on the CEFT was higher on the tent trials than the house trials in study 1, which somewhat supports findings with adults where symmetrical shapes were easier to disembed than asymmetrical shapes (De-Wit et al., 2017; Huygelier et al., 2018). The eye-tracking CEFT task in study 3 only used the tent stimuli and, contrary to study 2, revealed no significant unique contribution to accuracy by any EF. Together, this suggests that where targets are more easily disembedded, there may be a reduced need to draw on domain-general resources of IQ and EF. Therefore, it is likely that an ability to overcome the influence of automatic perceptual processes and the ability to analyse and decontextualise information may in fact be describing similar domain-general processes (at least in part).

Illustrating this further, the fixation patterns revealed in study 3 found that those who looked longer and more frequently at the embedded target shape, and less frequently at the distractor shapes, were more likely to achieve a higher score. This in itself is perhaps unsurprising, but the processes driving this pattern are of great interest. Those with better semantic IC looked at distractor areas less often, however, it was better response IC that associated with higher CEFT accuracy. Disembedding performance can therefore be conceptualised as the extent to which IC guides attention across a stimulus towards relevant areas and away from distractor areas, and ensures a response is not made until the correct target shape has been found. There was some indication that the fixation patterns varied between FI group as a function of stimulus difficulty. This suggests that the more challenging stimuli, where the target is more integrated into the complex stimulus based on Gestalt principles, require the engagement of IC processes (and possibly other domain-general processes) to a greater extent than stimuli where the target is more easily disembedded. This study has made an important first step into understanding how domain-general processes contribute to behavioural outcomes on the CEFT. However, further research is needed to investigate associations with fixation patterns and executive functions (EF) on stimuli with differing levels of embeddedness according to Gestalt principles.

6.2.4 Systemizing development

There was no age effect identified in the children's systemizing scores, and systemizing scores were not associated with empathizing or autism traits. This was contrary to expectation, but may be due to the smaller sample here than in previous studies (Auyeung & Wheelwright, 2009; Wakabayashi, 2013). Higher scores on the autism traits questionnaire did relate to lower empathizing scores, which was in line with our prediction. This indicates that some social skills may be a challenge for those without a formal autism diagnosis but who exhibit some autistic traits.

6.2.5 Whole-part construct associations

There were few cross-construct associations identified here, which is supportive of studies undertaken with adults (Chamberlain et al., 2017; Dale & Arnell, 2013; Milne & Szczerbinski, 2009; Pletzer et al., 2017) and with children and adolescents (Van Eylen et al., 2018). This may be due to a lack of common underlying processes involved in the tasks, or because the differing task demands mask any overall associations. Although this thesis cannot conclude which of these options is more likely, it does highlight that further research is required to fully comprehend the interaction of perceptual and cognitive processes involved in each of the tasks. The fact that study 1 revealed quite distinct patterns of response in each of the global and local processing tasks, underscores the fact that even small changes to the task design can elicit different measures of global or local advantage. This is why it was considered important in this thesis to measure whole-part responses using a number of tasks. Nonetheless, higher DOT scores did associate with better local processing on the selective and big-small divided attention Navon tasks, as well as global processing on the divided attention task. This may reflect common segmenting processes involved in these tasks, as well as the need to switch between whole and part levels to succeed on both the DOT and the big-small divided attention task. The associations between a greater local processing advantage on the divided attention Navon task and both the autism trait score and D-score (relative strength in systemizing over empathizing) were in line with expectation but other expected associations were not revealed. There is currently no clear explanation for the specificity of this inter-construct association. However, the fact that the associations were with the Navon task which included a switching component, is in line with previous research where autistic participants found it harder to widen their attentional field after focussing on local elements (Katagiri et al., 2013; Mann & Walker, 2003; White et al., 2009). It is possible that a difficulty in widening spatial attention is also an explanatory factor in this study.

6.3 Objective 2: Whole-part processing and mathematics and science

The literature has revealed a series of associations which, taken together, suggest that whole-part perceptual processing may be relevant to mathematics and science achievement. There is a trend towards a local processing advantage, enhanced disembedding abilities, and greater systemizing in autistic individuals (Almeida et al., 2014; Baron-Cohen et al., 2003; Happé & Frith, 2006; Jarrold et al., 2005). A common feature of these abilities is a focus on individual elements to a greater extent than the whole. Achievement on mathematics and science has been found to associate with FI (Alamolhodaei, 2002; Azari et al., 2013; Leo-Rhynie, 1985; Roberge & Flexer, 1983; Tinajero & Páramo, 1998) and systemizing ability (Billington et al., 2007; Bressan, 2018; Escovar et al., 2016; Focquaert et al., 2007) in neurotypical and atypical populations. Therefore, it is possible that local processing may also associate with mathematics and science. This has not previously been explored, and it is this gap in our understanding that led to the second objective of this thesis. Mathematics and science achievement are important to individuals (Cragg & Gilmore, 2014; Morgan et al., 2016; Walker & Zhu, 2013) and to the economy as a whole (Centre for Economics and Business Research, 2015; Deloitte, 2013). Therefore, it is pertinent that we further our understanding of factors influencing achievement in these subjects.

There is a substantial body of research examining mathematical cognition and factors affecting achievement, including EF (Alloway & Alloway, 2010; Alloway & Passolunghi, 2011; Cragg & Gilmore, 2014; De Smedt et al., 2009; Formoso et al., 2018; Merkley et al., 2016) and spatial cognition (Gilligan et al., 2017, 2018). These factors are increasingly being evaluated for each distinct mathematical components such as factual knowledge, procedural skill, and conceptual development (Cragg, Keeble, et al., 2017; Gilmore et al., 2017; Rittle-Johnson et al., 2001). Comparisons have also been made between tests which measure computation and fact retrieval, and those which present mathematical content in real-world contexts through word problems and mathematical reasoning (Meyer et al., 2010; Oh, Glutting, Watkins, Youngstrom, & McDermott, 2004). Other studies have divided mathematical tasks by topic, such as geometry and algebra (Zhang, 2004). This thesis contributes to the literature with analyses of Mathematical Reasoning and Numerical Operations responses, as well as comparisons between topic subsets of mathematics; Number, Word, Shape, and Graph. This approach of dividing mathematics into more specific subdivisions is important, as the domain-general and domain-specific processes involved in each are quite different. By using an overall mathematics score, some contributory factors might be overlooked. In contrast, the science test devised for studies 4 and 6 was not separated into particular skills, such as conceptual understanding, reasoning and scientific inquiry, and therefore scores reflected a general science ability. This allowed for an initial investigation of associations with primary science as a whole.

Mathematics and science associates were examined in study 6 with a cross-sectional and longitudinal sample of Years 1, 3, and 5 children. Mathematics also formed part of study 7, involving Year 5 children. The predicted associations between the academic tasks and local processing and systemizing were not identified here. In fact, the only significant association after controlling for age and gender was that a higher proportion of global responses on the free choice Navon task weakly related to higher scores on the Mathematical Reasoning task. However, this was non-significant after controlling for domain-general factors. The prediction that higher FI scores would associate with higher scores on the mathematics and science tests was seen in both study 6 and, to a lesser extent, study 7. Although much of the overlapping variance between the tasks could be explained by domain-general abilities, the different FI tasks explained additional unique variance after accounting for these. In study 6, the CEFT explained additional variance on the Mathematical Reasoning and science tests. Both these academic tasks require important features of the question to be separated and analysed, and for conceptual understanding to be applied appropriately to the context of the question. It is possible that the CEFT associated with these specific tasks as these processes mirror disembedding, where relevant features of the complex figure are recognised and analysed, enabling the target to be isolated from the context. Conversely, the DOT explained additional variance on the Numerical Operations test, which required participants to segment written arithmetic questions into chunks in order to complete the calculations. In study 7, the CEFT explained additional variance on the Shape task after controlling for domain-general factors. Although it is possible this is related to the visuospatial nature of the task, it is also likely that processes involved with identifying and isolating features of the stimulus in order to answer the question may have led to the association, for example, by isolating an angle of a shape from the distracting whole or other distracting features.

In sum, studies 6 and 7 have replicated the expected positive association between FI and mathematics and science, whereby those with better FI achieved higher scores on the academic tasks. Additionally, by measuring IQ and EF, the studies here have begun to explain in more detail the processes underlying these associations. There was evidence that much of the better mathematics and science achievement in more field independent participants could be explained by overlapping variance with domain-general abilities.

6.4 Limitations and future research

6.4.1 Development

The results of study 1 revealed the impact of differing attentional task demands on global and local processing responses over development. However, the study did not include tasks to specifically measure attention, limiting the extent to which conclusions can be drawn about how attentional and perceptual processes interact in Navon tasks. In future research, attentional tasks could be included in the test battery, particularly in developmental studies as it would facilitate an understanding of how the maturity of attentional processes may impact on global and local responses. It would also be beneficial to test adolescents and adults on the study 1 tasks to establish how the developmental changes identified here continue to full maturity. This is particularly important for the divided attention Navon tasks which did not reveal a global advantage within the current age group. This would additionally allow for a more complete understanding of how associations between task and between construct change over time.

Due to time limitations, the mathematics and science tasks were only presented to children in Study 1 at Time 1 (T1). It would have been more informative to repeat all the tasks at Time 2 (T2), allowing for longitudinal predictors to be identified. Future longitudinal research would ideally include all tasks at both time-points. Further, the science task was designed for children to complete at the end of the academic year, so that they would have covered much of the content through their lessons during the year. Unfortunately, due to initial difficulties with finding a school able to participate, the study was delayed and children completed the battery of tasks at the beginning of the school year. Therefore, the children were unlikely to have been taught the concepts in the test. There were no floor effects, but the test may not be measuring recall and application of formally-learnt concepts but instead may be measuring more informal learning which is likely to vary between participants. Although it would be useful for this test to be used again for replication purposes, it would also be interesting to present it to the children at the end of the academic year, in the way it was originally intended, even though it may limit its comparability with the current study.

The data collected in studies 4 and 5 revealed some Year effects on associations between EF and mathematics and science achievement. This leads to difficulty in interpreting the behavioural results, as this differential pattern of association may be due to the content of the task, or differences in domain-general development; for example, the domain-general requirements of the Numerical Operations task that is accessible to Year 1 children are likely to be very different from the processes involved in questions suitable for Year 5 children. It therefore becomes difficult to disentangle the effect of mathematical content from domain-general development. Future research could include age-appropriate mathematical content which measures performance across a variety of topics, including mental mathematics, specific subsets similar to the Year 5 task described in **Section 4.3.4**, and different domain-specific components of conceptual development, number facts, and procedural understanding (Cragg & Gilmore, 2014; Gilmore, Keeble, Richardson, & Cragg, 2015).

6.4.2 Navon tasks

A key difficulty in assessing the impact of differing attentional task demands on global and local responses is that additional confounds can be introduced. In Study 1, the selective attention Navon stimuli were presented in one of four corners of the screen. This design aimed to reduce the possibility of participants narrowing or widening their focus for the whole block which could artificially reduce the impact of the non-target level on responses (Gerlach & Krumborg, 2014; Wilkinson et al., 2001). However, this then adds an additional step where participants have to locate the stimulus before processing it, leading to difficulties in making true comparisons with the free choice and divided attention tasks. At T2, an additional selective attention task was included in the task battery where the stimuli remained in the centre of the screen, which would have enabled a comparison with both the moving selective attention task and the other Navon tasks. However, due to administration difficulties, it was felt that the data from the new selective attention task were not sufficiently reliable, and therefore were excluded from this thesis. A future examination of global and local tasks with varying attentional demands could include a comparison of responses of both stationary and moving stimuli to further our understanding of whether a widening or narrowing of attention does indeed modulate responses on selective attention Navon tasks in children.

Whilst Navon figures provide control over global and local processing comparisons, they are not representative of stimuli encountered on a day-to-day basis, and are therefore not ecologically valid. One study with adults aimed to address this by using global silhouettes of cats and dogs, and local patterns on their body (Hübner & Studer, 2009). A key difference between these stimuli and Navon figures is that the global level had an identifiable boundary determined by the animal silhouette. In contrast, the individual parts in a Navon stimulus have to be grouped using the spatial relationships between the local elements before the global shape can be perceived. One possible explanation for a local advantage in younger children is that the global level is less meaningful as it does not match their early representation of that shape or letter, whereas older children may have encountered shapes and letters in many forms and therefore find it easier to make sense of the global level. This is particularly important as more meaningful stimuli have been shown to be more salient than non-meaningful information (Harrison & Stiles, 2009; Poirel, Mellet, et al., 2008; Poirel et al., 2006; Poirel, Pineau, & Mellet, 2008). Further exploration using real-life global and local stimuli would help to determine whether they are comparable to children's responses on Navon stimuli, and therefore establish the extent to which responses to Navon stimuli are generalisable to other hierarchical stimuli.

6.4.3 Field independence tasks

Some work has already been undertaken with adults to understand further the characteristics of embedded figures stimuli which associate with performance (De-Wit et al., 2017; Huygelier et al., 2018). This involved comparing responses on features such as shape of the target, number of shared lines with the background, and symmetry, enabling clarification of how and why disembedding varies across stimuli. These designs control for Gestalt principles across stimuli, and any individual differences in automatic grouping that takes place at the early stages of visual processing. Currently, these Gestalt grouping factors vary across the CEFT, and therefore may be an unmeasured contributory factor to accuracy. Future examination of the impact of Gestalt grouping factors on CEFT performance would provide further information about the mechanisms involved in the task. Therefore, although Gestalt processing was outside the scope of this thesis in terms of direct involvement with mathematics and science, a consideration of grouping principles is undoubtedly important for developing a mechanistic understanding of whole-part tasks, and the role of domain-general measures in overcoming automatic Gestalt processes.

This was the first time the DOT was used with a sample of children. There was one design which the children found particularly difficult to segment as the overall design (stripes) distracted from the identification of the elements. This artificially created a ceiling amongst those children reaching the row of more challenging stimuli. Based on this, it would be interesting for future studies if the order of the stimuli was changed as this could lead to a greater spread of scores in children who were better able to segment.

6.4.4 Intervention study

The studies discussed in this thesis have been associational in nature. As FI has been shown to associate with mathematics and science, future studies could explore whether mathematics and science achievement could be enhanced through training of abilities relating to FI such as identifying elemental details and isolating them from contextual noise. One training study aimed to draw children's attention to less salient perceptual features of a control-of-variables task (where participants carried out an experiment to find an answer to a question by keeping all variables constant apart from one), such as looking more closely at the height of two balls that were dropped, even when there initially appeared to be no discernible height difference. They found that this did indeed improve outcomes post-test relative to pre-test for children with lower FI scores, although this improvement was only identified in children aged 8 years, not the younger age groups (Globerson et al., 1985). This suggests that in some age groups, it may be possible to improve children's mathematics and science achievement through a FI training intervention, which would encourage children to identify and focus upon the most relevant features of a task or stimulus.

6.5 Conclusion

The understanding of whole-part processing has been fairly fragmented to date, with little clarity of how the constructs of global and local processing, FI, Gestalt processing, and systemizing interrelate. This thesis has revealed far closer associations between tasks within the same construct than across constructs. This indicates that there is not a dominant, common whole or part processing ability which determines how individuals respond across constructs. There are distinct patterns of development across tasks. For global and local processing, there was a notable change taking place between Years 1 and 3 (ages 6 to 7 years), although the specifics of this change varied according to the attentional demands of the Navon task. The general trajectory of performance on the Navon tasks is supportive of the local to global change typically reported in childhood. For FI, the data revealed that as children develop, they became more field independent. This may reflect the gradual maturity of domain-general processes which explained 50% and 65% of performance on the two tasks examined here.

