Transdiagnostic Approaches to Neurodevelopment

Silvana Mareva

December 2021

Medical Research Council Cognition and Brain Sciences Unit
University of Cambridge
St Catharine’s College

This thesis is submitted for the degree of Doctor of Philosophy.
Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee.
Abstract

Neurodevelopmental difficulties can have a substantial impact on children’s lives and often have lasting effects in adult life. The scientific study of the causes and consequences of neurodevelopmental difficulties has arguably been slowed by the overreliance on case-control designs, which fail to capture the overlap across different neurodevelopmental disorders and the heterogeneity within them. The work presented in this thesis applies a transdiagnostic framework to large developmental datasets to advance our understanding of neurodevelopmental diversity. Three empirical studies were conducted with a cohort of children with a range of diagnosed and undiagnosed needs. The first study explored the interrelationships between language, communication, cognitive, and behavioural difficulties using network science (Chapter I). A subsequent study investigated how children vary in their relative strengths and weakness across these domains, and how this variation relates to socio-emotional functioning, academic skills, and neural white matter organisation (Chapter II). The third study compared the pattern of interrelationships between cognitive and academic skills observed in this cohort to those found in a community sample (Chapter III). Collectively, the studies presented in this thesis demonstrate the value of studying neurodevelopmental diversity transdiagnostically. I argue that the field needs to move away from studying groups of children with restricted difficulties and dedicate more effort towards understanding how multiple developmental factors interact to shape individual trajectories over developmental time.
Two of the chapters have been published and one has been submitted for publication. In all cases, I am the first author and was responsible for the study design, data analyses and interpretation, and write-up in collaboration with my supervisor Professor Joni Holmes.

- Chapter I has not been published. Many of the arguments presented in the chapter provide the rationale for subsequent empirical chapters. Some of its content is therefore included in the published versions of the empirical chapters.

- Chapter II has been published and appears modified to improve thesis coherence:
  
  

- Chapter III has been for publication and appears modified to improve thesis coherence:
  

  [https://doi.org/10.1101/2021.11.29.21267002](https://doi.org/10.1101/2021.11.29.21267002)

- Chapter IV has been published and appears modified:
  

  [https://doi.org/10.1016/j.jarmac.2021.08.005](https://doi.org/10.1016/j.jarmac.2021.08.005)

- Chapter V has not been published. Many of the discussion points presented in the chapter are related to the results of the empirical chapters. Therefore, some of its content is included in the published versions of the empirical chapters.
The Centre for Attention, Learning, and Memory (CALM) dataset, which is analysed in all three empirical chapters is not yet available as the study is still ongoing. The data will be made available via managed open-access once the study is complete.

The Nathan-Klein Institute Rockland sample (NKI-RS) dataset analysed in Chapter IV is freely available from the Collaborative Informatics and Neuroimaging Suite (COINS).
Acknowledgements

First and foremost, I would like to express my immense gratitude to my supervisor Professor Joni Holmes. Her indispensable support, kindness, and encouragement have been instrumental in enabling me to write this thesis and to develop the technical skills which were required for the work I wanted to do. She has been a true role model and an inspiring mentor who taught me a lot about neurodevelopment, cognitive science, work-life balance, and so much more. It was a true privilege to work with her and I can only aspire to one day mentor others in the same inspiring way. Here, I would also like to thank the Cognition, Education, & Emotion (CEE) lab for the welcoming environment, valuable feedback, and many insightful discussions.

I would also like to express my gratitude to Dr Thomas Bak, Dr Candice Morey, Dr Jenny Gibson, Dr Elian Fink, Dr Pablo Torres, and Dr Kai Ruggeri. They have all been brilliant mentors to me in the years before my PhD and inspired me to pursue this path. Special thanks to Dr Nicolas Chevalier who taught the developmental modules during my undergraduate degree and first introduced me to the topic of neurodevelopment and got me interested in this field. Similarly, I would like to thank all the children and volunteers whom I met while I worked with the charity Children’s Holiday Venture; they have all been an inspiration for this work.

Very special thanks to the Medical Research Council for their financial support and the Cognition and Brain Sciences Unit (CBU) for being a truly unprecedented academic home. The CBU provided a friendly and stimulating environment in which I was supported and challenged to grow and learn. Here, I met many inspiring colleagues and friends, some of whom played a key role in my journey. I would like to thank my PhD cohort, and particularly Sinéad, Andrea, Alex, and Giacomo, for how we supported one another. You made this journey incredibly special. I would like to thank Sally Butterfield for her kindness, genuine approach, and inspiring wisdom. Sally was the first person who welcomed me into the unit and always
provided the most generous support and guidance. Sally, I will always be grateful to you. I would also like to thank Dr Jacalyn Guy, who was a postdoctoral associate in the CEE lab and quickly became an invaluable advisor and a friend. I would also like to acknowledge Dr Duncan Astle, Dr Danyal Akarca, Dr Joe Bathelt, and Dr Jonathan Jones for advising me on the technical aspects of the research and especially the neuroimaging analyses; Dr Rogier Kievit whose work inspired some of my empirical investigations; and Professor Susan Gathercole for her key role in setting up the Centre for Attention, Learning, and Memory (CALM) project.

The CALM project is central to the work presented in this thesis, which would not have been possible without the brilliant team of researchers who enable this collaborative effort. The CALM Team includes lead investigators Duncan Astle, Kate Baker, Susan Gathercole, Joni Holmes, Rogier Kievit and Tom Manly. Data collection is assisted by a team of researchers and PhD students that includes Danyal Akarca, Joe Bathelt, Marc Bennett, Giacomo Bignardi, Sarah Bishop, Erica Bottacin, Lara Bridge, Diandra Brkic, Annie Bryant, Sally Butterfield, Elizabeth Byrne, Gemma Crickmore, Edwin Dalmaijer, Fánchea Daly, Tina Emery, Laura Forde, Grace Franckel, Delia Furhmann, Andrew Gadie, Sara Gharooni, Jacalyn Guy, Erin Hawkins, Agnieszka Jaroslawska, Sara Joeghan, Amy Johnson, Jonathan Jones, Elise Ng-Cordell, Sinead O'Brien, Cliodhna O'Leary, Joseph Rennie, Ivan Simpson-Kent, Roma Siugzdaitė, Tess Smith, Stephani Uh, Maria Vedechkina, Francesca Woolgar, Natalia Zdorovtsova, and Mengya Zhang. I also wish to thank the many professionals working in children’s services in the South-East and East of England for their support, and to the children and their families for giving up their time to visit the clinic.

Finally, I would like to thank my wonderful friends and family who have shown me the most generous love and support. And my grandparents, who always said: Умната. Горе главата. I am so fortunate to have you and will always be grateful.
# Table of contents

- Declaration .................................................................................................................. 2
- Abstract ......................................................................................................................... 3
- Data Availability Statement .......................................................................................... 5
- Acknowledgements ........................................................................................................ 6
- Table of contents .......................................................................................................... 8
- Preface ........................................................................................................................... 16

## CHAPTER I ..................................................................................................................... 17

1. Understanding Neurodevelopmental Difficulties .......................................................... 17

1.1. Traditional Approach to Neurodevelopmental Difficulties ....................................... 18

1.2. Challenges to the Traditional Approach to Neurodevelopmental Difficulties ............. 19

1.2.1. Heterogeneity ........................................................................................................ 19

1.2.2. Non-specific Contributions of Biopsychosocial Processes ..................................... 21

1.2.3. Little Evidence for Qualitative Discontinuity ......................................................... 21

1.2.4. Comorbidity ......................................................................................................... 22

1.2.5. From Single to Multiple Deficit Models .................................................................. 23

1.2.6. Circular Dynamics ............................................................................................... 24

1.2.7. Developmental Cascades ...................................................................................... 26

1.3. Summary ................................................................................................................... 27

2. Rethinking the Traditional Approach to Neurodiversity: Looking Beyond the Label

......................................................................................................................................... 28

2.1. New Recruitment Strategies ..................................................................................... 29

2.2. Diverse Measurement Practices ................................................................................ 30

2.3 Analytical Strategies ................................................................................................. 32

3. Summary ....................................................................................................................... 43
4. Thesis objectives and outline................................................................................................. 43

CHAPTER II ........................................................................................................................ 45

1. Introduction ......................................................................................................................... 46

2. Method ............................................................................................................................... 50
   2.1. Recruitment ................................................................................................................. 50
   2.2. Assessments ............................................................................................................... 50
   2.3. Participants ............................................................................................................... 51
       Table II.1 ..................................................................................................................... 52
   2.4. Statistical Analysis ................................................................................................. 53
       Table II.2 ..................................................................................................................... 57

3. Results ............................................................................................................................. 58
   3.1. Descriptive Statistics ............................................................................................. 58
   3.2. Subscale Validity Checks ....................................................................................... 58
       Figure II.1 .................................................................................................................. 59
   3.3. Network Estimation and Stability ......................................................................... 60
       Figure II.2 .................................................................................................................. 61
   3.4. Community Detection ............................................................................................. 62
       Figure II.3 .................................................................................................................. 63
   3.5. Symptom Centrality and Predictability .................................................................. 64
       Figure II.4 .................................................................................................................. 65

4. Discussion ......................................................................................................................... 66
   4.1. Limitations & Future directions ............................................................................. 69
   4.2. Conclusion ............................................................................................................... 70

CHAPTER III ......................................................................................................................... 72

1. Introduction ......................................................................................................................... 73
   1.1. Transdiagnostic Approaches to Neurodevelopmental Comorbidities.................... 75
   1.2. The current study ....................................................................................................... 76
2. Method ................................................................................................................. 77
   2.1. Recruitment ................................................................................................. 77
   2.2. Participants .................................................................................................. 77
   2.3. Assessments ................................................................................................. 78
   2.4. Magnetic resonance imaging (MRI) .............................................................. 80
   2.5. Statistical Analysis Overview ...................................................................... 83

3. Results ............................................................................................................... 88
   3.1. Community Detection .................................................................................. 88
   3.3. Data Reduction ............................................................................................. 94
   3.2. Subgroup Profiles ....................................................................................... 92
   3.3. Data Reduction ............................................................................................. 94
   3.4. Profile Comparisons .................................................................................... 96
   3.5. Subgroup Profiles: Behavioural Validation ............................................... 100
   3.6. Subgroup Profiles: Neural White Matter .................................................... 102

4. Discussion .......................................................................................................... 112
   4.1. Subgroups ..................................................................................................... 112
   4.2. Subgroup 1 .................................................................................................. 113
   4.3. Subgroup 2 .................................................................................................. 114
   4.4. Subgroup 3 .................................................................................................. 115
   4.5. Summary ...................................................................................................... 116
   4.6. Limitations & Future directions ................................................................. 117
   4.7. Conclusion ................................................................................................... 117
CHAPTER IV

1. Introduction ........................................................................................................... 120
   1.1. Challenges of the Unidirectional One-to-one Mapping Perspective ............ 120
   1.2. Bidirectional Dynamics ................................................................................. 121
   1.3. Cognitive and Academic Skills: New Approaches ..................................... 122
   1.4. Objectives .................................................................................................. 124

2. Method ................................................................................................................. 124
   2.1. Recruitment and Participants ...................................................................... 124
   2.2. Assessments ................................................................................................. 126
   2.3. Analysis Plan ............................................................................................... 128

3. Results .................................................................................................................. 128
   3.1. Descriptive Statistics ................................................................................... 128
       Table IV.1 ........................................................................................................ 130
       Table IV.2 ........................................................................................................ 131
       Figure IV.1 ....................................................................................................... 132
       Figure IV.2 ....................................................................................................... 133
       Figure IV.3 ....................................................................................................... 134
   3.2. Age Effects .................................................................................................. 135
   3.3. Network Estimation and Stability ............................................................... 135
       Figure IV.4 ....................................................................................................... 137
       Figure IV.5 ....................................................................................................... 138
   3.4. Network Comparison .................................................................................. 139
   3.5. Community detection .................................................................................. 139
   3.6. Node Centrality and Predictability .............................................................. 140
       Figure IV.6 ....................................................................................................... 141

4. Discussion ............................................................................................................. 142
   4.1. Limitations and Future Directions ............................................................... 143
   4.2. Conclusion ................................................................................................. 144
CHAPTER V

1. Introduction ................................................................. 145

1.1. Chapter II: Summary ......................................................... 145

1.2. Chapter III: Summary ......................................................... 146

1.3. Chapter IV: Summary ......................................................... 148

2. Key takeaways ............................................................. 149

2.1. Diagnostic Labels Provide Limited Information about Individual Strengths and Difficulties 149

2.2. Multiple Pathways to Similar Difficulties ................................ 150

2.3. The Unlikely Case of Isolated Difficulties ............................. 150

2.4. Shared and Unique Neural Substrates .................................. 151

3. Implications ................................................................. 152

3.1. Practical Implications ......................................................... 152

3.2. Theoretical Implications .................................................... 156

4. Limitations and Future Directions ...................................... 157

4.1. Understanding Age Effects .................................................. 157

4.2. Studying Developmental Change ....................................... 158

4.3. Focus on Mechanisms Across Levels .................................. 158

4.4. A Life-Course Perspective .................................................. 159

4.5. Measurement Practices ...................................................... 160

4.6. Consider Context ............................................................ 160

4.7. Summary ................................................................. 160

4.8. Conclusion ....................................................................... 162

Appendix I ................................................................. 164

1. Alternative Network Estimation ......................................... 164

Figure AI.1 ........................................................................... 165

2. Centrality Measures Derived from Alternative Estimation Methods 166
3. Robustness Analyses .........................................................................168

Figure AI.2 ..............................................................................................167

Figure AI.3 ..............................................................................................169

Figure AI.4 ..............................................................................................172

Appendix II .............................................................................................175

Table AII.1 ..............................................................................................175

Table AII.2 ..............................................................................................176

Figure AII.1 ..............................................................................................177

Figure AII.2 ..............................................................................................178

Appendix III ..........................................................................................179

1. Software and Library Versions ..........................................................179

2. Results Following Missing Data Imputation with MICE ......................180

Table AIII.1 ..............................................................................................181

3. Homogeneity of Variance ..................................................................182

Table AIII.2 ..............................................................................................183

4. Investigating Age Effects: Network Moderation Analysis ..................184

Figure AIII.1 ............................................................................................185

Figure AIII.2 ............................................................................................186

5. Network Stability ..............................................................................187

Figure AIII.3 ............................................................................................188

Figure AIII.4 ............................................................................................189

Figure AIII.5 ............................................................................................190
Figure AIII.6 ..............................................................................................................191
Figure AIII.7 ..............................................................................................................191
Figure AIII.8 ..............................................................................................................193
6. Network Comparison via Confirmatory Fit .........................................................194
Figure AIII.9 ..............................................................................................................195
Table AIII.3..............................................................................................................196
7. Joint Network Estimation with Fused Graphical Lasso...........................................197
Figure AIII.10 .........................................................................................................198

References ...............................................................................................................199
“Diversity is the best evidence that exists of the potential for change in the states and conditions of human life “

-Brim & Kagan, 1980

*Constancy and Change in Human Development*
Preface

This thesis was written amid a very active and arguably transformational period in developmental science related to the study of neurodiversity, which some have recently described as a *transdiagnostic revolution* (Astle et al., 2021). The starting point of this thesis coincided with the impetus of this movement, which shaped the empirical work presented in later chapters. The introductory chapter provides a brief overview of the historical backdrop and some of the conceptual shifts and methodological advances that have enabled this so-called *transdiagnostic revolution*. I outline the challenges of the earlier literature and provide recent examples of how transdiagnostic approaches have been used to address them. The empirical work presented in later chapters is part of the wider transdiagnostic movement and adds to a growing body of evidence supporting a shift away from diagnosis-centred approaches. It instead emphasises a more inclusive research agenda focused on the child and their individual strengths and weaknesses. I argue that in due time, this framework has the potential to guide the support required to enable all children to thrive.
CHAPTER I

Towards a Transdiagnostic Approach to Neurodevelopmental Difficulties

1. Understanding Neurodevelopmental Difficulties

Most individual differences in adult life arise from a developing system. Consistent with this, the origins of many adult-life difficulties are rooted in neurodevelopment. Developmental diversity inevitably results in some children struggling with school or aspects of social daily functioning more than others, and a minority of this struggling population receive a clinical diagnosis of a developmental disorder (Astle & Fletcher-Watson, 2020). Neurodevelopmental problems including difficulties with cognitive skills, academic achievement, behaviour, and socio-emotional functioning affect between 15% and 30% of school children (Department for Education, 2018; National Center for Education Statistics, 2021). Long-term outcomes for this heterogeneous group of children include low employment rates, increased risk of mental health difficulties and other adverse outcomes related to health, education, the legal system, personal finances, and social functioning (de Beer et al., 2014; Emerson & Hatton, 2007; Zendarski et al., 2022). The risk for negative adult outcomes is not only observed in groups that meet diagnostic criteria but also affects individuals experiencing subthreshold difficulties in early life (Copeland et al., 2015; Zendarski et al., 2022). Consequently, there is a widespread recognition that characterising developmental diversity is key to appreciating how the variation that exists in human brains and behaviour relates to risk and resilience for psychopathology and adverse outcomes. Building this understanding has fundamental implications for social, educational, and clinical policies and investment strategies.
1.1. Traditional Approach to Neurodevelopmental Difficulties

Historically, the prevailing approach to studying neurodevelopmental difficulties has been grounded in the nosology outlined in international diagnostic systems such as the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, American Psychiatric Association, 2013). The DSM-5 includes neurodevelopmental conditions such as specific learning disorders, autism spectrum disorder (hereafter ASD or autism), attention-deficit hyperactivity disorder (ADHD), and social pragmatic communication disorder (SPCD). Diagnostic manuals such as the DSM-5 have been very influential in shaping how we address and study neurodevelopmental diversity. Clinical professionals rely on its taxonomy to categorise difficulties and to make decisions about additional support and intervention strategies. In research settings, the nosology provides the foundation for much of our conceptual thinking, shaping every step of the scientific endeavour including hypotheses formulation, study design, recruitment, analytical approaches, and interpretation.

This is illustrated in the prevailing approach to studying neurodevelopmental difficulties, which involves recruiting highly selective groups of children who meet the criteria for a specific condition (e.g., dyslexia, ADHD) or who fall below a pre-defined age-expected threshold of ability level in the domain of interest (e.g., two standard deviations below the age-expected mean). These groups are typically compared to children who do not meet diagnostic criteria and whose abilities are within the age-expected range and/or groups matched on key characteristics such as age and IQ. Case-control designs usually rely on group-level statistics to test hypotheses about the neuropsychological origins of different neurodevelopmental and learning difficulties. Their results have provided foundation and support for the “core-deficit” hypotheses: the idea that a single mechanistic impairment can explain the profile of a particular diagnostic/struggling group. Examples include the core theory of mind deficit in autism (Baron-Cohen et al., 1985; Frith et al., 1991); the hypothesis that dyslexia is underpinned by a
primary deficit in phonological processing (Snowling, 1998); the central role of executive dysfunction in ADHD (Pennington & Ozonoff, 1996); and the notion that mathematical learning disabilities can be attributed to a principal problem with basic number processing skills (e.g., Mazzocco et al., 2011). The common feature across these theories is the assumption that there is a one-to-one mapping from a single restricted cognitive deficit (also referred to as an “impaired module”) to a specific diagnosis/area of difficulty. The simplicity of the ‘common underlying cause’ thinking affords advantages because it can easily be translated to intervention strategies (e.g., phonological interventions with struggling readers). However, a substantial body of evidence challenging these theories presents multiple conceptual problems that cannot be easily accommodated within a strictly categorical/core-deficit framework. Below I outline some of these key challenges, including the rampant heterogeneity within diagnostic groupings, evidence that common biopsychosocial processes shape outcomes and trajectories across diagnosed and undiagnosed populations, the considerable overlap across the difficulties experienced by children with different diagnostic labels, and the overly static and simplistic outlook on the relationship between cause and outcome.

1.2. Challenges to the Traditional Approach to Neurodevelopmental Difficulties

1.2.1. Heterogeneity

One challenge to the diagnostic framework and associated core-deficit accounts arises from studies demonstrating overwhelming heterogeneity within diagnostic/struggling groups. For example, not all children with reading difficulties have phonological processing problems (Snowling, 2008). Phonological processing difficulties are implicated in the application of letter to sound correspondences to decode words and are common in struggling readers. However, some poor readers do not have difficulties with decoding but instead struggle with reading comprehension, an aspect of reading closely supported by domain-general skills such as working memory (Cain et al., 2004; Nation, 1999). Similarly, a theory of mind deficit cannot
explain all difficulties experienced by autistic individuals, such as sensory sensitivities, consuming interests, or executive difficulties (Tager-Flusberg, 2007). For ADHD, it has been shown that not all children with the diagnosis have executive function impairments, and there is substantial variability among those who do (Kofler et al., 2019). In terms of mathematical difficulties, large-scale investigations fail to support the idea that mathematical problems can be explained by a core deficit with either domain-specific numerosity skills or domain-general cognitive skills (Mammarella et al., 2021). Overall, contemporary evidence does not support the proposition that a single core deficit can explain all patterns of difficulties within groups of children with the same label/difficulty (Peters & Ansari, 2019).

The examples outlined above suggest that children with the same diagnostic label can sometimes have different difficulties, casting doubt on the idea that diagnostic categories represent homogeneous groups, which can be used as a reliable guide to study recruitment or intervention strategies. Consistent with this, Bathelt et al. (2018) observed poor correspondence between diagnostic labels and children’s behavioural problems in domains such as inattention, hyperactivity, and aggression. Similarly, data-driven explorations of the links between cognitive and academic skills indicate that children can arrive at similar profiles of learning impairment through multiple etiological routes: children with comparable difficulties in both reading and mathematics can have different cognitive profiles, such as relatively more severe problems with phonological processing, spatial abilities, and/or working memory (Astle et al., 2019). These results support the idea that similar profiles of difficulties may stem from different causes (i.e., equifinality; Cicchetti & Rogosch, 1996) and similarly that the same underlying causes can have different consequences for functional outcomes across individuals (i.e., multifinality).
1.2.2. Non-specific Contributions of Biopsychosocial Processes

The empirical evidence against specific one-to-one mappings is supplemented by analysis at the level of genetic or neural variants and environmental influences. Attempts to link a specific diagnosis or behavioural profile/outcome to a single brain region or gene have failed to identify robust one-to-one causal pathways (Haworth et al., 2009; Peters et al., 2018).

The same genetic variants are implicated in a range of neurodevelopmental disorders (e.g., Price et al., 2020, Plomin & Kovacs, 2005) and non-specific heritability is evident in the observation that family members of individuals with ASD and ADHD have an increased risk for a range of other neurodevelopmental and psychiatric conditions (e.g., Ghirardi et al., 2021; Jokiranta-Olkonemi et al., 2016). At the neural level, the few studies that have compared groups with different diagnoses point towards substantial similarities across groups at both the structural and functional level (Bethlehem et al., 2017; Dajani et al., 2019; Moreau et al., 2018; Peters et al., 2018; Qian et al., 2021). Brain regions that have been implicated in multiple neurodevelopmental disorders include the precuneus (e.g., Di Martino et al., 2013), corpus callosum (e.g., Ameis et al., 2016; Aoki et al., 2017), anterior cingulate cortex (e.g., Langer et al., 2019), and cerebellum (Stoodley, 2014). In terms of environmental influences, risk factors such as premature birth, low birth weight, increased parental age, early medical events, and lower socioeconomic status have been associated with multiple neurodevelopmental disorders (e.g., Carlsson et al., 2021; Myers et al., 2021; although some of these influences may still share genetic variance; for discussion, see Taylor, 2021). Consequently, there is a growing consensus that the search for neural or genetic atypicalities or environmental risk factors specific to a given disorder/difficulty may not be fruitful.

1.2.3. Little Evidence for Qualitative Discontinuity

In fact, research across levels of analysis (e.g., genes, brain, and cognition) has consistently demonstrated that the risk factors implicated in neurodevelopmental conditions
are also related to normal variation in the relevant traits (Haworth et al., 2009; Peters & Ansari, 2019; Plomin & Kovas, 2005; Szűcs, 2016). In other words, there is little evidence for qualitative differences between low and high performers. Instead, individual differences are related to the same genetic, neural, cognitive, and environmental causal factors and their interactions, which influence functioning across the entire spectrum. This is supported by evidence from all levels. The genetic variants implicated in specific disorders such as dyslexia and mathematical learning problems are also linked to variation in the normal range of reading and arithmetic ability (e.g., Haworth et al., 2009). Similarly, the neural patterns linked to ASD are also dimensionally related to variance in social responsiveness and pragmatic communication (e.g., Aoki et al., 2017). At the cognitive level, meta-analyses have shown that individual differences in phonological processing relate not only to dyslexia but also to reading ability in unselected samples (Melby-Lervåg et al., 2012). Dimensional continuity is also favoured for both symptoms of inattention and hyperactivity/impulsivity that are more typically associated with ADHD (e.g., Frazier et al., 2007). Notably, the overwhelming support for qualitative continuity suggests that studying restricted groups defined as falling below an arbitrary threshold may be counterproductive, because arbitrarily specified cut-off points can obscure true associations between variables and may yield misleading results (MacCallum et al., 2002).

1.2.4. Comorbidity

Traditional taxonomic approaches are also challenged by evidence that there are remarkably high comorbidity rates between different diagnoses and difficulties. For instance, it has been shown that many autistic children meet the clinical criteria for an ADHD diagnosis, and vice-versa. This finding has been consistently reported in the literature (Ronald et al., 2008; Salazar et al., 2015), but the precise comorbidity estimates vary widely across studies, with some investigations reporting that up to 76% of youth diagnosed with ASD also meets the DSM
criteria for ADHD (Joshi et al., 2017). There is also substantial overlap between different learning difficulties: maths and reading problems co-occur in between 44% to 65% of cases (Barbaresi et al., 2005). Similarly, learning problems are common in ADHD, with co-occurring ADHD and reading difficulties estimates ranging between 18–45% (Germanò et al., 2010). In autism, learning difficulties have been reported in up to 67% of cases (Mayes & Calhoun, 2006). Generally, it seems that co-occurring neurodevelopmental difficulties are the rule rather than the exception.

1.2.5. From Single to Multiple Deficit Models

In line with the overwhelming evidence for comorbidities between different difficulties, there has been an increasing interest in the search for the cognitive factors that account for the overlap (e.g., Fias & Menon, 2013; Willcutt et al., 2019). ‘Additive models’ suggest that different areas of difficulties have distinct domain-specific origins and that comorbid presentations arise from the additive effect of two independent causes. In the case of reading and maths difficulties, for example, a phonological deficit underlies the reading problem and poor number processing skills cause the maths difficulty. The comorbidity between the two learning problems then results from the additive effect of two separate domain-specific skills being impaired within the same individual (Landerl et al., 2009; Willburger et al., 2008).

Another account, known as the ‘multiple-deficit model’, proposes that risk factors spanning several levels (genes, brain, cognition, behaviour, and environment) contribute probabilistically to neurodevelopmental difficulties, and that shared risk factors contribute to comorbidity (McGrath et al., 2020; Pennington, 2006). For example, verbal comprehension, phonological processing, and visual perception have all been found to contribute to the comorbidity between problems with reading and maths (Cheng et al., 2018; Peterson et al., 2017; Slot et al., 2016). Similarly, domain-general skills such as processing speed and working memory have been implicated in the comorbidity between learning difficulties and ADHD.
(Chu et al., 2016; Gooch et al., 2019; Kibby et al., 2021; McGrath et al., 2011). Generally, there is wider recognition that multiple cognitive factors need to be studied together. Nonetheless, the search for the common factors accounting for the overlap among disorders has been mixed and inconclusive and evidence about how different factors work in combination is still limited (Zoccolotti et al., 2020).