There was little evidence that variation in mathematics and science achievement can be explained by global and local processing or systemizing, however, the predicted association between better FI and higher academic achievement was observed. A large proportion of these FI associations were explained by domain-general factors of general intelligence and EF, but specific associations between FI and mathematics and science measures remained. This indicates that while domain-general factors play an important role in explaining FI and academic correlations, there may be additional overlapping processes involved which are unique to each FI task.

This thesis has made an important contribution to furthering our understanding of whole-part processing and the whole-part associations with domain-general and academic tasks. However, there remain gaps in our understanding. Further research is needed to fully understand the processes that explain developmental changes in global and local processing, and also how Navon tasks relate to real-world stimuli. Based on the findings in this thesis, future studies should be carried out using a number of global and local tasks. Using this thesis as a starting point, future research could also gain a more complete understanding of factors contributing to CEFT accuracy, and how these vary as a function of stimuli characteristics. This may clarify and explain associations with academic achievement, which in turn could lead to the development of intervention studies.

References

Aagten-Murphy, D., Attucci, C., Daniel, N., Klaric, E., Burr, D. C., & Pellicano, E. (2015). Numerical estimation in children with autism. *Autism Research*, *8*(6), 668–681. https://doi.org/10.1002/aur.1482

Aguinis, H., Gottfredson, R., & Joo, H. (2013). Best-practice recommendations for defining, identifying, and handling outliers. *Organizational Research Methods*, *16*(2), 270–301. https://doi.org/10.1177/1094428112470848

Akshoomoff, N. A., & Stiles, J. (1996). The influence of pattern type on children's Block Design performance. *Journal of the International Neuropsychological Society*, *2*(5), 392–402. https://doi.org/10.1017/S1355617700001466

Alahyane, N., Lemoine-Lardonnois, C., Tailhefer, C., Collins, T., Fagard, J., & Dore-Mazars, K. (2016). Development and learning of saccadic eye movements in 7- to 42-month-old children. *Journal of Vision*, *16*(1), 1–12. https://doi.org/10.1167/16.1.6.doi

Alamolhodaei, H. (2002). Students' cognitive style and mathematical word problem solving. *Journal of the Korea Society of Mathematical Education*, *6*(2), 171–182.

Alamolhodaei, H. (2011). Students' field-dependency and their mathematical performance based on Bloom's cognitive levels. *Journal of the Korea Society of Mathematical Education*, *15*(4), 373–386.

Allen, M. (2014). *Misconceptions in primary science*. McGraw-Hill education (UK).

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(2), 202–212. https://doi.org/10.1016/j.jaac.2011.11.003

Allison, C., Baron-Cohen, S., Stone, M. H., & Muncer, S. J. (2015). Rasch modeling and confirmatory factor analysis of the systemizing quotient-revised (SQ-R) scale. *The Spanish Journal of Psychology*, *18*, E16. https://doi.org/10.1017/sjp.2015.19

Alloway, T., & Alloway, R. (2010). Investigating the predictive roles of working memory and IQ in academic attainment. *Journal of Experimental Child Psychology*, *106*(1), 20–29. https://doi.org/10.1016/j.jecp.2009.11.003

Alloway, T., Banner, G., & Smith, P. (2010). Working memory and cognitive styles in adolescents' attainment. *British Journal of Educational Psychology*, *80*(4), 567–581. https://doi.org/10.1348/000709910X494566

Alloway, T., Gathercole, S., Kirkwood, H., & Elliott, J. (2008). Evaluating the validity of the Automated Working Memory Assessment. *Educational Psychology*, *28*(7), 725–734. https://doi.org/10.1080/01443410802243828

Alloway, T., & Passolunghi, M. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*(1), 133–137. https://doi.org/10.1016/j.lindif.2010.09.013

Almeida, R., Dickinson, J., Maybery, M., Badcock, J., & Badcock, D. (2014). Enhanced global integration of closed contours in individuals with high levels of autistic-like traits. *Vision Research*, *103*, 109–115. https://doi.org/10.1016/j.visres.2014.08.015

Amador-Campos, J. A., & Kirchner-Nebot, T. (1997). Relations of scores on Children's Embedded Figures Test with age, item difficulty and internal consistency. *Perceptual and Motor Skills*, (1954), 675–682.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders.* Arlington, VA: American Psychiatric Publication, 5th ed.

Amso, D., & Scerif, G. (2015). The attentive brain: Insights from developmental cognitive neuroscience. *Nature Reviews Neuroscience*, *16*(10), 606–619. https://doi.org/10.1038/nrn4025

Angeli, C., Valanides, N., Polemitou, E., & Fraggoulidou, E. (2016). An interaction effect between young children's field dependence-independence and order of learning with glass-box and black-box simulations: Evidence for the malleability of cognitive style in computer-supported learning. *Computers in Human Behavior*, *61*, 569–583. https://doi.org/http://dx.doi.org/10.1016/j.chb.2016.03.060

Aring, E., & Grönlund, M. A. (2007). Visual fixation development in children, 1659–1665. https://doi.org/10.1007/s00417-007-0585-6

Arnup, J., Murrihy, C., Roodenburg, J., & McLean, L. (2013). Cognitive style and gender differences in children's mathematics achievement. *Educational Studies*, *39*(3), 355–368. https://doi.org/10.1080/03055698.2013.767184

Assouline, S., Foley Nicpon, M., & Dockery, L. (2012). Predicting the academic achievement of gifted students with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *42*(9), 1781–1789. https://doi.org/10.1007/s10803-011-1403-x

Aubrey, C., Dahl, S., & Godfrey, R. (2006). Early mathematics development and later achievement: Further evidence. *Mathematics Education Research Journal*, *18*(1), 27–46.

Auyeung, B., & Wheelwright, S. (2009). The children's empathy quotient and systemizing quotient: Sex differences in typical development and in autism spectrum conditions. *Journal of Autism & Developmental Disorders*, *39*.

Awasthi, B., Williams, M. A., & Friedman, J. (2016). Examining the role of red background in magnocellular contribution to face perception. *PeerJ*, *4*, e1617. https://doi.org/10.7717/peerj.1617

Azari, S., Radmehr, F., & Mohajer, M. (2013). A study on the relationship between students' cognitive style and mathematical word and procedural problem solving while controlling for students' intelligent quotient and math anxiety. *European Journal of Child Development*, *1*(2), 59–73.

Bae, Y. S., Chiang, H.-M., & Hickson, L. (2015). Mathematical word problem solving ability of children with autism spectrum disorder and their typically developing peers. *Journal of Autism and Developmental Disorders*, *45*(7), 2200–2208. https://doi.org/10.1007/s10803-015-2387-8

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6), 248–254. https://doi.org/10.1016/S1364-6613(02)01904-6

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: An investigation of adults with Asperger syndrome or high−functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(2), 361–374. https://doi.org/10.1098/rstb.2002.1206

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome / high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism & Developmental Disorders*, *31*(1), 5–17.

Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T., & Chakrabarti, B. (2009). Talent in autism: Hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1522), 1377–1383. https://doi.org/10.1098/rstb.2008.0337

Baron-Cohen, S., Wheelwright, S., Burtenshaw, A., & Hobson, E. (2007). Mathematical talent is linked to autism. *Human Nature*, *18*(2), 125–131. https://doi.org/10.1007/s12110-007-9014-0

Baron-Cohen, S. (2009). Autism: the empathizing–systemizing (E-S) theory. *Annals of the New York Academy of Sciences*, *1156*, 68–80. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.2009.04467.x/pdf

Baron-Cohen, S., Bolton, P., Wheelwright, S., Scahill, V., Short, L., Mead, G., & Smith, A. (1998). Does autism occur more often in families of physicists, engineers, and mathematicians? *Autism*, *2*(3), 296–301.

Baroody, A. (2000). Does mathematics instruction for three-to five-year-olds really make sense? *Young Children*, *55*(4), 61–67. Retrieved from http://www.jstor.org/stable/42727827

Basso, M., & Lowery, N. (2003). Global-local visual biases correspond with visual-spatial orientation. *Journal of Clinical and Experimental Neuropsychology*, *26*(1), 24–30. https://doi.org/10.1076/jcen.26.1.24.23939

Basso, M., Schefft, B., Ris, M., & Dember, W. (2009). Mood and global-local visual processing. *Journal of the International Neuropsychological Society*, *2*(03), 249. https://doi.org/10.1017/s1355617700001193

Bates, K.E., Farran, E. K. (2019), Mental imagery in typical development and ADHD. *Poster Presentation. European Society for Cognitive Psychology Conference, Tenerife, September 2019*

Baumann, N., & Kuhl, J. (2005). Positive affect and flexibility: Overcoming the precedence of global over local processing of visual information. *Motivation and Emotion*, *29*(2), 123–134. https://doi.org/10.1007/s11031-005-7957-1

Beaucousin, V., Cassotti, M., Simon, G., Pineau, A., Kostova, M., Houdé, O., & Poirel, N. (2011). ERP evidence of a meaningfulness impact on visual global/local processing: When meaning captures attention. *Neuropsychologia*, *49*(5), 1258–1266. https://doi.org/10.1016/j.neuropsychologia.2011.01.039

Beaucousin, V., Simon, G., Cassotti, M., Pineau, A., Houdé, O., & Poirel, N. (2013). Global interference during early visual processing: ERP evidence from a rapid global/local selective task. *Frontiers in Psychology*, *4*(AUG), 1–6. https://doi.org/10.3389/fpsyg.2013.00539

Becker, K., Plessow, F., Coniglio, K., Tabri, N., Franko, D., Zayas, L., Germine, L., Thomas, J., & Eddy, K. (2017). Global/local processing style: Explaining the relationship between trait anxiety and binge eating. *International Journal of Eating Disorders*, *50*(11), 1264–1272. https://doi.org/10.1016/j.biomaterials.2016.06.015.Tunable

Behrmann, M., & Kimchi, R. (2003). What does visual agnosia tell us about perceptual organization and its relationship to object perception? *Journal of Experimental Psychology: Human Perception and Performance*, *29*(1), 19–42. https://doi.org/10.1037/0096-1523.29.1.19

Bennett, M. (2016). *What is the UK doing about its STEM skills shortfall?* Retrieved from: https://www.telegraph.co.uk/business/ready-and-enabled/stem-skills-shortfall/

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

Betts, J., Mckay, J., Maruff, P., & Anderson, V. (2006). The development of sustained attention in children: The effect of age and task load. *Child Neuropsychology*, *12*(3), 205–221. https://doi.org/10.1080/09297040500488522

Bigelow, G. S. (1971). Field dependence-field independence in 5- to 10-year-old children. *Journal of Educational Research*, *64*(9), 397–400. https://doi.org/10.1080/00220671.1971.10884202

Billington, J., Baron-Cohen, S., & Bor, D. (2008). Systemizing influences attentional processes during the Navon task: an fMRI study. *Neuropsychologia*, *46*(2), 511–520. https://doi.org/10.1016/j.neuropsychologia.2007.09.003

Billington, J., Baron-Cohen, S., & Wheelwright, S. (2007). Cognitive style predicts entry into physical sciences and humanities: Questionnaire and performance tests of empathy and systemizing. *Learning and Individual Differences*, *17*(3), 260–268. https://doi.org/10.1016/j.lindif.2007.02.004

Boccia, M., Piccardi, L., Di Marco, M., Pizzamiglio, L., & Guariglia, C. (2016). Does field independence predict visuo-spatial abilities underpinning human navigation? Behavioural evidence. *Experimental Brain Research*, *234*(10), 2799–2807. https://doi.org/10.1007/s00221-016-4682-9

Bölte, S., Holtmann, M., Poustka, F., Scheurich, A., & Schmidt, L. (2007). Gestalt perception and local-global processing in high-functioning autism. *Journal of Autism and Developmental Disorders*, *37*(8), 1493–1504. https://doi.org/10.1007/s10803-006-0231-x

Booth, R. (2006). *Local-global processing and cognitive style in autism spectrum disorders and typical development*. *King's College London doctoral thesis*.

Booth, R., & Happé, F. (2010). " Hunting with a knife and ... fork": Examining central coherence in autism, attention deficit/hyperactivity disorder, and typical development with a linguistic task. *Journal of Experimental Child Psychology*, *107*(4), 377–393. https://doi.org/10.1016/j.jecp.2010.06.003

Bott, N. T., Lange, A., Rentz, D., Buffalo, E., Clopton, P., & Zola, S. (2017). Web camera based eye tracking to assess visual memory on a visual paired comparison task. *Frontiers in Neuroscience*, *11*(JUN), 1–9. https://doi.org/10.3389/fnins.2017.00370

Bouvet, L., Rousset, S., Valdois, S., & Donnadieu, S. (2011). Global precedence effect in audition and vision: Evidence for similar cognitive styles across modalities. *Acta Psychologica*, *138*, 329–335. https://doi.org/10.1016/j.actpsy.2011.08.004

Bressan, P. (2018). Systemisers are better at maths. *Scientific Reports*, *8*(1), 1–5. https://doi.org/10.1038/s41598-018-30013-8

Brian, J., & Bryson, S. (1996). Disembedding performance and recognition memory in autism/ PDD. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *37*(7), 865–872.

Briskman, J., Happé, F., & Frith, U. (2001). Exploring the cognitive phenotype of autism: weak “central coherence” in parents and siblings of children with autism: II. Real-life skills and preferences. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(3), 309–316. https://doi.org/10.1017/S0021963001006916

Brookman-Byrne, A., Mareschal, D., Tolmie, A., & Dumontheil, I. (2018). Inhibitory control and counterintuitive science and maths reasoning in adolescence, 1–19. https://doi.org/10.1371/journal.pone.0198973

Brooks, J. (2015). Traditional and new principles of perceptual grouping. In *Oxford Handbook of Perceptual Organization*. https://doi.org/10.1080/00369220601100075

Brosnan, M., Ashwin, C., Gamble, T., Daggar, R., Collomosse, J., Ashwin, C., & Gamble, T. (2011). Greater empathizing and reduced systemizing in people who show a jumping to conclusions bias in the general population: Implications for psychosis. *Psychosis*, *5*(1), 37–41. https://doi.org/10.1080/17522439.2011.626526

Brosnan, M., Daggar, R., & Collomosse, J. (2010). The relationship between systemising and mental rotation and the implications for the extreme male brain theory of autism. *Journal of Autism and Developmental Disorders*, *40*(1), 1–7. https://doi.org/10.1007/s10803-009-0815-3

Brosnan, M., Gwilliam, L., & Walker, I. (2012). Brief report: The relationship between visual acuity, the Embedded Figures Test and systemizing in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(11), 2491–2497. https://doi.org/10.1007/s10803-012-1505-0

Brosnan, M., Scott, F., Fox, S., & Pye, J. (2004). Gestalt processing in autism: Failure to process perceptual relationships and the implications for contextual understanding. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *45*(3), 459–469. https://doi.org/10.1111/j.1469-7610.2004.00237.x

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). Perceptual organisation. In *Visual Perception: physiology, psychology and ecology* (1st ed.). Hove: Psychology Press.

Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*(1). https://doi.org/10.1111/cdep.12059

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children’s mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology*, *19*(3), 273–293. https://doi.org/10.1207/S15326942DN1903

Buriel, R. (1978). Relationship of three field-dependence measures to the reading and math achievement of Anglo American and Mexican American children. *Journal of*

Educational Psychology, *70*(2), 167–174. https://doi.org/10.1037/0022-0663.70.2.167

Busch, J., Watson, J., Brinkley, V., Howard, J., & Nelson, C. (1993). Preschool Embedded Figures Test performance of young children: Age and gender differences. *Perceptual and Motor Skills*, *77*, 491–496.

Buschman, T. J., & Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science (New York, N.Y.)*, *315*(5820), 1860–1862. https://doi.org/10.1126/science.1138071

Cairns, E., Malone, S., Johnston, J., & Cammock, T. (1985). Sex differences in children's Group Embedded Figures Test performance. *Personality and Individual Differences*, *6*(5), 653–654. https://doi.org/10.1016/0191-8869(85)90018-2

Caldwell-Harris, C. L., & Jordan, C. J. (2014). Systemizing and special interests: Characterizing the continuum from neurotypical to autism spectrum disorder. *Learning and Individual Differences*, *29*, 98–105. https://doi.org/10.1016/j.lindif.2013.10.005

Canobi, K. H. (2004). Individual differences in children's addition and subtraction knowledge. *Cognitive Development*, *19*, 81–93. https://doi.org/10.1016/j.cogdev.2003.10.001

Caparos, S., Ahmed, L., Bremner, A., de Fockert, J., Linnell, K., & Davidoff, J. (2012). Exposure to an urban environment alters the local bias of a remote culture. *Cognition*, *122*(1), 80–85. https://doi.org/10.1016/j.cognition.2011.08.013

Caparos, S., Linnell, K. J., Bremner, A. J., de Fockert, J. W., & Davidoff, J. (2013). Do local and global perceptual biases tell us anything about local and global selective attention? *Psychological Science*, *24*(2), 206–212. https://doi.org/10.1177/0956797612452569

Carlson, J. S., & Jensen, C. M. (1981). Reliability of the Raven Colored Progressive Matrices test: Age and ethnic group comparisons. *Journal of Consulting and Clinical Psychology*, *49*(3), 320–322. https://doi.org/10.1037/0022-006X.49.3.320

Carlson, S., Zelazo, P., & Faja, S. (2014). Executive Function. In *The Oxford Handbook of Developmental Psychology, Volume 1: Body and mind* (Vol. 1). https://doi.org/10.1016/B978-0-12-385157-4.01147-7

Carpenter, R., & Alloway, T. (2018). Computer versus paper-based testing: Are they equivalent when it comes to working memory? *Journal of Psychoeducational Assessment*. https://doi.org/10.1177/0734282918761496

Carroll, J. M., & Chiew, K. Y. (2006). Sex and discipline differences in empathising, systemising and autistic symptomatology: Evidence from a student population. *Journal of Autism and Developmental Disorders*, *36*(7), 949–957.

https://doi.org/10.1007/s10803-006-0127-9
Centre for Economics and Business Research. (2015). The contribution of engineering to the UK economy - A report for EngineeringUK. *A Report for EngineeringUK*, (January).
Chamberlain, R., Van der Hallen, R., Huygelier, H., Van de Cruys, S., & Wagemans, J. (2017). Local-global processing bias is not a unitary individual difference in visual processing. *Vision Research*, 1–11. https://doi.org/10.1016/j.visres.2017.01.008
Chen, K., & Choi, H. (2008). Visual attention and eye movements. *Attention*, 119–153.
Chen, L., Meier, K., Blair, M., Watson, M., & Wood, M. (2013). Temporal characteristics of overt attentional behavior during category learning. *Attention, Perception, and Psychophysics*, *75*(2), 244–256. https://doi.org/10.3758/s13414-012-0395-8
Chiang, H.-M., & Lin, Y.-H. (2007). Mathematical ability of students with Asperger syndrome and high-functioning autism: a review of literature. *Autism : The International Journal of Research and Practice*, *11*(6), 547–556. https://doi.org/10.1177/1362361307083259
Coates, S. (1974). Sex differences in field dependence-independence between the ages of 3 and 6. *Perceptual and Motor Skills*, *39*, 1307–1310.
Cohen, L. (1992). Power Primer. *Psychological Bulletin, 112(1) 155-159*
Colom, R., García, L., Juan-Espinosa, M., & Abad, F. (2002). Null sex differences in general intelligence: Evidence from the WAIS-III. *The Spanish Journal of Psychology*, *5*(1), 29–35.
Conklin, R., Muir, W., & Boersma, F. (1968). Field dependency-independency and eye-movement patterns. *Perceptual and Motor Skills*, *26*, 59–65.
Cook, C. M., & Saucier, D. M. (2010). Mental rotation, targeting ability and Baron-Cohen's empathizing-systemizing theory of sex differences. *Personality and Individual Differences*, *49*(7), 712–716. https://doi.org/10.1016/j.paid.2010.06.010
Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. https://doi.org/10.1038/nrn755
Cousineau, D., & Chartier, S. (2011). Outliers detection and treatment: a review. *International Journal of Psychological Research*, *3*(1), 58–67.
Cragg, L. (2016). The development of stimulus and response interference control in midchildhood. *Developmental Psychology*, *52*(2), 242–252. https://doi.org/10.1037/dev0000074
Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, *3*(2), 63–68. https://doi.org/10.1016/j.tine.2013.12.001

Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, *162*, 12–26. https://doi.org/10.1016/j.cognition.2017.01.014

Cragg, L., & Nation, K. (2008). Go or no-go? Developmental improvements in the efficiency of response inhibition in mid-childhood. *Developmental Science*, *11*(6), 819–827. https://doi.org/10.1111/j.1467-7687.2008.00730.x

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). https://doi.org/10.1371/journal.pone.0188693

D'Souza, D., Booth, R., Connolly, M., Happé, F., & Karmiloff-Smith, A. (2016). Rethinking the concepts of "local or global processors": Evidence from Williams syndrome, Down syndrome, and autism spectrum disorders. *Developmental Science*, *19*(3), 452–468. https://doi.org/10.1111/desc.12312

Dajani, D., & Uddin, L. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *TRENDS in Neuroscience*, *38*(9), 571–578. https://doi.org/10.1016/j.tins.2015.07.003.Demystifying

Dale, G., & Arnell, K. M. (2013). Investigating the stability of and relationships among global/local processing measures. *Attention, Perception, & Psychophysics*, *75*(3), 394–406. https://doi.org/10.3758/s13414-012-0416-7

Dalrymple, K. A., Manner, M. D., Harmelink, K. A., Teska, E. P., & Elison, J. T. (2018). An examination of recording accuracy and precision from eye tracking data from toddlerhood to adulthood. *Frontiers in Psychology*, *9*(MAY), 1–12. https://doi.org/10.3389/fpsyg.2018.00803

Daniels, H., & Moore, D. (2000). Interaction of cognitive style and learner control of presentation mode in a hypermedia environment. *International Journal of Instructional Media*, *27*(4), 369–383. https://doi.org/10.1017/CBO9781107415324.004

Davey, B. (1990). Field dependence-independence and reading comprehension questions: Task and reader interactions. *Contemporary Educational Psychology*, *15*(3), 241–250. https://doi.org/10.1016/0361-476X(90)90021-R

Davidoff, J., Fonteneau, E., & Fagot, J. (2008). Local and global processing: Observations from a remote culture. *Cognition*, *108*(3), 702–709. https://doi.org/10.1016/j.cognition.2008.06.004

Davidoff, J., Fonteneau, E., & Goldstein, J. (2008). Cultural differences in perception: Observations from a remote culture. *Journal of Cognition and Culture*, *8*, 189–209.

Davies, J., & Graff, M. (2006). Wholist-analytic cognitive style: A matter of reflection. *Personality and Individual Differences*, *41*(6), 989–997.

https://doi.org/10.1016/j.paid.2005.09.011

De-Wit, L., Huygelier, H., Van der Hallen, R., Chamberlain, R., & Wagemans, J. (2017). Developing the Leuven Embedded Figures Test (L-EFT): Testing the stimulus features that influence embedding. *PeerJ*, 1–16. https://doi.org/10.7717/peerj.2862

De-Wit, L., & Wagemans, J. (2015). *Individual Differences in Local and Global Perceptual Organization*. (J. Wagemans, Ed.), *The Oxford Handbook of Perceptual Organization* (Vol. 1). Oxforrd Handbook. https://doi.org/10.1093/oxfordhb/9780199686858.013.028

De Jonge, M., Kemner, C., Naber, F., & Van Engeland, H. (2009). Block Design reconstruction skills: Not a good candidate for an endophenotypic marker in autism research. *European Child and Adolescent Psychiatry*, *18*(4), 197–205. https://doi.org/10.1007/s00787-008-0708-6

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, *103*(2), 186–201. https://doi.org/10.1016/j.jecp.2009.01.004

Deary, I. J., Strand, S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, *35*(1), 13–21. https://doi.org/10.1016/j.intell.2006.02.001

Deloitte. (2013). *Measuring the Economic Benefits of Mathematical Science Research in the UK. Measuring the economic benefits of mathematical science research in the UK*. Retrieved from https://www.lms.ac.uk/sites/lms.ac.uk/files/Report EconomicBenefits.pdf

Deruelle, C., Rondan, C., Gepner, B., & Fagot, J. (2006). Processing of compound visual stimuli by children with autism and Asperger syndrome. *International Journal of Psychology*, *41*(2), 97–106. https://doi.org/10.1080/00207590500184610

Derussy, E. A., & Futch, E. (1971). Field dependence-independence as related to college curricula. *Perceptual and Motor Skills*, *33*, 1235–1237.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1016/j.biotechadv.2011.08.021.Secreted

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. *Child and Adolescent Mental Health*, *20*(2), 116–125. https://doi.org/10.1111/camh.12072

Dore, P., Dumani, A., Wyatt, G., & Shepherd, A. J. (2018). Links between global and local shape perception, coloured backgrounds, colour discrimination, and non-verbal IQ. *Vision Research*, *151*(January), 31–40. https://doi.org/10.1016/j.visres.2018.02.004

Dukette, D., & Stiles, J. (1996). Children's analysis of hierarchical patterns: Evidence from a similarity judgment task. *Journal of Experimental Child Psychology*, *63*(1), 103–140. https://doi.org/10.1006/jecp.1996.0044

Dukette, D., & Stiles, J. (2001). The effects of stimulus density on children's analysis of hierarchical patterns. *Developmental Science*, *4*(2), 233–251. https://doi.org/10.1111/1467-7687.00168

Duncan, J., Chylinski, D., Mitchell, D. J., & Bhandari, A. (2017). Complexity and compositionality in fluid intelligence. *Proceedings of the National Academy of Sciences*, 201621147. https://doi.org/10.1073/pnas.1621147114

Dunn, L. M., Dunn, D. M., Sewel, J., Styles, B. (2009) The British Picture Vocabulary Scale (BPVS-3). 3rd edition. London: GL Assessment Ltd.

Eckstein, M. K., Guerra-carrillo, B., Singley, A. T. M., & Bunge, S. A. (2016). Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development? *Developmental Cognitive Neuroscience*, 1–23. https://doi.org/10.1016/j.dcn.2016.11.001

Escovar, E., Rosenberg-Lee, M., Uddin, L. Q., & Menon, V. (2016). The empathizing-systemizing theory, social abilities, and mathematical achievement in children. *Scientific Reports*, *6*(February), 23011. https://doi.org/10.1038/srep23011

Evans, C., Richardson, J., & Waring, M. (2013). Field independence: Reviewing the evidence. *British Journal of Educational Psychology*, *83*(2), 210–224. https://doi.org/10.1111/bjep.12015

Evers, K., Steyaert, J., Noens, I., & Wagemans, J. (2015). Reduced recognition of dynamic facial emotional expressions and emotion-specific response bias in children with an autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(6), 1774–1784. https://doi.org/10.1007/s10803-014-2337-x

Farran, E. K., Jarrold, C., & Gathercole, S. E. (2003). Divided attention, selective attention and drawing: Processing preferences in Williams Syndrome are dependent on the task administered. *Neuropsychologia*, *41*(6), 676–687. https://doi.org/10.1016/S0028-3932(02)00219-1

Ferrari, P. L. (2003). Abstraction in mathematics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*(1435), 1225–1230. https://doi.org/10.1098/rstb.2003.1316

Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics (4th edition)*. London: Sage Publications Ltd.

Fletcher-Watson, S., Findlay, J., Leekam, S., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, *37*, 571–583. https://doi.org/10.1068/p5705

Flevaris, A., Martinez, A., & Hillyard, S. (2014). Attending to global versus local stimulus features modulates neural processing of low versus high spatial frequencies: an analysis with event-related brain potentials. *Frontiers in Psychology*, *5*(April), 1–11. https://doi.org/10.3389/fpsyg.2014.00277

Flexer, B., & Roberge, J. (1980). IQ, field dependence-independence, and the development of formal operational thought. *The Journal of General Psychology*, *103*(2), 191–201. https://doi.org/10.1080/00221309.1980.9920998

Flexer, B., & Roberge, J. (1983). A longitudinal investigation of dependence-independence and the development of formal operational thought. *British Journal of Educational Psychology*, *53*(53), 195–204.

Focquaert, F., Steven, M., Wolford, G., Colden, A., & Gazzaniga, M. (2007). Empathizing and systemizing cognitive traits in the sciences and humanities. *Personality and Individual Differences*, *43*(3), 619–625. https://doi.org/10.1016/j.paid.2007.01.004

Formoso, J., Injoque-Ricle, I., Barreyro, J., Calero, A., Jacubovich, S., & Burin, D. (2018). Mathematical cognition, working memory, and processing speed in children. *Cognition, Brain, Behavior. An Interdisciplinary Journal*, *22*(2), 59–84. https://doi.org/10.24193/cbb.2018.22.05

Franceschini, S., Bertoni, S., Gianesini, T., Gori, S., & Facoetti, A. (2017). A different vision of dyslexia: Local precedence on global perception. *Scientific Reports*, *7*(1), 1–10. https://doi.org/10.1038/s41598-017-17626-1

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, *19*(3), 313–332. https://doi.org/10.1080/02699930441000238.Positive

Friedman, N., Miyake, A., Corley, R., Young, S., Defries, J., & Hewitt, J. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172–179.

Frith, U., & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology*, *1*(July 2016), 329–342. https://doi.org/10.1111/j.2044-835X.1983.tb00906.x

Gasper, K., & Clore, G. L. (2002). Attending to the big picture: Mood and global versus local processing of visual information. *Psychological Science*, *13*(1), 34–40. https://doi.org/10.1111/1467-9280.00406

Gathercole, S. E., Pickering, S. J., Knight, C., & Stegmann, Z. (2004). Working memory skills and educational attainment: Evidence from National Curriculum assessments at 7 and 14 years of age. *Applied Cognitive Psychology*, *18*(1), 1–16. https://doi.org/10.1002/acp.934

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A five year longitudinal study. *Developmental Psychology*, *47*(6), 1539–1552. https://doi.org/10.1037/a0025510.Cognitive

Gerlach, C., & Krumborg, J. R. (2014). Same, same - but different: On the use of Navon derived measures of global/local processing in studies of face processing. *Acta Psychologica*, *153*, 28–38. https://doi.org/10.1016/j.actpsy.2014.09.004

Gerlach, C., & Poirel, N. (2018). Navon's classical paradigm concerning local and global processing relates systematically to visual object classification performance. *Scientific Reports*, *8*(1), 1–9. https://doi.org/10.1038/s41598-017-18664-5

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., Hodgkiss, A., Thomas, M. S. C., & Farran, E. K. (2018). The developmental relations between spatial cognition and mathematics in primary school children. *Developmental Science*, (November). https://doi.org/10.1111/desc.12786

Gilmore, C., Clayton, S., Cragg, L., McKeaveney, C., Simms, V., & Johnson, S. (2018). Understanding arithmetic concepts: The role of domain-specific and domain-general skills. *PLoS ONE*, *13*(9), 1–20. https://doi.org/10.1371/journal.pone.0201724

Gilmore, C., Keeble, S., Richardson, S., & Cragg, L. (2015). The role of cognitive inhibition in different components of arithmetic. *ZDM - Mathematics Education*, *47*(5), 771–782. https://doi.org/10.1007/s11858-014-0659-y

Gilmore, C., Keeble, S., Richardson, S., & Cragg, L. (2017). The interaction of procedural skill, conceptual understanding and working memory in early mathematics achievement. *Journal of Numerical Cognition*, *3*(2), 400–416. https://doi.org/10.5964/jnc.v3i2.51

Globerson, T., Weinstein, E., & Sharabany, R. (1985). Teasing out cognitive development from cognitive style : A training study. *Developmental Psychology*, *21*(4), 682–691.

Glynn, M. A., & Stoner, S. (1987). Construct validity of the Children's Embedded Figures Test. *Perceptual and Motor Skills*, *64*, 1035–1038.

Goldenfeld, N., Baron-cohen, S., Wheelwright, S., Ashwin, C., & Chakrabarti, B. (2007). Empathizing and systemizing in males, females, and autism: A test of the neural competition theory. In T. Farrow & P. Woodruff (Eds.), *Empathy in Mental Illness* (pp. 322–334). Cambridge University Press. Retrieved from http://opus.bath.ac.uk/19162/

Goodenough, D., & Karp, S. (1961). Field dependence and intellectual functioning. *Journal of Abnormal and Social Psychology*, *63*(2), 241–246. https://doi.org/10.1037/h0046524

Goodenough, D. R., & Eagle, C. J. (1963). A modification of the Embedded-Figures Test for use with young children. *Journal of Genetic Psychology*, *103*(1), 67–74. https://doi.org/10.1017/CBO9781107415324.004

Goodenough, D., & Witkin, H. (1977). Origins of the field-dependent and field-independent cognitive styles. *ETS Research Bulletin Series*, *July*, i–80. https://doi.org/10.1002/j.2333-8504.1977.tb01134.x

Goodwin, L. D., & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of r. *Journal of Experimental Education*, *74*(3), 249–266. https://doi.org/10.3200/JEXE.74.3.249-266

Greenberg, D. M., Warrier, V., Allison, C., & Baron-Cohen, S. (2018). Testing the empathizing-systemizing theory of sex differences and the extreme male brain theory of autism in half a million people. https://doi.org/10.1073/pnas.1811032115

Griswold, D. E., Barnhill, G. P., Myles, B. S., Hagiwara, T., & Simpson, R. L. (2002). Asperger syndrome and academic achievement. *Focus on Autism and Other Developmental Disabilities*, *17*(2), 94–102. https://doi.org/10.1177/10883576020170020401

Gropen, J., Clark-Chiarelli, N., Hoisington, C., & Ehrlich, S. B. (2011). The importance of executive function in early science education. *Child Development Perspectives*, *5*(4), 298–304. https://doi.org/10.1111/j.1750-8606.2011.00201.x

Guisande, M., Páramo, M., Tinajero, C., & Almeida, L. (2007). Field dependence-independence (FDI) cognitive style: An analysis of attentional functioning. *Psicothema*, *19*(4), 572–577.

Guisande, M., Rodríguez, M., Almeida, L., Tinajero, C., & Páramo, M. (2008). Digits backward and the Children's Embedded Figures Test among school-age Spanish children. *Perceptual and Motor Skills*, *107*(3), 923–932. https://doi.org/10.2466/PMS.107.3.923-932

Hadad, B. S., & Ziv, Y. (2015). Strong bias towards analytic perception in ASD does not necessarily come at the price of impaired integration skills. *Journal of Autism and Developmental Disorders*, *45*(6), 1499–1512. https://doi.org/10.1007/s10803-014-2293-5

Han, S., & Humphreys, G. (2002). Segmentation and selection contribute to local processing in hierarchical analysis. *The Quarterly Journal of Experimental Psychology*, *55A*(1), 5–21. https://doi.org/10.1080/0272498014300012

Han, S., & Humphreys, G. W. (1999). Interactions between perceptual organization based on Gestalt laws and those based on hierarchical processing. *Perception & Psychophysics*, *61*(7), 1287–1298. https://doi.org/10.3758/BF03206180

Hansen, D., & Ji, Q. (2010). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 478–500. https://doi.org/10.1109/TPAMI.2009.30

Hansen, A., Drews, D., Dudgeon, J., Lawton, F., & Surtees, L. (2017). *Children's errors in mathematics*. London, UK: Learning Matters

Hao, X., Wang, K., Li, W., Yang, W., Wei, D., Qiu, J., & Zhang, Q. (2013). Individual differences in brain structure and resting brain function underlie cognitive styles: Evidence from the Embedded Figures Test. *PLoS ONE*, *8*(12), 1–9. https://doi.org/10.1371/journal.pone.0078089

Happé, F. (1997). Central coherence and theory of mind in autism: Reading homographs in context. *British Journal of Developmental Psychology*, *15*(1), 1–12. https://doi.org/10.1111/j.2044-835X.1997.tb00721.x

Happé, F. (1999). Autism: cognitive deficit or cogntive style? *Trends in Cogntive Science*, *3*(6), 216–222. https://doi.org/10.1016/S1364-6613(99)01318-2

Happé, F., & Booth, R. (2008). The power of the positive: revisiting weak coherence in autism spectrum disorders. *Quarterly Journal of Experimental Psychology (2006)*, *61*(1), 50–63. https://doi.org/10.1080/17470210701508731

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5–25. https://doi.org/10.1007/s10803-005-0039-0

Harrison, T. B., & Stiles, J. (2009). Hierarchical forms processing in adults and children. *Journal of Experimental Child Psychology*, *103*(2), 222–240. https://doi.org/10.1016/j.jecp.2008.09.004

Hautus, M. J. (1995). Corrections for extreme proportions and their biasing effects on estimated values of d′. *Behavior Research Methods, Instruments, & Computers*, *27*(1), 46–51. https://doi.org/10.3758/BF03203619

Hayward, D. A., Shore, D. I., Ristic, J., Kovshoff, H., Iarocci, G., Mottron, L., & Burack, J. A. (2012). Flexible visual processing in young adults with autism: The effects of implicit learning on a global–local task. *Journal of Autism and Developmental Disorders*, *42*(11), 2383–2392. https://doi.org/10.1007/s10803-012-1485-0

Hasenäcker, J., Verra, L., & Schroeder, S. (2019). Comparing length and frequency effects in children across modalities. *Quarterly Journal of Experimental Psychology*, *72*(7), 1682–1691. https://doi.org/10.1177/1747021818805063

Hebart, M. N., & Hesselmann, G. (2012). What visual information is processed in the human dorsal stream? *Journal of Neuroscience*, *32*(24), 8107–8109. https://doi.org/10.1523/JNEUROSCI.1462-12.2012

Hegdé, J. (2008). Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, *84*(4), 405–439. https://doi.org/10.1016/j.pneurobio.2007.09.001

Helo, A., Pannasch, S., Sirri, L., & Rämä, P. (2014). The maturation of eye movement behavior: Scene viewing characteristics in children and adults. *Vision Research*, *103*, 83–91. https://doi.org/10.1016/j.visres.2014.08.006

Henderson, J. M. (2017). Gaze control as prediction. *Trends in Cognitive Sciences*, *21*(1), 15–23. https://doi.org/10.1016/j.tics.2016.11.003

Hessels, R., Kemner, C., van den Boomen, C., & Hooge, I. (2016). The area-of-interest problem in eyetracking research: A noise-robust solution for face and sparse stimuli. *Behavior Research Methods*, *48*(4), 1694–1712. https://doi.org/10.3758/s13428-015-0676-y

Hessels, R., Niehorster, D., Andersson, R., & Hooge, I. (2018). Is the eye-movement field confused about fixations and saccades ? A survey among 124 researchers. *Royal Society Open Science*.

Hodgkiss, A., Gilligan, K. A., Tolmie, A. K., Thomas, M. S. C., & Farran, E. K. (2018). Spatial cognition and science achievement: The contribution of intrinsic and extrinsic spatial skills from 7 to 11 years. *British Journal of Educational Psychology*. https://doi.org/10.1111/bjep.12211

Holmes, J., & Adams, J. W. (2006). Working memory and children's mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology*, *26*(3), 339–366. https://doi.org/10.1080/01443410500341056

Holmqvist, K., Nystrom, M., Andersson, R., Dewhurst, R., Jarodzka, H., van de Weijer, J. (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press, New York.

Hooge, I., Holmqvist, K., & Nyström, M. (2016). The pupil is faster than the corneal reflection (CR): Are video based pupil-CR eye trackers suitable for studying detailed dynamics of eye movements? *Vision Research*, *128*, 6–18. https://doi.org/10.1016/j.visres.2016.09.002

Hubber, P. J., Gilmore, C., & Cragg, L. (2014). The roles of the central executive and visuospatial storage in mental arithmetic: A comparison across strategies. *Quarterly Journal of Experimental Psychology*, *67*(5), 936–954. https://doi.org/10.1080/17470218.2013.838590

Hübner, R., & Studer, T. (2009). Functional hemispheric differences for the categorization of global and local information in naturalistic stimuli. *Brain and Cognition*, *69*(1), 11–18. https://doi.org/10.1016/j.bandc.2008.04.009

Hughes, C. (2011). Changes and challenges in 20 years of research into the development of executive functions. *Infant and Child Development*, *20*, 251–271. https://doi.org/10.1002/icd

Hughes, C., & Ensor, R. (2011). Individual differences in growth in executive function across the transition to school predict externalizing and internalizing behaviors and self-perceived academic success at 6 years of age. *Journal of Experimental Child Psychology*, *108*(3), 663–676. https://doi.org/10.1016/j.jecp.2010.06.005

Hugrass, L., Verhellen, T., Morrall-Earney, E., Mallon, C., & Crewther, D. P. (2018). The effects of red surrounds on visual magnocellular and parvocellular cortical processing and perception. *J Vis*, *18*(4), 8. https://doi.org/10.1167/18.4.8

Huizinga, Mariëtte, Baeyens, D., & Burack, J. A. (2018). Editorial: Executive function and education. *Frontiers in Psychology*, *9*(August), 1–3. https://doi.org/10.3389/fpsyg.2018.01357

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

Huntsinger, J., Clore, G., & Bar-Anan, Y. (2010). Mood and global-local focus: Priming a local focus reverses the link between mood and global-local processing. *Emotion*, *10*(5), 722–726. https://doi.org/10.1037/a0019356

Huttermann, S., Bock, O., & Memmert, D. (2014). Subliminal primes for global or local processing influence judgments of vehicular traffic. *Consciousness and Cognition*, *29*, 230–234. https://doi.org/10.1016/j.concog.2014.08.007

Huygelier, H., Van der Hallen, R., Wagemans, J., De-Wit, L., & Chamberlain, R. (2018). The Leuven Embedded Figures Test (L-EFT): Measuring perception, intelligence or executive function? *PeerJ*, *6*, e4524. https://doi.org/10.7717/peerj.4524

Iarocci, G., Burack, J. A., Shore, D. I., Mottron, L., & Enns, J. T. (2006). Global-local visual processing in high functioning children with autism: Structural vs. implicit task biases. *Journal of Autism and Developmental Disorders*, *36*(1), 117–129. https://doi.org/10.1007/s10803-005-0045-2

Imanaka, M., Kakigi, R., & Nakata, H. (2017). The relationship between cognitive style and event-related potentials during auditory and somatosensory go/no-go paradigms. *NeuroReport*, *28*, 822–827. https://doi.org/10.1097/WNR.0000000000000833

Isbell, E., Fukuda, K., Neville, H., Vogel, E., & Conklin, H. (2015). Visual working memory continues to develop through adolescence. *Frontiers in Psychology*, *6*(May), 1–10. https://doi.org/10.3389/fpsyg.2015.00696

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews. Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Jäkel, F., Singh, M., Wichmann, F. A., & Herzog, M. H. (2016). An overview of quantitative approaches in Gestalt perception. *Vision Research*, *126*, 3–8. https://doi.org/10.1016/j.visres.2016.06.004

James, I. (2010). Autism and mathematical talent. *Mathematical Intelligencer*, *32*(1), 56–58. https://doi.org/10.1007/s00283-009-9086-2

Jantan, H. (2014). Relationship between students' cognitive style (field-dependent and field-independent cognitive styles) with their mathematic achievement in primary school. *International Journal of Humanities, Social Sciences and Education (IJHSSE)*, *1*(10), 88–93.

Jarrold, C., Gilchrist, I. D., & Bender, A. (2005). Embedded figures detection in autism and typical development: Preliminary evidence of a double dissociation in relationships with visual search. *Developmental Science*, *8*, 344–351. https://doi.org/10.1111/j.1467-7687.2005.00422.x

Jia, S., Zhang, Q., & Li, S. (2014). Field dependence-independence modulates the efficiency of filtering out irrelevant information in a visual working memory task. *Neuroscience*, *278*, 136–143. https://doi.org/10.1016/j.neuroscience.2014.07.075

Jones, C., Happé, F., Golden, H., Marsden, A., Tregay, J., Simonoff, E., Pickles, A., Baird, G. & Charman, T. (2009). Reading and arithmetic in adolescents with autism spectrum disorders: Peaks and dips in attainment. *Neuropsychology*, *23*(6), 718–728. https://doi.org/10.1037/a0016360

Karmiloff-Smith, A. (1994). Précis of Beyond modularity: A developmental perspective on cognitive science. *Behavioral and Brain Sciences*, *17*, 693–745. https://doi.org/10.3109/13682829409041485

Karp, S. A., & Konstadt, N. L. (1963). Manual for the Children's Embedded Figures Test. Oxford, England: Cognitive Tests.