1.2.6. Circular Dynamics

A further challenge to both single and multiple deficit models is the relatively static outlook on the relationship between constructs. To take phonological processing and reading as an example, it is traditionally considered that good phonological awareness supports reading development. However, there is also substantial evidence that while early phonological skills contribute to reading development, they also benefit from reading instruction and exposure to text (Cunningham et al., 2021; Huettig et al., 2018; Nation & Hulme, 2011). Similar reciprocal effects, where the development of a given skill typically modelled as the outcome predicts growth in skills typically considered foundational and vice-versa, have also been reported in other domains; for example for working memory and reading/maths (Miller-Cotto & Byrnes, 2019). Bidirectional links are also commonplace outside the cognitive-academic realm and have been observed across domains such as inattention and emotion difficulties, and conduct problems and prosocial behaviour (Han et al., 2020; Speyer et al., 2021).

Although the evidence for bidirectional dynamics is relatively recent, the concept of bidirectionality has for decades been emphasised by (neuro)constructivist models of development (e.g., Karmiloff-Smith, 1998). Proponents of this framework reject the ‘core-deficit’ idea and the associated assumption that the mind is comprised of separate ‘modules’ dedicated to specific abilities/functions (Karmiloff-Smith, 2009; Westermann et al., 2007). Instead, (neuro)constructivism takes a much more dynamic outlook, whereby early atypicalities in one part of the system have ramifications for development in other parts of the
system. In contrast to the traditional view that causes operate in a single direction along the axis gene-brain-cognition-behavioural phenotype, (neuro)constructivism proposes that two-way interactions operate within and across these levels. Focusing on the neural level, it puts forward the principle of interactive specialisation suggesting that if an early developing region develops atypically, this is likely to have knock-on consequences for later developing regions (Johnson, 2011). Crucially, the specific developmental pathway triggered by the initial liability will not be set or static but will instead be an adaptive response informed by input from all levels (i.e., genetic, neural, environmental). This dynamic adaptive response can thereby explain the heterogeneity and the multiple different causal factors implicated in developmental disorders (Johnson et al., 2021).

While the (neuro)cognitivist framework has traditionally been studied at the neural level, the same principles apply behaviourally. If a child has a predisposition for language difficulties and their language develops atypically this is likely to trigger downstream consequences in other areas of functioning where language skills are important such as comprehending classroom instructions, learning from written material, or building friendships (Conti-Ramsden et al., 2009; Durkin & Conti-Ramsden, 2010). Difficulties in any of these areas can in turn contribute to anxiety, low mood, and/or disruptive behaviour, with the specific trajectories likely impacted by the child context and characteristics (Botting et al., 2016; Goh et al., 2021). Notably, some of the downstream consequences of the initial predisposition for a language difficulty may further impede language development, for example by limiting the child’s opportunities to enrich their verbal skills/vocabulary through written text or peer interactions. Given the dynamic nature of development, changes in any one domain of function can have varying system-wide reverberations (Thelen, 1989). This dynamic process of cascading effects can amplify initial difficulties to produce substantial heterogeneity in the emerging neurodevelopmental phenotypes. Under this conceptualisation, individual
differences arise from the action of a developing system, and in turn, can shape the system in manners that further elaborate diversity (Lerner et al., 2010). The dynamic interaction of systems across development is commonly referred to as developmental cascades, but terms such as amplification, snowballing, mutualism, and coupling are also used depending on the processes assumed to be involved (e.g., Cicchetti & Tucker, 1994; Deserno et al., 2020; Dodge & Pettit, 2003; Murray et al., 2020; Sameroff, 2000).

1.2.7. Developmental Cascades

Evidence for developmental cascades comes from longitudinal studies tracking the same individuals over long periods of time. For example, it has been shown that externalising behavioural problems in childhood undermine academic outcomes/development in adolescence, which subsequently contributes to internalising problems in young adulthood (Masten et al., 2005). In another investigation, early difficulties with hyperactivity/impulsivity have been observed to play a role in developmental cascades related to later socio-emotional functioning (Speyer et al., 2021). This may be because attentional difficulties cause children to miss social cues and/or because impulsivity makes it hard for them to respect others’ turns in social interactions. Consequently, children with high levels of inattention and/or hyperactivity often experience peer exclusion, dislike, and victimisation, which can, in turn, trigger further hyperactive and disruptive behaviours (Roy et al., 2015; Yu et al., 2018). Similar to academic difficulties, poor peer relationships have also been observed to partially mediate the increased longitudinal risk for mental health problems observed in this group of children (van Lier et al., 2012). Longitudinal studies also support the idea that individual trajectories are moderated by contexts and environmental characteristics, as proposed by the (neuro)constructivist framework. For instance, even after accounting for genetic risk, school characteristics still moderate the relationship between ADHD difficulties and academic achievement (Cheesman et al., 2021). The association is context-dependent, with the school-level variance being more
strongly predictive of children’s academic abilities at higher ADHD severity. Similarly, school-level factors also explain variance in student mental health (Ford et al., 2021).

1.3. Summary

Bidirectional influences and cross-domain developmental cascades likely contribute to neurodiversity in multiple ways. While researchers are beginning to characterise some of these processes, several methodological and analytical challenges related to the longstanding influence of core-deficit thinking have arguably impeded progress (Astle & Fletcher-Watson, 2020). For example, despite high comorbidity estimates across different diagnoses and difficulties, traditional case-control designs, which treat comorbidity as confound, remain the norm. It is typical for studies to screen out known co-occurring problems/disorders at the point of recruitment. The premise of this approach is grounded in one-to-one mapping thinking; it is favoured precisely because it enables researchers to focus on groups with isolated difficulties within a domain of interest. However, studying such pure presentations limits generalisability, because it fails to capture the substantial proportion of struggling children who experience multiple problems and/or whose difficulties may not fit traditional diagnostic labels (Astle et al., 2021; Holmes et al., 2020). Relatedly, behavioural and cognitive domains of measurement are often chosen based on the difficulty considered characteristic for the studied group (e.g., executive functions are primarily studied in ADHD, social communication in ASD, phonological difficulties in dyslexia). If other domains of function are included, they are often matched across groups or modelled as covariate. Both the overreliance on studying groups of children with specific difficulties and the choice of restricted sets of measures can bias the results towards interpretations in line with the one-to-one mapping/core-deficit perspectives (see Astle & Fletcher-Watson, 2020 for discussion). The elimination of alternative possibilities at the point of study conception and design inevitably perpetuates and reinforces core-deficit accounts. This poses a challenge to the pursuit of characterising how different risk and
protective factors work in combination to produce both positive and negative cascades, which would be of key practical significance for the remediation of neurodevelopmental difficulties in a much wider population. Finally, emerging evidence for bidirectional influences challenges unidirectional assumptions and analytical strategies in which cognitive skills are uniquely modelled as predictors and academic/behavioural difficulties as outcomes. It instead invites researchers to consider the dynamic nature of development.

2. Rethinking the Traditional Approach to Neurodiversity: Looking Beyond the Label

The empirical evidence presented in Section 1 suggests that developmental difficulties and their associated genetic, neural, cognitive, and environmental correlates may not be as specific and separable as previously assumed. Similar conclusions have been reached in the adult psychopathology literature, where there is converging evidence that diagnostic systems do not adequately represent individual variation in psychopathology (Cuthbert & Insel, 2013; Dalgleish et al., 2020; Insel, 2014; Kotov, et al., 2017). In this field, researchers are increasingly reconsidering the reliance on the diagnostic nosology in guiding research agendas and study designs. Most prominently, the National Institute of Mental Health has endorsed an alternative transdiagnostic approach, which is reflected in the Research Domain Criteria (RDoC; Insel, 2014). This framework shifts the focus away from discrete categorical entities with poorly specified boundaries and instead aims to characterise the factors that contribute to risk and resilience across individuals irrespective of diagnosis (Caspi & Moffitt, 2018). This approach embraces the individual variability within the population and avoids classification based on arbitrary cut-offs by drawing on substantial evidence that there is limited discontinuity in the distribution from struggling to thriving. This framework has since been extended to neurodevelopmental diversity (Astle et al., 2021; Casey et al., 2014; Coghill & Sonuga-Barke, 2012; Fonagy & Luyten, 2018; Holmes et al., 2019; Sonuga-Barke et al., 2016 - note that the majority of these applications occurred during this PhD). To study neurodevelopment
transdiagnostically we need fundamentally different approaches to research design, analyses, and interpretation. In the section below I review opportunities raised by the transdiagnostic approach and some of its remaining challenges.

2.1. New Recruitment Strategies

To adopt a truly transdiagnostic approach we need to rethink our reliance on the case-control design. One strategy is to recruit diverse samples based on functionally defined needs rather than arbitrary thresholds or pure presentations of a restricted problem (Astle et al., 2021; Holmes et al., 2019). This could mean including children with different or multiple diagnoses or children whose difficulties do not fit the current diagnostic classification. Needs-based recruitment is much more inclusive, does not artificially inflate the homogeneity within groups and thereby has the potential to capture the true diversity within the population. This approach can result in samples that are more representative of the large group of struggling children routinely presenting at health and education services, which can increase the practical value and generalisability of the findings. Alternatively, sampling can be targeted to capture the full spectrum of abilities along dimensions of interest, for example through community sampling or demographically representative cohort studies.

One example of an early study adopting functionally defined recruitment explored the language profiles of children attending specialist language units in England (Conti-Ramsden & Botting, 1999). These units support children for whom language difficulties represent a barrier to mainstream education irrespective of diagnostic status, and therefore provide a window into the diversity of children’s language needs. Six subgroups were identified within this cohort, characterised by different combinations of speech, phonological processing, syntactic and pragmatic difficulties (Conti-Ramsden et al., 1997), thereby providing early evidence of the heterogeneity within language development.
2.2. Diverse Measurement Practices

Another important feature key for the success of transdiagnostic designs relates to the need to consider multiple facets simultaneously, including their potential interactions (Boulton et al., 2021). The focus should be on understanding variation across domains of functioning related to risk and resilience. This means that large samples of children should be studied, and their profiles should be characterised across different areas including academic and socio-emotional functioning, as well as domain-specific and domain-general cognitive performance. Assessments of genetic, environmental, and neural variation should also be included where feasible, and children should be followed longitudinally to enable a better understanding of the developmental dynamics across time and levels of description. In the following section, I outline a study that adopted a functionally defined recruitment strategy, used multiple assessments across different levels, and included a longitudinal phase. I have been heavily involved in this project and use the resulting cohort data throughout the empirical chapters in this thesis. For this reason, I provide a detailed description of the recruitment and assessment details for the cohort here.

2.2.1. The Centre for Attention, Learning, and Memory (CALM)

The Centre for Attention, Learning, and Memory (CALM, Holmes et al., 2019) provides a recent example of the use of a functionally defined recruitment strategy with a broader focus across multiple domains of function. The Centre was established in 2014 to collect data from a large heterogeneous cohort of struggling children (N = 805). Children were recruited into the study in partnership with local health and educational practitioners who referred children to the study when they considered that the child had difficulties in the areas of attention, learning, and/or memory. The referring professionals included clinical psychologists, special educational needs coordinators, child psychiatrists, speech and language therapists and related professionals. The children were enrolled into the study irrespective of
diagnostic status, provided they met the inclusion criteria of (1) being a native English speaker; (2) not having a known genetic condition that affects cognition or (3) uncorrected sensory impairment. The study was restricted to school-age children and adolescents (i.e., 5-18 years old). The recruitment strategy captured children with both undiagnosed and diagnosed needs, including children with multiple diagnosed difficulties.

Children completed extensive neuropsychological assessments capturing multiple cognitive domains that have been implicated in learning and neurodevelopmental disorders (e.g., working memory, sustained attention, executive functioning, phonological skills, nonverbal reasoning, processing speed). Standardised assessments tapping reading, spelling, and maths were also included in the testing protocol. Each child was individually tested by a researcher in a quiet room while parents/carers completed questionnaires capturing aspects of daily functioning relating to executive function, communication, socio-emotional needs, and mental health. Participants were invited to participate in two optional aspects of the study; to provide a saliva sample for genetic analysis and to participate in a magnetic resonance neuroimaging session so structural and functional imaging data could be acquired. A smaller comparison group of children, who were not referred for difficulties, was tested following the same protocol ($N = 158$). To ensure that the two groups were demographically matched, the comparison sample was recruited from the same schools attended by the children in the struggling group. The longitudinal follow-up of the cohort, in which each child is re-tested approximately 5 years after their initial assessments, is currently ongoing. Further details on the study recruitment, cohort profile, and testing procedures are available in the study's protocol (Holmes et al., 2019).

The CALM cohort provides a novel way to enhance our understanding of the common and interrelated needs of children who struggle neurodevelopmentally. It applies the multi-systems approach advocated by RDoC and provides a rich resource for both exploratory and
theory-driven investigations of neurodevelopmental difficulties. The data from this unique cohort is therefore well-suited to the empirical work conducted within this thesis. Before I present the objectives of these investigations, I turn to another critical issue for contemporary studies of neurodevelopment, which relates to the analytical strategies used in the field and how they can be adapted to address the conceptual challenges outlined in previous sections.

2.3 Analytical Strategies

2.3.1. Latent Variable Approaches

Much of the literature generated by the standard one-to-one mapping perspective has relied on univariate statistics (Astle & Fletcher-Watson, 2020). Univariate approaches align with the categorical perspective in which performance between groups can be compared to identify deficits characteristic to a specific group. However, to study the complexity of neurodevelopmental difficulties we need alternative statistical tools that are better suited to describe the inter-related and multifaceted nature of neurodevelopment. The RDoC approach advocates for a focus on dimensional analyses, which can capture shared variance across multiple measures, arguably resulting from common underlying latent constructs (Caspi et al., 2014; Kotov et al., 2017). An intuitive solution might be to formulate discrete disorders as dimensional scales (e.g., ADHD scale, autism scale, etc.). This approach has for example been used to identify a higher-order latent ADHD dimension, which could be separated into two lower-order dimensions capturing hyperactivity and inattention (Martel et al., 2010). However, such attempts remain restricted to the characteristics thought relevant to conventional disorder categories and are unlikely to serve the purpose of illuminating how difficulties cascade and overlap across domains and levels of description.

A more useful approach might be to delineate relevant dimensions through hypothesis-free data-driven analyses of broad sets of measures (Astle et al., 2021; Holmes et al., 2020, 2021). Data-driven dimensional approaches afford a focus on multiple overlapping continuities.
rather than singular deficits or pre-existing classifications of disorders. Through such techniques, the dimensional characteristics of the data can be mapped and the inter-relationships between several dimensions understood. Individuals can then be located along the continuous distribution of each studied dimension. This approach enables researchers to study the full breadth of interindividual variability, which may be more insightful than simply observing isolated difficulties. Following a dimensional approach can also help researchers to identify characteristics linked to multiple neurodevelopmental outcomes, and discover how broad dimensions from different levels (e.g., behavioural, cognitive, and environmental) relate to one another and whether they can be linked to genetic or neural patterns of variance (Astle et al., 2021; Peters & Ansari, 2019).

The dimensional approach does not come without challenges. It is increasingly recognised that underlying dimensions do not easily generalise across samples, measurement instances, or types of assessments. This lack of generalisability stems from the fact that the resulting dimensions, their labels, and interrelationships will depend on the type and quality of the data that one is analysing. The available data will dictate the number and specificity of the latent dimensions, such that with increasing numbers of measures broad dimensions could be split into more specific components. For example, language ability could be decomposed into more specific dimensions of oral, written, expressive, receptive, and pragmatic language ability with increased measures (e.g., Geurts et al., 2009; Lonigan & Milburn, 2017; Ramus et al., 2013). More complex relationships such as hierarchical structures are also likely and require more advanced dimensional approaches than those typically adopted in the field (see the following for examples of these approaches, Holmes et al., 2021; Michelini et al., 2019). Unintended sources of common variance, such as issues around common reporter bias (e.g., Boer et al., 2002; Bornstein et al., 2006), the type of measurement (e.g., speed or accuracy; performance-based or rating-based; (Toplak et al., 2013), or how ‘purely’ assessments measure
the intended construct, can influence the observed dimensions and potentially hamper the interpretation and generalisability of the results (Friedman & Miyake, 2017; van der Sluis et al., 2007). The relationships between specific measures and latent dimensions can also change developmentally, such that tasks can tap different constructs (or subcomponents of the same construct) with increasing age (e.g., Huizinga et al., 2006; Lee & Bull, 2016). Additionally, potentially important insights about mechanisms or developmental cascades may be missed when the focus is on latent variables and their interrelationships. In some cases, specific relationships between the constituent parts (i.e., observed measures) of each dimension may provide additional important information beyond the latent level associations alone. Finally, the assumption that a latent construct causes the covariance between a set of measures may not always be warranted (e.g., Van Der Maas et al., 2017).

2.3.2. Network Approaches

An alternative approach that has gained considerable attention both in the adult psychopathology and cognitive development literature focuses on the interrelationships between relevant features (e.g., symptoms of mental disorders or cognitive skills) without assuming latent dimensions (Borsboom, 2016; Cramer et al., 2016; McNally et al., 2015; Van Der Maas et al., 2006). The traditional view that the relationships between features are manifestations of underlying constructs is replaced with the proposition that the covariance between features emerges because they dynamically interact to amplify and reinforce one another. In the adult psychopathology literature, this idea provides the basis of the network theory of psychopathology (Borsboom, 2016). Contrary to the medical model of psychopathology, which assumes that symptoms are consequences of an underlying common cause (e.g., major depressive disorder), network theory proposes that symptoms co-occur owing to causal interactions between the symptoms themselves. Causal interactions between symptoms are assumed to act as recursive loops that can give rise to a specific disorder (e.g.,
insomnia leads to increased fatigue; compulsive rituals alleviate the distress of obsessions; avoidant behaviours are the consequence of recurrent panic attacks, etc.). The precise chain of activation across symptoms and the specific maintenance loops between them, may vary across individuals and different circumstances. For example, adverse life events may trigger depressed mood, which can activate rumination or insomnia, which can in turn cause concentration problems or fatigue. Such dynamics of symptom-symptom causation may ultimately culminate into a full-blown major depressive episode (Borsboom et al., 2021).

This conceptualisation of mental disorders accounts for comorbidities through so-called *bridge symptoms*: symptoms that can activate another symptom(s) network (Blanken et al., 2018). This term is used to capture the possibility that a symptom considered characteristic for a specific condition can drive or activate the development of symptoms or problems in other domains. While specific networks (or clusters within them) might be given labels such as depression or anxiety, the central assumption of this framework is that all mental health domains are interconnected. In this way, the network theory of psychopathology overcomes the unsupported idea that disorders are distinct entities with separate causes and instead conceptualises them as emergent properties of dynamic systems.

A conceptually similar idea in the cognitive development literature is represented by the theory of mutualism (Peng & Kievit, 2020; Van Der Maas et al., 2006). In brief, the mutualism hypothesis suggests that different abilities bidirectionally interact across the course of development to reinforce one another. In this way, mutualism provides an alternative explanation of the positive manifold (i.e., Spearman's original observation that various cognitive abilities correlate positively with another; Spearman, 1904). Compared to traditional models assuming that the positive manifold can be explained by unobserved latent variables such as the general intelligence factor (i.e., ‘g’), mutualism proposes that cognitive abilities are instead directly related to each other through dynamic reciprocal casual interactions (Van Der
Maas et al., 2006). In this conceptualisation, the growth of a given cognitive skill is not only governed by its own capacity limitations but it is also affected by the levels of other cognitive abilities. This idea is analogous to the basic principles of (neuro)constructivism, which also emphasise dynamic reciprocal causation, whereby small perturbations in early development can carry downstream consequences (Karmiloff-Smith, 2009). In fact, simulations of mutualistic models support this position and have demonstrated that disruptions to this dynamic system can amplify over time to create substantial differences in phenotypes (Savi et al., 2019, 2021; Van Der Maas et al., 2006, 2017).

Due to the shared goal of conceptualising phenomena of interest as dynamic causal systems, the network theory of psychopathology and the mutualism model commonly rely on the same analytical tool, known as psychometric network analysis. Applied to cross-sectional datasets, network analysis focuses on the patterns of pairwise conditional dependencies across variables (for an overview, see Borsboom et al., 2021). The technique is particularly well-suited for capturing multivariate patterns of dependency, doing so by representing all measures of interest (e.g., symptoms, cognitive tasks) as network nodes. The links between each pair of nodes (i.e., network edges) are then drawn to represent the strength of the partial correlation coefficients between them that remain after taking into account all other measures included in the network. In the resulting network, a conditional association between two nodes (i.e., measures) holds when these variables are probabilistically dependent, conditional on all other variables included in the network. Conversely, if the association between two measures can be explained by other variables in the network (i.e., an association between A and B is not supported when other variables are controlled for), then the corresponding network nodes will be disconnected. In this way, network analysis offers a visualisation tool for patterns of statistical association and can thereby be used to generate hypotheses about potential causal dynamics and/or developmental cascades.
Notably, while this method is traditionally associated with network and mutualism theories, it can facilitate important insights regardless of the researcher’s assumptions about the generating process behind the observed relationships (e.g., multiple or single latent variables vs. dynamic causal systems). In the context of neurodevelopment, it provides a powerful technique for exploring multivariate patterns of association within and across domains. For instance, network analysis has been used to demonstrate that a variety of cognitive, demographic, and environmental factors relate to reading development and to highlight that their contributions are associated with different aspects of the reading process (Verwimp et al., 2021). This example demonstrates how network analysis can be useful in evaluating where specific features sit within the network structure. Additionally, tools from network science can be applied to examine the topological properties of the network and the attributes of specific nodes (i.e., features of interest such as cognitive or behavioural difficulties). For example, nodal strength centrality reflects how strongly a given network feature is conditionally associated with the other features included in the model; and nodal predictability quantifies how much of a given node’s variance is explained by connected nodes. Further, clustering algorithms (also referred to as community detection algorithms) can be applied to the network structure to explore whether certain nodes cluster in communities (i.e., groups of nodes, which have stronger connections to one another relative to their connections to the rest of the network nodes). The presence of communities may provide clues about the dimensionality of the system and any potential underlying unobserved causes. Finally, while network analysis is typically used in an entirely exploratory and data-driven fashion, confirmatory models can also be specified and investigated as a tool to compare competing theoretical ideas (Kan et al., 2019).

Network models can capture the inherent complexity of the data, which may be particularly informative when trying to understand fine-grained patterns of conditional
dependence between symptoms or difficulties, or when trying to map developmental cascades. Nonetheless, many of the challenges related to dimensional methods are also applicable to network analysis. This includes difficulties in specifying what constitutes a system and relatedly in determining which elements belong to a system (Fried & Cramer, 2017). It is at the discretion of the researcher to select relevant variables to include and to evaluate the extent to which these variables capture the phenomena of interest. Issues around the type and quality of data and potential unwanted sources of common variance can also affect the results and interpretations of network models. Finally, minority subpopulations, which may be best characterised by different intervariable relationships can be masked within the broader sample. An alternative complementary analytical approach often described as clustering or subgrouping can address the latter issue. It should be noted that although clustering is described as an alternative approach in the context of identifying clusters of children/individuals, clustering techniques can be applied to networks of measured constructs to identify clusters of variables, as discussed briefly above.

2.3.3. Clustering Approaches

Clustering or subgrouping approaches, in the context of neurodevelopment, primarily refer to identifying data-driven groupings of children with similar profiles. This technique explores how different dimensions or observed variables intersect within groups of individuals. In the context of neurodevelopmental diversity, clustering algorithms can be used to search for participant groupings that have similar profiles across a constellation of relevant features (e.g., cognitive, behavioural, environmental, or a combination). These groupings can then be used to study the variability in the pathways that lead to specific profiles/trajectories and may be key for understanding how various genetic and neural liabilities contribute to heterogeneity. Identifying data-driven groups can also assist in the stratification of individuals to suitable support strategies.
Cluster analysis is a highly versatile tool that can be applied to different types of data to answer various questions related to neurodevelopment. For example, clustering techniques can be deployed within a diagnostic group to identify different presentation profiles within the same disorder. Such approaches are often described as subtyping. For instance, Karalunas et al. (2014) looked at temperament dimensions in a sample of children diagnosed with ADHD and applied a subtyping approach based on community detection. They identified three ADHD subtypes, which did not align with DSM-5 ADHD presentation specifications (i.e., inattentive vs. hyperactive/impulsive vs. combined). The identified subtypes were stable over time and could be distinguished by their cardiac physiological responses, resting-state functional brain connectivity, and clinical outcomes at follow-up (Karalunas et al., 2014). Similarly, Stevens et al. (2019) searched for subtypes in a large sample of children diagnosed with ASD using data from several domains including social, language, executive functioning, motor skills, and play. They identified five subtypes, which were related to children’s responses to an applied behavioural analysis intervention.

Clustering techniques can similarly be used in samples of children without diagnoses or with different diagnostic labels. In such cases, the identified groupings are usually referred to as subgroups or clusters. The goal of such clustering approaches may be to test whether the identified subgroups align with diagnostic categories, and/or to explore whether subgroups can parse out heterogeneity and facilitate the discovery mechanisms (which could be both unique to each subgroup or shared across several). For example, Cordova et al. (2020) looked at executive functioning in a group of children diagnosed with ADHD and/or ASD. They proposed that if diagnoses aligned with the data-driven subgroups derived based on executive function skills, there would be evidence that the mechanisms leading to different conditions are potentially distinct. Inversely, they hypothesised that if there was no agreement between subgroups and diagnoses, this could indicate that underlying executive functioning
mechanisms are likely to be shared across conditions. Their results supported the latter; they identified multiple mechanistic subgroups, which did not correspond to diagnoses. It was also observed that different combinations of executive function difficulties led to similar symptom presentations, suggesting that there are several shared executive function-related mechanisms implicated in ASD and ADHD.

Other studies that have also identified subgroups of children based on variation in executive functioning skills have similarly demonstrated that data-driven groups show links to both shared and specific neurobiological mechanisms (e.g., Jones et al., 2021). Furthermore, relative to diagnostic categories, subgroups identified by executive function profiles have more similarities in their symptom profiles and these groupings are more reliable predictors of patterns of brain function (Bathelt et al., 2018; Vaidya et al., 2020). Overall, these subtyping/subgrouping examples demonstrate the potential of clustering approaches to link behavioural phenotypes to relevant biological mechanisms, which may aid clinical decision-making about appropriate interventions and likely prognoses.

The application of clustering methods may also be specifically targeted towards identifying subgroups that are likely to respond to a specific intervention (Fletcher & Vaughn, 2009). For example, subgrouping methods have been used to establish which participants are likely to benefit from cognitive training (Rennie et al., 2020). In a similar vein, subgroup characterises can be used to predict whether children with developmental language disorder are likely to respond to an intervention targeting phonology or semantics (Best et al., 2021). In this context, clustering methods provide an opportunity to use response to intervention as a tool for understanding heterogeneity and/or informing practitioner choices for support strategies (Catts & Petscher, 2021). Within a response to intervention framework, these methods allow focusing on the strengths of subgroups, rather than their difficulties (Burns et al., 2016). This approach can accommodate compensatory accounts stressing the possibility that children might
compensate weaknesses with their strengths, which could lead to learning gains in the affected domains. Characterising groups in terms of both advantages and disadvantages may provide not only clues about where their needs are but further what tools could be used to target them (Astle & Fletcher-Watson, 2020). An example of how considering both strengths and weaknesses can inform interventions is the practice of using image-based rather than written materials for children who struggle with reading (Mortimore & Dupree, 2008).

Understanding who is likely to respond to a given intervention may carry substantial practical value, but notably, evidence for reliable improvements within a group following treatment does not necessarily imply that their problems have the same origins (e.g., to use a medical analogy, ibuprofen can alleviate headaches caused by a variety of factors). Consequently, clustering approaches targeting response to intervention and those aiming to elucidate underlying liabilities may not always converge on the same subgroupings. Nonetheless, both approaches are likely to facilitate important insights and offer a useful tool for characterising heterogeneity.

Finally, it is important to recognise the limitations of clustering methods. Akin to diagnostic approaches, most clustering algorithms assign each child to a category, assuming set boundaries between subgroups. The applied boundaries are based on the data, and therefore the quality and type of data are key for the success of this approach. Nonetheless, even with high-quality data and large sample sizes, borderline participants who do not fit the identified categories remain likely and represent a challenge for such methods.

2.3.4. Combining Methods

Clustering approaches offer a tool for starting to tease apart some of the heterogeneity in neurodevelopment. Once clusters are identified, a crucial next step is understanding the mechanisms that lead to the phenotypical differences defining the subtypes. Here, I use the word mechanism in a very broad sense to refer to the dynamic and interactive processes within
and across levels that drive and shape the path from an early liability and/or risk to a later phenotype (Benton, 2021; Karmiloff-Smith, 2009; Oakes & Rakison, 2020). Under this conceptualisation, no one specific aspect is considered the mechanism; rather, the entire dynamic and active process consisting of multiple steps from a start state to an outcome state is seen as the mechanism (Benton, 2021). One possibility illustrated in the examples discussed above is to focus on the physiological features associated with these data-driven subgroups. Aetiological liabilities likely set the path for developmental cascades, but these cascades can be shaped by other child characteristics and aspects of their environment. Therefore, it is important to also study the multivariate patterns of associations within subgroups to explore hypotheses about whether and how specific factors might moderate developmental cascades. This could include potentially leveraging complementary analytical strategies such as network analysis or dimensional models.

Overall, data-driven dimensions, networks, and clusters, when used appropriately, provide tools to advance our understanding of neurodevelopmental diversity beyond diagnostic categorisations and core-deficit accounts. The call for more complex systems thinking in developmental psychology requires us to conceptualise developmental outcomes as the result of cross-domain interplay over time: early liabilities interact with the environment, giving rise to both difficulties and compensatory adaptations, which could in turn trigger negative and/or positive developmental cascades with ramifications in later life. In other words, variation at any level of organisation is expected to be structured in relation to variance at other levels, such that the development of any feature is shaped by its relations to features both within and across levels (G. H. Lerner, 1991; R. M. Lerner et al., 2006). The fundamental difference that separates this complex systems framework from traditional approaches is that the focus shifts from unilevel units of analyses (such as the presence/absence of difficulty) to making relationships themselves the focus of the analysis. The ultimate goal should therefore be to model individual
differences in multivariate space. To meet this goal, different analytical approaches such as latent dimensions, network models, and clustering techniques should be seen as complementary tools, which can elucidate the mechanisms involved in neurodevelopmental processes. Integrating the insights afforded by these approaches is likely to be the key to characterising complexity.

3. Summary

In the preceding sections, I have outlined several challenges associated with the diagnostic categorical approach for understanding neurodevelopmental difficulties, including between category overlap, substantial within-diagnosis heterogeneity, and the unsupported assumptions of qualitative discontinuity and one-to-one mappings. Subsequently, I discussed some of the methods and analytical tools that have been devised to help address these limitations including dimensional, network, and clustering approaches. This provides the backdrop for the empirical work reported in subsequent chapters. All three empirical chapters adopt a transdiagnostic approach to neurodevelopment and leverage recent methodological advances, triangulating across them to gain complementarity insights about the multifaceted character of neurodevelopmental diversity.

4. Thesis objectives and outline

The overarching aim of the current thesis is to attempt to advance our understanding of the complex nature of neurodiversity by investigating transdiagnostic associations between symptoms/difficulties across different domains of function. The CALM cohort described in Section 2.2.1 was the focus of all three empirical chapters. In Chapter II, I explored the interrelationships between language, communication, cognitive and behavioural difficulties using network science. I interrogated direct and indirect patterns of associations (i.e., conditional dependencies) across these commonly comorbid areas of difficulties. In Chapter III, I used a subgrouping approach to explore the heterogeneity in children’s relative strengths
and weaknesses across the domains identified in Chapter II and how this variation relates to socio-emotional functioning, academic skills, and neural white matter organisation. In the third empirical chapter, Chapter IV, I focused on cognitive and academic skills. Network analyses were used to model the patterns of interrelationships between cognitive and academic skills observed in the CALM cohort and to compare them to the patterns observed in a community sample.

The rationale for much of the empirical work in this thesis is covered in this introductory chapter. To aid thesis cohesion, the short introductions to each chapter reiterate some of the key arguments presented here, focusing on examples relevant to the specific chapter. For all empirical chapters, all statistical analyses are reported within the respective chapter apart from analyses testing the sensitivity/robustness of the methods (e.g., evaluating the influence of data transformations, approaches for handling missing data, etc.). These analyses are instead presented as an associated appendix for each chapter. In cases where these supplementary analyses produced a different result to the one presented in the chapter; the discrepancies are discussed within the chapter text.

The final chapter, Chapter V, provides a summary of the main insights from each empirical investigation and attempts to integrate them. It discusses both the strengths and limitations of the work presented in the empirical chapters and highlights potential implications and future directions. Collectively, the studies presented in this thesis demonstrate the value of studying neurodevelopmental diversity transdiagnostically. The thesis concludes with a reflection on the crucial next steps that need to be taken for the field to progress towards an understanding of how multiple developmental factors interact to shape individual outcomes and trajectories.
CHAPTER II

Transdiagnostic Associations across Communication, Cognitive, and Behavioural Problems: A Network Approach

Chapter II has been published, but it appears modified to ensure coherence:


First-pronouns are used throughout the chapter, but the work is the result of a collaborative research project. My contributions to the publication are:

1. I led the conceptualisation of the study research questions, methods, analyses and interpretation supported by J. Holmes.
2. I contributed to the data collection, scoring and curation, which was done in a collaborative effort including a large team of researchers (the CALM team).
3. I performed all statistical analyses.
4. I wrote the manuscript draft and led revisions aided by J. Holmes.
1. Introduction

Difficulties with social communication, executive functions (EF), and behaviour are common across a range of neurodevelopmental conditions. Pragmatic or social communication impairments are among the diagnostic criteria for ASD and social (pragmatic) communication disorder (SPCD); difficulties with attention and/or behaviour are characteristic of ADHD; and structural language problems are the hallmark of developmental language disorder (DLD, American Psychiatric Association, 2013). Each set of symptoms is commonly associated with a specific developmental disorder and is consequently typically studied in groups of children in whom such problems are characteristic. For example, behavioural problems such as hyperactivity are typically studied in children with ADHD, while communication difficulties such as poor speech are often studied in children with DLD.

However, these difficulties commonly co-occur across different diagnoses. There is substantial evidence that children with ADHD often have pragmatic language difficulties (Green et al., 2014), and structural communication problems are common in both children with ADHD (Korrel et al., 2017) and ASD (Mandy et al., 2017). Similarly, behavioural and attentional difficulties occur in both ASD and DLD (Henry et al., 2012; Rommelse et al., 2011). There are practical advantages to this categorical approach. It defines clear symptom-based criteria to inform practitioner decision-making about diagnoses and interventions. However, it fails to accommodate high rates of comorbidity across developmental disorders (Coghill & Sonuga-Barke, 2012; Faraone et al., 1998; Kessler et al., 2006; Kotov et al., 2017; Willcutt & Pennington, 2000) and substantial heterogeneity within disorders (Duinmeijer et al., 2012; Germanò et al., 2010; Rommelse et al., 2011; Willcutt & Pennington, 2000; see Chapter I, Section 1.2 for a discussion). The current study aimed to move away from studying links between problems of behaviour, communication, and cognition in discrete groups. Instead,
Symptom associations were explored in the CALM cohort presented in Chapter I, which is a large heterogeneous sample of children with clinical and subclinical levels of difficulties.

As discussed in Chapter I, Section 2, over the past decade there has been a broad shift away from diagnosis-specific deficits toward identifying dimensions that cut across disorders conventionally considered to be distinct (Cuthbert & Insel, 2013; Owen, 2014). This approach has been applied most widely to adult psychiatric conditions (R.-Mercier et al., 2018; Reininghaus et al., 2019; Titov et al., 2011), but there is widespread recognition of its value for characterising developmental disorders in terms of underlying dimensions of symptoms (Casey et al., 2014; Snowling & Hulme, 2012; Sonuga-Barke et al., 2016b; Sonuga-Barke & Coghill, 2014; Zhao & Castellanos, 2016). One of the most common methods for understanding how symptoms are related uses latent variable models, a statistical method that groups variables based on shared variance to derive underlying dimensions of difficulties (Kotov et al., 2017; see Chapter I, Section 2.3.1 for a detailed description of this approach) This technique has been used to identify dimensions of phonological and non-phonological skills in children with diagnosed DLD and dyslexia (Ramus et al., 2013); and separate latent constructs for inattention and hyperactivity in children with ADHD (Martel et al., 2010).

Network analysis offers an alternative approach to understanding symptom interrelations. Instead of identifying underlying dimensions, network models focus on symptom-level associations, allowing for the possibility that symptoms could interact to causally affect and activate one another (Cramer et al., 2010; Kendler et al., 2011; see Chapter I, Section 2.3.2 for a detailed description of network approaches and their applications). This framework is suited to conceptualising and evaluating the potential origins of comorbidities. For example, pragmatic difficulties characteristic of ASD, and hyperactivity problems common in ADHD often co-occur (Henry et al., 2012; Leonard et al., 2011; Rommelse et al., 2011; Willcutt & Pennington, 2000). From a network model perspective, this co-occurrence can be
conceptualised as arising in a dynamic system in which symptoms traditionally linked with one developmental disorder might trigger and/or maintain symptoms commonly associated with a different disorder.

In the current study network science is used to understand comorbidities between problems with communication, EFs, and behaviour. As previously stated, symptoms of communication and behavioural problems co-occur in the general population and in children with developmental disorders such as ADHD (Henry et al., 2012; Korrel et al., 2017; Rommelse et al., 2011; Willcutt & Pennington, 2000), ASD, DLD, and reading difficulties (Henry et al., 2012; Korrel et al., 2017; Snowling & Hulme, 2012; Yeniad et al., 2013). Additional deficits in EFs, the cognitive abilities regulating thoughts and behaviour (Castellanos et al., 2006), are often also present in individuals with communication and/or behavioural problems. EFs can be broadly construed as two distinct but related neurodevelopmental systems. Cool cognitive-based EFs encompass working memory, planning, and cognitive inhibition, and are associated with academic learning (Yeniad et al., 2013) and attention (Holmes et al., 2014); hot executive processes are associated with stronger affective valence and involve the regulation of emotional responses and social awareness (Castellanos et al., 2006).

It is not clear why communication, cognitive, and behavioural problems co-occur in childhood. These areas of difficulty likely share common environmental and genetic influences. Nonetheless, there are several possibilities about how they may interact dynamically. One is that deficits in cool EFs might underlie both behavioural and communication difficulties. Consistent with this, poor working memory has been shown to underpin problems in attention, behaviour, and structural language (Green et al., 2014; Holmes et al., 2014), and also accounts for the relationship between hyperactivity and pragmatic communication problems (Nilsen et al., 2015). Alternatively, associations between cool EFs and pragmatics could be mediated by
difficulties with hot EFs. Difficulties in inhibitory control could lead to hyperactive-impulsive behaviour and consequently poor social communication skills. Children who regularly display difficult behaviour may have limited opportunities to socialise, and thereby fail to develop good communication skills. In line with this, hyperactive-impulsive behaviours have been shown to account for the relationship between inhibition and the ability to apply pragmatic rules in everyday situations (Rints et al., 2014). Another possibility is that language deficits may directly and/or indirectly impact the ability to regulate cognition, behaviour, and communication (Botting et al., 2017; Petersen et al., 2015). For example, difficulties with language may lead to peer rejection and academic difficulties, which in turn may trigger behavioural problems.

The current study uses network modelling to estimate associations and conditional dependencies across symptoms. A community detection algorithm is then used to identify closely inter-connected symptoms that may, or may not, correspond to areas of difficulty commonly associated with categorical diagnoses such as ADHD or DLD. This novel approach provides the opportunity to identify where symptoms sit in the network (i.e., which symptoms sit together), and to quantify the importance of different symptoms within the network. This includes identifying bridging symptoms that have multiple strong links across domains. Symptoms linking different domains may reflect causal processes and/or shared aetiological influences. The conceptual interpretation of bridging symptoms identified in cross-sectional networks is not straightforward but one possibility is that such symptoms may spread activation across the system and that they may be the origins of the comorbidities observed on a dimensional level (Blanken et al., 2018; see Chapter I, Section 2.3.2, for more information). All analyses are cross-sectional, exploratory, and data-driven. However, based on reports of co-morbid symptom presentations in the literature, it is predicted that multiple direct associations will emerge across pragmatic and behavioural difficulties (hot EFs), with
additional connections between everyday cognitive abilities (cool EFs) and structural language skills.

2. Method

2.1. Recruitment

The data included are those collected between 2014 and 2018 at the Centre for Attention, Learning, and Memory (CALM). The CALM cohort is already described in Chapter I (Section 2.2.1.) and the full recruitment and testing procedures are described in the study protocol paper (Holmes et al., 2019). Details about the assessments included in the current analysis are presented below.

2.2. Assessments

The analysis presented in this chapter is based on the parent ratings of behaviour and communication described below. These checklists are routinely administered in health and educational settings in the United Kingdom to capture the child’s natural behaviour and communication in day-to-day situations when there is no external direction or regulation. Notably, rating-based measures are empirically separate from performance-based assessments of the same constructs which instead focus on efficiency of processing under structured and controlled conditions (Toplak et al., 2013). Raw scores are used across all analyses. Analyses based on age-referenced norm scores retained substantial similarity with the results presented here and are reported in Appendix I, Figure AI.1.

2.2.1. Brief Rating Inventory of Executive Function (BRIEF)


2.2.2. Conners-3
The Conners Parent Rating Short Form 3rd Edition (Conners, 2008) assesses ADHD-related difficulties across six subscales: Inattention, Hyperactivity/Impulsivity, Learning problems, Executive functioning, Aggression, and Peer relations. The Executive functioning subscale was dropped from the analysis as it was closely related to BRIEF scores (BRIEF Global Executive Score and Conners-3 EF: \( r_s (687) = .67, p < .0001, \) two-tailed) and because the eight BRIEF subscales provide a more comprehensive measure of EF.

2.2.3. Children’s Communication Checklist (CCC-2)

The Children’s Communication Checklist (CCC-2: Bishop, 2003) contains ten subscales assessing communication. The subscales Speech, Syntax, Semantics, and Coherence measure structural communication skills. Inappropriate initiation, Stereotyped language, Use of context, and Nonverbal communication address pragmatic communication. The Social relations and Interests subscales assess communication problems commonly observed in ASD.

2.3. Participants

Parent ratings were available for 720 children (i.e., all available data from the CALM study at the time of analysis). Outlier treatment was conducted as follows: three individual subscale scores were replaced with missing values due to falling \( \pm 3.5 \) standard deviations (SD) from the sample mean; six cases were detected as multivariate outliers based on Mahalanobis’ \( D^2 \) and omitted from the analyses. The characteristics of the final sample of 714 children, including diagnostic status and referral route, are available in Table II.1.
**Table II.1**  

Demographics, diagnoses, and referral route

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD)</th>
<th>Min-Max</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>9.4 (2.33)</td>
<td>5.17-18.58</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>N male</td>
<td>% male</td>
</tr>
<tr>
<td><strong>Diagnosis</strong></td>
<td>Total</td>
<td>%</td>
</tr>
<tr>
<td>None</td>
<td>427</td>
<td>59</td>
</tr>
<tr>
<td>ADHD/ADD</td>
<td>187</td>
<td>26</td>
</tr>
<tr>
<td>Learning deficit</td>
<td>67</td>
<td>9</td>
</tr>
<tr>
<td>ASD</td>
<td>57</td>
<td>7</td>
</tr>
<tr>
<td>Other</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>65</td>
<td>9</td>
</tr>
<tr>
<td><strong>Referrer</strong></td>
<td>Total</td>
<td>%</td>
</tr>
<tr>
<td>SENCo</td>
<td>395</td>
<td>55.32</td>
</tr>
<tr>
<td>Specialist Teacher</td>
<td>16</td>
<td>2.24</td>
</tr>
<tr>
<td>Educational Psychologist</td>
<td>7</td>
<td>0.98</td>
</tr>
<tr>
<td>Speech &amp; Language Therapist</td>
<td>32</td>
<td>4.48</td>
</tr>
<tr>
<td>Clinical Psychologist</td>
<td>30</td>
<td>4.20</td>
</tr>
<tr>
<td>Child Psychiatrist</td>
<td>9</td>
<td>1.26</td>
</tr>
<tr>
<td>Paediatrician</td>
<td>183</td>
<td>25.63</td>
</tr>
<tr>
<td>ADHD nurse practitioner</td>
<td>30</td>
<td>4.2</td>
</tr>
<tr>
<td>Family worker locality team</td>
<td>10</td>
<td>1.4</td>
</tr>
<tr>
<td>Private Tutor</td>
<td>2</td>
<td>0.28</td>
</tr>
</tbody>
</table>

*Note.* ADHD = attention deficit/hyperactivity disorder; ASD = autism spectrum disorder; Learning deficit = Primary diagnosis of developmental language disorder, dyslexia, dyscalculia, or dysgraphia; Other = Obsessive compulsive disorder, depression, anxiety; SENCo = special educational needs coordinator; Comorbidity = presence of more than one diagnosis.
2.4. Statistical Analysis

The analyses were conducted in four steps: network estimation, network stability, community detection, and network inference based on centrality indices and predictability. All analyses were performed in R (R Core Team, 2020) using the packages qgraph (Epskamp et al., 2012), bootnet (Epskamp et al., 2018), mgm (Haslbeck & Waldorp, 2015), igraph (Csardi & Nepusz, 2006), and networktools (Jones, 2018).

2.4.1. Missing Data and Overly Negative Scores

The number of participants with complete data for each scale is presented in Table II.2. The network was estimated based on complete pairwise correlations to allow for the largest possible sample size (see Figure AI.1 and AI.2 in Appendix I for alternative estimations). As a robustness check, the analyses were conducted with and without ratings flagged as overly negative or inconsistent by the indices provided in each scale.

2.4.2. Network Estimation

The regularised partial correlation network of the BRIEF, CCC-2, and Conner’s 3 subscales was estimated. Subscales had excellent to good internal consistency, suggesting that the items within each scale were capturing the same construct (Table II.2). Therefore, the analyses focused on subscales interrelations to avoid potential bias introduced by the inclusion of multiple items assessing the same construct (for a discussion of such biases see Fried & Cramer, 2017). The distribution of the ratings on multiple subscales deviated from normality (see Table II.2 for skewness and kurtosis). For this reason, the network was estimated with a non-paranormal transformation (Liu et al., 2009). This transformation uses cumulative distributions to transform the observed variable to the distribution of the latent normally distributed variable (for details see Epskamp & Fried, 2018; Liu et al., 2009). The consequences of the transformation were empirically evaluated by comparing the network structure derived from the transformed data to the network estimated based on non-transformed data. The same
conclusions were supported across procedures (this analysis is reported in Appendix I, Section 1). All raw scores were transformed such that higher values represented more difficulties and were centred on the sample mean to put them on the same scale. Age was included in the network estimation but omitted from figures and the calculation of the centrality indices. In the final network, each node represented a subscale rating and each edge corresponded to the regularised partial correlation coefficient across the two subscales, controlling for the influence of age and all other ratings. The network was estimated using the graphical variant of the least absolute shrinkage and selection operator (glasso: Friedman et al., 2008) to avoid including spurious edges. The best-fitting model was selected based on the Extended Bayesian Information Criterion (EBIC: Chen & Chen, 2008). This method is reported to accurately retrieve the true network structure (Foygel & Drton, 2010) and is described in detail elsewhere (Epskamp et al., 2018). To ensure good specificity (including only edges that are truly present) the estimated elements of the inverse variance-covariance matrix were first thresholded using the theoretical bound proposed by Jankova and van de Geer (2018). This procedure reflects a strong assumption of sparsity and could result in a loss of sensitivity to detect small edges that are truly present. In the current sample, the minimum estimated absolute edge weight in the non-thresholded network was 0.0007; and 0.067 in the thresholded network. The implications of this analytical decision were evaluated by comparing the thresholded, non-thresholded, and unregularised solutions, all of which retained acceptable similarity (see Appendix I, Figure AI.1 for details).

2.4.3. Network Stability

The robustness of the network solution was scrutinised by calculating bootstrapped non-parametric 95% confidence intervals for all edges weights, together with the percentage of bootstrapped networks ($N = 2000$), in which the edge was estimated as different from zero (% non-zero) (Epskamp et al., 2018).
2.4.4. **Community Detection**

Symptoms that are closely related to one another may cluster together in communities. Communities may represent key reinforcement loops of reciprocal inter-symptom influence or may be part of the same latent variable or dimension (see Chapter I, Section 2.3.2, for details). The Walktrap algorithm was applied to the network to evaluate the presence of symptom clusters (Pons & Latapy, 2006). The algorithm recursively takes random walks between pairs of nodes to define communities as densely connected parts of the network. Simulation work suggests that Walktrap can retrieve the true generating structure (i.e., number of communities/clusters) across a range of conditions (Gates et al., 2016; Golino & Epskamp, 2017; Pons & Latapy, 2006). The quality of the cluster separation was quantified using the modularity ($Q$) index, which increases when within-community links are stronger than what would be expected at random. Modularity values above 0.3 are considered evidence for sufficient community separation (Newman & Girvan, 2004).

2.4.5. **Symptom Centrality and Predictability**

Centrality measures quantify the inter-connectedness of each node, potentially revealing the relative importance of nodes within the network (Epskamp et al., 2018). The centrality indices presented here were chosen based on correlation stability coefficients (CS): indices were considered stable if at least 50% of children could be dropped while maintaining a 95% probability of a 0.7 correlation between the centrality indices based on the full sample and those derived from subsamples (Epskamp & Fried, 2018). The case-dropping analysis ($N$ bootstraps = 2000) suggested stable node strength (CS (cor = 0.7) = 0.69), expected influence (CS (cor = 0.7) = 0.75), and bridge strength CS (cor=0.7) = 0.52). Strength represents the sum of all edge weights (regularised partial correlations) directly linked to a given node; expected influence is based on the same formula while taking negative relationships into account (Robinaugh et al., 2016); and bridge strength represents the sum of edge weights of given node...
to all nodes of a different cluster. In other words, bridge strength captures symptoms with multiple strong conditional relationships with symptoms from other clusters. Due to the cross-sectional nature of the data the conceptual interpretation of these metrics is unclear. One possibility is that symptoms of high centrality may be the cause or the consequence of the symptoms they are related to. To ensure centrality indices are not biased by item properties such as differences in the nodal variance, the associations across node centrality measures and the standard deviations of each subscale included in the network was estimated (a procedure commonly used in the literature, see Fried et al., 2018). They were weak and non-significant (strength: $r_s = 0.23, p = 0.30$; expected influence: $r_s = 0.22, p = 0.31$; bridge strength: $r_s = 0.02, p = 0.92$), suggesting that centrality metrics were not overly influenced by differences in variance.

The predictability of individual symptoms was examined to quantify the shared variance between a given node and all of the symptoms connected to it (i.e., the proportion of variance explained) (Haslbeck & Waldorp, 2015). Pairwise missing data handling was not feasible for this analysis. However, the network estimated based on listwise correlations was very similar to the network based on pairwise complete correlations (see Appendix I, Figure AI.1), therefore predictability was estimated based on listwise complete data ($N = 668$).
Table II.2

Descriptive statistics, reliability, and percentage of children rated as experiencing clinical levels of difficulties. The means and standard deviations are based on raw scores.