Katagiri, M., Kasai, T., Kamio, Y., & Murohashi, H. (2013). Individuals with Asperger's disorder exhibit difficulty in switching attention from a local level to a global level. *Journal of Autism and Developmental Disorders*, *43*(2), 395–403. https://doi.org/10.1007/s10803-012-1578-9

Katsuki, F., & Constantinidis, C. (2014). Bottom-up and top-down attention: Different processes and overlapping neural systems. *The Neuroscientist*, *20*(5), 509–521. https://doi.org/10.1177/1073858413514136

Kemps, E., De Rammelaere, S., & Desmet, T. (2000). The development of working memory: Exploring the complementarity of two models. *Journal of Experimental Child Psychology*, *109*(77), 89–109. https://doi.org/10.1006/jecp.2000.2589

Kepner, M. D., & Neimark, E. D. (1984). Test-retest reliability and differential patterns of score change on the Group Embedded Figures Test. *Journal of Personality and Social Psychology*, *46*(6), 1405–1413. https://doi.org/10.1037/0022-3514.46.6.1405

Kidron, R., Kaganovskiy, L., & Baron-Cohen, S. (2018). Empathizing-systemizing cognitive styles: Effects of sex and academic degree. *PLoS ONE*, *13*(3), 1–17. https://doi.org/10.1371/journal.pone.0194515

Killgore, W. D. S., & Gogel, H. (2014). The Design Organization Test: Further demonstration of reliability and validity as a brief measure of visuospatial ability. *Applied Neuropsychology: Adult*, *21*(4), 297–309. https://doi.org/10.1080/23279095.2013.811671

Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, *112*(1), 24–38.

Kimchi, R. (1994). The role of wholistic/configural properties versus global properties in visual form perception. *Perception*, *23*(5), 489–504. https://doi.org/10.1068/p230489

Kimchi, R. (2015). The perception of hierarchical structure. *Oxford Handbook of Perceptual Organization*, (March 2018), 129–149. https://doi.org/10.1093/oxfordhb/9780199686858.013.025

Kimchi, R., Amishav, R., & Sulitzeanu-Kenan, A. (2009). Gender differences in global-local perception? Evidence from orientation and shape judgments. *Acta Psychologica*, *130*(1), 64–71. https://doi.org/10.1016/j.actpsy.2008.10.002

Kimchi, R., Hadad, B., Behrmann, M., & Palmer, S. (2005). Microgenesis and ontogenesis of perceptual organization. *Psychological Science*, *16*(4), 282–290. https://doi.org/10.1111/j.0956-7976.2005.01529.x

Kimchi, R., Yeshurun, Y., Spehar, B., & Pirkner, Y. (2016). Perceptual organization, visual attention, and objecthood. *Vision Research*, *126*, 34–51. https://doi.org/10.1016/j.visres.2015.07.008

Kinchla, R., & Wolfe, J. (1979). The order of visual processing:“Top-down,”“bottom-up,” or “middle-out.” *Perception & Psychophysics*, *25*(3), 225–231. https://doi.org/10.3758/BF03202991

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2010). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology*, *20*(1), 407–428. https://doi.org/10.1207/S15326942DN2001

Koldewyn, K., Jiang, Y., Weigelt, S., & Kanwisher, N. (2013). Global/local processing in autism: Not a disability, but a disinclination. *Journal of Autism and Developmental Disorders*, *43*(10), 2329–2340. https://doi.org/10.1007/s10803-013-1777-z

Kovacs, I. (2000). Human development of perceptual organization. *Vision Research*, *40*, 1301–1310.

Krakowski, C. S., Poirel, N., Vidal, J., Roëll, M., Pineau, A., Borst, G., & Houdé, O. (2016). The forest, the trees, and the leaves: Differences of processing across development. *Developmental Psychology*, *52*(8), 1262–1272. https://doi.org/10.1037/dev0000138

Kramer, J., Ellenberg, L., Leonard, J., & Share, L. (1996). Developmental sex differences in global-local perceptual bias. *Neuropsychology*, *10*(3), 402–407.

Kveraga, K., Boshyan, J., & Bar, M. (2007). Magnocellular projections as the trigger of top-down facilitation in recognition. *The Journal of Neoroscience*, *27*(48), 13232–13240. https://doi.org/10.1523/JNEUROSCI.3481-07.2007

Kyttala, M., & Lehto, J. (2008). Some factors underlying mathematical performance: The role of visuospatial working memory and non-verbal intelligence. *European Journal of Psychology of Education*, *23*(1), 77–94. https://doi.org/10.1007/978-3-662-45657-6_34

Lai, M. C., Lombardo, M. V., Chakrabarti, B., Ecker, C., Sadek, S. A., Wheelwright, S. J., Murphy, D., Suckling, J., Bullmore, E., MRC AIMS Consortium, & Baron-Cohen, S. (2012). Individual differences in brain structure underpin empathizing-systemizing cognitive styles in male adults. *NeuroImage*, *61*(4), 1347–1354. https://doi.org/10.1016/j.neuroimage.2012.03.018

Lai, M. L., Tsai, M. J., Yang, F. Y., Hsu, C. Y., Liu, T. C., Lee, S. W. Y., Chiou, G. & Tsai, C. C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*(88), 90–115. https://doi.org/10.1016/j.edurev.2013.10.001

Lamb, M. R., & Robertson, L. C. (1990). The effect of visual angle on global and local reaction times depends on the set of visual angles presented. *Perception & Psychophysics*, *47*(5), 489–496. https://doi.org/10.3758/BF03208182

Lao, J., Vizioli, L., & Caldara, R. (2013). Culture modulates the temporal dynamics of global/local processing. *Culture and Brain*, *1*(2–4), 158–174. https://doi.org/10.1007/s40167-013-0012-2

Lawson, A. (1976). Formal operations and field independence. *Perceptual and Motor*

Skills, *42*, 981–982.

Lawson, A. (1983). Predicting science achievement: The role of developmental level, disembedding ability, mental capacity, prior knowledge, and beliefs. *Journal of Research in Science Teaching*, *20*(2), 117–129.

Leat, S. J., Yadav, N. K., & Irving, E. L. (2009). Development of visual acuity and contrast sensitivity in children. *Journal of Optometry*, *2*(1), 19–26. https://doi.org/10.3921/joptom.2009.19

Lee, K, & Bull, R. (2016). Developmental changes in working memory, updating, and math achievement. *Journal of Educational Psychology*, *108*(6), 869–882.

Lee, Kyoung-min, & Kang, S. (2002). Arithmetic operation and working memory: Differential suppression in dual tasks. *Cognition*, *83*, 63–68.

Lee, P. S., Foss-Feig, J., Henderson, J. G., Kenworthy, L. E., Gilotty, L., Gaillard, W. D., & Vaidya, C. J. (2007). Atypical neural substrates of Embedded Figures Task performance in children with autism spectrum disorder. *NeuroImage*, *38*(1), 184–193. https://doi.org/10.1016/j.neuroimage.2007.07.013

Leo-Rhynie, E. (1985). Field independence, academic orientation, and achievement. *Current Psychology*, *4*(1), 22–27. https://doi.org/10.1007/BF02686563

Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on pre-registration. *International Review of Social Psychology*, *32*(1), 1–10.

Li, Y., Wu, S., Zhu, J., & O'Boyle, M. W. (2014). Sex and ability differences in neural activation for disembedding figures: An EEG investigation. *Learning and Individual Differences*, *35*(October 2014), 142–146. https://doi.org/10.1016/j.lindif.2014.07.012

Ling, J., Burton, T., Salt, J., & Muncer, S. (2009). Psychometric analysis of the systemizing quotient (SQ) scale. *British Journal of Psychology*, *100*(3), 539–552. https://doi.org/10.1348/000712608X368261

Liu, L., Wang, F., Zhou, K., Ding, N., & Luo, H. (2017). Perceptual integration rapidly activates dorsal visual pathway to guide local processing in early visual areas. *PLOS Biology*, *15*(11), e2003646. https://doi.org/10.1371/journal.pbio.2003646

Lubin, A., Vidal, J., Lanoë, C., Houdé, O., & Borst, G. (2013). Inhibitory control is needed for the resolution of arithmetic word problems: A developmental negative priming study. *Journal of Educational Psychology*, *105*(3), 701–708. https://doi.org/10.1037/a0032625

Ludwig, C. J. H., Davies, J. R., & Eckstein, M. P. (2014). Foveal analysis and peripheral selection during active visual sampling. *Proceedings of the National Academy of Sciences*, *111*(2), E291–E299.

https://doi.org/10.1073/pnas.1313553111

Luna, B., Velanova, K., & Geier, C. F. (2008). Development of eye-movement control. *Brain and Cognition*, *68*(3), 293–308. https://doi.org/10.1016/j.bandc.2008.08.019

McLeod, S. A. (2008). *Likert scale*. Retrieved from https://www.simplypsychology.org/likert-scale.html

Manjaly, Z. M., Bruning, N., Neufang, S., Stephan, K. E., Brieber, S., Marshall, J. C., Kamp-Becker, I., Remschmidt, H., Herpertz-Dahlmann, B., Konrad, K. & Fink, G. R. (2007). Neurophysiological correlates of relatively enhanced local visual search in autistic adolescents. *NeuroImage*, *35*(1), 283–291. https://doi.org/10.1016/j.neuroimage.2006.11.036

Mann, T. A., & Walker, P. (2003). Autism and a deficit in broadening the spread of visual attention. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*(2), 274–284. https://doi.org/10.1111/1469-7610.00120

Männamaa, M., Kikas, E., Peets, K., & Palu, A. (2012). Cognitive correlates of math skills in third-grade students. *Educational Psychology*, *32*(1), 21–44. https://doi.org/10.1080/01443410.2011.621713

Mareschal, D. (2016). The neuroscience of conceptual learning in science and mathematics. *Current Opinion in Behavioral Sciences*, *10*, 114–118. https://doi.org/10.1016/j.cobeha.2016.06.001

Marini, F., & Marzi, C. A. (2016). Gestalt perceptual organization of visual stimuli captures attention automatically: Electrophysiological evidence. *Frontiers in Human Neuroscience*, *10*(August), 1–10. https://doi.org/10.3389/fnhum.2016.00446

Massonnié, J., Rogers, C. J., Mareschal, D., & Kirkham, N. Z. (2019). Is classroom noise always bad for children? The contribution of age and selective attention to creative performance in noise. *Frontiers in Psychology*, *10*(FEB), 1–12. https://doi.org/10.3389/fpsyg.2019.00381

May, T., Rinehart, N., Wilding, J., & Cornish, K. (2013). The role of attention in the academic attainment of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *43*(9), 2147–2158. https://doi.org/10.1007/s10803-013-1766-2

May, T., Rinehart, N., Wilding, J., & Cornish, K. (2015). Attention and basic literacy and numeracy in children with autism spectrum disorder: A one-year follow-up study. *Research in Autism Spectrum Disorders*, *9*, 193–201. https://doi.org/10.1016/j.rasd.2014.10.010

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities.

Learning and Instruction, *29*, 43–55. https://doi.org/10.1016/j.learninstruc.2013.07.005

Mayes, S., & Calhoun, S. (2003). Analysis of WISC-III, Stanford-Binet:IV, and academic achievement test scores in children with autism. *Journal of Autism & Developmental Disorders*, *33*(3).

Mazzocco, M. M. M., & Kover, S. T. (2007). A longitudinal assessment of executive function skills and their association with math performance. *Child Neuropsychology*, *13*(1), 18–45. https://doi.org/10.1080/09297040600611346

McKone, E., Aimola Davies, A., Fernando, D., Aalders, R., Leung, H., Wickramariyaratne, T., & Platow, M. J. (2010). Asia has the global advantage: Race and visual attention. *Vision Research*, *50*(16), 1540–1549. https://doi.org/10.1016/j.visres.2010.05.010

Merkley, R., Thompson, J., & Scerif, G. (2016). Of huge mice and tiny elephants: Exploring the relationship between inhibitory processes and preschool math skills. *Frontiers in Psychology*, *1903*(6), 1–14. https://doi.org/10.3389/fpsyg.2015.01903

Meyer, M. L., Salimpoor, V. N., Wu, S. S., Geary, D. C., & Menon, V. (2010). Differential contribution of specific working memory components to mathematics achievement in 2nd and 3rd graders. *Learning and Individual Differences*, *20*(2), 101–109. https://doi.org/10.1016/j.lindif.2009.08.004

Michimata, C., Okubo, M., & Mugishima, Y. (1999). Effects of background color on the global and local processing of hierarchically organized stimuli. *Journal of Cognitive Neuroscience*, *11*(1), 1–8. https://doi.org/10.1162/089892999563201

Milne, E., & Szczerbinski, M. (2009). Global and local perceptual style, field-independence, and central coherence: An attempt at concept validation. *Advances in Cognitive Psychology*, *5*(1), 1–26. https://doi.org/10.2478/v10053-008-0062-8

Mitchelmore, M., & White, P. (2007). Abstraction in mathematics learning. *Mathematics Education Research Journal*, *19*(2), 1–9.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “Frontal Lobe” tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49–100. https://doi.org/10.1006/cogp.1999.0734

Miyake, A., Witzki, A., & Emerson, M. (2001). Field dependence-independence from a working memory perspective: a dual-task investigation of the Hidden Figures Test. *Memory*, *9*(4–6), 445–457. https://doi.org/10.1080/09658210143000029

Miyamoto, Y., Nisbett, R., & Masuda, T. (2006). Culture and the physical environment. *Psychological Science*, *17*(2), 113–119. https://doi.org/10.1111/j.1467-

9280.2006.01673.x

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., Houts, R., Poulton, R., Ross, S., Sears, M., Murray Thompson, W., & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, *108*(7), 2693–2698. https://doi.org/10.1073/pnas.1010076108

Mondloch, C. J., Geldart, S., Maurer, D., & de Schonen, S. (2003). Developmental changes in the processing of hierarchical shapes continue into adolescence. *Journal of Experimental Child Psychology*, *84*, 20–40. https://doi.org/10.1016/S0022-0965(02)00161-3

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Science achievement gaps begin very early, persist, and are largely explained by modifiable factors. *Educational Researcher*, *45*(1), 18–35. https://doi.org/10.3102/0013189X16633182

Morsanyi, K., Primi, C., Handley, S., Chiesi, F., & Galli, S. (2012). Are systemizing and autistic traits related to talent and interest in mathematics and engineering? Testing some of the central claims of the empathizing–systemizing theory. *British Journal of Psychology*, *103*, 472–496. https://doi.org/10.1111/j.2044-8295.2011.02089.x

Mousavi, S., Radmehr, F., & Alamolhodaei, H. (2012). The role of mathematical homework and prior knowledge on the relationship between students' mathematical per- formance, cognitive style and working memory capacity. *Electronic Journal of Research in Educational Psychology*. Retrieved from http://www.investigacion-psicopedagogica.org/revista/new/ContadorArticulo.php?757

Mumma, G. (1993). The Embedded Figures Test: internal structure and development of a short form. *Personality and Individual Differences*, *15*(2), 221–224. Retrieved from http://www.sciencedirect.com/science/article/pii/0191886993900293

National Audit Office. (2018). *Delivering STEM Science-Technology-Engineering-Maths for the economy-summary*. Retrieved from https://www.nao.org.uk/wp-content/uploads/2018/01/Delivering-STEM-Science-technology-engineering-and-mathematics-skills-for-the-economy-Summary.pdf

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Navon, D. (1981). The forest revisited: More on global precedence. *Psychological Research*, *43*(1), 1–32. https://doi.org/10.1007/BF00309635

Navon, D. (2003). What does a compound letter tell the psychologist's mind? *Acta Psychologica*, *114*(3), 273–309. https://doi.org/10.1016/j.actpsy.2003.06.002

Nayar, K., Franchak, J., Adolph, K., & Kiorpes, L. (2015). From local to global processing: The development of illusory contour perception. *Journal of Experimental Child Psychology*, *131*, 38–55. https://doi.org/10.1016/j.jecp.2014.11.001

Nayfeld, I., Fuccillo, J., & Greenfield, D. B. (2013). Executive functions in early learning: Extending the relationship between executive functions and school readiness to science. *Learning and Individual Differences*, *26*, 81–88. https://doi.org/10.1016/j.lindif.2013.04.011

Naylor, B., Naylor, S., & Mitchell, G. (2000). *The Snowman's Coat and Other Science Questions*. Hodder Children's.

Nersessian, N. (1989). Conceptual change in science and in science education. *Synthese*, *80*, 163–183.

Nisbett, R., & Masuda, T. (2003). Culture and point of view. *PNAS*, *100*(19), 11163–11170. https://doi.org/10.4324/9780203774908

Nisiforou, E., & Laghos, A. (2013). Do the eyes have it? Using eye tracking to assess students cognitive dimensions. *Educational Media International*, *50*(4), 247–265. https://doi.org/10.1080/09523987.2013.862363

Nisiforou, E., & Laghos, A. (2016). Field dependence-independence and eye movement patterns: Investigating users' differences through an eye tracking study. *Interacting with Computers*, *28*(4), 407–420. https://doi.org/10.1093/iwc/iwv015

Norbury, C., & Nation, K. (2011). Understanding variability in reading comprehension in adolescents with autism spectrum disorders: Interactions with language status and decoding skill. *Scientific Studies of Reading*, *15*(3), 191–210. https://doi.org/10.1080/10888431003623553

Noudoost, B., Chang, M., Steinmetz, N. a., & Moore, T. (2010). Top-down control of visual attention. *Current Opinion in Neurobiology*, *20*(2), 183–190. https://doi.org/10.1016/j.conb.2010.02.003.TOP-DOWN

Noyes, J. M., & Garland, K. J. (2008). Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, *51*(9), 1352–1375. https://doi.org/10.1080/00140130802170387

Oh, H., Glutting, J., Watkins, M., Youngstrom, E., & McDermott, P. (2004). Correct interpretation of latent versus observed abilities: Implications from structural equation modeling applied. *The Journal of Special Education*, *38*(3), 159–173.

Oishi, S., Jaswal, V. K., Lillard, A. S., Mizokawa, A., Hitokoto, H., & Tsutsui, Y. (2014).

Cultural variations in global versus local processing: A developmental perspective. *Developmental Psychology*, *50*(12), 2654–2665. https://doi.org/10.1037/a0038272

Onyekuru, B. U. (2015). Field dependence-field independence cognitive style, gender, career choice and academic achievement of secondary school students in Emohua local government area of Rivers State. *Journal of Education and Practice*, *6*(10), 76–86.

Ozonoff, S., Strayer, D., McMahon, W., & Filloux, F. (1994). Executive function abilities in autism and Tourette syndrome: an information processing approach. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *35*(6), 1015–1032. https://doi.org/10.1111/j.1469-7610.1994.tb01807.x

Panis, S., De Winter, J., Vandekerckhove, J., & Wagemans, J. (2008). Identification of everyday objects on the basis of fragmented versions of outlines. *Perception*, *37*(2), 271–289. https://doi.org/10.1068/p5429

Passolunghi, M., & Cornoldi, C. (2008). Working memory failures in children with arithmetical difficulties. *Child Neuropsychology*, *14*(5), 387–400. https://doi.org/10.1080/09297040701566662

Pellicano, E., Maybery, M., Durkin, K., & Maley, A. (2006). Multiple cognitive capabilities/deficits in children with an autism spectrum disorder: “Weak” central coherence and its relationship to theory of mind and executive control. *Development and Psychopathology*, *18*(1), 77–98. https://doi.org/10.1017/S0954579406060056

Pellizzoni, S., & Passolunghi, M. C. (2017). Convergent evaluation of working memory and arithmetic ability in a child with autism spectrum disorder without intellectual impairment. *Frontiers in Psychology*, *8*(JUL), 1–5. https://doi.org/10.3389/fpsyg.2017.01278

Peng, P., & Miller, A. C. (2016). Does attention training work? A selective meta-analysis to explore the effects of attention training and moderators. *Learning and Individual Differences*, *45*, 77–87. https://doi.org/10.1016/j.lindif.2015.11.012

Pereg, M., & Meiran, N. (2018). Evidence for instructions-based updating of task-set representations: the informed fadeout effect. *Psychological Research*, *82*(3), 549–569. https://doi.org/10.1007/s00426-017-0842-1

Petersen, S. E., & Posner, M. I. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. https://doi.org/10.1146/annurev-neuro-062111-150525

Pinto, Y., van der Leij, A., Sligte, I., Lamme, V., & Scholte, H. (2013). Bottom-up and top-down attention are independent. *Journal of Vision*, *13*(3), 1–14. https://doi.org/10.1167/13.3.16.doi

Pithers, R. (2002). Cognitive learning style: A review of the field dependent-field independent approach. *Journal of Vocational Education & Training*, *54*(1), 117–132. https://doi.org/10.1080/13636820200200191

Plaisted, K., O'Riordan, M., & Baron-Cohen, S. (1998). Enhanced visual search for a conjunctive target in autism: A research note. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *39*(5), 777–783. https://doi.org/10.1111/1469-7610.00376

Plaisted, K., Swettenham, J., & Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *40*(5), 733–742. https://doi.org/10.1111/1469-7610.00489

Pletzer, B., & Harris, T. (2018). Sex hormones modulate the relationship between global advantage, lateralization, and interhemispheric connectivity in a Navon paradigm. *Brain Connectivity*, *8*(2), 106–118. https://doi.org/10.1089/brain.2017.0504

Pletzer, B., Petasis, O., & Cahill, L. (2014). Switching between forest and trees: Opposite relationship of progesterone and testosterone to global-local processing. *Hormones and Behavior*, *66*(2), 257–266. https://doi.org/10.1016/j.yhbeh.2014.05.004

Pletzer, B., Scheuringer, A., & Scherndl, T. (2017). Global-local processing relates to spatial and verbal processing: Implications for sex differences in cognition. *Scientific Reports*, (May), 1–9. https://doi.org/10.1038/s41598-017-11013-6

Poirel, N., Krakowski, C., Sayah, S., Pineau, A., Houdé, O., & Borst, G. (2014). Do you want to see the tree? Ignore the forest. Inhibitory control during local processing: A negative priming study of local-global processing. *Experimental Psychology*, *61*(3), 205–214. https://doi.org/10.1027/1618-3169/a000240

Poirel, N., Mellet, E., Houdé, O., & Pineau, A. (2008). First came the trees, then the forest: Developmental changes during childhood in the processing of visual local-global patterns according to the meaningfulness of the stimuli. *Developmental Psychology*, *44*(1), 245–253. https://doi.org/10.1037/0012-1649.44.1.245

Poirel, N., Pineau, A., Jobard, G., & Mellet, E. (2008). Seeing the forest before the trees depends on individual field-dependency characteristics. *Experimental Psychology*, *55*(5), 328–333. https://doi.org/10.1027/1618-3169.55.5.328

Poirel, N., Pineau, A., & Mellet, E. (2006). Implicit identification of irrelevant local objects interacts with global/local processing of hierarchical stimuli. *Acta Psychologica*, *122*(3), 321–336. https://doi.org/10.1016/j.actpsy.2005.12.010

Poirel, N., Pineau, A., & Mellet, E. (2008). What does the nature of the stimuli tell us about the global precedence effect? *Acta Psychologica*, *127*(1), 1–11. https://doi.org/10.1016/j.actpsy.2006.12.001

Poirel, N., Simon, G., Cassotti, M., Leroux, G., Perchey, G., Lanoë, C., Lubin, A., Turbelin, M., Rossi, S., Pineau, A., & Houdé, O. (2011). The shift from local to global visual processing in 6-year-old children is associated with grey matter loss. *PLoS ONE*, *6*(6), 1–5. https://doi.org/10.1371/journal.pone.0020879

Pollet, T. V, & Meij, L. Van Der. (2017). To remove or not to remove: The impact of outlier handling on significance testing in testosterone data. *Adaptive Human Behavior and Physiology*, *3*, 43–60. https://doi.org/10.1007/s40750-016-0050-z

Quinlan, P. T., & Wilton, R. N. (1998). Competition between the Gestalt principles in vision. *Perception*, *27*, 417–430.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110–122. https://doi.org/10.1016/j.lindif.2009.10.005

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, UK: Oxford Psychologists Press; San Antionio, TX: The Psychological Corporation.

Rayner, K. (2009). *The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search*. *Quarterly Journal of Experimental Psychology* (Vol. 62). https://doi.org/10.1080/17470210902816461

Razpurker-Apfeld, I., & Kimchi, R. (2007). The time course of perceptual grouping: The role of segregation and shape formation. *Perception & Psychophysics*, *69*(5), 732–743. https://doi.org/10.3758/BF03193775

Razumnikova, O. M., & Volf, N. V. (2011). Information processing specialization during interference between global and local aspects of visual hierarchical stimuli in men and women. *Human Physiology*, *37*(2), 137–142. https://doi.org/10.1134/S0362119711020186

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*(1), 1–41. https://doi.org/10.1016/j.cogpsych.2004.11.001

Rémy, L., & Gilles, P. Y. (2014). Relationship between field dependence-independence and the g factor: What can problem-solving strategies tell us? *Revue Europeene de Psychologie Appliquee*, *64*(2), 77–82. https://doi.org/10.1016/j.erap.2014.02.001

Rhodes, S. M., Booth, J. N., Palmer, L. E., Blythe, R. A., Delibegovic, M., & Wheate, N.

J. (2016). Executive functions predict conceptual learning of science. *British Journal of Developmental Psychology*, *34*(2). https://doi.org/10.1111/bjdp.12129

Richardson, J., & Turner, T. (2000). Field dependence revisited I: Intelligence. *Educational Psychology*, *20*(3), 255–270. https://doi.org/10.1080/713663747

Richardson, K. (2002). What IQ tests test. *Theory & Psychology*, *12*(3), 283–314.

Riekki, T., Salmi, J., Svedholm-Häkkinen, A. M., & Lindeman, M. (2018). Intuitive physics ability in systemizers relies on differential use of the internalizing system and long-term spatial representations. *Neuropsychologia*, *109*(June 2017), 10–18. https://doi.org/10.1016/j.neuropsychologia.2017.11.030

Rinehart, N. J., Bradshaw, J. L., Moss, S. A., Brereton, A. V., & Tonge, B. J. (2000). Atypical interference of local detail on global processing in high-functioning autism and Asperger's disorder. *Journal of Child Psychology and Psychiatry*, *41*(6), 769–778. https://doi.org/10.1111/1469-7610.00664

Ritchie, J. B., & de Beeck, H. O. (2019). Using neural distance to predict reaction time for categorizing the animacy, shape, and abstract properties of objects. *Scientific Reports*, *9*(1), 1–8. https://doi.org/10.1038/s41598-019-49732-7

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics : An iterative process. *Journal of Educational Psychology*, *93*(2), 346–362. https://doi.org/10.1037//0022-0663.93.2.346

Rittschof, K. (2010). Field dependence-independence as visuospatial and executive functioning in working memory: Implications for instructional systems design and research. *Educational Technology Research and Development*, *58*(1), 99–114. https://doi.org/10.1007/s11423-008-9093-6

Roalf, D., Lowery, N., & Turetsky, B. I. (2006). Behavioral and physiological findings of gender differences in global-local visual processing. *Brain and Cognition*, *60*(1), 32–42. https://doi.org/10.1016/j.bandc.2005.09.008

Roberge, J. J., & Flexer, B. K. (1983). Cognitive style, operativity, and mathematics achievement. *Journal for Research in Mathematics Education*, *14*(5), 344–353.

Rönnlund, M., & Nilsson, L. G. (2006). Adult life-span patterns in WAIS-R Block Design performance: Cross-sectional versus longitudinal age gradients and relations to demographic factors. *Intelligence*, *34*(1), 63–78. https://doi.org/10.1016/j.intell.2005.06.004

Rozencwajg, P., & Corroyer, D. (2002). Strategy development in a Block Design task. *Intelligence*, *30*(1), 1–25. https://doi.org/10.1016/S0160-2896(01)00063-0

Salvucci, D., & Goldberg, J. (2000). Identifying fixations and saccades in eye-tracking protocols. *Proceedings of the Eye Tracking Research and Applications*

Symposium, 71–78.
Scherf, S., Behrmann, M., & Luna, B. (2009). Emergence of global shape processing continues through adolescence. *Child Development*, *80*(1), 162–177. https://doi.org/10.1111/j.1467-8624.2008.01252.x.Emergence
Scherf, S., Luna, B., Kimchi, R., Minshew, N., & Behrmann, M. (2008). Missing the big picture: Impaired development of global shape processing in autism. *Autism Research*, *1*(2), 114–129. https://doi.org/10.1002/aur.17
Scheuringer, A., & Pletzer, B. (2016). Sex differences in the Kimchi-Palmer task revisited: Global reaction times, but not number of global choices differ between adult men and women. *Physiology & Behavior*, *165*, 159–165. https://doi.org/10.1016/j.physbeh.2016.07.012
Schielzeth, H., Dingemanse, N. J., Nakagawa, S., Westneat, D. F., Allegue, H., Teplitsky, C., Reale, D., Dochtermann, N.A., Garamszegi, L. Z., & Araya-Ajoy, Y. G. (2020). Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, *2020* (January), 1–12. https://doi.org/10.1111/2041-210x.13434
Schorr, D., Bower, G. H., & Kiernan, R. J. (1982). Stimulus variables in the Block Design task. *Journal of Consulting and Clinical Psychology*, *50*(4), 479–487. https://doi.org/10.1037/0022-006X.50.4.479
Schuch, S., & Pütz, S. (2019). Mood state and conflict adaptation: an update and a diffusion model analysis. *Psychological Research*, (0123456789). https://doi.org/10.1007/s00426-019-01258-3
Shah, A., & Frith, U. (1983). An islet of ability in autistic children: A research note. *Journal of Child Psychology and Psychiatry*, *24*(4), 613–620. https://doi.org/10.1111/1469-7610.ep11422546.
Shah, A., & Frith, U. (1993). Why do autistic individuals show superior performance on the Block Design task? *Journal of Child Psychology & Psychiatry*, *34*(8), 1351–1364.
Shapiro, H. M., Wong, L. M., & Simon, T. J. (2013). A cross-sectional analysis of the development of response inhibition in children with chromosome 22q11.2 deletion syndrome. *Frontiers in Psychiatry*, *4*(AUG), 1–10. https://doi.org/10.3389/fpsyt.2013.00081
Shimi, A., Nobre, A., 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
Solman, G. J. F., Foulsham, T., & Kingstone, A. (2017). Eye and head movements are complementary in visual selection. *Royal Society Open Science*, *4*(1).

https://doi.org/10.1098/rsos.160569

Soltanlou, M., Pixner, S., & Nuerk, H. (2015). Contribution of working memory in multiplication fact network in children may shift from verbal to visuo-spatial: A longitudinal investigation. *Frontiers in Psychology*, *6*(July), 1–10. https://doi.org/10.3389/fpsyg.2015.01062

Spector, R. (1990). Visual fields. In H. Walker, W. Hall, & H. J (Eds.), *Clinical Methods: The historical, physical, and laboratory examinations* (3rd ed., pp. 565–572). Boston: Butterworths.