<table>
<thead>
<tr>
<th>Conners-3</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt</th>
<th>Min</th>
<th>Max</th>
<th>α</th>
<th>ω</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inattention</td>
<td>702</td>
<td>11.68</td>
<td>3.44</td>
<td>90.48</td>
<td>-1.11</td>
<td>0.54</td>
<td>15</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>Hyperactivity/Impulsivity</td>
<td>704</td>
<td>10.96</td>
<td>5.63</td>
<td>75.77</td>
<td>-0.39</td>
<td>-1.11</td>
<td>0</td>
<td>18</td>
<td>0.95</td>
</tr>
<tr>
<td>Learning Problems</td>
<td>704</td>
<td>9.83</td>
<td>3.48</td>
<td>87.54</td>
<td>-0.42</td>
<td>-0.56</td>
<td>0</td>
<td>15</td>
<td>0.78</td>
</tr>
<tr>
<td>Aggression</td>
<td>701</td>
<td>3.55</td>
<td>3.99</td>
<td>49.58</td>
<td>1.21</td>
<td>0.55</td>
<td>0</td>
<td>15</td>
<td>0.93</td>
</tr>
<tr>
<td>Peer Relations</td>
<td>696</td>
<td>5.50</td>
<td>4.47</td>
<td>67.37</td>
<td>0.47</td>
<td>-0.97</td>
<td>0</td>
<td>15</td>
<td>0.93</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BRIEF</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt</th>
<th>Min</th>
<th>Max</th>
<th>α</th>
<th>ω</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inhibition</td>
<td>707</td>
<td>21.51</td>
<td>6.23</td>
<td>60.5</td>
<td>-0.24</td>
<td>-1.21</td>
<td>10</td>
<td>30</td>
<td>0.97</td>
</tr>
<tr>
<td>Shifting</td>
<td>708</td>
<td>17.09</td>
<td>4.28</td>
<td>66.11</td>
<td>-0.27</td>
<td>-0.90</td>
<td>8</td>
<td>24</td>
<td>0.91</td>
</tr>
<tr>
<td>Emotional Control</td>
<td>707</td>
<td>22.12</td>
<td>5.59</td>
<td>61.9</td>
<td>-0.37</td>
<td>-0.93</td>
<td>10</td>
<td>30</td>
<td>0.95</td>
</tr>
<tr>
<td>Initiation</td>
<td>707</td>
<td>17.69</td>
<td>3.33</td>
<td>70.59</td>
<td>-0.37</td>
<td>-0.33</td>
<td>8</td>
<td>24</td>
<td>0.82</td>
</tr>
<tr>
<td>Working Memory</td>
<td>704</td>
<td>25.81</td>
<td>3.94</td>
<td>89.08</td>
<td>-1.14</td>
<td>0.89</td>
<td>12</td>
<td>30</td>
<td>0.93</td>
</tr>
<tr>
<td>Planning</td>
<td>698</td>
<td>29.06</td>
<td>4.97</td>
<td>85.15</td>
<td>-0.79</td>
<td>0.18</td>
<td>13</td>
<td>36</td>
<td>0.90</td>
</tr>
<tr>
<td>Organisation</td>
<td>708</td>
<td>14.62</td>
<td>3.22</td>
<td>56.58</td>
<td>-0.81</td>
<td>-0.26</td>
<td>6</td>
<td>18</td>
<td>0.92</td>
</tr>
<tr>
<td>Monitoring</td>
<td>706</td>
<td>18.98</td>
<td>3.53</td>
<td>68.21</td>
<td>-0.56</td>
<td>-0.36</td>
<td>9</td>
<td>24</td>
<td>0.86</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CCC-2</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt</th>
<th>Min</th>
<th>Max</th>
<th>α</th>
<th>ω</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech</td>
<td>702</td>
<td>5.18</td>
<td>5.03</td>
<td>62.75</td>
<td>1.10</td>
<td>0.55</td>
<td>0</td>
<td>21</td>
<td>0.90</td>
</tr>
<tr>
<td>Syntax</td>
<td>704</td>
<td>5.11</td>
<td>4.62</td>
<td>66.67</td>
<td>1.00</td>
<td>0.36</td>
<td>0</td>
<td>20</td>
<td>0.87</td>
</tr>
<tr>
<td>Semantics</td>
<td>702</td>
<td>8.16</td>
<td>4.74</td>
<td>78.15</td>
<td>0.37</td>
<td>-0.43</td>
<td>0</td>
<td>21</td>
<td>0.84</td>
</tr>
<tr>
<td>Coherence</td>
<td>705</td>
<td>8.57</td>
<td>5.13</td>
<td>77.17</td>
<td>0.21</td>
<td>-0.87</td>
<td>0</td>
<td>21</td>
<td>0.85</td>
</tr>
<tr>
<td>Inappropriate Initiation</td>
<td>702</td>
<td>10.79</td>
<td>5.71</td>
<td>68.21</td>
<td>-0.04</td>
<td>-1.07</td>
<td>0</td>
<td>21</td>
<td>0.88</td>
</tr>
<tr>
<td>Stereotyped Language</td>
<td>703</td>
<td>6.10</td>
<td>4.27</td>
<td>64.99</td>
<td>0.62</td>
<td>-0.25</td>
<td>0</td>
<td>20</td>
<td>0.80</td>
</tr>
<tr>
<td>Use of Context</td>
<td>704</td>
<td>9.58</td>
<td>5.48</td>
<td>77.87</td>
<td>0.05</td>
<td>-0.94</td>
<td>0</td>
<td>21</td>
<td>0.87</td>
</tr>
<tr>
<td>Nonverbal Communication</td>
<td>703</td>
<td>7.71</td>
<td>5.05</td>
<td>73.81</td>
<td>0.34</td>
<td>-0.81</td>
<td>0</td>
<td>21</td>
<td>0.84</td>
</tr>
<tr>
<td>Social Relations</td>
<td>703</td>
<td>6.61</td>
<td>4.94</td>
<td>70.87</td>
<td>0.49</td>
<td>-0.67</td>
<td>0</td>
<td>20</td>
<td>0.85</td>
</tr>
<tr>
<td>Interests</td>
<td>702</td>
<td>9.19</td>
<td>4.61</td>
<td>68.21</td>
<td>0.21</td>
<td>-0.62</td>
<td>0</td>
<td>21</td>
<td>0.80</td>
</tr>
</tbody>
</table>

**Note.** If children’s age was outside the standardisation range, the closest age match was used. BRIEF = Behaviour Rating Inventory of Executive Function; CCC-2 = Children’s Communication Checklist 2. M = mean, SD = standard deviation, Kurt = kurtosis, Skew = Skewness, α = Cronbach’s alpha, ω = McDonald’s omega. For BRIEF and Conners-3, higher raw and T-scores values indicate presence of more difficulties. For CCC-2, lower raw scores indicate fewer difficulties, whereas high scaled scores indicate fewer difficulties.
3. Results

3.1. Descriptive Statistics

Means, SDs, and internal consistency indices (Cronbach’s alpha and McDonald’s omega) based on polychoric correlations for all untransformed measures are presented in Table II.2. Acceptable internal consistency was observed across all subscales. Pearson correlations across all measures following the data transformation are displayed in Figure II.1.

3.2. Subscale Validity Checks

The questionnaires included in this study provide indices to detect potential inconsistencies in parent ratings and/or negative response styles. The Conners-3 contains a scale to detect overly negative responses and potential bias. In the current sample, the validity scale suggested an overly negative response style for 70 responses, which may be indicative of extreme difficulties or negative reporter bias. The validity indices of the Behavioural Rating Inventory of Executive Function (BRIEF) identified two cases as inconsistent and 67 cases as having overly negative response styles. The Children’s Communication Checklist 2 (CCC-2) consistency check score identified 46 potentially inconsistent ratings. Analyses were carried out with and without the ratings flagged as highly negative or inconsistent. The estimated regularised partial correlation network obtained after omitting all potentially biased scores retained high similarity to the network estimated with them (correlation of adjacency matrices: $r = .92$). Therefore, the analyses reported below are based on the full sample.
Figure II.1

Pearson correlations after nonparanormal transformation across all variables in the network.

Note. Conners-3: Inatt = Inattention; Hyp = Hyperactivity/Impulsivity; Learn = Learning Problems; Agg = Aggression; Peer = Peer Relations; BRIEF (Behaviour Rating Inventory of Executive Function): Inhib = Inhibition; Shift = Shifting; Emot = Emotional control; Initi = Initiation; WM = Working memory; Plan = Planning/Organisation; Org = Organisation of Materials; Monit = Monitoring; CCC-2 (Children’s Communication Checklist): Synt = Syntax; Seman = Semantics; Coher = Coherence; Nonver = Nonverbal Communication; Social = Social Relations; Interest = Interests.
3.3. Network Estimation and Stability

The network is displayed in Figure II.2. Stability analyses indicated that the network was estimated with sufficient accuracy, as the corresponding confidence intervals (CIs) were small to moderate (Figure AI.3, Appendix I). All edges featured in the final network were included in the majority of the 2000 bootstrapped samples. Notably, due to the regularisation method, the confidence intervals are not to be interpreted as significance tests to zero (Epskamp et al., 2018). Instead, Figure II.3 shows the proportion of bootstraps for which a given edge was set as different from zero. For edges discussed in the chapter, the following information is provided in parentheses: edge weight, 95% bootstrapped confidence interval for times the edge weight was not set to zero, and a percentage indicating how often the parameter was estimated as different from zero across the 2000 bootstraps. For all other edges, this information is available in Figure II.4 and Appendix I, Figure AI.4.
Figure II.2

The network of behavioural and communication problems in children struggling at school.

Note. The thickness of an edge corresponds to the magnitude of the partial correlation between two nodes after adjusting for all other nodes in the network. Green edges depict positive associations and red edges depict negative associations. The node colours correspond to the clusters identified by the Walktrap algorithm. The blue ring around each node corresponds to the proportion of variance explained in the node by directly related nodes. BRIEF = Behaviour Rating Inventory of Executive Function; CCC-2 = Children’s Communication Checklist.
3.4. Community Detection

The Walktrap algorithm identified four clusters corresponding to: 1) pragmatic language and peer relationships; 2) structural language and learning; 3) cognitive skills (cool EFs); and 4) emotional and behavioural problems (hot EFs). The modularity index ($Q$) was 0.43, suggesting a good level of separation between the clusters. Nonetheless, these clusters were interrelated. Several direct paths were observed across pragmatic and behavioural difficulties (e.g., inappropriate initiation – hyperactivity (0.13, 95% CI [0.07 - 0.18], non-zero $= 97$%); inappropriate initiation – inhibition (0.19, 95% CI [0.12 - 0.26], non-zero $= 100$%). No direct relationships were observed between cool EFs and structural language; these clusters were indirectly connected through learning.
Figure II.3

The upper triangle represents estimated edge weights, where darker shades correspond to stronger absolute edge weight, green corresponds to positive and red corresponds to negative associations. The values in the lower triangle represent how often an edge was estimated to be non-zero in the 2000 bootstraps. Values of 1 shaded in the darkest green indicate that an edge was included in all networks.

Note. Conners-3: Inatt = Inattention; Hyp = Hyperactivity/Impulsivity; Learn = Learning Problems; Agg = Aggression; Peer = Peer Relationships; BRIEF: Inhib = Inhibition; Shift = Shifting; Emot = Emotional Control; Initi = Initiation; WM = Working memory; Plan = Planning/Organisation; Org = Organisation of Materials; Monit = Monitoring; CCC-2: Synt = Syntax; Seman = Semantics; Coher = Coherence; Inap. Initi = Inappropriate Initiation; Stereo = Stereotyped Language; Context = Use of Context; Nonver = Nonverbal Communication; Social = Social Relations; Interest = Interests.
3.5. Symptom Centrality and Predictability

Strength, expected influence, and bridge strength centrality are shown in Figure II.4. In the current network, working memory, coherence, use of context, and hyperactivity had the highest strength and expected influence. Coherence and shifting had the highest bridge strength (Figure II.4), reflecting multiple strong relationships with symptoms from other clusters.

Predictability analysis (Figure II.2) suggested that the symptoms explained an acceptable amount of variance across the network ($M = 0.66$, $SD = 0.10$, $Min = 0.36$ (Organisation), $Max = 0.78$ (Inhibition)). A close relationship was observed between the network model and symptom predictability: if a symptom was connected to only a few other symptoms, the explained variance was lower; inversely, the more connections a symptom had, the higher the estimated predictability.
Figure II.4

Centrality indices bridge strength, strength, and expected influence across all measures in the network.

Notes. Conners-3: Inatt = Inattention; Hyp = Hyperactivity/Impulsivity; Learn = Learning Problems; Agg =Aggression; Peer =Peer Relations; BRIEF (Behaviour Rating Inventory of Executive Function): Inhib = Inhibition; Shift = Shifting; Emot = Emotional control; Initi = Initiation; WM = Working memory; Plan = Planning/Organisation; Org = Organisation of Materials; Monit = Monitoring; CCC-2 (Children’s Communication Checklist): Synt = Syntax; Seman = Semantics; Coher = Coherence; Inap.Initi = Inappropriate Initiation; Stereo = Stereotyped language; Context = Use of Context; Nonver = Nonverbal Communication; Social = Social Relations; Interest = Interests.
4. Discussion

The study presented in this chapter used network analysis to explore the co-occurrence of symptoms commonly observed across developmental disorders such as ADHD, DLD, ASD, and reading difficulties. Inter-symptom associations across communication, behavioural, and cognitive difficulties were modelled in the large heterogeneous sample of children (the CALM cohort). Four empirically-derived clusters of symptoms emerged corresponding to: 1) structural language and learning; 2) pragmatic abilities and peer relations; 3) behavioural and emotional difficulties (hot EFs); and 4) cognitive skills (cool EFs). Hot and cool EFs were directly related, as were structural and pragmatic language skills. Cool EFs were directly linked with learning but not with the formal use of language in communication.

Problems with pragmatic communication and peer relationships were directly connected to emotional and behavioural difficulties (hot EFs). In turn, emotional and behavioural difficulties were directly related to cool EFs. This was manifested in the high bridge centrality rankings of shifting and hyperactivity, both of which were in the hot EFs cluster. Their centrality in the network reflected associations with symptoms in the pragmatic communication cluster (e.g., hyperactivity – inappropriate initiation; shifting – interests) and the cognitive skills cluster (e.g., hyperactivity – inattention; shifting – working memory). These findings are consistent with hypotheses suggesting that social communication difficulties can arise as a downstream consequence of hyperactive-impulsive behaviours that are themselves partially underpinned by poor cool EFs (Green et al., 2014). The network structure cannot provide evidence for the direction of these associations, but the observed paths may imply that cool EFs (e.g., inattention, working memory) do not directly influence peer relations and pragmatic skills, but can lead to difficulties in these domains via the activation of behavioural problems such as hyperactivity (Holmes et al., 2016; Rints et al., 2014).

Contrary to expectations, structural language skills and cool EFs were not directly
linked. It has been previously reported that cool EFs such as verbal short-term and working memory underpin language development (Archibald & Gathercole, 2006). However, no such links were found in the network, potentially because all measures were rating-based rather than performance-based. Another possibility is that the development of structural communication skills is mostly constrained by phonological processing, which was not included in the current study, and not by executive function (Melby-Lervåg et al., 2012; Ramus et al., 2013). Links between both clusters and learning may reflect how cognitive deficits (e.g., working memory, inattention) and structural language skills (e.g., syntax, semantics) limit learning in different ways. Language impairments might be more closely related to literacy problems (Ramus et al., 2013), with cool EF problems impacting on broader aspects of learning (Yeniad et al., 2013). An alternative possibility is that cool EF deficits constrain learning, and this impairs the acquisition of structural language abilities (Korrel et al., 2017).

Multiple direct paths linked the structural and pragmatic communication clusters, but their associations with other symptoms in the network differed. This supports the view that these dimensions of communication are linked but could have distinct origins (Bishop, 2014). As suggested in the network, impairments in pragmatic language may arise through both structural language problems (Redmond & Rice, 1998) and social/behavioural difficulties (Staikova et al., 2013). Structural language abilities appear to be more closely tied to learning, and indirectly to cool EFs.

Centrality indices highlighted important roles for working memory, language coherence, and the appropriate use of context in communication. Working memory and language coherence had multiple connections both within and outside their own clusters. Symptoms with multiple connections across problem areas may potentially interact with other areas of difficulty and may be the origin/consequence of co-occurrences (Blanken et al., 2018). Working memory bridged the cool EF and structural language clusters via learning. The role of working
memory in learning is well-established (Gathercole et al., 2004). Working memory also linked the cool EF and pragmatic clusters via coherence and shared direct paths with hot EFs (shifting). Working memory did not rank highest on bridge strength, but its highest overall strength and expected influence, together with its direct associations with all symptoms within the cool EF cluster and direct links with symptoms from all other clusters, provide evidence for its potential role as an area of difficulty that may spread activation across the network. Although causal conclusions are not warranted due to the correlational nature of network models, working memory is important for holding information in mind, focusing attention, and ignoring distractions. Impairments in these abilities might activate difficulties in other areas such as producing coherent speech/narratives and shifting on to novel activities. Considering working memory as a transdiagnostic risk factor for developmental difficulties fits with data from the mental health field (Huang-Pollock et al., 2017).

Coherence in communication (pragmatic abilities and peer relations cluster) was associated with all three symptoms of structural communication and four symptoms in its own cluster. Poor language coherence may therefore activate, or be the consequence of, multiple other communication problems. In contrast, the high centrality of the use of context when communicating reflected multiple links within the pragmatics and peer relations cluster with a single association with the structural cluster (semantics). Symptoms with strong connections in the same cluster may be core symptoms of this problem area (Blanken et al., 2018). Consistent with this, the appropriate use and interpretation of language in relation to context is central to the definition of pragmatics, and difficulties in this specific ability are reported to differentiate across children with primarily structural versus pragmatic difficulties (Bishop & Baird, 2001).
4.1. Limitations & Future directions

There are several limitations to this study. First, the symptom relationships identified are based on parent ratings that are designed to capture aspects of functioning distinct/supplementary from those measured in lab-based assessments of cognition (Toplak et al., 2013). Performance-based measures focus on the efficiency of cognitive abilities in structured situations, whereas parent ratings provide ecologically valid assessments of children’s day-to-day functioning across different situations (Toplak et al., 2013), but are known to be subject to reporter bias. To test whether bias affected the results, the current analyses were conducted both including and excluding reports flagged as overly negative or inconsistent. The overall pattern of results did not differ between analyses, indicating that such reporter bias is unlikely to be the cause of the overall network structure. In addition, the identification of clusters of symptoms that align with theoretical constructs provides further validation that the ratings were meaningful as they effectively distinguished between different aspects of functioning. In the future, it should be explored whether a similar network structure emerges with performance-based assessments of language and cognition.

The communities of symptoms identified depend on the chosen community detection algorithm and the stability of the input network. The algorithm applied here is reported to identify network communities with an acceptable level of accuracy, and the input network was stable in terms of the strength and number of connections across symptoms. Furthermore, the problem areas present in the network map on to broad latent constructs identified in previous studies of these comorbidities (Hawkins et al., 2016), suggesting some level of agreement across methods.

The heterogeneous nature of the sample was suited for investigating the possibility of transdiagnostic symptom-level associations across comorbid difficulties. Nonetheless, it is also possible that there might be subgroups within the larger sample with alternative patterns of
inter-symptom associations, a possibility which merits further investigation but would require larger sample sizes. An important direction for future research will be to compare the structure of the network across children with different profiles of difficulties.

The network model revealed bridging symptoms, which may provide important insights about the origins of comorbidities observed on a dimensional level, and which could be interpreted as candidate targets for interventions. However, for such interventions to be successful the temporal order of activation of the connected symptoms is of key importance. The targeted difficulty should be the cause rather than the effect of other symptoms. The current analytical framework does not afford such conclusions – central symptoms may cause other symptoms or may be the consequence of those other symptoms. Furthermore, the estimated relationships may not signal interacting areas of difficulties and may instead reflect shared item content or aetiological influences that were not included in the model. To disentangle these possibilities, it is important to evaluate whether bridging symptoms play a causal role in the co-occurrence of problem domains and to test whether interventions targeting these symptoms reduce the activation of difficulties in other areas of functioning. The incorporation of neurological, genetic, and environmental factors into network models is another important step that could provide important insights into the origins of comorbidities and their potential dynamic interactions.

4.2. Conclusion

The co-occurrence of pragmatic communication and behavioural problems was observed in a conditional dependence network estimated in a large heterogeneous sample of children with a broad range of difficulties. This suggests these comorbidities extend beyond specific diagnostic groups (e.g., ADHD, ASD). On a practical level, these findings highlight the importance of considering potential language and communication problems among children presenting with difficult behaviour and vice versa.
Pragmatic communication skills might be indirectly influenced by cognitive skills through the mediating role of behavioural regulation. In line with this, working memory was identified as a bridging symptom, suggesting it may spread difficulties across different domains of functioning. Working memory, inattention, and structural language difficulties were associated with learning, in line with many previous reports.

The data presented here provide one of the first large-scale applications of network modelling to reported symptoms associated with a range of developmental disorders. Akin to developments in psychiatry, this investigation suggests that there might be utility in shifting away from conceptualising developmental disorders as nosological entities. Transdiagnostic approaches can enable the discovery of shared liabilities and are suited for investigating the possibility that different developmental difficulties may cause one another: each problem may be the starting point for the activation of other symptoms. Using network modelling to conceptualise developmental comorbidities as arising in dynamic causal systems can provide insights into the nature of these comorbidities and may help researchers and clinicians to formulate specific hypotheses about potential causal mechanisms and intervention strategies.

In the following chapter, I use a similar network approach to identify subgroups of children with common profiles of strength and weaknesses across the clusters of symptoms identified in this study. I then explore how these subgroups differ from the comparison group of non-referred CALM children (for details see Chapter I, Section 2.2.1.) on measures of academic, socio-emotional functioning and neural white matter organisation.
CHAPTER III

Transdiagnostic Profiles of Behaviour and Communication
Relate to Academic and Socio-emotional Functioning and Neural
White Matter Organisation

Chapter III has been accepted for publication, and a preprint is available but appears
modified to ensure thesis cohesion.

Profiles of Behaviour and Communication Relate to Academic and Socio-emotional
Functioning and Neural White Matter Organisation. Journal of Child Psychology and
Psychiatry. https://doi.org/10.1101/2021.11.29.21267002

First-pronouns are used throughout the chapter, but the work is the result of a collaborative
research project. My contributions to the publication are:

1. I led the conceptualisation of the study research questions, methods, analyses and
   interpretation supported by J. Holmes.
2. I contributed to the data collection, scoring and curation, which was achieved in a
   collaborative effort including a large team of researchers (the CALM team).
3. I performed all statistical analyses. The pre-processing of the neuroimaging data
   was completed by D. Akarca.
4. I wrote the manuscript draft aided by J. Holmes with input from D. Akarca.
1. Introduction

The results of Chapter II align with a wide literature suggesting that behavioural difficulties, poor communication skills, and everyday cognitive problems are common in children and that they often co-occur (Hawkins et al., 2016; Ketelaars et al., 2010; Loe & Feldman, 2007; Lui & Tannock, 2007; Pimperton & Nation, 2014). The high prevalence of these comorbidities and their dense interrelations suggests that difficulties across these domains may not be independent and disorder-specific. Instead, they likely share common aetiological origins and/or may interact dynamically, such that difficulties in one of these areas of functioning may trigger developmental cascades leading to more widespread problems (see Masten et al., 2005).

As discussed in Chapter I, Section 1.2., comorbidity is one of the main challenges to the current diagnostic system, which is further compounded by the high levels of symptom variability within diagnostic categories: very different profiles of strengths and difficulties are common among children with the same diagnostic label (e.g., Astle et al., 2019; Kushki et al., 2019). This has led to widespread recognition of the limitations of diagnostic frameworks for guiding research and support strategies (Casey et al., 2014), and an increase in the application of transdiagnostic approaches for understanding neurodevelopment (see Astle et al., 2021, for a review). Transdiagnostic approaches adopt alternative recruitment and analytical strategies, as discussed in Chapter I. To re-cap here, one analytical goal is to identify dimensional characteristics that can predict or explain difficulties experienced across individuals irrespective of diagnostic status (e.g., Brislin et al., 2020). The relations between different dimensions and their links to neural and genetic factors can then be studied to identify potential mechanisms of shared or unique variance (e.g., Holmes et al., 2020; Parkes et al., 2021; see Chapter I, Section 2.3.1. for a full discussion of dimensional methods). Network approaches, such as the analyses presented in Chapter II, support this effort and provide additional insights.
by exploring how the constituent parts of dimensions relate to one another. A complementary analytical approach, often described as *clustering* or *subtyping*¹, focuses on deriving homogeneous data-driven groupings of children with similar profiles of relative strengths and weaknesses along the studied dimensions (e.g., Feczko et al., 2018, 2019, van Rentergem et al., 2021). These novel groupings can then be similarly used to facilitate the discovery of common aetiological pathways to specific difficulties, or to stratify individuals to support strategies appropriate for their needs (see Chapter I, Section 2.3.3, for a fuller explanation of clustering approaches).

Dimensional and subgrouping approaches can provide complementary insights into the aetiology of transdiagnostic symptoms and their implications for functioning. As one example, consider EFs. Dimensional approaches provide support for two dimensions of EFs linked to separate neurodevelopmental pathways (Zelazo, 2020): cool EFs refer to the ability to regulate behaviour and cognition in emotionally neutral contexts, whereas hot EF skills are implicated in situations of stronger motivational and emotional valence (Castellanos et al., 2006). Studies using dimensional analysis reveal these dimensions of EFs are related, but are associated with relatively distinct neural networks (see Salehinejad et al., 2021 for an overview of the neuroimaging literature) and have somewhat different predictive validity in terms of cool EFs being more closely related to academic achievement and hot EFs more predictive of disruptive behaviours (e.g., Willoughby et al., 2011).

Subgrouping approaches similarly reveal that different EF profiles have different neural correlates and are associated with different outcomes. For example, Vaidya et al. (2020)

---

¹ Subtyping is usually used in the context of applying clustering techniques within a sample of children with the same diagnosis to uncover different presentations of the same disorder (e.g., ADHD-combined, ADHD-inattentive, etc.). For clarity, I use the term subgrouping from this point forward to refer to data-driven clustering approaches applied in diverse samples rather than within a diagnostic category.
identified three data-driven EF-subgroups across typically developing, autistic, and ADHD children. These subgroups explained more variance in frontal-parietal engagement (inferred via functional MRI) during a sustained attention task than diagnostic groupings. Similarly, Bathelt et al. (2018) identified three distinct EF-subgroups in a transdiagnostic sample of struggling youth. Each subgroup was associated with different variations in white matter connections in prefrontal regions and there was more within-group homogeneity in the behavioural profiles of the subgroups than in groupings based on diagnoses. Both examples demonstrate that studying subgroups in transdiagnostic samples can facilitate the linking of behavioural phenotypes to neurophysiological mechanisms.

1.1. Transdiagnostic Approaches to Neurodevelopmental Comorbidities

The examples of transdiagnostic designs described so far in this chapter, and in the majority of the literature, focus on one domain of function (e.g., EFs). An important next step is to apply these methods to multiple areas of functioning to advance our understanding of the co-occurrence of difficulties across multiple domains. This was attempted in Chapter II to explore how symptoms of communication, behaviour, and EF difficulties relate to one another in the CALM cohort. Using a network approach, I was able to investigate how specific symptoms clustered together and start to characterise how these clusters of difficulties were linked.

As a reminder, four densely interrelated clusters of symptoms were identified relating to structural language and learning, hot EFs, pragmatic communication and peer relationships, and cool EFs. The symptoms within each cluster did not align with the diagnostic features of any diagnosis. Moreover, specific symptoms bridged these clusters (e.g., inappropriate initiation of communication linked the pragmatic communication and hot EF clusters), providing some insight into the co-occurrence of symptoms across domains of difficulty. It may be the case that these bridging symptoms can trigger difficulties through developmental
cascades (Masten et al., 2005) or, alternatively, they may be particularly susceptible to the influence of symptoms in other domains. It is also possible that bridging symptoms do not always signal interacting areas of difficulties but may instead also reflect aspects of shared neural and/or genetic underpinnings.

Aetiological factors likely underlie the co-occurrence of behavioural symptoms across different domains, but their influence can be masked in traditional diagnostic comparisons. In the current study, I was interested in extending the transdiagnostic approach presented in Chapter II to capture potential links to brain structure. Diffusion tensor imaging (DTI) allowed me to estimate the macroscopic organisation of white-matter connectivity of the brain in vivo, which is known as a structural connectome (Basser et al., 2000). This offers a window into exploring how the brain’s connectivity may vary between individuals and how such variation relates to individual differences across the domains of interest. Research insights from case-control designs suggest there are both shared and distinct differences at the structural neural level across children with different neurodevelopmental disorders (Ameis et al., 2016; Baribeau et al., 2019; Moreau et al., 2018). Examples of structural neural substrate variation implicated in multiple neurodevelopmental disorders include the corticostriatal (Arnsten & Rubia, 2012; Beauchaine et al., 2019) and interhemispheric circuits (e.g., Ameis et al., 2016; Aoki et al., 2017).

1.2. The current study

The current study aimed to identify homogeneous subgroups of children with similar behavioural, EF, and communication strengths/difficulties in the CALM cohort. Subsequently, I explored how the profiles of these transdiagnostic subgroups related to measures of brain structure, and academic and socio-emotional functioning. The previous chapter included a subsample of the currently studied population (i.e., those for whom data were available at the time) and focussed on identifying clusters of symptoms that characterised the sample. Here, I
use a complementary subgrouping approach to study how this space is occupied by participants. In other words, I searched for individuals with similar profiles across the four previously identified clusters of difficulties. To explore the external validity of the subgroups and further characterise their profiles, I investigated how they differed on independent (i.e., not included in the subgroup identification) measures of nonverbal cognitive ability, academic, and socio-emotional functioning, which are theorised to be related to interindividual differences in communication and EFs (e.g., Arnold et al., 2020; Harpin et al., 2016; Helland & Helland, 2017). I did not formulate a hypothesis about the number of groups, their specific profiles, or the neural and external behavioural features that would differentiate them. However, based on previous literature I expected that both difficulties with structural language and cool EFs would be related to difficulties with academic performance, while pragmatic and hot EF difficulties were expected to relate more strongly to socio-emotional functioning.

2. Method

The data presented here were collected as part of a cohort study at the Centre for Attention, Learning, and Memory (CALM), described in Chapter I, Section 2.2.1. The data included in the current analyses was collected between 2014 and 2021.

2.1. Recruitment

Full details about the CALM cohort are available in the study protocol (Holmes et al., 2019) and an overview is presented in Chapter I, Section 2.2.1. This study included both groups of children recruited to CALM, the referred sample with difficulties with attention, memory, and/or learning and the comparison/non-referred group. The two groups were recruited from the same schools to ensure they had similar socio-demographic profiles.

2.2. Participants

The referred cohort included 805 children (CALM800: 69% male, \( M_{age} = 9.48, SD_{age} = 2.38 \)) and the comparison 158 children (CALM200: 56% male, \( M_{age} = 10, SD_{age} = 2.33 \)).
The majority (63%) of the struggling cohort were referred by education practitioners (e.g., educational psychologist, special educational needs coordinator, etc.); 33% were referred by health professionals (e.g., clinical psychologist, child psychiatrist, etc.); and 4% by speech and language therapists. Most of the sample (60%) were undiagnosed, despite being recognised by a professional as having additional needs, and 8% had more than one diagnosis. Among those with diagnoses ADHD was the most common ($N = 197$), followed by ASD ($N = 57$), and diagnoses of learning disorders (e.g., dyslexia, dyscalculia, developmental language disorder, $N=62$). The non-referred sample consisted mostly of children without a diagnosis (96%), and among those with diagnoses one had ADHD, two had DLD, and two had dyslexia.

2.3. Assessments

Each child completed a battery of standardised neuropsychological assessments following the procedures documented in the assessment’s testing kits (see Holmes et al., 2019 for details). Parents/caregivers provided ratings of their child’s behaviour, communication, and socio-emotional functioning. All children were invited to participate in an optional neuroimaging session within six months of their behavioural assessment (CAM800: $N = 313$, age at scan: $M = 10.23$, $SD = 2.32$; CALM200: $N = 77$, age at scan: $M = 10.75$, $SD = 2.01$). The assessments included in the current study, and the MRI data acquisition procedure, are described below.

2.3.1. Measures included in Community Detection

The community detection analysis, used to identify the subgroups, was based on the same measures analysed in Chapter II. For cohesion, details of these assessments are briefly outlined again.

2.3.1.1. Conners-3

The Conners Parent Rating Short Form 3rd Edition (Conners-3, Conners, 2008) asks about the child’s ADHD-related difficulties in the past month. Item ratings are summarised
into six subscales: Inattention, Hyperactivity/Impulsivity, Learning problems, Executive function, Aggression, and Peer relations.

2.3.1.2. Brief Rating Inventory of Executive Function


2.3.1.3. Children’s Communication Checklist

The Children’s Communication Checklist (CCC-2, Bishop, 2006; Norbury, Nash, Baird, & Bishop, 2004) focuses on communication strengths and weaknesses. Items are organised into ten subscales: Speech, Syntax, Semantics, Coherence; Inappropriate initiation, Stereotyped language, Use of context, Nonverbal communication, Social relations, and Interests.

2.3.2. Measures used for Subgroup Validation

The following measures were used to validate and further characterise the profiles identified by the community detection procedure.

2.3.2.1. Strengths and Difficulties Questionnaire

The Strengths and Difficulties Questionnaire (SDQ, Goodman, 1997; Goodman, 2001) measures everyday social and emotional functioning. Items are organised into five subscales, but only four were analysed in the current study: Emotional Problems (to index of the tendency to internalise problems); Conduct Problems (to index of the tendency to externalise problems), Peer Problems (to index the ability to build and maintain social relationships), and Prosocial (index of social strengths and resources, rather than difficulties). The focus on these subscales was motivated by extensive literature showing that communication and EF skills are related to
individual differences in these realms. The final SDQ subscale was omitted because it captures hyperactivity and conceptually overlaps with the Hyperactivity/Impulsivity subscale of the Conners that was included in the subgrouping. Age-uncorrected SDQ scores were used in all analyses. As a sensitivity check, the analyses were re-run using age-regressed scores, and the same conclusions were supported (see Table III.3).

2.3.2.2. Reading and Mathematics

To assess children’s academic abilities the Word Reading and Numerical Operations subtests of the Wechsler Individual Achievement Test-II (WIAT-II, Wechsler, 2005) were administered. Raw scores for each subtest were converted to standard scores ($M = 100, SD = 15$).

2.3.2.3. Nonverbal Cognitive Ability

The Matrix reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (WASI-II, Wechsler, 2011) was administered as an index of non-verbal cognitive ability. The raw score, corresponding to the number of correctly completed matrices, was converted into a T-score ($M = 50, SD = 10$).

2.4. Magnetic resonance imaging (MRI)

2.4.1. Data Acquisition

The magnetic resonance imaging (MRI) data were collected at the Medical Research Council Cognition and Brain Sciences Unit, Cambridge, United Kingdom on a Siemens 3 T Prisma-fit system using a 32-channel quadrature head coil. T1-weighted volume scans were acquired using a whole-brain coverage 3D Magnetization Prepared Rapid Acquisition Gradient Echo (MP RAGE) sequence acquired using 1 mm isometric image resolution. Echo time was 2.98ms, and repetition time was 2,250ms. Diffusion scans were obtained using echo-planar diffusion-weighted images with an isotropic set of 68 noncollinear directions, using a
weighting factor of $b = 1,000\text{s} \times \text{mm}^{-2}$ interleaved with 4 T2-weighted ($b = 0$) volume. Whole-brain coverage was based on 60 contiguous axial slices and an isometric image resolution of 2 mm. Echo time was 90ms and repetition time was 8500ms. Both MRI pre-processing and reconstruction were performed using QSIPrep 0.13.0RC1, which is based on Nipype 1.6.0 (Gorgolewski et al., 2011). All pre-processing steps are described below.

2.4.2. MRI pre-processing

2.4.2.1. Anatomical data pre-processing.

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) using N4BiasFieldCorrection (Tustison et al., 2010) and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped using antsBrainExtraction.sh (ANTs 2.3.1), using OASIS as target template. Spatial normalisation to the ICBM 152 Nonlinear Asymmetrical template version 2009c (RRID:SCR_008796, Fonov et al., 2009) was performed through nonlinear registration with antsRegistration (ANTs 2.3.1, RRID:SCR_004757, Avants et al., 2008), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid, white matter, and grey matter was performed on the brain-extracted T1w using FAST (FSL 6.0.3:b862cdd5, RRID:SCR_002823, Zhang et al., 2001).

2.4.2.2. Diffusion data pre-processing

Any images with a $b$-value less than 100 s/mm$^2$ were treated as a $b=0$ image. MP-PCA denoising as implemented in Mrtrix3’s dwidenoise (Veraart et al., 2016) was applied with a 5-voxel window. After MP-PCA, B1 field inhomogeneity was corrected using dwibiascorrect from Mrtrix3 with the N4 algorithm (Tustison et al., 2010). After B1 bias correction, the mean intensity of the DWI series was adjusted so all the mean intensity of the $b=0$ images matched across each separate DWI scanning sequence. FSL’s (version 6.0.3:b862cdd5) eddy was used for head motion correction and Eddy current correction (Andersson & Sotiropoulos, 2016). Eddy was configured with a $q$-space smoothing factor of
10, a total of 5 iterations, and 1000 voxels used to estimate hyperparameters. A linear first-level model and a linear second-level model were used to reorganise Eddy current-related spatial distortion. $Q$-space coordinates were forcefully assigned to shells. Field offset was attempted to be separated from subject movement. Shells were aligned post-eddy. Eddy’s outlier replacement was run (Andersson et al., 2016). Data were grouped by slice, only including values from slices determined to contain at least 250 intracerebral voxels. Groups deviating by more than 4 standard deviations from the prediction had their data replaced with imputed values. Final interpolation was performed using the “jac” method (Jenkinson et al., 2012). Framewise displacement was calculated using the implementation Nipype following the definitions by Power et al. (2014). Many internal operations of the QSIPrep pipeline use Nilearn 0.7.0 (Abraham et al., 2014) and Dipy (Garyfallidis et al., 2014). More details are available online at: [https://qsiprep.readthedocs.io](https://qsiprep.readthedocs.io).

2.4.2.3. DSI Studio Reconstruction

Diffusion orientation distribution functions were reconstructed using a generalised $q$-sampling imaging (GQI, Yeh et al., 2010) with a ratio of mean diffusion distance of 1.25. The standard GQI plus deterministic tractography pipeline was used (dsi_studio_gqi) (Yeh et al., 2013). 5 million streamlines were created with a maximum length of 250mm, minimum length of 30mm, random seeding with a step size of 1mm.

2.4.2.4. Connectome construction

Whole-brain white matter connectivity matrices (i.e., connectomes) were constructed for each child based on the Brainnetome atlas (Fan et al., 2016). For each pairwise combination of regions ($N = 246$), the number of streamlines intersecting them was estimated and transformed to a 246 x 246 streamline matrix.
2.5. Statistical Analysis Overview

A data-driven community detection algorithm was applied to the child-by-child associations across the behaviour, EF, and communication ratings in the referred sample (CALM800) to identify subgroups of children with similar profiles of behavioural and communication strengths and difficulties. The demographic and diagnostic characteristics, academic performance, socio-emotional ratings, and neural white matter connections were then compared across the derived subgroups and the comparison sample. The details for each step of the analysis are presented below.

2.5.1. Missing Data Handling

Missing data was estimated only for the variables used for community detection. The decision was made to keep all other data independent and suitable for external validation. In the referred sample, the missingness across the 24 parent rating subscales used in the community detection analysis ranged from 0.5% to 2.11%. Data was complete for 94% of participants. Missingness levels were similar across boys and girls ($\chi^2 (1) = 0.06, p = 0.80$) and across diagnosed and undiagnosed participants ($\chi^2 (1) = 0.0054, p = 0.94$). Age did not significantly predict missingness ($F (1,803) = 1.59, \beta = -.0004, p = 0.21$). For the non-referred sample, 97% of participants had complete data and missingness across variables ranged from 0.63% to 2.53%. Missingness levels were again similar across boys and girls ($\chi^2 (1) = 0.39, p = 0.53$) and age was not a significant predictor of missingness ($F (1,156) = 3.38, \beta = -0.0009, p = 0.07$). Missing data was separately estimated for each cohort using an identical procedure based on nonparametric missing value imputation using random forests with the R package `missForest`, (version 1.4; Stekhoven & Bühlmann, 2012) implemented with a maximum number of iterations set at 10 and number of trees set at 100 (i.e., package defaults). Participant sex and age were also added to the imputation model to improve estimates. The imputed datasets were used in the community detection (referred cohort only) and the principal
components analyses used to reduce the data (the two separately imputed datasets were combined).

2.5.2. Normality

Across the 24 subscales used in the community detection, three cases fell outside of ±3.5 SD of the sample mean, and skewness ranged between -1.18 and 1.20. Age was regressed from each subscale using a robust linear model fitting via the R package MASS (Venables & Ripley, 2002). Age-regressed residuals were saved. To limit the influence of univariate outliers, approximate normal distribution percent-ranks ranging from zero to one were estimated for each participant across tasks and the resulting ranks were mapped onto a standard normal distribution using a quantile function (Bignardi et al., 2021; Gregory, 2014). Following this procedure, no cases fell outside the ±3.5 SD of the sample mean range and no variable had a skewness value outside the ±1 range. The normalised residuals were used for all subsequent analyses. As a robustness check the community detection and principal component analyses were also conducted with the non-normalised residuals. No substantial differences regarding the number of subgroups and their characteristics or principal components were noted across the two approaches.

2.5.3. Community Detection

A network was built to represent the child-by-child correlations across the 24 subscales of the CCC-2, Conners-3, and BRIEF questionnaires (using age-regressed normalised residuals) for those in the referred sample. Child-by-child correlations were the focus of the community detection to capture similarities and differences in children’s profiles across the subscales. This method has been previously used to identify subgroups of children with similar neuropsychological profiles, temperaments, and EF-related behavioural problems (Bathelt et al., 2018; Fair et al., 2012; Karalunas et al., 2014). The community detection was based on the Louvain weight-conserving algorithm (gamma set at the default value of 1), as implemented in the Brain Connectivity Toolbox (BCT: Rubinov & Sporns, 2011). The strength of the
community separation was quantified using the asymmetric modularity ($Q$) index, which conceptually represents the overall segregation between identified communities, with values above 0.3 considered evidence for sufficient community separation (Newman & Girvan, 2004). This method performs well in recovering the true number of communities across a range of conditions (Gates et al., 2016). To reach a stable community assignment, the consensus clustering method was applied with 1000 iterations (Lancichinetti & Fortunato, 2012), a procedure that has demonstrated robustness even at substantial levels of noise in the data (Bathelt et al., 2018).

2.5.4. Subgroup Profiles and Behavioural Validation

A series of two-tailed t-tests with Holm corrected $p$-values were run to compare the profiles of the derived subgroups to each other, and the comparison group, to quantify variation in the severity of difficulties across data-driven groups and to capture differences relative to the non-referred children. To facilitate the interpretation of the subgroup profiles, the communication and behavioural data used in the community detection was reduced via Principal Component analyses (PCA) and component scores were extracted and compared across the groups. Data from both the referred and comparison samples was combined for the PCA, and the optimal number of components to retain was chosen based on the results of parallel analyses (note that repeating the PCA separately for the referred and non-referred samples supported the same conclusions). Additional group comparisons were conducted on measures not included in the community detection to externally validate and further characterise the profiles of the subgroups.

2.5.6 Neuroimaging Analysis

Neural white matter connectome data was analysed using graph theory. The nodes of the network were the 246 regions of interest (ROI-s) of the Brainnetome atlas (Fan et al., 2016) and the connections (or edges) corresponded to the number of streamlines intersecting each
pair of regions. All connectome analyses are based on weighted unthresholded matrices. This choice was motivated by observing that the application of consistency-thresholding (Roberts et al., 2017) to retain all but the top 30% or 20% of most consistent edges (on the assumption that the connections with the highest inter-subject variability are spurious) fully reproduced the results based on unthresholded matrices. I explored whether and how each derived subgroup differed from the comparison sample and did not compare the data-driven subgroups to one another. This approach was chosen to investigate both common and subgroup-unique differences among the referred children relative to the non-referred children. Due to the presence of outliers in the data, the metrics of the data-driven subgroups were compared to the comparison group using 10% trimmed-means pairwise $t$-tests with false-discovery rate correction, and robust explanatory measure of effect size based on trimmed means is reported ($\xi$; Wilcox & Tian, 2011).

2.5.6.1. Global and Local Metrics.

At the whole-brain level, I focussed on global efficiency and global clustering coefficients, metrics that have previously been linked to variation in children’s cognitive and academic skills (Bathelt et al., 2018). Global efficiency describes the potential for information exchange in the network and is the average inverse distance from any node (i.e., brain region) to any other node (Sporns et al., 2007). Global clustering quantifies the fraction of each node’s neighbours that are also neighbours of each other. Subsequently, to explore the local properties of the connectomes with the minimal number of comparisons, the 246 brain regions in the Brainnetome atlas (Fan et al., 2016) were grouped according to their corresponding intrinsic connectivity networks (ICN) as defined by Yeo et al. (2011): Default mode, Dorsal attention, Frontoparietal, Limbic, Somatomotor, and Visual networks Subcortical regions were also grouped together. The sum of all connections within each network (and the subcortex) of the connectome was calculated. This approach is based on the known structural and functional
organisation properties of the brain and was therefore favoured as a balanced alternative to hypothesis-driven pre-specified regions of interest approach or a fully data-driven reduction of the connectomes. For each data-driven subgroup, networks that showed a significant difference from the comparison group were selected for further analyses. Subsequently, the node strength (i.e., the sum of all connections) in each region within the networks flagged as significantly different was tested against the comparison group. The functional characterisation of these regions was based on the behavioural domain metadata labels of the BrainMap Database (www.brainmap.org/taxonomy), which uses both forward and reverse inferences (Eickhoff & Grefkes, 2011; Fox et al., 2014). The global clustering metric was normalised according to the average of 1000 random graphs using the algorithms for weighted undirected networks implemented in the GRETNA toolbox (Wang et al., 2015). Prior to the analysis, age, gender, and motion artefacts (average frame displacement) were regressed from each graph metric using a robust regression approach to account for outliers.

2.5.6.2. Network-Based Statistics.

To investigate differences in the connections between regions network-based statistics was used (Zalesky et al., 2010). This involved comparing each data-driven subgroup to the comparison group in a series of between-group *t*-tests for any pair of regions (i.e., edges). For each group comparison, a pre-determined threshold of Cohen’s *d* effect size ≥ 0.4 (i.e., *t* = 2.8) was used to select which pairs of regions would be chosen for further analyses. Pairs of regions that met this condition were then used to identify subnetworks for which a path could be found between any two regions. A family-wise error (FWE)-corrected *p*-value was then obtained for each subnetwork via permutation tests (*N* = 1000). For each permutation, the children were randomly exchanged between the data-driven subgroup and the comparison group. The size of the subnetwork can be measured in terms of its extent (i.e., the total number of connections) or intensity (i.e., the sum of test statistic values across all connections). The probability of
observing a subnetwork of a given extent or intensity is estimated as the proportion of permutations for which the largest subnetwork was the same size or greater. In the current analyses, the focus was on intensity and subnetworks which met the FWE-corrected threshold of \( p < 0.01 \) were considered significantly different between groups. Notably, the FWE-correction does not allow inferences to be drawn at the level of individual connections. Age, gender, and average frame displacement were used as covariates in all analyses to mitigate their potential confounding effects. To ensure the robustness of the results, I repeated all procedures by varying the target \( t \)-threshold, exploring the extent-based measure of size, and setting a more liberal FWE-corrected \( p \)-value threshold (\( p < .05 \)).

3. Results

3.1. Community Detection

Pearson correlations across the 24 subscales are presented in Figure III.1 (see Figure III.2 for Pearson correlations of age-regressed normalised residuals). Community detection applied to the child-by-child correlation network based on the Louvain algorithm identified three subgroups of children (\( Q = .43 \); see Figure III.3 for community separation). The descriptive statistics of each cluster are displayed in Table III.2.
Figure III.1

Pearson correlations across all subscales of Conners, BRIEF, and CCC-2

Note. Conners (Conners Parent Rating Short Form 3rd Edition) subscales; INT = Inattention; HYP = Hyperactivity/Impulsivity; LRN = Learning Problems; AGG = Aggression; PEER = Peer Relationships; BRIEF (Brief Rating Inventory of Executive Function) subscales: INH = Inhibition; SHIF = Shifting; EMO = Emotional Control; INIT = Initiation; WM = Working memory; PLAN = Planning/Organisation; ORG = Organisation of Materials; MONT = Monitoring; CCC-2 (Children’s Communication Checklist) subscales: SYN = Syntax; SEM = Semantics; COH = Coherence; INAP = Inappropriate Initiation; STER = Stereotyped Language; CONT = Use of Context; I = Nonverbal Communication; SOC = Social Relations; INTR = Interests.
Note. Conners (Conners Parent Rating Short Form 3rd Edition) subscales; INT = Inattention; HYP = Hyperactivity/Impulsivity; LRN = Learning Problems; AGG = Aggression; PEER = Peer Relationships; BRIEF (Brief Rating Inventory of Executive Function) subscales: INH = Inhibition; SHIF = Shifting; EMO = Emotional Control; INIT = Initiation; WM = Working memory; PLAN = Planning/Organisation; ORG = Organisation of Materials; MONT = Monitoring; CCC-2 (Children’s Communication Checklist) subscales: SYN = Syntax; SEM = Semantics; COH = Coherence; INAP = Inappropriate Initiation; STER = Stereotyped Language; CONT = Use of Context; I= Nonverbal Communication; SOC = Social Relations; INTR = Interests.
Figure III.3

Child-by-child correlation matrix in Fruchterman-Reingold layout colour-coded according to the results of the community detection algorithm.

Note. Correlation coefficients below absolute value .5 were removed for visualisation purposes.
3.2. Subgroup Profiles

The profiles of the subgroups identified by the community detection and the comparison group across the 24 subscales are displayed in Figure III.5a. Descriptive statistics for each measure included in the community detection for each subgroup and the non-referred comparison group are presented in Table III.1, along with the group comparisons and Cohen’s $d$ effect sizes. Across the 24 subscales, subgroups showed substantial differences (often with large effect sizes) from the non-referred sample; and in some cases, large effect sizes were observed in their differences from one another (Table III.1). Broadly, Subgroup 1 (S1, $N = 300$) mostly differed from the other groups on the CCC-2 subscales measuring structural language skills (Speech, Syntax, Semantics). Subgroup 2 (S2, $N = 261$) mostly differed on the subscales EF, Learning (Conners-3), and Working memory (BRIEF). And Subgroup 3 (S3, $N = 244$) showed consisted differences on subscales Emotional control (BRIEF), Aggression (Conners-3), and Social relations (CCC-2). To capture group differences more concisely and to reduce the number of comparisons between groups, data reduction analysis was used in the next steps.
Table III.1
Descriptive statistics of raw scores derived from the CCC-2, Conners-3, and BRIEF questionnaires and Cohen’s d effect size for comparisons across data-driven subgroups and non-referred sample.

<table>
<thead>
<tr>
<th>Scale</th>
<th>S1 M</th>
<th>S1 SD</th>
<th>S2 M</th>
<th>S2 SD</th>
<th>S3 M</th>
<th>S3 SD</th>
<th>Ref. M</th>
<th>Ref. SD</th>
<th>S1 vs. S2</th>
<th>S1 vs. S3</th>
<th>S2 vs. S3</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>10.5</td>
<td>3.9</td>
<td>12.7</td>
<td>2.7</td>
<td>12.0</td>
<td>3.0</td>
<td>4.1</td>
<td>4.0</td>
<td>-0.2</td>
<td>0.4</td>
<td>-1.6</td>
<td>-2.6</td>
</tr>
<tr>
<td>HYP</td>
<td>8.9</td>
<td>5.6</td>
<td>10.7</td>
<td>5.4</td>
<td>13.6</td>
<td>4.5</td>
<td>4.4</td>
<td>4.5</td>
<td>0.3</td>
<td>0.6</td>
<td>-0.9</td>
<td>-1.3</td>
</tr>
<tr>
<td>LRN</td>
<td>10.4</td>
<td>3.5</td>
<td>10.4</td>
<td>3.1</td>
<td>8.2</td>
<td>3.6</td>
<td>2.6</td>
<td>3.0</td>
<td>0.0</td>
<td>-0.7</td>
<td>-0.6</td>
<td>-2.4</td>
</tr>
<tr>
<td>EF</td>
<td>8.6</td>
<td>3.8</td>
<td>11.8</td>
<td>2.9</td>
<td>10.4</td>
<td>3.2</td>
<td>4.3</td>
<td>3.7</td>
<td>0.9</td>
<td>-0.4</td>
<td>0.5</td>
<td>-1.1</td>
</tr>
<tr>
<td>AGG</td>
<td>2.7</td>
<td>3.4</td>
<td>2.4</td>
<td>3.0</td>
<td>5.9</td>
<td>4.5</td>
<td>1.3</td>
<td>2.3</td>
<td>-0.1</td>
<td>0.9</td>
<td>0.8</td>
<td>-0.4</td>
</tr>
<tr>
<td>PEER</td>
<td>5.1</td>
<td>4.2</td>
<td>3.8</td>
<td>3.8</td>
<td>8.0</td>
<td>4.2</td>
<td>1.6</td>
<td>2.9</td>
<td>-0.3</td>
<td>1.1</td>
<td>0.7</td>
<td>-0.9</td>
</tr>
<tr>
<td>INH</td>
<td>19.3</td>
<td>5.9</td>
<td>20.5</td>
<td>6.0</td>
<td>25.6</td>
<td>4.5</td>
<td>14.4</td>
<td>4.8</td>
<td>0.2</td>
<td>1.0</td>
<td>1.2</td>
<td>-0.9</td>
</tr>
<tr>
<td>SHIF</td>
<td>16.3</td>
<td>4.3</td>
<td>16.0</td>
<td>4.0</td>
<td>19.2</td>
<td>3.6</td>
<td>11.9</td>
<td>3.7</td>
<td>-0.1</td>
<td>0.8</td>
<td>0.7</td>
<td>-1.1</td>
</tr>
<tr>
<td>EMO</td>
<td>20.6</td>
<td>5.4</td>
<td>20.8</td>
<td>5.5</td>
<td>25.5</td>
<td>4.3</td>
<td>16.1</td>
<td>4.7</td>
<td>0.0</td>
<td>1.0</td>
<td>1.0</td>
<td>-0.9</td>
</tr>
<tr>
<td>INIT</td>
<td>17.0</td>
<td>3.6</td>
<td>18.0</td>
<td>3.1</td>
<td>18.3</td>
<td>3.0</td>
<td>12.8</td>
<td>3.5</td>
<td>0.3</td>
<td>0.1</td>
<td>0.4</td>
<td>-1.2</td>
</tr>
<tr>
<td>WM</td>
<td>24.6</td>
<td>4.5</td>
<td>27.0</td>
<td>3.3</td>
<td>25.7</td>
<td>3.8</td>
<td>15.6</td>
<td>5.1</td>
<td>0.6</td>
<td>-0.4</td>
<td>0.3</td>
<td>-1.9</td>
</tr>
<tr>
<td>PLAN</td>
<td>27.0</td>
<td>5.4</td>
<td>30.6</td>
<td>4.2</td>
<td>29.6</td>
<td>4.4</td>
<td>19.3</td>
<td>6.0</td>
<td>0.8</td>
<td>-0.2</td>
<td>0.5</td>
<td>-1.4</td>
</tr>
<tr>
<td>ORG</td>
<td>13.1</td>
<td>3.5</td>
<td>16.0</td>
<td>2.4</td>
<td>15.1</td>
<td>2.8</td>
<td>12.0</td>
<td>3.3</td>
<td>1.0</td>
<td>-0.4</td>
<td>0.6</td>
<td>-0.3</td>
</tr>
<tr>
<td>MONT</td>
<td>17.7</td>
<td>3.9</td>
<td>19.3</td>
<td>3.0</td>
<td>20.4</td>
<td>2.9</td>
<td>13.0</td>
<td>3.8</td>
<td>0.5</td>
<td>0.4</td>
<td>0.8</td>
<td>-1.2</td>
</tr>
<tr>
<td>SPE</td>
<td>7.8</td>
<td>5.4</td>
<td>3.7</td>
<td>4.1</td>
<td>3.5</td>
<td>3.9</td>
<td>1.6</td>
<td>2.7</td>
<td>-0.8</td>
<td>-0.1</td>
<td>-0.9</td>
<td>-1.3</td>
</tr>
<tr>
<td>SYN</td>
<td>8.1</td>
<td>5.1</td>
<td>3.3</td>
<td>3.5</td>
<td>3.5</td>
<td>3.2</td>
<td>1.2</td>
<td>2.3</td>
<td>-1.1</td>
<td>0.1</td>
<td>-1.1</td>
<td>-1.6</td>
</tr>
<tr>
<td>SEM</td>
<td>10.6</td>
<td>4.7</td>
<td>6.6</td>
<td>4.4</td>
<td>6.7</td>
<td>3.9</td>
<td>1.9</td>
<td>2.5</td>
<td>-0.9</td>
<td>0.0</td>
<td>-0.9</td>
<td>-2.2</td>
</tr>
<tr>
<td>COH</td>
<td>10.3</td>
<td>5.1</td>
<td>6.6</td>
<td>4.7</td>
<td>8.6</td>
<td>4.7</td>
<td>2.2</td>
<td>3.0</td>
<td>-0.8</td>
<td>0.4</td>
<td>-0.4</td>
<td>-1.8</td>
</tr>
<tr>
<td>INAP</td>
<td>10.6</td>
<td>5.6</td>
<td>8.6</td>
<td>5.3</td>
<td>13.5</td>
<td>5.1</td>
<td>3.7</td>
<td>3.4</td>
<td>-0.4</td>
<td>1.0</td>
<td>0.5</td>
<td>-1.4</td>
</tr>
<tr>
<td>STER</td>
<td>7.2</td>
<td>4.6</td>
<td>4.1</td>
<td>3.4</td>
<td>6.9</td>
<td>4.0</td>
<td>1.9</td>
<td>2.5</td>
<td>-0.8</td>
<td>0.8</td>
<td>-0.1</td>
<td>-1.3</td>
</tr>
<tr>
<td>CONT</td>
<td>11.0</td>
<td>5.3</td>
<td>6.9</td>
<td>4.9</td>
<td>10.9</td>
<td>5.2</td>
<td>2.7</td>
<td>3.3</td>
<td>-0.8</td>
<td>0.8</td>
<td>0.0</td>
<td>-1.8</td>
</tr>
<tr>
<td>NVER</td>
<td>8.1</td>
<td>5.0</td>
<td>5.5</td>
<td>4.4</td>
<td>9.9</td>
<td>4.8</td>
<td>2.5</td>
<td>3.3</td>
<td>-0.6</td>
<td>0.9</td>
<td>0.4</td>
<td>-1.2</td>
</tr>
<tr>
<td>SOC</td>
<td>6.6</td>
<td>5.0</td>
<td>4.4</td>
<td>4.0</td>
<td>9.2</td>
<td>4.5</td>
<td>2.0</td>
<td>2.9</td>
<td>-0.5</td>
<td>1.1</td>
<td>0.6</td>
<td>-1.1</td>
</tr>
<tr>
<td>INTR</td>
<td>9.2</td>
<td>4.5</td>
<td>7.1</td>
<td>3.9</td>
<td>11.5</td>
<td>4.2</td>
<td>4.7</td>
<td>3.3</td>
<td>-0.5</td>
<td>1.1</td>
<td>0.5</td>
<td>-1.1</td>
</tr>
</tbody>
</table>

Note. Conners (Conners Parent Rating Short Form 3rd Edition) subscales; INT = Inattention; HYP = Hyperactivity/Impulsivity; LRN = Learning Problems; AGG = Aggression; PEER = Peer Relationships; BRIEF (Brief Rating Inventory of Executive Function) subscales: INH = Inhibition; SHIF = Shifting; EMO = Emotional Control; INIT = Initiation; WM = Working memory; PLAN = Planning/Organisation; ORG = Organisation of Materials; MONT = Monitoring; CCC-2 (Children’s Communication Checklist) subscales: SYN = Syntax; SEM = Semantics; COH = Coherence; INAP = Inappropriate Initiation; STER = Stereotyped Language; CONT = Use of Context; I= Nonverbal Communication; SOC = Social Relations; INTR = Interests. S1 = Subgroup 1, S2 = Subgroup 2, S3 = Subgroup 3, Nref. = Non-referred group. Cohen’s d > |1| presented in bold.
3.3. Data Reduction

To aid the interpretation of the differences across the groups, data was reduced with PCAs. The parent rating data from the CCC-2, Conners, and BRIEF ratings were combined across the two samples (referred and comparison). PCA was then used separately on the communication (i.e., all CCC-2 ratings) and behavioural data (i.e., Conners and BRIEF ratings). An orthogonal rotation (i.e., varimax) was used to maximise differences between the extracted components. The maximum number of components to retain was chosen based on the results of parallel analysis. All component structures from one to the maximum suggested in the parallel analysis were considered. Component structures were considered interpretable if each component had at least three primary loadings (highest loading ≥0.35 and at least 0.10 greater than all other loadings). For both the behavioural and communication data, the two-component solution was the one that met these criteria (see Figure III.4 for output of parallel analyses and Figure III.5b and Figure III.5c for loadings). For the CCC-2 data, the two principal components in combination explained 80% of the variance. The first component captured pragmatic communication (accounting for 46% of variance), and the second structural language skills (accounting for 34% of variance). In the behavioural data, the two components cumulatively explained 68% of the variance. The first component (accounting for 37% of variance) was labelled cool EF, and the second component (accounting for 31% of variance) hot EF. The components resembled the clusters of difficulties identified in Chapter II (Figure II.2). Component scores were found using the regression method (for details see Revelle, 2019) and were saved for subsequent analyses.
Figure III.4

Parallel analysis of the CCC-2 subscales (right) and Conners and BRIEF (left)

Note. The simulated line is the top of 95% confidence interval around simulated eigenvalues.