Srinivasan, N., & Hanif, A. (2010). Global-happy and local-sad: Perceptual processing affects emotion identification. *Cognition & Emotion*, *24*(6), 1062–1069. https://doi.org/10.1080/02699930903101103

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology (2006)*, *59*(4), 745–759. https://doi.org/10.1080/17470210500162854

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods*, *31*(1), 137–149.

Steele, A., Karmiloff-Smith, A., Cornish, K., & Scerif, G. (2012). The Multiple subfunctions of attention: Differential developmental gateways to literacy and numeracy. *Child Development*, *83*(6), 2028–2041. https://doi.org/10.1111/j.1467-8624.2012.01809.x

Sternberg, R. J. (2008). Increasing fluid intelligence is possible after all. *Proceedings of the National Academy of Sciences*, *105*(19), 6791–6792. https://doi.org/10.1073/pnas.0803396105

Stoesz, B., Jakobson, L., Kilgour, A., & Lewycky, S. (2007). Local processing advantage in musicians: Evidence from disembedding and constructional tasks. *Music Perception: An Interdisciplinary Journal*, *25*(2), 153–165.

Svedholm-Häkkinen, A. M., & Lindeman, M. (2016). Testing the empathizing-systemizing theory in the general population: Occupations, vocational interests, grades, hobbies, friendship quality, social intelligence, and sex role identity. *Personality and Individual Differences*, *90*, 365–370. https://doi.org/10.1016/j.paid.2015.11.044

Swyter, L., & Michael, W. (1982). The interrelationships of four measures hypothesized to represent the field dependence-field independence construct. *Educational and Psychological Measurement*, *42*, 877–888.

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2014). Cognitive components of a mathematical processing network in 9-year-old children.

Developmental Science, *17*(4), 506–524. https://doi.org/10.1111/desc.12144

Tai, R., Loehr, J., & Brigham, F. (2006). An exploration of the use of eye-gaze tracking to study problem-solving on standardized science assessments. *International Journal of Research & Method in Education*, *29*(2), 185–208. https://doi.org/10.1080/17437270600891614

Thomas, C., Kveraga, K., Huberle, E., Karnath, H. O., & Bar, M. (2012). Enabling global processing in simultanagnosia by psychophysical biasing of visual pathways. *Brain*, *135*(5), 1578–1585. https://doi.org/10.1093/brain/aws066

Thomas, M., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*, *52*(April), 336–358.

Thomason, M., Race, E., Burrows, B., Whitfield-gabrieli, S., Glover, H., & Gabrieli, J. (2009). Development of spatial and verbal working memory capacity in the human brain. *Journal of Cognition and Neuroscience*, *21*(2), 316–332. https://doi.org/10.1162/jocn.2008.21028.Development

Thompson, F., & Logue, S. (2006). An exploration of common student misconceptions in science. *International Education Journal*, *7*(4), 553–559.

Tiego, J., Testa, R., Bellgrove, M., Pantelis, C., Whittle, S., & Flynn, R. (2018). A hierarchical model of inhibitory control. *Frontiers in Psychology*, *9*(August), 1–25. https://doi.org/10.3389/fpsyg.2018.01339

Tinajero, C., & Paramo, M. (1997). Field dependence-independence and academic achievement: A re-examination of their relationship. *British Journal of Educational Psychology*, *67*, 199–212.

Tinajero, C., & Páramo, M. (1998). Field dependence-independence cognitive style and academic achievement: A review of research and theory. *European Journal of Psychology of Education*, *13*(2), 227–251. https://doi.org/10.1007/BF03173091

Tinajero, C., & Paramo, M. F. (1998). Educational implications of field dependence-independence: In answer to Bagley and Mallick. *British Journal of Educational Psychology*, *68*, 589–593. https://doi.org/10.1111/j.2044-8279.1998.tb01314.x

Tobii. (2014). User's manual: Tobii X2-60 eye tracker. Retrieved August 20, 2005, from https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/tobii-pro-x2-60-eye-tracker-user-manual.pdf/?v=1.0.3

Tobii. (2015). Tobii Studio user's manual; Version 3.4.5. https://doi.org/10.1017/CBO9781107415324.004

Tobii. (2019). *Types of eye movement*. Retrieved from: https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/types-of-eye-movements/

Tobii. (2019b). *The human eye*. Retrieved from: https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/the-human-eye/

Tobii. (2019c). *What happens during the eye tracker calibration?* Retrieved from: https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/what-happens-during-the-eye-tracker-calibration/

Trick, L., & Enns, J. (1997). Clusters precede shapes in perceptual organization. *Psychological Science*, *8*(2), 124–129. https://doi.org/10.2519/jospt.2014.5065

Tzuriel, D., & Egozi, G. (2010). Gender differences in spatial ability of young children: The effects of training and processing strategies. *Child Development*, *81*(5), 1417–1430. https://doi.org/10.1111/j.1467-8624.2010.01482.x

Underwood, G., Chapman, P., Berger, Z., & Crundall, D. (2003). Driving experience, attentional focusing, and the recall of recently inspected events. *Transportation Research Part F: Traffic Psychology and Behaviour*, *6*(4), 289–304. https://doi.org/10.1016/j.trf.2003.09.002

Van der Hallen, R., Chamberlain, R., De-Wit, L., & Wagemans, J. (2018). Superior disembedding in children with ASD: New tests using abstract, meaningful, and 3D contexts. *Journal of Autism and Developmental Disorders*, *48*(7), 2478–2489. https://doi.org/10.1007/s10803-018-3508-y

Van der Hallen, R., Evers, K., Brewaeys, K., Van den Noortgate, W., & Wagemans, J. (2015). Global processing takes time: A meta-analysis on local–global visual processing in ASD. *Psychological Bulletin*, *141*(3), 549–573. https://doi.org/10.1037/bul0000004

Van der Ven, S., Kroesbergen, E., Boom, J., & Leseman, P. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, *82*(1), 100–119. https://doi.org/10.1111/j.2044-8279.2011.02035.x

van der Ven, S., van der Maas, H., Straatemeier, M., & Jansen, B. (2013). Visuospatial working memory and mathematical ability at different ages throughout primary school. *Learning and Individual Differences*, *27*, 182–192. https://doi.org/10.1016/j.lindif.2013.09.003

Van Eylen, L., Boets, B., Steyaert, J., Wagemans, J., & Noens, I. (2018). Local and global visual processing in autism spectrum disorders: Influence of task and sample characteristics and relation to symptom severity. *Journal of Autism and Developmental Disorders*, *48*(4), 1359–1381. https://doi.org/10.1007/s10803-015-2526-2

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods*, *49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Vandierendonck, A. (2018). Further Tests of the Utility of Integrated Speed-Accuracy Measures in Task Switching. *Journal of Cognition*, *1*(1), 1–16. https://doi.org/10.5334/joc.6

Vanegas, S. B., & Davidson, D. (2015). Investigating distinct and related contributions of weak central coherence, executive dysfunction, and systemizing theories to the cognitive profiles of children with autism spectrum disorders and typically developing children. *Research in Autism Spectrum Disorders*, *11*, 77–92. https://doi.org/10.1016/j.rasd.2014.12.005

Vanmarcke, S., Mullin, C., van der Hallen, R., Evers, K., Noens, I., Steyaert, J., & Wagemans, J. (2016). In the eye of the beholder: Rapid visual perception of real-life scenes by young adults with and without ASD. *Journal of Autism and Developmental Disorders*, *46*(8), 1–18. https://doi.org/10.1007/s10803-016-2802-9

Verbruggen, F., Liefooghe, B., & Vandierendonck, A. (2004). The interaction between stop signal inhibition and distractor interference in the flanker and Stroop task. *Acta Psychologica*, *116*(1), 21–37. https://doi.org/10.1016/j.actpsy.2003.12.011

Vernon, P. (1972). The distinctiveness of field independence. *Journal of Personality*, *40*(3), 366–391. https://doi.org/10.1111/j.1467-6494.1972.tb00068.x

Vetter, P., & Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Consciousness and Cognition*, *27*(1), 62–75. https://doi.org/10.1016/j.concog.2014.04.007

Vidal-López, J., & Romera-Vivancos, J. A. (2009). Is manipulation of color effective in study of the global precedence effect? *Perceptual and Motor Skills*, *108*(2), 631–635. https://doi.org/10.2466/pms.108.2.631-635

Vinter, A., Puspitawati, I., & Witt, A. (2010). Children's spatial analysis of hierarchical patterns: Construction and perception. *Developmental Psychology*, *46*(6), 1621–1631. https://doi.org/10.1037/a0020615

Vosniadou, S. (2019). The development of students' understanding of science. *Frontiers in Education*, *4*(32), 1–6. https://doi.org/10.3389/feduc.2019.00032

Vosniadou, S., Pnevmatikos, D., Makris, N., Lepenioti, D., Eikospentaki, K., Chountala, A., & Kyrianakis, G. (2018). The recruitment of shifting and inhibition in on-line science and mathematics tasks. *Cognitive Science*, *42*(6), 1860–1886. https://doi.org/10.1111/cogs.12624

Vosniadou, S. (2013). Model based reasoning and the learning of counter-intuitive science concepts. *Infancia y Aprendizaje*, *36*(1), 5–33.

https://doi.org/10.1174/021037013804826519

Wagemans, J. (2015). Historical and conceptual background. In J. Wagemans (Ed.), *The Oxford Handbook of Perceptual Organization*. Oxford: Oxford Handbooks. https://doi.org/10.1093/oxfordhb/9780199686858.013.026

Wagemans, J., Elder, J., Kubovy, M., Palmer, S., Peterson, M., Singh, M., & von der Heydt, R. (2012a). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. https://doi.org/10.1037/a0029333

Wagemans, J., Elder, J., Kubovy, M., Palmer, S., Peterson, M., Singh, M., & von der Heydt, R. (2012b). A century of Gestalt psychology in visual perception: II. Perceptual grouping and figure–ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. https://doi.org/10.1037/a0029333

Wakabayashi, A, Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., & Weil, L. (2006). Development of short forms of the Empathy quotient (EQ-Short) and the systemizing quotient (SQ-Short). *Personality and Individual Differences*, *41*(5), 929–940. https://doi.org/10.1016/j.paid.2006.03.017

Wakabayashi, A. (2013). Individual differences in empathizing and systemizing in Japanese children: Psychometric properties of the children's versions of the Empathy Quotient (EQ) and Systemizing Quotient (SQ). *Japanese Psychological Research*, *55*(1), 12–19. https://doi.org/10.1111/j.1468-5884.2012.00537.x

Walker, I., & Zhu, Y. (2013). The benefit of STEM skills to individuals, society, and the economy. *Report to Royal Society's Vision for Science and Mathematics*, 1–38.

Walter, E. (2007). Visuospatial contextual processing: Illusions, hidden figures and autistics traits. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, *68*(4-B), 2683. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc5&NEWS=N&AN=2007-99200-029

Walter, E. (2009). A specific autistic trait that modulates visuospatial illusion susceptibility. *Journal of Autism and Developmental Disorders*, *39*(2), 339–349. https://doi.org/10.1007/s10803-008-0630-2.A

Wang, L., Mottron, L., Peng, D., Berthiaume, C., & Dawson, M. (2007). Local bias and local-to-global interference without global deficit: A robust finding in autism under various conditions of attention, exposure time, and visual angle. *Cognitive Neuropsychology*, *24*(5), 550–574. https://doi.org/10.1080/13546800701417096

Wass, S. V., Smith, T. J., & Johnson, M. H. (2013). Parsing eye-tracking data of variable quality to provide accurate fixation duration estimates in infants and adults. *Behavior Research Methods*, *45*(1), 229–250.

https://doi.org/10.3758/s13428-012-0245-6
Wechsler, D. (1997). *WAIS-III. Administration and scoring manual*. San Antonio, TX: The Psychological Association.
Wechsler, D. (2002). *Wechsler individual achievement test (2nd ed.)*. San Antonio, TX: Psychological Corp.
Wei, W., Guo, L., Georgiou, G. K., Tavouktsoglou, A., & Jordan, J. A. (2018). Different subcomponents of executive functioning predict different growth parameters in mathematics: Evidence from a 4-year longitudinal study with chinese children. *Frontiers in Psychology*, *9*(June), 1–10. https://doi.org/10.3389/fpsyg.2018.01037
Wei, X., Christiano, E. R., Yu, J. W., Wagner, M., & Spiker, D. (2015). Reading and math achievement profiles and longitudinal growth trajectories of children with an autism spectrum disorder. *Autism*, *19*(2), 200–210. https://doi.org/10.1177/1362361313516549
Wei, X., Yu, J. W., Shattuck, P., McCracken, M., & Blackorby, J. (2013). Science, technology, engineering, and mathematics (STEM) participation among college students with an autism spectrum disorder. *Journal of Autism & Developmental Disorders*, *43*(7), 1539–1546. https://doi.org/10.1016/j.micinf.2011.07.011.Innate
Weinbach, N., & Henik, A. (2014). Alerting enhances attentional bias for salient stimuli: Evidence from a global/local processing task. *Cognition*, *133*(2), 414–419. https://doi.org/10.1016/j.cognition.2014.07.005
Wheelwright, S., & Baron-Cohen, S. (2006). Predicting autism spectrum quotient (AQ) from the systemizing quotient-revised (SQ-R) and empathy quotient (EQ). *Brain Research*, 47–56. Retrieved from http://www.sciencedirect.com/science/article/pii/S0006899306000722
White, S., O'Reilly, H., & Frith, U. (2009). Big heads, small details and autism. *Neuropsychologia*, *47*(5), 1274–1281. https://doi.org/10.1016/j.neuropsychologia.2009.01.012
White, S., & Saldaña, D. (2011). Performance of children with autism on the Embedded Figures Test: A closer look at a popular task. *Journal of Autism and Developmental Disorders*, *41*(11), 1565–1572. https://doi.org/10.1007/s10803-011-1182-4
Wilkinson, D., Halligan, P., Marshall, J., Büchel, C., & Dolan, R. (2001). Switching between the forest and the trees: Brain systems involved in local/global changed-level judgments. *NeuroImage*, *13*(1), 56–67. https://doi.org/10.1006/nimg.2000.0678
Wilkinson, H. R., Smid, C., Morris, S., Farran, E. K., Dumontheil, I., Mayer, S., Tolmie, A., Bell, D., Porayska-Pomsta, K., Homes, W., Mareschal, D., Thomas, M., & The

UnLocke Team. (2019). Domain-specific inhibitory control training to improve children’s learning of counterintuitive concepts in mathematics and science. *Journal of Cognitive Enhancement*.