CCC-2 = Children communications checklist 2; Conners = Conners Parent Rating Short Form 3rd Edition; BRIEF = Brief Rating Inventory of Executive Function.
3.4. Profile Comparisons

Group profiles across the four components are displayed in Figure III.5d. The data-driven subgroups had significantly more difficulties across all components relative to the comparison sample, except for S2 who had similar hot EFs to the comparison sample. The children in each subgroup had different patterns of specific strengths and difficulties. Children in S1 were characterised by poor structural language, children in S2 had weaknesses in cool EF, and children in S3 had pronounced difficulties with hot EFs and pragmatics (for descriptive statistics see Table III.3). The same patterns were evident in the subgroup profiles on the subscales that most strongly loaded on the respective principal components (see Table III.1 and Figure III.5a). There were no significant differences in age across the subgroups ($F(1,803) = 3.138, p = 0.08$, see Table III.2). Using a chi-square test, which compared the overall proportion of boys and girls in the non-referred sample to the proportion observed in each subgroup, it was observed that girls were overrepresented in S1, while boys were overrepresented in S3 (Table III.2). The same chi-square analyses were used to investigate whether there were differences in the prevalence of different diagnoses across subgroups. The focus was on children with ADHD and ASD because overall numbers for other diagnoses were too low to make meaningful comparisons. Children with ADHD and ASD were represented in each subgroup, but children with ADHD were overrepresented in S3 and underrepresented S1, and children with ASD were underrepresented in S2 and overrepresented in S3 (Table III.2). Finally, children who had recently received Speech and Language Therapy were overrepresented in S1 and underrepresented in the other two subgroups (Table III.2).
Table III.2

Subgroup demographics.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Age</th>
<th>Sex</th>
<th>ADHD</th>
<th>ASD</th>
<th>SLT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M</td>
<td>SD</td>
<td>boys</td>
<td>ch²</td>
</tr>
<tr>
<td>S1</td>
<td>300</td>
<td>9.6</td>
<td>2.5</td>
<td>184</td>
<td>7.3</td>
</tr>
<tr>
<td>S2</td>
<td>261</td>
<td>9.7</td>
<td>2.4</td>
<td>174</td>
<td>0.4</td>
</tr>
<tr>
<td>S3</td>
<td>244</td>
<td>9.2</td>
<td>2.2</td>
<td>194</td>
<td>13.5</td>
</tr>
</tbody>
</table>

Note. ADHD = Attention deficit hyperactivity disorder; ASD = Autism spectrum disorders; SLT = Speech and Language therapy in the last two years or ongoing. S1 = Subgroup 1 (most severe structural language difficulties); S2 = Subgroup 2 (most severe cool executive difficulties); S3 = Subgroup 3 (most severe difficulties with hot executive skills & pragmatic communication).
### Table III.3

*Descriptive statistics across the data-driven subgroups and the comparison sample for all behavioural measures*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Subgroup 1</th>
<th></th>
<th></th>
<th></th>
<th>Subgroup 2</th>
<th></th>
<th></th>
<th></th>
<th>Subgroup 3</th>
<th></th>
<th></th>
<th>Non-referred</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$N$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$N$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$N$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$N$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$N$</td>
</tr>
<tr>
<td>SDQ: Emotion problems (raw)</td>
<td>297</td>
<td>4.22</td>
<td>2.70</td>
<td>259</td>
<td>3.96</td>
<td>2.76</td>
<td>240</td>
<td>5.17</td>
<td>2.76</td>
<td>157</td>
<td>1.86</td>
<td>2.27</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ: Conduct problems (raw)</td>
<td>297</td>
<td>2.84</td>
<td>2.29</td>
<td>259</td>
<td>2.95</td>
<td>2.34</td>
<td>241</td>
<td>4.93</td>
<td>2.46</td>
<td>157</td>
<td>1.47</td>
<td>1.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ: Peer problems (raw)</td>
<td>297</td>
<td>3.27</td>
<td>2.59</td>
<td>259</td>
<td>2.40</td>
<td>2.30</td>
<td>241</td>
<td>4.74</td>
<td>2.42</td>
<td>157</td>
<td>1.24</td>
<td>1.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ: Prosocial (raw)</td>
<td>297</td>
<td>7.14</td>
<td>2.41</td>
<td>259</td>
<td>7.42</td>
<td>2.20</td>
<td>241</td>
<td>5.76</td>
<td>2.17</td>
<td>157</td>
<td>8.15</td>
<td>1.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ: Emotion problems (res.)</td>
<td>297</td>
<td>0.22</td>
<td>2.68</td>
<td>259</td>
<td>-0.05</td>
<td>2.75</td>
<td>240</td>
<td>1.2</td>
<td>2.74</td>
<td>157</td>
<td>-2.17</td>
<td>2.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ: Conduct problems (res.)</td>
<td>297</td>
<td>-0.33</td>
<td>2.29</td>
<td>259</td>
<td>-0.23</td>
<td>2.34</td>
<td>241</td>
<td>1.75</td>
<td>2.46</td>
<td>157</td>
<td>-1.7</td>
<td>1.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ: Peer problems (res.)</td>
<td>297</td>
<td>0.2</td>
<td>2.53</td>
<td>259</td>
<td>-0.69</td>
<td>2.27</td>
<td>241</td>
<td>1.72</td>
<td>2.37</td>
<td>157</td>
<td>-1.89</td>
<td>1.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SDQ: Prosocial (res.)</td>
<td>297</td>
<td>0.11</td>
<td>2.4</td>
<td>259</td>
<td>0.39</td>
<td>2.19</td>
<td>241</td>
<td>-1.29</td>
<td>2.17</td>
<td>157</td>
<td>1.13</td>
<td>1.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WIAT-II: Numerical operations</td>
<td>265</td>
<td>81.25</td>
<td>15.74</td>
<td>233</td>
<td>85.38</td>
<td>15.02</td>
<td>227</td>
<td>90.58</td>
<td>17.91</td>
<td>157</td>
<td>115.52</td>
<td>18.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WIAT-II: Word reading</td>
<td>294</td>
<td>80.75</td>
<td>15.64</td>
<td>255</td>
<td>88.49</td>
<td>16.07</td>
<td>236</td>
<td>93.17</td>
<td>16.62</td>
<td>156</td>
<td>108.68</td>
<td>12.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WASI-II: Matrix reasoning</td>
<td>300</td>
<td>40.96</td>
<td>8.77</td>
<td>261</td>
<td>43.48</td>
<td>9.33</td>
<td>242</td>
<td>45.14</td>
<td>10.26</td>
<td>158</td>
<td>53.28</td>
<td>9.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural language (fs.)</td>
<td>300</td>
<td>-0.82</td>
<td>0.78</td>
<td>261</td>
<td>0.08</td>
<td>0.85</td>
<td>244</td>
<td>0.47</td>
<td>0.79</td>
<td>158</td>
<td>0.7</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pragmatic Communication (fs.)</td>
<td>300</td>
<td>0.01</td>
<td>0.9</td>
<td>261</td>
<td>0.24</td>
<td>0.79</td>
<td>244</td>
<td>-0.83</td>
<td>0.78</td>
<td>158</td>
<td>0.87</td>
<td>0.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cool EF (fs.)</td>
<td>300</td>
<td>-0.02</td>
<td>0.8</td>
<td>261</td>
<td>-0.8</td>
<td>0.78</td>
<td>244</td>
<td>0.12</td>
<td>0.72</td>
<td>158</td>
<td>1.19</td>
<td>0.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hot EF (fs.)</td>
<td>300</td>
<td>0.16</td>
<td>0.89</td>
<td>261</td>
<td>0.41</td>
<td>0.87</td>
<td>244</td>
<td>-0.95</td>
<td>0.74</td>
<td>158</td>
<td>0.48</td>
<td>0.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. WIAT-II = Wechsler Individual Achievement Test-II; WASI-II = Wechsler Abbreviated Scale of Intelligence II; SDQ = Strengths and difficulties questionnaire, parent-report. Res. = age-regressed residual; fs. = factor score.*
Figure III.5
Profiles of the data-driven subgroups and the comparison group.

Panel A shows group profiles across all BRIEF, Conners, and CCC-2 subscales used in community detection. To put all measures on the same scale, age-referenced subscale scores were converted to z-scores where higher values indicate strengths and lower values indicate difficulties. Panel B shows the varimax loadings from Principal component analyses (PCA) of the CCC-2 subscales based on combined data from the referred and comparison samples. The first principal component (PC1) explained 46% of the variance and was labelled Pragmatic communication, the second principal component (PC2) explained an additional 34% of the variance and was labelled Structural language. Panel C shows the loadings from the same analyses applied to the BRIEF and Conners data. PC1 explained 37% of the variance and was labelled Cool executive function (EF) and PC2 explained an additional 31% of the variance and was labelled Hot EF. Panel D shows group performance across the four components identified in the PCA: higher values indicate strengths and lower values indicate difficulties. Comparisons are based on two-tailed t-tests, p-values are Holm-corrected. *** p < .001; ** p < .01, * p < .05.

S1 = Subgroup 1; S2 = Subgroup 2; S3 = Subgroup 3; Conners (Conners Parent Rating Short Form 3rd Edition) subscales; INT = Inattention; HYP = Hyperactivity/Impulsivity; LRN = Learning Problems; AGG = Aggression; PEER = Peer Relationships; BRIEF (Brief Rating Inventory of Executive Function) subscales; INH = Inhibition; SHIF = Shifting; EMO = Emotional Control; INIT = Initiation; WM = Working memory; PLAN = Planning/Organisation; ORG = Organisation of Materials; MONT = Monitoring; CCC-2 (Children’s Communication Checklist) subscales; SYN = Syntax; SEM = Semantics; COH = Coherence; INAP = Inappropriate Initiation; STER = Stereotyped Language; CONT = Use of Context; NVER = Nonverbal Communication; SOC = Social Relations; INTR = Interests.
3.5. Subgroup Profiles: Behavioural Validation

The groups were compared on external measures not used in the community detection (see Figure III.6 and Table III.3). In terms of socio-emotional functioning and cognitive and academic performance, all three subgroups derived from the referred sample had significantly more difficulties than the comparison sample (Figure III.6). There were differences in the patterns of severity among the subgroups. Children in S1 had the most pronounced difficulties with maths, reading, and nonverbal reasoning, scoring significantly lower than the other two data-driven groups. Children in S3 had higher performance than the other two subgroups on maths and reading, but significantly more difficulties with emotion, conduct, peer relations, and prosocial behaviour (Figure III.6). Overall, there was a consistent pattern in which children in S3 had more pronounced socio-emotional difficulties, whereas children in S1 and S2 showed more difficulties with cognitive and academic skills.
Figure III.6

Group comparisons across measures of socio-emotional functioning, cognitive, and academic skills.

Note. The top panel includes the Emotion problems, Conduct problems, Peer problems, and Prosocial subscales derived from the Strengths and difficulties questionnaire (SDQ, parent-report). Higher scores indicate more difficulties, except the Prosocial scale, which has the reverse interpretation. The bottom panel shows performance on the Word reading and Numerical operations subsets of the Wechsler Individual Achievement Test-II, and the Matrix reasoning subset of Wechsler Abbreviated Scale of Intelligence II. All scores are based on age-referenced norms: higher scores indicate better performance. The grey-dotted line represents the age-expected mean. For both panels, the observed mean in each group is represented by the large dot within each boxplot. All comparisons are based on two-tailed t-tests, p-values are Holm-corrected. S1 = Subgroup 1 (most severe structural language difficulties); S2 = Subgroup 2 (most severe cool executive difficulties); S3 = Subgroup 3 (most severe difficulties with hot executive skills & pragmatic communication).

*** p < .001; ** p < .01; * p < .05
3.6. Subgroup Profiles: Neural White Matter

3.6.1. Graph Measures.

At the global level, S1 and S2 had significantly lower global efficiency relative to the comparison sample (S1, \( N_{\text{imaging}} = 110; p_{\text{corrected}} = .04 \); S2, \( N_{\text{imaging}} = 121; p_{\text{corrected}} = .04 \)). The difference between the comparison sample and S3 (\( N_{\text{imaging}} = 82 \)) was not significant (\( p_{\text{corrected}} = .10 \)). There were no differences in global clustering between the comparison sample and any of the subgroups (uncorrected \( p\)-s > 0.25). These results are presented in Figure III.7. At the ICN-level, all subgroups had weaker overall connections in the limbic network and subcortical areas relative to the comparison sample (see Figure III.8). Follow-up analyses showed that relative to the non-referred sample all subgroups had reduced regional strength in the same sub-regions of the basal ganglia, hippocampus, thalamus, inferior temporal gyrus, and fusiform gyrus (see Figure III.9). Some unique patterns of reduced regional strength were observed for each subgroup relative to the comparison group (Figure III.9 and Appendix II Tables AII.1 and AII.2 for descriptive statistics, \( p\)-values and effect sizes). For children in S1, unique differences were observed in sub-regions of the basal ganglia and inferior temporal gyrus, which the BrainMap taxonomy labels as primarily implicated in cognition, emotion, action execution, memory, and semantics. For S2, unique differences were within parahippocampal gyrus, superior temporal gyrus, inferior temporal gyrus, amygdala, areas that have previously been implicated in cognition, emotion, social cognition, and perception. Finally, for children in S3, these differences were in the BG, cingulate gyrus, and thalamus, which are linked to action execution, perception, somesthesis, and cognition.
Figure III.7

Comparison of the global organisation of the white matter connectome between data-driven subgroups and the non-referred group

![Graph showing Global efficiency and Global clustering across different groups](image)

**Note.** Comparisons are based on 10% trimmed means to account for the presence of outliers and *p*-values are adjusted with false discovery rate correction. S1 = Subgroup 1 (most severe structural language difficulties); S2 = Subgroup 2 (most severe cool executive difficulties); S3 = Subgroup 3 (most severe difficulties with hot executive skills & pragmatic communication). Ns = not significant.

* *p corrected* < 0.05
For each subgroup, the overall network connectivity strength within the intrinsic connectivity networks (defined by Yeo et al., 2011) and the subcortex was compared to connectivity observed in the non-referred sample (CALM200). The location of the dots shows the estimated effect size based on a 10% trimmed mean comparison used to account for the presence of outliers, the color of the dots denotes whether the corresponding false-discovery-rate adjusted $p$-value was significant. The width of the lines corresponds to the 95% Confidence intervals.

**Note.** Subgroup 1: most severe structural language difficulties; Subgroup 2: most severe cool executive difficulties; Subgroup 3: most severe difficulties with hot executive skills & pragmatic communication.
Figure III.9
Comparison of the regional strength of connections across groups

Panel A shows regions of interest (ROI-s) within the subcortex and the limbic network that showed reduced connection strength in all data-driven subgroups relative to the comparison group. Panel B shows subcortical and limbic ROI-s that had reduced connection strength in two subgroups relative to the comparison group. Note that non-significant comparisons are omitted from the figure. Panel C1 shows subcortical and limbic ROI-s that were significantly different between S1 and the comparison group. Panel C2 shows subcortical and limbic ROI-s that were significantly different between S2 and the comparison group. Panel C3 shows subcortical and limbic ROI-s that were significantly different between S3 and the comparison group. S1 = Subgroup 1 (most severe structural language difficulties); S2 = Subgroup 2 (most severe cool executive difficulties); S3 = Subgroup 3 (most severe difficulties with hot executive skills & pragmatic communication). See Tables AII.1 and AII.2 in Appendix II for descriptive statistics, p-values and effect sizes.
3.6.2. Network-based Statistics

Subnetworks that were significantly different for each subgroup relative to the comparison group are displayed in Figure III.10. There was greater connectivity in the non-referred group relative to S1 in a subnetwork comprising five connections across limbic—subcortical and subcortical—subcortical areas ($p_{corrected} = .001$). This subnetwork involved six ROI-s located within the basal ganglia, orbital gyrus, and thalamus (see Table III.7 for ROI labels). For S2, there was reduced connectivity relative to the comparison group in a subnetwork spanning nine connections across ten regions ($p_{corrected} = .003$). This subnetwork included connections across the following networks: somatomotor — dorsal attention, limbic — dorsal attention, dorsal attention — visual, frontoparietal — somatomotor, visual — visual, and dorsal attention — subcortical. The involved ROI-s fell within the lateral occipital cortex, medioventral occipital cortex, orbital gyrus, parahippocampal gyrus, precuneus, superior parietal lobule, superior temporal gyrus, and thalamus (Table III.8). Finally, S3 had reduced connection strength relative to the comparison group in a subnetwork comprised of eight links ($p_{corrected} = .002$) across limbic — subcortical, limbic — visual, visual — subcortical, and subcortical — subcortical ROI-s. The eight ROI-s involved formed part of the basal ganglia, lateral occipital cortex, middle frontal gyrus, orbital gyrus, precuneus (the same ROI was also part of the network that differentiated S2), and thalamus (Table III.9). Sensitivity analyses supported the robustness of the findings: taking the extent-based measure of size fully reproduced the results; and setting the $t$-threshold in the range 2.6 to 3, produced similar results, which followed the expected tendency to discover larger subnetworks at lower $t$-thresholds (see Appendix II, Figure AII.2; for an investigation of the influence of the $t$-threshold see Beare et al., 2017). Results based on the more liberal FWE-corrected threshold of $p < 0.05$ suggested further differences in connectivity for S3 (see Appendix II, Figure AII.1).
Overall, the three subgroups showed a mix of similar and subgroup-unique patterns of differentiation from the comparison sample. At the global level, the two groups (S1 and S2) with the most pronounced difficulties with academic achievement showed reduced global efficiency compared to the non-referred group. At the ICN-level, all subgroups showed reduced connection strengths in the limbic network and the subcortex relative to the comparison sample. There were both common and subgroup-specific differences in the connection strength of regions within the limbic network and the subcortex. Finally, at a whole-connectome level, we identified subnetworks that differentiated each subgroup from the comparison sample.
Figure III.10

Subnetworks of the neural white matter connectome identified as significantly weaker in each data-driven subgroup relative to the comparison group.

Note. Subgroup 1: most severe structural language difficulties; Subgroup 2: most severe cool executive difficulties; Subgroup 3: most severe difficulties with hot executive skills & pragmatic communication. For nodal labels, see Tables III.7 – III.9.
Table III.7

*Edges within significant subnetwork identified for Subgroup 1*

<table>
<thead>
<tr>
<th>Regions of interest (Gyrus, Intrinsic connectivity network)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Left medial area 11 (Orbital gyrus, Limbic)</td>
<td>Left nucleus accumbens (Basal ganglia, Subcortical)</td>
</tr>
<tr>
<td>Left medial area 11 (Orbital gyrus, Limbic)</td>
<td>Left medial pre-frontal thalamus (Thalamus, Subcortical)</td>
</tr>
<tr>
<td>Left medial area 11 (Orbital gyrus, Limbic)</td>
<td>Left occipital thalamus (Thalamus, Subcortical)</td>
</tr>
<tr>
<td>Left medial pre-frontal thalamus (Thalamus, Subcortical)</td>
<td>Left caudal temporal thalamus (Thalamus, Subcortical)</td>
</tr>
<tr>
<td>Left medial area 11 (Orbital gyrus, Limbic)</td>
<td>Left lateral pre-frontal thalamus (Thalamus, Subcortical)</td>
</tr>
</tbody>
</table>

*Note.* Subgroup 1: most severe structural language difficulties.
### Table III.8

*Edges within significant subnetwork identified for Subgroup 2*

<table>
<thead>
<tr>
<th>Regions of interest (Gyrus, Intrinsic connectivity network)</th>
<th>Edges within significant subnetwork identified for Subgroup 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right lateral area 11 (Orbital gyrus, Frontoparietal)</td>
<td>Right TE1.0/TE1.2 (Superior temporal Gyrus, Somatomotor)</td>
</tr>
<tr>
<td>Right TE1.0/TE1.2 (Superior temporal gyrus, Somatomotor)</td>
<td>Right caudal area 7</td>
</tr>
<tr>
<td>Right caudal area 22 (Superior temporal gyrus, Somatomotor)</td>
<td>(Superior parietal lobule, Dorsal attention)</td>
</tr>
<tr>
<td>Right caudal area 7 (Superior parietal lobule, Dorsal attention)</td>
<td>(Superior parietal lobule, Dorsal attention)</td>
</tr>
<tr>
<td>Right area 28/34 (EC entorhinal cortex)</td>
<td>Right caudal area 7</td>
</tr>
<tr>
<td>(Parahippocampal gyrus, Limbic)</td>
<td>(Superior parietal lobule, Dorsal attention)</td>
</tr>
<tr>
<td>Right area TI [temporal agranular insular cortex]</td>
<td>Right caudal area 7</td>
</tr>
<tr>
<td>(Parahippocampal gyrus, Limbic)</td>
<td>(Superior parietal lobule, Dorsal attention)</td>
</tr>
<tr>
<td>Right caudal area 7 (Superior parietal lobule, Dorsal attention)</td>
<td>Left dorsomedial parietooccipital sulcus</td>
</tr>
<tr>
<td>Right caudal area 7 (Superior parietal Lobule, Dorsal attention)</td>
<td>(Precuneus, Visual)</td>
</tr>
<tr>
<td>Right rostral lingual gyrus (Medioventral occipital cortex, Visual)</td>
<td>Left inferior occipital gyrus</td>
</tr>
<tr>
<td>Right caudal area 7 (Lateral occipital cortex, Visual)</td>
<td>Right posterior parietal thalamus</td>
</tr>
<tr>
<td>(Superior parietal lobule, Dorsal attention)</td>
<td>(Thalamus, Subcortical)</td>
</tr>
</tbody>
</table>

*Note. Subgroup 2: most severe cool executive difficulties.*
Table III.9

**Edges within significant subnetwork identified for Subgroup 3**

<table>
<thead>
<tr>
<th>Regions of interest (Gyrus, Intrinsic connectivity network)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Left lateral area 10 (Middle frontal gyrus, Limbic)</td>
<td>Left middle occipital gyrus (Lateral occipital cortex, Visual)</td>
</tr>
<tr>
<td>Left lateral area 10 (Middle frontal gyrus, Limbic)</td>
<td>Left ventral caudate (Basal ganglia, Subcortical)</td>
</tr>
<tr>
<td>Left medial area 11 (Orbital gyrus, Limbic)</td>
<td>Left ventral caudate (Basal ganglia, Subcortical)</td>
</tr>
<tr>
<td>Left medial area 11 (Orbital gyrus, Limbic)</td>
<td>Left nucleus accumbens (Basal ganglia, Subcortical)</td>
</tr>
<tr>
<td>Left dorsomedial parietooccipital sulcus (Precuneus, Visual)</td>
<td>Left nucleus accumbens (Basal ganglia, Subcortical)</td>
</tr>
<tr>
<td>Left lateral area 10 (Middle frontal gyrus, Limbic)</td>
<td>Left rostral temporal thalamus (Thalamus, Subcortical)</td>
</tr>
<tr>
<td>Left medial area 11 (Orbital gyrus, Limbic)</td>
<td>Left rostral temporal thalamus (Thalamus, Subcortical)</td>
</tr>
<tr>
<td>Left ventral caudate (Basal ganglia, Subcortical)</td>
<td>Left rostral temporal thalamus (Thalamus, Subcortical)</td>
</tr>
</tbody>
</table>

*Note.* Subgroup 3: most severe difficulties with *hot* executive skills & pragmatic communication.
4. Discussion

The current study adopted a data-driven approach to identify subgroups of children with homogeneous profiles across different domains of function as measured by ratings of communication, behaviour, and executive function (EF). As in Chapter II, and following the arguments outlined in Chapter I, a transdiagnostic approach was used: enrolment was based on cognitive and academic needs rather than diagnostic status. Differences in brain structure and behaviour were compared across the subgroups in relation to a comparison group of non-referred children. Three subgroups of children were identified. Each performed poorer than the comparison group across measures of communication, behaviour, and EF, and was distinguished from the other subgroups by different profiles of strengths and weaknesses in these areas. These differences extended to measures of cognitive, academic, and socio-emotional functioning that were not included in the identification of the subgroups. Shared and specific patterns of differences in neural white matter organisation were observed across the groups. These results are discussed below.

4.1. Subgroups

Three data-driven subgroups were identified based on parent/carer ratings of behaviour, communication, and EFs in the referred sample of the CALM cohort, a heterogeneous group of children with a range of neurodevelopmental difficulties (see Chapter I, Section 2.2.1 for more details about the CALM cohort). One subgroup (S1) was characterised by relative difficulties in structural language use, a second by cool cognitive difficulties (S2), and a third by co-occurring pragmatic communication difficulties and hot affective cognitive problems (S3). All three subgroups had greater difficulties in behaviour, communication, EFs, socio-emotional functioning, and academic attainment relative to the demographically matched comparison group of non-referred children in the CALM cohort.
There were both similarities and differences in the way neural white matter was organised in the three subgroups relative to the comparison group. The two subgroups with the most pronounced difficulties in cognitive and academic skills (S1 & S2) showed reduced global efficiency, which has been previously linked to educational attainment (Bathelt et al., 2019; Lou et al., 2019). However, no differences were observed in global clustering coefficients. At the intrinsic connectivity networks-level, all groups showed reduced connectivity within the subcortex and the limbic network relative to the comparison group. Exploring differences in the connectivity of specific ROI-s within these areas, all three subgroups showed reduced regional strength relative to the comparison sample in subregions of the hippocampus, basal ganglia, inferior temporal gyrus, thalamus, and fusiform gyrus, which the BrainMap taxonomy labels as being involved in cognition, memory, action execution, language, interoception, and emotion. These findings demonstrate that the subgroups of referred children identified by the data-driven approach have shared and distinct behavioural and neural features relative to a non-referred comparison group. In the following sections, the detailed profiles of the three subgroups are considered.

4.2. Subgroup 1

The children in S1 were characterised by elevated difficulties with structural language skills relative to the other subgroups, and those who had attended Speech and Language Therapy in the past two years were overrepresented in this group. They also had the most severe difficulties in learning. Finding concurrent difficulties in structural language use and academic achievement is consistent with previous reports (Dockrell et al., 2011; Mareva & Holmes, 2019), and might reflect underlying phonological processing difficulties that have been implicated in language, reading, and maths difficulties (e.g., De Smedt et al., 2010; Ramus et al., 2013).
The neural subregions that showed significantly reduced connectivity only for children in S1 relative to the comparison group were in the right inferior temporal gyrus (lateral area 20), the left ventromedial putamen, and lateral amygdala. A left-lateralised subnetwork of primarily limbic-subcortical connections that involved subregions of medial orbitofrontal cortex, thalamus, and basal ganglia also distinguished this group. Consistent with their language and learning difficulties, corticostriatal and thalamocortical pathways are involved in procedural learning, language development, goal-directed behaviour, and reward processing (Arnsten & Rubia, 2012; Krishnan et al., 2016). Furthermore, most connections in the identified subnetwork involved left medial area 11, which the BrainMap taxonomy labels as functionally implicated in language and orthography.

4.3. Subgroup 2

The most pronounced area of weakness for the second subgroup (S2) was in cool EFs where their scores were lower than those of the children in the other two subgroups. These difficulties encompassed everyday difficulties in attention, planning, and working memory – skills that have been implicated in classroom learning (Holmes et al., 2020; Peng et al., 2018; Yeniad et al., 2013). Consistent with this, children in this subgroup had difficulties in maths and reading, although these were not as severe as those for children in S1. Relative to the other two subgroups, children in S2 had relative strengths in social skills and affective cognition, with comparable hot EF ratings to the comparison group.

Subregions that uniquely deviated in S2 included ROI-s within caudate, parahippocampal gyrus, and superior temporal gyrus (medial area 38). The subnetwork that differentiated them from the comparison group was relatively wide-spread and involved right-lateralised temporal-parietal pathways and medioventral and lateral occipital ROI-s. Some of the implicated subregions of these dorsal attention and visual networks are known to interact
to suppress attention to irrelevant stimuli (Castellanos & Proal, 2012; Shulman et al., 2009), and as such reduced connectivity in this neural circuit might contribute to the cool EF-related difficulties characteristic of this subgroup. In particular, seven of the nine connections within the identified subnetwork involved right caudal area 7, which is implicated in cool EFs such as attention, working memory, inhibition, and spatial cognition (for details see http://atlas.brainnetome.org/, Fan et al., 2016).

4.4. Subgroup 3

The children in S3 were characterised by having the most severe difficulties with hot EFs and pragmatic communication, but relative strengths in structural language skills, maths and reading, compared to the children in the other two subgroups. Strong associations between social communication skills and affective behavioural problems have been reported previously (e.g., Hawkins et al., 2016; Mareva & Holmes, 2019), and difficulties in these areas commonly co-occur in ADHD and ASD (Geurts & Embrechts, 2008; Green et al., 2014; Helland et al., 2014) alongside socio-emotional difficulties (Staikova et al., 2013). Indeed, a disproportionate number of autistic children and children with an ADHD diagnosis were assigned to this subgroup, and they had substantial problems with behavioural conduct, emotion, peer relationships, and prosocial behaviour. Identifying a subgroup with this profile and composition suggests the intersection of pragmatic communication, hot EF, and socio-emotional difficulties may be relevant for understanding some of the comorbidity between ADHD and ASD. Girls were underrepresented in this subgroup, consistent with literature suggesting a higher boy: girl ratio in ADHD and ASD (Maenner et al., 2020; Polanczyk & Rohde, 2007), and previous reports observing that relative to struggling boys, girls struggling at school are rated as experiencing less externalising behavioural difficulties (for a discussion of sex differences in the CALM cohort, see Guy et al., 2021).
The neural characteristics that specifically differentiated children in S3 from the comparison group were ROI-s within the putamen, thalamus, and cingulate gyrus (rostroventral area 24). They had reduced connectivity strength in a left-lateralised, primarily frontostrialat subnetwork, which also included regions of the visual network (precuneus and lateral occipital cortex). Corresponding to the severity of their hot EF difficulties, these circuits play role in goal-directed behaviours related to rewards, affect, and motivation (Arnsten & Rubia, 2012; Castellanos & Proal, 2012)

4.5. Summary

In summary, three subgroups of children with distinct communication, behavioural, and EF profiles were identified. These subgroups were characterised by primary difficulties in structural language, cool EFs, or hot EFs and pragmatics, respectively, and provide initial evidence for three alternative but related pathways to academic and socio-emotional difficulties. While a greater number of children receiving speech and language therapy were assigned to the structural language subgroup, and more autistic children and those with ADHD to the hot EFs and pragmatics subgroup, none of the subgroup profiles aligned with the diagnostic features of a particular disorder, and children with each of these diagnoses were present in each of the three subgroups. This finding suggests the subgroups were not synonymous with disorder-based categories, adding to growing support for transdiagnostic approaches to understanding neurodevelopment (Astle et al., 2021) and extending the findings of Chapter II that showed symptoms of neurodevelopmental difficulties did not neatly cluster according to diagnostic features.

The three subgroups were further distinguished by patterns of differences in the connectivity of circuits previously implicated in procedural learning, visual attention, and reward processing. These differences partially correspond to previously reported neurobiological correlates of the behavioural difficulties of the subgroups, suggesting they may
be distinguished at the neural level. That said, not all regions that uniquely differentiated the subgroups from the non-referred sample had clear links to their behavioural profiles. It should also be noted that all subgroups had reduced connection strength within the limbic network and the subcortex, relative to the comparison group, and shared several atypicalities within the same subcortical and temporal subregions. Furthermore, their profiles of behavioural weaknesses were all relative, meaning the correspondence between brain and behaviour was not one-to-one.

4.6. Limitations & Future directions

There were several caveats to the current subgrouping approach. First, the community detection was based on parent ratings, which are prone to subjective bias, and are not synonymous with performance-based measures of the same constructs (e.g., Toplak et al., 2013). However, the differences between the subgroups identified through these ratings were reflected in differences in performance-based measures of cognition and learning, providing some validity to the ratings and suggesting the algorithm was not overfitting the data. Second, the wide age range and cross-sectional nature of the cohort did not allow us to explore questions about age-related heterogeneity or developmental continuity, which remain key issues to be addressed in future research.

4.7. Conclusion

This study demonstrates the value of data-driven subgrouping approaches for understanding common, complex and co-occurring neurodevelopmental difficulties across multiple domains and their relationships to behavioural outcomes and neurobiology. It shows that homogeneous groups can be identified and differentiated in terms of distinct profiles of relative strengths and difficulties. The identified subgroups begin to characterise some of the heterogeneity and provide initial evidence for three alternative but related developmental pathways to difficulties with academic and socio-emotional functioning. Together with Chapter
II, these outcomes suggest that the diagnostic rubrics used to classify children’s difficulties across clinical and research settings do not neatly align with children’s profiles of strengths and weaknesses.

Having focused primarily on behaviour to this point, I move on to consider the learning and cognitive performance of the referred sample of the CALM cohort in more detail in the following chapter. I compare their performance to that of a population-representative cohort. It was not possible to include the comparison sample from the CALM cohort in these analyses as the data were not available at the time that I completed the study.
CHAPTER IV

Cognitive and Academic Skills in Two Developmental Cohorts of Different Ability Level: A Mutualistic Network Perspective

Chapter IV has been published, but appears modified to ensure thesis coherence:


*Journal of Applied Research in Memory and Cognition*,

[https://doi.org/10.1016/j.jarmac.2021.08.005](https://doi.org/10.1016/j.jarmac.2021.08.005)

First-pronouns are used throughout the chapter, but the work is the result of a collaborative research project. My contributions to the publication are:

5. I led the conceptualisation of the study research questions, methods, analyses and interpretation supported by J. Holmes.

6. I contributed to the data collection, scoring and curation, which was achieved in a collaborative effort including a large team of researchers (the CALM team).

7. I performed all statistical analyses.

8. I wrote the manuscript draft and led revisions aided by J. Holmes.
1. Introduction

Multiple theories attempt to explain academic difficulties. The traditional, and arguably still dominant assumption, is that learning problems are caused by specific cognitive deficits (e.g., Bishop & Snowling, 2004; Melby-Lervåg et al., 2012; Pennington & Ozonoff, 1996; Szucs et al., 2013). While this unidirectional approach has generated insights that have helped guide practice and policy (e.g., Reid & Fawcett, 2008), there are several challenges that it cannot easily accommodate (see Chapter I, Sections 1.2.6 and 1.2.7 for a detailed discussion of these challenges). More contemporary theories acknowledge that there is a more complex developmental interplay between cognitive and academic skills (e.g., Karmiloff-Smith, 2009; Peng & Kievit, 2020).

1.1. Challenges of the Unidirectional One-to-one Mapping Perspective

Traditional theoretical models of learning difficulties have assumed that cognitive abilities provide the foundation for academic development. For example, phonological processing deficits observed among children with reading difficulties provide the basis for the hypothesis that reading difficulties are caused by poor phonological processing (Bishop & Snowling, 2004). Similarly, working memory deficits observed among children with specific maths problems provide support for the hypothesis that memory systems are critical for maths (Szucs et al., 2013). Such theories are appealing due to the simplicity of assuming a “core deficit” (Astle & Fletcher-Watson, 2020), and their direct implications for intervention (e.g., phonological interventions for struggling readers, Bowyer-Crane et al., 2008).

As discussed in Chapter I, Sections 1.2.1., 1.2.6. and 1.2.7, these theories are challenged by evidence suggesting that similar learning difficulties may stem from different causes (i.e., equifinality). For example, reading difficulties are not wholly explained by phonological deficits. Some struggling readers do not have phonological impairments, but instead show problems with executive skills such as working memory (e.g., Cain et al., 2004). Similarly,
data-driven explorations of the links between cognitive and academic skills suggest that children can arrive at similar profiles of learning impairment through multiple etiological routes: children with comparable difficulties in both reading and maths can have different cognitive profiles, such as relatively more severe problems with phonological processing or working memory (Astle et al., 2019). Consistent with equifinality, multiple deficit theories suggest that risk factors spanning multiple levels (genes, brain, cognition and behaviour) contribute probabilistically to neurodevelopmental difficulties (Pennington, 2006), and that shared risk factors contribute to comorbidity (see Chapter I, Section 1.2.5). This idea is supported by evidence that processing speed explains some of the comorbidity between dyslexia, dyscalculia, and ADHD, while language difficulties contribute to comorbidity between dyscalculia and dyslexia (McGrath et al., 2020; Peterson et al., 2017).

1.2. Bidirectional Dynamics

A challenge to both single and multiple deficit models comes from studies showing that while cognitive skills influence academic skills, academic skills also shape cognitive development. Take the earlier example of the association between phonological processing and reading. Phonological awareness contributes to reading development, but it also benefits from reading instruction and exposure to text (Cunningham et al., 2021; Huettig et al., 2018; Nation & Hulme, 2011). Similar reciprocal effects, where the development of a given academic skill predicts growth in cognitive performance and vice-versa, have been reported between working memory and reading and maths (Miller-Cotto & Byrnes, 2019), IQ and reading (Ferrer et al., 2007, 2010), nonverbal reasoning and vocabulary (Kievit et al., 2017, 2019), and executive functions and maths (Schmitt et al., 2017; Van der Ven et al., 2012). These examples suggest that difficulties in one domain may have downstream effects on other abilities and challenge the use of analytic strategies in which cognitive deficits are uniquely modelled as predictors and academic skills as outcomes.
An alternative view of the relationship between cognitive and academic skills, consistent with evidence for reciprocal benefits, is mutualism. As discussed in Chapter I, Section 2.3.2., mutualism proposes that different abilities interact bidirectionally to reinforce one another during development (Peng & Kievit, 2020; Van Der Maas et al., 2006). In other words, the mastering of a skill supports and amplifies the development of other abilities. The mutualism model is neuro-constructivist in nature, acknowledging that specialised abilities likely emerge developmentally through a process of multidirectional interactions between genes, brain, cognition, and environment (Karmiloff-Smith, 2009; Van Der Maas et al., 2006).

In the context of cognitive-academic coupling, mutualistic transactions might be driven by experience, and particularly by educational experiences. Fundamental cognitive resources support the development of academic skills while performing academic tasks uses and trains cognitive abilities, and over time, these relationships become mutually beneficial (Peng & Kievit, 2020). For example, fluid reasoning skills aid the use of analogies and abstract schema in academic tasks, while concrete knowledge (e.g., verbal skills) supports the decomposition of complex reasoning tasks (Kievit et al., 2017). The type and strength of these transactions might be moderated by experience. Weaker or absent bidirectional relations have been reported in children with learning difficulties (e.g., Ferrer et al., 2010; Quinn et al., 2019). This could be because difficulties with specific skills operate as a bottleneck to the development of the wider cognitive system and/or because struggling learners may choose to avoid exercises that tax their area(s) of weakness. In the latter case, the net result could be that poor learners may engage in fewer activities that develop positively reinforcing associations, which over time might constrain both cognitive and academic development.

1.3. Cognitive and Academic Skills: New Approaches

To address the challenges outlined above, it is necessary to rethink the recruitment and analysis strategies typically used to study learning difficulties. The prevailing sampling
approach involves recruiting highly selective samples of children based on the presence of a specific diagnosis or difficulty, with comorbidity often treated as a confound (e.g., Szucs et al., 2013). This runs counter to a wealth of evidence reviewed in Chapter I showing that disorders are highly comorbid, heterogeneous, and explained by multiple causes (Astle et al., 2019; Peters & Ansari, 2019). Overstating the “purity” of learning problems at the point of recruitment biases outcomes towards core-deficit accounts. To avoid such biases, learning-related problems are increasingly studied using transdiagnostic approaches (Astle et al., 2021; Casey et al., 2014; Holmes et al., 2019; Mareva & Holmes, 2019). These approaches aim to understand the processes and causes of difficulties that occur across individuals irrespective of diagnosis or group membership, as discussed in the preceding chapters.

In terms of methods, unidirectional associations between cognition and academic performance are typically modelled using latent variable approaches whereby cognitive and academic factors are derived separately, and the relations between them are mapped as directional paths from cognitive to academic skills (e.g., Peterson et al., 2016). One alternative, which can better accommodate the possibility of equifinality without assuming causational directionality, comes from network science. Network models are relatively new to cognitive science but have already proved useful in tackling similar challenges in the field of psychopathology (Borsboom, 2016). Network models have already been introduced in Chapters I (Section 2.3.2.) and II. To re-cap briefly here, network models use partial correlations to depict how each observed variable relates to all other observed variables. In this way, they offer a modelling framework that is more consistent with mutualism and the possibility that reciprocal mutual transactions are the generating process behind the relationships observed across cognitive and academic skills. Applied to cross-sectional data, these models offer a tool to explore or test specific hypotheses about whether and how the
complex developmental interplay between academic and cognitive skills may differ across time points and/or groups of children.

1.4. Objectives

The current study aimed to apply networks models to explore and compare the interrelations between cognitive and academic abilities in a community sample and the referred sample of struggling learners from the CALM cohort (see Chapter I, Section 2.2.1, for an overview of the CALM cohort). The comparison cohort was an age-matched group of children selected as being nationally representative. It was not possible to use the comparison sample from the CALM cohort in these analyses as the data were still being collected, and not available at the time I completed this study. So as an alternative I identified a community sample with the same measures as the CALM study. For academic skills, the focus was on literacy and maths. For cognitive skills, assessments of processing speed, working memory, executive function, and nonverbal reasoning were included, all of which have been previously linked to academic performance in both typical and atypical learners (Altemeier et al., 2008; Booth et al., 2010; Gathercole et al., 2004; Geary, 2011; Green et al., 2017; Holmes et al., 2020; Mayes & Calhoun, 2007; Peng & Fuchs, 2016; Taub et al., 2008; Yeniad et al., 2013). To my knowledge, this was the first application of network science to cognitive-academic interrelationships in learners of different abilities. Consistent with mutualism and equifinality, I anticipated there would be multiple direct links between cognitive and academic skills in both cohorts. I was agnostic as to whether and how task interrelationships would differ across cohorts.

2. Method

2.1. Recruitment and Participants

Children from two cohorts were included in this study: the referred sample from CALM, and those from the Nathan-Klein Institute Rockland sample (NKI-RS). These two cohorts were
chosen because they include identical tasks while having different ability levels. The CALM cohort has already been discussed in previous chapters: the data included in the current analyses was collected between 2014 and 2018 and included only those who had been referred for problems related to cognition and learning (see earlier chapters and Holmes et al., 2019 for more details about recruitment and the cohort profile). The NKI-RS cohort is a community-based lifespan sample in which key demographics (age, ethnicity, and socioeconomic status) are representative of Rockland County, New York. The county was chosen because its ethnic and economic demographics broadly resemble those of the United States. The project used zip code based recruitment (e.g., advertisements through mail and community hubs). Enrolment began in 2012 and was monitored to avoid over-representation of specific demographic groups. Later phases of the project oversampled younger and older participants to increase statistical power for ages characterised by the greatest changes. Additional details are available in the study protocol (Nooner et al., 2012).

Testing sessions in the NKI-RS sample were split across several days, which were sometimes more than a year apart. To ensure that each child was roughly within the same developmental period for all assessments, any participants whose assessments were completed more than six months apart were excluded (in CALM all participants met this criterion). Six months is a relatively liberal estimate of developmental stage, but it is consistent with other studies using the NKI-RS cohort (e.g., Simpson-Kent et al., 2020). The youngest participants included in the NKI-RS cohort were 8 years old and 5 years old in CALM. Therefore, to circumvent the age difference, participants aged 8 or higher were subsampled from each cohort ($N_{\text{CALM}} = 566, N_{\text{NKI-RS}} = 350$). Subsequently, to match the sample size of the cohorts, CALM participants were age-matched to NKI-RS children using propensity matching based on the nearest neighbour method (Ho et al., 2011). The final age-matched samples included 350 participants each (CALM: $M_{\text{age}} = 11.26, SD_{\text{age}} = 2.21$, 69% male; NKI-RS: $M_{\text{age}} = 11.99, SD_{\text{age}} =$
2.89, 56% male). As expected, children with diagnosed neurodevelopmental problems were overrepresented in CALM (based on parent report: ADHD: 32% CALM and 14% NKI-RS; Learning problems (developmental language disorder, dyslexia, dysgraphia, or dyscalculia): 11% CALM and 6% NKI-RS; Autism Spectrum Disorder (ASD): 10% CALM and 0.3% NKI-RS). The referral routes of the CALM children in the current subsample matched the proportions observed in the full cohort: 60% education (e.g., Special Educational Needs Coordinators, Educational Psychologists), 36% health (e.g., Child Psychiatrist, ADHD nurse, Paediatrician, Clinical Psychologist), and 4% speech and language therapy (for the full CALM cohort, the proportions were 63%, 33%, and 4%, respectively).

2.2. Assessments

Assessments that were available for both cohorts were included. All tasks were taken from standardised test batteries. The psychometric properties for each assessment, together with the standardised administration procedures, can be found in the associated testing kits. Brief descriptions of each assessment are provided below.

2.2.1. Forward and Backward Digit Recall

The forward and backward digit recall tasks from the Automated Working Memory Assessment (Alloway, 2007) were administered to CALM participants, and the same tasks from the Wechsler Intelligence Scale for Children - Revised (WISC-R, Wechsler, 1974) were administered to NKI-RS participants. In both cases, forward digit recall involved the serial recall of sequences of spoken digits, while backward digit recall required children to recall the digits in reverse serial order. The number of trials per sequence length was six for the AWMA and two for the WISC-R. The number of trials correct was scored and used in all analyses, except for the permutation tests. These required identical measurement scales, so span was used to index performance for these analyses.

2.2.2. Matrix Reasoning
The Matrix reasoning subtest of the Wechsler Abbreviated Scales of Intelligence II (WASI-II, Wechsler, 2011) was administered to children in both cohorts. Children are presented with nonverbal analogical reasoning problems in 2x2 matrices and are asked to select from a range of alternatives a shape that completes the pattern. The number of correctly solved matrices was scored.

**2.2.3. Motor Speed**

Children in both cohorts completed the Motor Speed test of the Delis Kaplan Executive Function System (D-KEFS, Delis et al., 2001). This measured time (in seconds) to trace a dotted line as quickly as possible. Completion time in seconds was scored.

**2.2.4. Tower**

The Tower subtest of the D-KEFS, completed by both cohorts, required children to move disks of different sizes around pegs from a start position to an end state shown on a picture following a set of rules. The number of correctly completed towers was scored.

**2.2.5. Trails Number-Letter Sequencing**

Both cohorts completed the Trails Number-Letter sequencing task of the D-KEFS, which required children to connect letters and numbers in a progressive alternating sequence (e.g., 1-A, 2-B, etc). Completion time in seconds was scored.

**2.2.6. Spelling, Reading, and Maths**

The Spelling, Word reading, and Numerical operations subtests of the Wechsler Individual Achievement Test II (WIAT-II, Wechsler, 2005) were administered to children in both cohorts to measure academic abilities. Numerical operations captured number identification, counting, and the ability to solve simple and complex maths problems. Spelling required children to write down individual letters and spell single words. Word reading involved identifying letters, matching sounds to letters, and reading single words of increasing
complexity. To account for the large degree of overlap in the latter two scores (see Figure IV.2), they were combined into a single variable called literacy.

2.2.7. Visual Scanning

The D-KEFS Visual Scanning subtest was administered to children in both cohorts. Children were required to cross out all the number threes on a response page of numbers and letters as quickly as possible. Completion time in seconds was scored.

2.3. Analysis Plan

Network models were estimated to characterise the interrelations between cognitive and academic skills in each cohort. The accuracy of these results was scrutinised via a bootstrapping procedure. Permutation testing was then used to compare the two networks. As a robustness check the networks were also compared in a confirmatory fashion (i.e., estimating the structure in the community sample, fitting the same model to the struggling cohort and exploring modification indices) and via a joint estimation procedure known as a fused graphical lasso. Node strength and predictability were also estimated and compared. All statistical tests were performed in R (R Core Team, 2020; see Appendix III, Section 1 for all library version details). In all analyses, raw scores were transformed such that higher values represented better performance. For most analyses apart from node predictability estimates, missing data was handled by estimating pairwise associations. To estimate node predictability and to check the robustness of the pairwise association method, all analyses were repeated following multiple imputation via chain equations. Both approaches produced similar outcomes. This robustness analysis is fully described in Appendix III, Section 2.

3. Results

3.1. Descriptive Statistics

Raw scores for each task are presented in Table IV.1 (for norm-referenced scores, see Table IV.2 and Figure IV.1). Following an adjustment for multiple comparisons, children in
NKI-RS significantly outperformed CALM participants across all measures. Pearson correlations across raw and age-regressed scores were all positive (Figures IV.2 and IV.3). For most tasks, there were significant differences in variance between the groups. As a robustness check, all analyses were repeated following a data transformation, which significantly reduced the differences in variance. The analyses of transformed data produced mostly similar outcomes and are fully reported in Appendix III, Section 3.
### Table IV.1

*Summary statistics for CALM and NKI-RS based on raw scores, including the number of complete cases and task abbreviations for all assessments.*

<table>
<thead>
<tr>
<th>Task</th>
<th>CALM</th>
<th>NKI-RS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Backward Digit Recall (BDR)*</td>
<td>348</td>
<td>11.44 (4.25)</td>
</tr>
<tr>
<td>Forward Digit Recall (FDR)</td>
<td>349</td>
<td>26.38 (5.08)</td>
</tr>
<tr>
<td>Numerical Operations (Math).</td>
<td>312</td>
<td>20.64 (8.87)</td>
</tr>
<tr>
<td>Matrix Reasoning (MxReas)</td>
<td>350</td>
<td>13.46 (5.52)</td>
</tr>
<tr>
<td>Spelling and Reading (Lit)</td>
<td>345</td>
<td>62.02 (13.18)</td>
</tr>
<tr>
<td>Visual Scanning (Scan)</td>
<td>301</td>
<td>32.13 (11.62)</td>
</tr>
<tr>
<td>Motor Speed (Speed)</td>
<td>299</td>
<td>40.14 (17.13)</td>
</tr>
<tr>
<td>Trails Number-Letter Sequencing (Switch)</td>
<td>270</td>
<td>154.31 (60.97)</td>
</tr>
<tr>
<td>Tower</td>
<td>279</td>
<td>14.44 (3.80)</td>
</tr>
</tbody>
</table>

*Note. *The descriptive statistics for backward and forward digit recall reflect the number of trials correct, which was used for all analyses apart from permutation tests. These require identical measurement scales, thus, for this analysis only, span was used to index performance.*
### Table IV.2

*Group comparisons (scaled scores) for CALM and NKI-RS.*

<table>
<thead>
<tr>
<th></th>
<th>CALM</th>
<th>NKI</th>
<th>p-value</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Digit Recall</strong></td>
<td>8.76</td>
<td>2.30</td>
<td>9.89</td>
<td>2.76</td>
</tr>
<tr>
<td><strong>Numerical Operations</strong></td>
<td>7.34</td>
<td>3.61</td>
<td>11.01</td>
<td>3.52</td>
</tr>
<tr>
<td><strong>Matrix Reasoning</strong></td>
<td>8.29</td>
<td>3.08</td>
<td>11</td>
<td>2.88</td>
</tr>
<tr>
<td><strong>Single Word Reading</strong></td>
<td>7.99</td>
<td>3.25</td>
<td>11.46</td>
<td>2.45</td>
</tr>
<tr>
<td><strong>Visual Scanning</strong></td>
<td>9.02</td>
<td>3.50</td>
<td>10.30</td>
<td>2.70</td>
</tr>
<tr>
<td><strong>Motor Speed</strong></td>
<td>9.99</td>
<td>2.81</td>
<td>11.13</td>
<td>2.34</td>
</tr>
<tr>
<td><strong>Spelling</strong></td>
<td>7.31</td>
<td>2.84</td>
<td>11.67</td>
<td>2.65</td>
</tr>
<tr>
<td><strong>Number-Letter Sequencing</strong></td>
<td>6.26</td>
<td>3.91</td>
<td>9.34</td>
<td>3.62</td>
</tr>
<tr>
<td><strong>Tower Achievement</strong></td>
<td>9.46</td>
<td>2.42</td>
<td>10.04</td>
<td>2.37</td>
</tr>
</tbody>
</table>

*Note.* Norm-referenced scores were a mix of *T*-scores, standard scores, and scaled scores. All scores were converted to scaled scores with a mean of 10 to facilitate interpretation. Higher scores indicate better performance. In NKI-RS, the digit recall score is based on the combined forward and backward digit recall norms provided in WISC-R. In CALM, this score reflects the average of the AWMA norm-referenced forward and backward digit recall scores. Holm-corrected *p*-values are shown, based on two-tailed *t*-tests (equal variances not assumed) and associated Cohen’s *d* effect size estimates.
Figure IV.1

Standardised performance across the tasks for CALM (yellow, left) and NKI-RS (green, right).