Witkin, H. A., & Asch, S. E. (1948). Studies in space orientation. III. Perception of the upright in the absence of a visual field. *Journal of Experimental Psychology*, *38*(5), 603–614. https://doi.org/10.1037/h0055372

Witkin, H., & Goodenough, D. (1977). Field dependence and interpersonal behavior. *Psychological Bulletin*, *84*(4), 661–689. https://doi.org/10.1037/0033-2909.84.4.661

Witkin, H., Goodenough, D., & Karp, S. (1967). Stability of cognitive style from childhood to young adulthood. *Journal of Personality and Social Psychology*, *7*(3), 291–300. https://doi.org/10.1037/h0025070

Witkin, H., Moore, C., Goodenough, D., & Cox, P. (1977). Field-dependent and field-independent cognitive styles and their educational implications. *Review of Educational Research*, *47*(1), 1–64. https://doi.org/10.1002/j.2333-8504.1975.tb01065.x

Witkin, H., Moore, C., Oltman, P., Goodenough, D., Friedman, F., Owen, D., & Raskin, E. (1977). Role of the field-dependent and field-independent cognitive styles in academic evolution: A longitudinal study. *Journal of Education & Psychology*, *69*(3), 197–211. https://doi.org/10.1037/0022-0663.69.3.197

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 0058. https://doi.org/10.1038/s41562-017-0058

Wright, D., & Skagerberg, E. (2012). Measuring empathizing and systemizing with a large US sample. *PLoS ONE*, *7*(2). https://doi.org/10.1371/journal.pone.0031661

Yarbus, A. (1967). *Eye movements and Vision*. New York: Plenum Press. https://doi.org/10.1016/0028-3932(68)90012-2

Ygge, J., Aring, E., Han, Y., Bolzani, R., & Hellström, A. (2005). Fixation stability in normal children. *Annals of the New York Academy of Sciences*, *483*, 480–483. https://doi.org/10.1196/annals.1325.049

Zachariou, V., Klatzky, R., & Behrmann, M. (2014). Ventral and dorsal visual stream contributions to the perception of object shape and object location. *Journal of Cognitive Neuroscience*, *26*(1), 189–209. https://doi.org/10.1162/jocn

Zaitchik, D., Iqbal, Y., & Carey, S. (2014). The effect of executive function on biological reasoning in young children: An individual differences study. *Child Development*, *85*(1), 160–175. https://doi.org/10.1111/cdev.12145

Zhang, L. F. (2004). Field-dependence/independence: Cognitive style or perceptual ability? - Validating against thinking styles and academic achievement. *Personality* *and Individual Differences*, *37*(6), 1295–1311. https://doi.org/10.1016/j.paid.2003.12.015

Appendix – scripts for introducing the Navon tasks

A.1 Free choice Navon task

"I'm going to show you some shapes-made-out-of-shapes. So you might see a big triangle made up of small squares, or a big circle made of little triangles (gesture). You will see one above the line and two below the line (gesture). What I'd like you to do is choose which of these under the line (gesture) is **most like** the one at the top (gesture).

"There's no right or wrong answer. Just choose whatever you think is most like the top one.

"If you think it's the one over here (gesture), press this key (key marked with a sticker) and if you think it's the one over here (gesture), press this key (key marked with a sticker).

"Ok, so have a look and choose which is most like the top one".

A.2 Selective attention Navon task

Block 1

"OK, now we're going to play a shape-matching game (**Figure A1A**).

"You're going to see some more shapes-made-out-of-shapes, and I'd like you to choose whether the **big** shape is a triangle or a square. So, have a look at the shape pitcture and decide is the big shape a triangle or a square.

"If you think it's a square, press this key (gesture – key marked with a sticker) and if you think it's a triangle, press this key (gesture – key marked with a sticker). OK?

"So fingers on the buttons – if you think it's a square, press this key (put a little post-it note with a square picture on the back of their hand so they can see which hand represents each shape) and if you think it's a triangle, press this key (put a little post-it note with a triangle picture on the back of their hand).

"The shape will appear in different places on the screen. Remember, you're looking at the big shape.

"Let's have a practice".

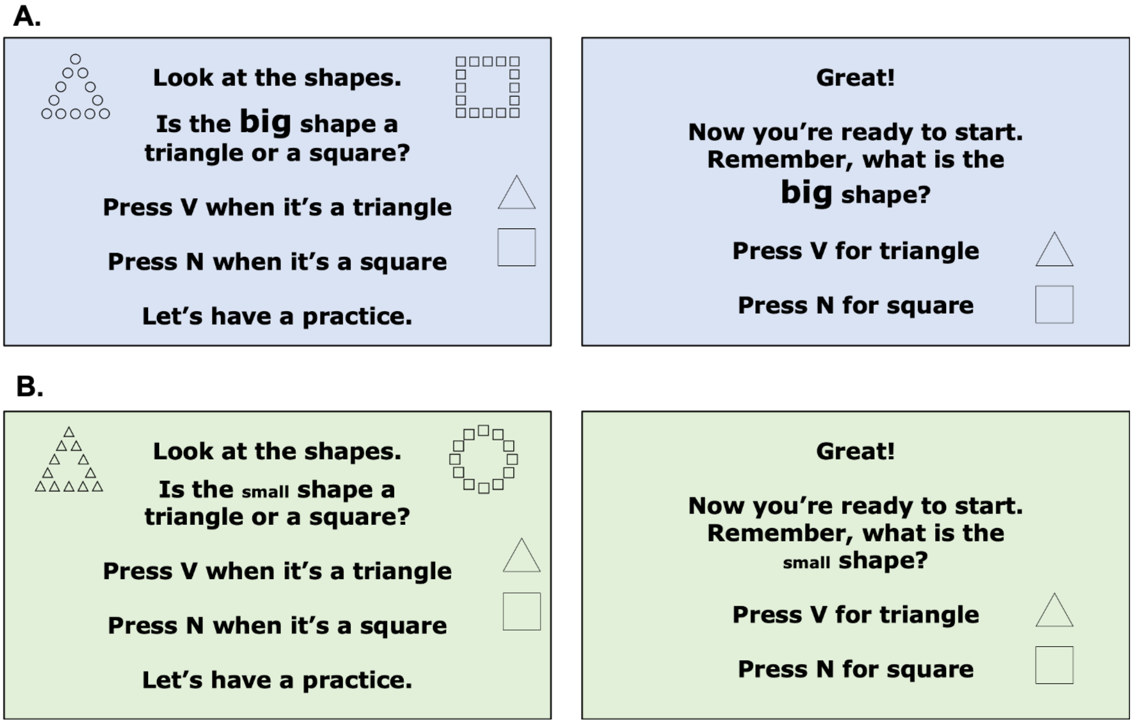


***Figure A1**: On-screen instructions for the selective attention Navon task. **A.** Instructions for the global level block. **B.** Instructions for the local level block.*

Block 2 – separated in the test battery

"Ok, now we're going to play the shape-matching game again, but this time, you're going to look at the **small** shape (**Figure A1B**). So have a look at the shapes-made-out-of-shapes and choose whether the small shape is a triangle or a square.

"If you think it's a square, press this key (gesture – key marked with a sticker) and if you think it's a triangle, press this key (gesture – key marked with a sticker). OK?

"So fingers on the buttons – if you think it's a square, press this key (put a little post-it note with a square picture on the back of their hand so they can see which hand represents each shape) and if you think it's a triangle, press this key (put a little post-it note with a triangle picture on the back of their hand).

"Remember, you're looking at the small shape.

"Let's have a practice".

A.3 Divided attention Navon task (big-small) – Time 1

"OK, we're going to play another shape game.

"This time, you're going to see whether the circle is the big shape or the small shape (**Figure A2**). So, each of the shapes-made-out-of-shapes will either have a circle as the big shape or a circle as the small shape.

"If you think the circle is the big shape, press this key (gesture – key marked with a sticker) and if you think the circle is the small shape, press this key (gesture – key marked with a sticker).

"So fingers on the buttons – if you think the circle is the big shape, press this key (put a little post-it note with 'big' written on it on the back of their hand) and if you think the circle is the small shape, press this key (put a little post-it note with 'small' written on it on the back of their hand).

"Remember, is the circle the big shape (gesture) or the small shape (gesture).

"Let's have a practice".

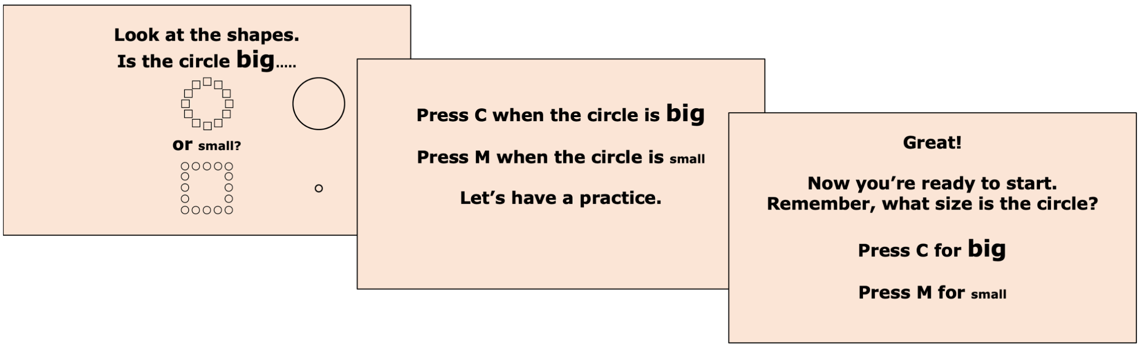


***Figure A2**: On-screen instructions for the big-small divided attention Navon task.*

A.4 Divided attention Navon tasks (big-small and yes-no) – Time 2

Talk through a powerpoint showing how the shapes-made-out-of-shapes have a circle at the big or small level, with some practice questions (verbal response) to test understanding of the big-small task and the yes-no task (**Figure A3**). This was introduced to ensure that the children understood the difference between the two similar divided attention tasks.

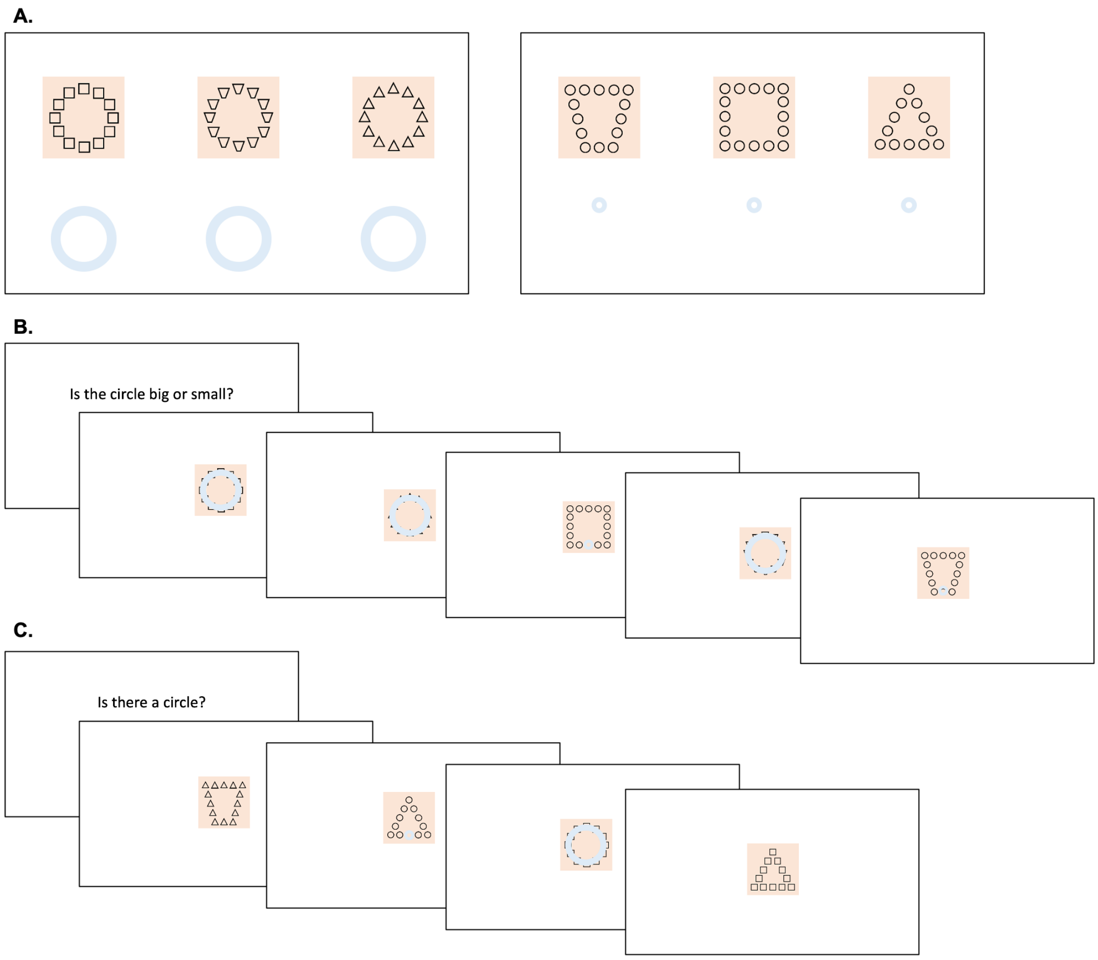


Figure A3: *Slides from the Powerpoint introduction to the divided attention Navon tasks. A. Introduction slides to explain the circle as the big and small shapes. Animation to move the shape onto the Navon figure. B. Examples of the circle being big and small. Blue solid line circle animation after a verbal response was given. C. Examples of the circle being present (yes) and absent (no). Blue solid line circle animation after a verbal response was given.*

Before they played each game, instructions were given as in **Section A.3** for the big-small task. For the yes-no task, the following instructions were given:

"OK, we're going to play another shape game.

"This time, you're going to see whether there is a circle in the shapes-made-out-of-shapes, or whether there is no circle (**Figure A4**). So, each of the shapes-made-out-of-shapes will either have a circle as the big or small shape, OR it will have no circle.

"If you think that YES, there is a circle, press this key (gesture – key marked with a sticker) and if you think NO, there is no circle, press this key (gesture – key marked with a sticker).

"So fingers on the buttons – if you think YES there is a circle, press this key (put a little post-it note with 'yes' written on it on the back of their hand) and if you think NO there is no circle, press this key (put a little post-it note with 'no' written on it on the back of their hand).

"Remember, is there a circle? Yes (gesture) or no (gesture).

"Let's have a practice".

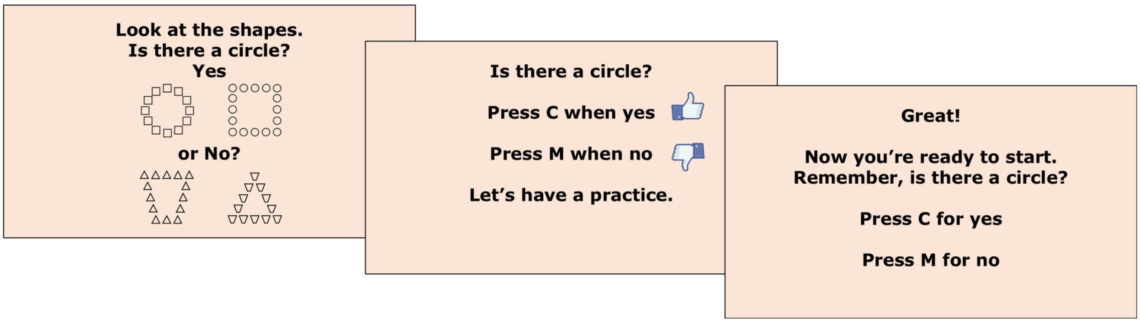


Figure A4: *On-screen instructions for the yes-no divided attention Navon task.*