Note. The large dots inside the boxplots show the mean of each cohort. The grey dotted line represents the age-expected mean. Norm-referenced scores were a mix of $T$-scores, standard scores, and scaled scores. All scores were converted to scaled scores for visualisation. Digit Recall = WISC-R/AWMA Combined forward and backward digit recall; Lit = Literacy (WIAT-II Word Reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix Reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower achievement score.
Figure IV.2

Pearson correlations across age, cognitive and learning skills for CALM (left) and NKI-RS (right).

Note. BDR = AWMA/ WISC-R Backward digit recall; FDR = AWMA/ WISC-R Forward digit recall; Math = WIAT-II Numerical Operations; MxReas = WASI-II Matrix reasoning; Read = WIAT-II Word reading; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Spell = WIAT-II Spelling; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
Figure IV.3

Pearson correlations of age-regressed residuals across all cognitive and learning tasks for CALM (left) and NKI-RS (right).

Note. BDR = AWMA/ WISC-R Backward digit recall; FDR = AWMA/ WISC-R Forward digit recall; Math = WIAT-II Numerical Operations; MxReas = WASI-II Matrix reasoning; Read = WIAT-II Word reading; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Spell = WIAT-II Spelling; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
3.2. Age Effects

Mutualistic theories predict that the relationship between academic achievement and cognitive skills should increase with age (Peng & Kievit, 2020). This prediction is supported by meta-analyses reporting small age-related increases in the strength of the associations between cognitive and academic skills (e.g., nonverbal reasoning and reading and mathematics (Peng et al., 2019); reading and working memory (Peng, Barnes, et al., 2018); and executive functions and academic achievement (Jacob & Parkinson, 2015)). The sample size of the cohorts included in the current analysis was too small to investigate age effects of the size reported in these meta-analyses (see Perugini et al., 2018; for a discussion of sample size and power in relation to moderation analysis). As a robustness check moderation analysis was run to ensure no large age moderation effects were missed (see Appendix III, Section 4 for details about the results of this analysis). No age-moderation was observed in either cohort (likely due to insufficient power). However, this should not be taken as evidence that age does not moderate the relationships in the networks: it should instead be formally investigated in appropriately powered datasets. For the purposes of the current analysis, age was included as a control variable (i.e., it was included in the network estimation) to mitigate its potential confounding effects on task associations.

3.3. Network Estimation and Stability

A regularised partial correlation network was estimated for each cohort. Age was included in the estimation but was omitted from plots and the calculation of centrality indices. In the final networks, nodes represented task performance and edge weights corresponded to the regularised partial correlation coefficient between any two tasks, controlling for age and all other scores. Each network was estimated using the graphical variant of the least absolute shrinkage and selection operator with Extended Bayesian Information Criterion used for model selection (Epskamp et al., 2018). The estimated networks are displayed in Figure IV.4. The
bootstrapped 95% confidence intervals (N boots = 1000) were small to moderate, suggesting acceptable stability (see Appendix III, Section 5, and Figures AIII.3-8). As discussed in Chapter II, due to the regularisation method, the confidence intervals (CIs) cannot be interpreted as a significance test to zero. Instead, CIs derived from the times the parameter was not set to zero, together with the proportion of bootstrapped samples in which edges were set as different from zero across the 1000 iterations are included whenever an edge is discussed. For all other edges, these details are presented in Figure IV.5 and Appendix III, Section 5, Figures AIII.5 (CALM) and AIII.6 (NKI-RS).
Figure IV.4

*Regularised partial correlation networks across cognitive and academic skills for CALM (left) and NKI-RS (right).*

*Note.* Age was included in the estimation but is omitted from figures. Thicker edges represent stronger associations. Red edges reflect negative associations. The rings around the nodes represent the proportion of variance explained in the respective node by all connected nodes. Digit Recall = WISC-R/AWMA Combined forward and backward digit recall; Lit = Literacy (WIAT-II Word Reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix Reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower achievement score.
Figure IV.5

The upper triangle represents estimated edge weights, where darker shades correspond to stronger edge weights. The values in the lower triangle represent the proportion of bootstraps where an edge was estimated as different from zero. Values of one shaded in the darkest blue indicate that an edge was included in all 1000 bootstrapped networks.

Note. BDR = WISC-R/AWMA Backward digit recall; FDR = WISC-R/AWMA Forward digit recall; Lit = Literacy (WIAT-II Single word reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
3.4. Network Comparison

As a first step to comparing the networks, the correlation between edge weights across networks was estimated. The correlation was \( r_s = .58 \), indicating moderate similarity. Second, a permutation test based on 1000 iterations was used to investigate differences in network organisation (van Borkulo et al., 2017). Global connectivity was operationalised as the sum of all absolute edge weights, and global structure reflected the highest absolute difference between two corresponding edges (\( M \)). Permutation testing detected no significant difference in global connectivity, suggesting a similar degree of task interrelatedness in both cohorts (CALM: 4.04; NKI-RS: 4.14; \( p = .65 \)). However, the network structure was not invariant across cohorts (\( M = 0.30, p < .001 \)). Following a false discovery rate correction, four relations between tasks differed between the cohorts. The estimated edge strength in each cohort together with the 95% CIs from the bootstraps when the edge was not set to zero (presented in square brackets) were the following: Matrix reasoning – Maths: \( r_{CALM} = .34 \) [non-zero 100%, CI: .25–0.42], \( r_{NKI-RS} = .04 \) [non-zero: 79%, CI: 0.01–0.14], \( p < .001 \); Literacy – Maths: \( r_{CALM} = .20 \) [non-zero 100%, CI: 0.1–0.29], \( r_{NKI-RS} = .41 \) [non-zero 100%, CI: 0.34–0.48], \( p < .001 \); Matrix reasoning – Switching: \( r_{CALM} = .09 \) [non-zero 92%, CI: 0.01–0.21], \( r_{NKI-RS} = .31 \) [non-zero 100%, CI: 0.19–0.40], \( p = .022 \); Maths – Backward digit span: \( r_{CALM} = .31 \) [non-zero 100%, CI: 0.21–0.39], \( r_{NKI-RS} = .10 \) [non-zero: 95%, CI: 0.01–0.21], \( p = .027 \). The latter difference was not significant following the data transformation applied to address differences in task variance: \( r_{CALM} = 0.20, r_{NKI-RS} = 0.06, p = .27 \) (see Appendix III, Section 3). Alternative methods of network comparison provided similar results and are reported in the Appendix III (see Sections 6-7, Figures AIII.9-10, and Table AIII.3).

3.5. Community detection

Finally, to compare how tasks clustered together across cohorts, the Walktrap community detection algorithm (Pons & Latapy, 2006) was applied to each network. The
algorithm is described in Chapter II (Section 2.4.4). The modularity index ($Q$) was again used to assess the quality of the partition, with higher values indicating stronger partitions and values around 0.3 considered evidence for sufficiently separated communities (Newman & Girvan, 2004). In both cohorts, the index was substantially below 0.3 ($Q_{CALM} = 0.15$ and $Q_{NKI-RS} = 0.08$), indicating the tasks did not form robust groups.

3.6. Node Centrality and Predictability

To explore and compare the relative importance of nodes within each network, node strength and predictability were estimated. Strength estimates, defined as the sum of all edge weights connected to a given node, are presented in Figure IV.6. Case-dropping analysis was used to assess the stability of these estimates: the metric was considered stable if at least half of the children within each sample could be dropped while still retaining a 95% probability of a 0.7 correlation between the strength estimates based on the full sample and those derived from subsamples (Epskamp & Fried, 2018), this recommendation was met in both samples: max. drop proportion CALM: 67% and NKI-RS: 75%. Across cohorts, node strength was strongly correlated ($r_s = 0.75, p = 0.03$) and permutation tests suggested no significant differences.

Node predictability is the proportion of shared variance between a given task and all tasks related to it (Haslbeck & Waldorp, 2018). It was estimated by averaging the results of graphical models fitted to each of the 100 imputed datasets per cohort. Nodewise predictability is displayed in Figure IV.4. For CALM node predictability ranged between 13% (Number-Letter switching) and 51% (Literacy), in NKI-RS the range was from 12% (Number-Letter switching) to 46% (Literacy). Across cohorts, average predictability estimates were similar (CALM: 32%; NKI-RS: 29%) and strongly correlated ($r_s = .70, p = 0.04$).
Figure IV.6

Centrality estimate strength across CALM and NKI-RS.

Note. BDR = WISC-R/AWMA Backward digit recall; FDR = WISC-R/AWMA Forward digit recall; Lit = Literacy (WIAT-II Single word reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
4. Discussion

The current study provides one of the first demonstrations of how network approaches can be used to model the complex relationships between cognitive and academic skills proposed by contemporary developmental theories. Associations between these two domains were compared across two developmental cohorts with different levels of ability: a community sample from the US and the sample from CALM comprised of children referred by health and educational professionals for difficulties related to learning. The children in the CALM cohort had significantly poorer performance across all tasks relative to the other cohort, which was recruited to be demographically representative. Despite substantial differences in performance across cohorts, the patterns of task interrelations were broadly similar. There were multiple direct links across cognitive and academic abilities for both cohorts: the matrix reasoning, backward digit recall, and switching tasks were directly related to both literacy and maths. The presence of multiple direct links is consistent with mutualistic and multiple deficit theories (McGrath et al., 2020; Peng & Kievit, 2020), and provides evidence against specific one-to-one mappings between a single cognitive ability and a specific academic skill.

Several differences in task interrelationships were observed across cohorts. The association between maths and literacy was stronger in the community sample than in the struggling learners. Within a mutualistic perspective, education simultaneously enhances maths and literacy knowledge, and over time these become mutually enriching (Ritchie & Tucker-Drob, 2018). In struggling learners, weaker skills in one domain might slow the accumulation of fluency, and over time limit the mutually beneficial exchanges between domains. Additionally, negative feedback and/or experiences in one domain may lead to school disengagement or poor motivation for learning, which could have consequences beyond the affected domain (Elliot & McGregor, 2001; Wery & Thomson, 2013).
The association between maths and matrix reasoning was stronger in the struggling learners. Through a mutualistic lens, the link between maths and fluid abilities reflects both the benefits of executive abilities for learning maths and the training effects of maths practice on these cognitive abilities (Peng et al., 2019). Learning maths is executively demanding, but practice makes maths knowledge fluent and more easily available for direct retrieval (Mussolin & Noel, 2008). The stronger link in struggling learners could potentially be due to weaker maths fluency, in which case automatic solution strategies such as direct retrieval will not be available, and mathematical problem solving will instead more strongly draw on executive resources.

Finally, the relationship between the matrix reasoning and switching tasks was weaker in the poor learners. Both tasks are typically considered measures of executive abilities. However, for switching task performance to rely on higher-order cognitive control, children need to be fluent in both the alphabet and counting. Consistent with this, it has been previously reported that insufficient automatisation of the alphabet can falsely impair performance on this task (Egeland & Follesø, 2020). Therefore, one explanation for the weaker link between tasks is the possibility that in struggling learners, poor alphanumeric knowledge constrained task performance more than executive control.

4.1. Limitations and Future Directions

The available data constrains the inferences that can be drawn from the current study. First, the assessments available in both cohorts did not capture the full breadth of children’s cognitive skills and academic performance. The inclusion of broader assessments such as reading comprehension and text-based maths problems remains an important avenue for future work. Second, the two cohorts were drawn from different countries. Despite relative similarities in the educational systems and cultural values of these countries, it remains possible that the observed differences were due to differences in demographic or school curricula factors.
Finally, longitudinal data are needed to understand whether the differences observed across cohorts arise because certain abilities operate as a bottleneck to the development of the wider cognitive system, emerge due to different environmental or educational experiences or a combination of the two.

This study was exploratory but provides another example, in addition to those presented in Chapters II and III, of the value of network approaches for the field of developmental science. They enable the simultaneous modelling of multiple routes to a specific outcome and can incorporate different levels of description (Simpson-Kent et al., 2020). They can therefore be used to test competing theories about how abilities are related across groups and/or time points, and to characterise how genetic, neural, and environmental factors influence these relationships (Isvoranu et al., 2020). Such applications will be crucial for building an understanding of not only how abilities are related, but also which mechanisms enable their wiring.

4.2. Conclusion

Using network science to model and compare the relationships between academic and cognitive skills across cohorts of children with different levels of ability, I found multiple and largely similar interrelationships, together with some key differences. These differences could suggest that weak (or absent) reciprocal links between and within academic and cognitive domains may contribute to learning difficulties. Crucially, this chapter demonstrates the potential value of network models for characterising the wiring of cognitive-academic systems across populations, extending their application beyond the work I completed in Chapters II and III. Mutualistic networks provide a promising new tool for capturing complexity in development, and in time may be useful for identifying time windows where interventions could enhance mutualistic coupling.
CHAPTER V

Transdiagnostic Approaches to Neurodevelopmental Difficulties: Findings, Implications, and Future Directions

1. Introduction

This thesis adopted a transdiagnostic approach to neurodevelopment, focusing on studying relationships across domains in the CALM cohort, a diverse group of children at developmental risk. A transdiagnostic approach was employed to account for evidence of the lack of specificity in the relationship between early risk factors, diagnoses/primary difficulties, and later life outcomes. In the three empirical chapters, contemporary methodological advances were leveraged to study the concurrent associations between abilities across different domains of function. The current chapter summarises the key findings and discusses their practical and theoretical implications. It further reflects on the strengths and limitations of the empirical work and considers important future directions.

1.1. Chapter II: Summary

In Chapter II, the focus was on communication, behavioural, and executive function problems, which often co-occur in childhood but are nonetheless predominantly studied separately. Previous attempts to identify the origins of these comorbidities have typically relied on comparisons of different deficit groups and/or latent variable models. To advance the field, a network approach was used in Chapter II to conceptualise these comorbidities as a dynamic system of interacting difficulties. Data from the CALM cohort was analysed, which at the time of analysis comprised 714 children referred for difficulties with attention, learning, and/or memory. The conditional association network of parent ratings of everyday behaviour, cognition, and communication was estimated to explore how skills across these domains relate to one another, and to identify the constituent parts linking different domains. Four
interconnected areas of difficulty were identified: (1) structural language and learning; (2) pragmatics and peer relationships; (3) behavioural and emotional problems (hot EFs); and (4) cognitive skills (cool EFs). These four clusters of symptoms did not correspond to any one set of diagnostic features for a specific disorder.

Emotional and behavioural symptoms shared multiple direct connections with pragmatic abilities and cognitive problems, but not with structural language skills or learning problems. Poor structural language and cognitive skills were associated with learning problems. Centrality indices highlighted working memory and language coherence as symptoms bridging different problem areas. This finding is corroborated by additional cross-sectional and longitudinal investigations from both typically developing children and those with ADHD, which have also identified working memory as a transdiagnostic risk factor bridging various function domains including ADHD-related, internalising, and externalising difficulties (Eadeh et al., 2021; Karalunas et al., 2021). Although the current data does not provide causal evidence, bridging symptoms such as working memory may be the origins of comorbidities; problems in these areas may cascade and activate problems in other areas. Alternatively, their bridging position may indicate that their development requires consolidation of skills across domains. Further research is needed to tease out these interpretations.

1.2. Chapter III: Summary

To follow up on these findings the subsequent chapter (Chapter III) aimed to identify distinct subgroups of children with common profiles of behavioural and language strengths and difficulties. The same sets of parental ratings of communication and behaviour studied in Chapter II were analysed again, this time with the goal of exploring how children occupy this multivariate space. The full CALM cohort was available for these analyses (referred: \( N = 805 \), non-referred: \( N = 158 \)). Within the referred cohort three distinct subgroups of children were identified, each with different levels of difficulties in structural language, pragmatic
communication, and hot and cool EFs. One subgroup had the most pronounced difficulties with structural language; another subgroup was characterised by the most severe cool EF problems; and a final subgroup had principal difficulties with hot EFs and pragmatic communication. In the next step, it was investigated whether and how these data-driven groupings could be distinguished from the comparison sample on academic, socio-emotional, and neural white matter characteristics.

All three subgroups struggled with academic and socio-emotional skills relative to the comparison sample, potentially representing three alternative but related developmental pathways to difficulties in these areas. The children with the weakest structural language skills had the most widespread difficulties with learning, whereas those with more pronounced difficulties with hot EF skills and pragmatic communication experienced the most severe difficulties in the socio-emotional domain. These findings were in line with the results of Chapter II where it was observed that learning and structural language formed part of the same network cluster; and the same was observed for pragmatic communication and peer relations. Further, the network structure in Chapter II suggested that hot EFs were indirectly linked to learning via cool EFs. Consistently, in Chapter III it was observed that the subgroup that had more severe cool EF difficulties had poorer academic performance relative to the group with the most pronounced hot EF problems. These converging results provide an example of the complementary insights that can be gained through different analytical approaches.

Finally, each data-driven subgroup could be distinguished from the comparison sample based on both shared and subgroup-unique patterns of neural white matter organisation. Relative to the non-referred group, all data-driven groups showed reduced connectivity within the subcortex and the limbic network, and particularly in subregions of the hippocampus, basal ganglia, inferior temporal gyrus, thalamus, and fusiform gyrus. The three subgroups could further be distinguished from the non-referred sample based on connectivity in circuits.
previously implicated in procedural learning, visual attention, and reward processing, which for the most part corresponded to the neurobiological correlates of their respective behavioural difficulties. Overall, despite some correspondence, the patterns that differentiated each subgroup from the non-referred sample both subsumed and extended those observed for the other subgroups, suggesting substantial neural overlap between groups and no neat direct behaviour to brain mapping. These findings advance our understanding of commonly co-morbid behavioural and language problems and their relationship to behavioural outcomes and neurobiological substrates.

1.3. Chapter IV: Summary

In Chapter IV the focus shifted to the relations between academic functioning and domain-general cognitive skills in both the CALM cohort and a comparison community cohort. Multiple theories attempt to explain why academic and cognitive performance are closely related. Typically, a one-way perspective is assumed whereby specific cognitive deficits constrain and impede learning (e.g., Peterson et al., 2017; Yeniad et al., 2013). However, the links between cognitive and academic skills are rarely specific, and there is evidence that the mastery of academic skills can influence cognitive development (Peng & Kievit, 2020).

To address these challenges, the work in this chapter adopted a mutualistic approach to conceptualise the associations between cognitive and academic performance as emergent properties of a network of cross-domain transactions. Conditional association network models, which make no assumptions about causality or one-to-one relationships, were used to map the interrelations between identical cognitive and academic tasks in two age-matched developmental cohorts. One cohort was drawn from the general school population and the other included struggling learners (CALM). The community sample outperformed the struggling learners across all tests of cognition and learning. Contrary to one-to-one mapping perspectives, multiple direct links between academic and cognitive skills were observed in both cohorts. In
particular, the matrix reasoning, backward digit recall, and switching tasks were directly related to both literacy and maths. Although the tasks were mostly similarly interrelated across cohorts, there were some notable differences in association strength: academic skills were more closely linked in the community sample, while maths was more strongly related to domain-general cognitive skills in the struggling learners.

Overall, the direct links between cognitive and academic skills were multiple and broadly similar for both groups of children, providing no evidence for qualitative discontinuity in these relationships across struggling and community samples. Finally, from a methodological perspective, this chapter demonstrates the utility of network models as an analytic framework that is consistent with contemporary developmental theories such as mutualism, which can also be used to address diverse research questions looking to characterise and compare interrelations between cognitive and academic skills across populations.

2. Key takeaways

The strengths of the work presented in the empirical chapters relate to the large transdiagnostic sample, the focus on a diverse set of functional domains, and the application of innovative methods suitable for capturing complexity. Below I consider the key takeaways from the empirical work and their conceptual and practical implications.

2.1. Diagnostic Labels Provide Limited Information about Individual Strengths and Difficulties

Findings across the three chapters were consistent with the key rationale for transdiagnostic research, which relates to the limited utility of diagnostic labels to capture the overlap across and heterogeneity within neurodevelopmental conditions. In Chapters II and IV it was observed that difficulties traditionally associated with specific diagnostic labels share multiple direct associations with one another; in Chapter III, the data-driven subgroups showed little direct correspondence to diagnoses (children with the same diagnostic label were represented in all
subgroups, and both children with ASD and ADHD were overrepresented within the same data-driven subgroup). These findings converge with previous evidence that diagnostic labels do not map neatly onto behaviourally defined profiles and reinforce the utility of investigations aiming to test the validity and variability of these categories.

2.2. Multiple Pathways to Similar Difficulties

Relatively, the results presented across the empirical chapters also demonstrate that one-to-one mappings from a specific difficulty to a restricted outcome are unlikely. This was supported by the multiple cross-domain links observed in the conditional association networks of cognitive and academic skills, and language and behaviour (Chapters II and IV). Similarly, data-driven subgroups of children characterised by different areas of principal difficulties all showed poorer academic and socio-emotional functioning relative to non-referred children. The mappings were not one-to-one, but the severity of a child’s difficulties in a given domain was related to the degree to which they struggled in others. Broadly, these patterns suggest that there are multiple pathways to the same type of problems.

2.3. The Unlikely Case of Isolated Difficulties

The densely interrelated conditional association networks observed in Chapters II and IV further call for increased recognition of the multiple ways in which difficulties could cascade across domains. If a child has a significant difficulty in one developmental domain (e.g., language), they are often at a far greater risk for difficulties in other domains (e.g., behaviour, academic). Some of these developmental cascades have not been sufficiently considered, especially in cases where they operate outside areas traditionally associated with specific disorders (e.g., language needs are not typically considered in children with ADHD). In other cases, the lack of empirical attention stems from commonplace assumptions about unidirectional causal relationships (e.g., from foundational cognitive skills to academic outcomes). Potentially valuable insights may have been missed by failing to address the
possibility that while cognitive skills can influence learning, the process of learning can also influence cognitive skills.

2.4. Shared and Unique Neural Substrates

Another important takeaway is the finding that there was poor evidence for neat direct mappings between isolated neural structures and behavioural phenotypes. In Chapter III, it was observed that most brain atypicalities associated with a given behavioural phenotype were not specific to the phenotype but were also observed in children with relatively distinct problems. This casts doubt on claims that focal neural deficits underlie a given neurodevelopmental disorder. Rather, the findings are consistent with the predictions of (neuro)constructivist theories that assume developmental difficulties have widespread effects that result from a brain that has developed atypically over a number of years (Johnson, 2011; Karmiloff-Smith, 2009). This highlights the importance of studying multivariate relations transdiagnostically across the behavioural and neural levels.

Triangulating approaches to ensure robustness

Collectively, Chapters II-IV provide examples of how tools such as network science and clustering can be applied to help researchers start to characterise the inherent complexity underpinning neurodevelopmental diversity. While each method has its limitations, the work presented here illustrates how insights gained from different approaches and datasets can be considered together to ensure the robustness of the findings. The triangulation of methods is particularly important in the context of exploratory investigations where data-driven models may be overfitting the data and thereby limiting the robustness and generalisability of the findings. In the next step, emerging patterns should be tested in a confirmatory fashion.
3. Implications

I now turn to consider the implications of the current findings, and the transdiagnostic approach more broadly, by discussing the potential practical and conceptual significance of the work.

3.1. Practical Implications

One of the key findings from Chapters II and III concerns the close relationship between behavioural, language, learning, and communication difficulties. From a practical perspective, this suggests that communication and language problems may need to be considered in children presenting with difficult behaviour, and vice-versa that behavioural problems should be assessed in children presenting with language difficulties. However, the somewhat fragmented nature of service provision may pose a barrier to such implementation. For example, in the United Kingdom if a child requires assessment for behavioural difficulties (e.g., suspected ADHD) this would fall within the remit of child mental health and paediatric services; language and social communication needs would require the attention of a speech and language therapist; and potential learning difficulties would be dealt with by education services. Cross-service communication and referrals are certainly commonplace, but nonetheless, the fact that relevant professionals are based in different services and clinical pathways are often organised around specific neurodevelopmental conditions, could mean that additional needs and sources of heterogeneity go undocumented. This might be most applicable to cases where such needs are relatively mild but could benefit from specialist attention. Generally, given that the co-occurrence of neurodevelopmental problems is the rule rather than the exception, better integration of various expertise and services is needed to ensure that overlaps are better recognised (Thapar et al., 2017).

To achieve impact, harmonised and integrated provision would need to be combined with comprehensive support strategies spanning several domains. Diagnosis-focused treatment
guidelines pose a challenge to this goal. Different problem areas often require different evidence-based management plans. For example, cognitive behavioural strategies are recommended for emotional problems, parenting interventions for behaviour difficulties, and phonological interventions for reading problems. Notably, the efficacy of these approaches may be altered by the presence of additional difficulties. For example, phonological training interventions targeting literacy are reported to be less effective for children with high levels of hyperactivity (Hurry et al., 2018). Similarly, children with language difficulties may not benefit from mental health interventions that are often delivered in a manner not comfortable for them due to high reliance on oral language skills (Hobson et al., 2021). Historically, such moderation effects have not received sufficient empirical attention and the available evidence base is quite mixed making it hard to draw reliable conclusions (e.g., Denton et al., 2020; Ollendick et al., 2008; Tannock et al., 2018). A strength-focused approach is likely to benefit support provision, and more intervention research on children experiencing multiple neurodevelopmental difficulties is needed to obtain a full picture of comorbidity and its effects on treatment outcomes and prognosis.

Relatedly, the findings of the three empirical studies presented here, and the broader literature reviewed in Chapter I, invite us to consider another key practical consideration: what is the utility of neurodevelopmental diagnostic labels? Despite their limitations, diagnostic classifications continue to have several practical advantages in applied settings. They can help simplify the key issues that a child is experiencing and can help direct focus towards a child’s most pressing needs. Diagnostic labels also provide a means for practitioners to easily communicate children’s profiles with each other, families, and the children themselves. They also help practitioners to easily consult the evidence base, which remains largely centred around specific conditions. Finally, diagnostic categories provide important utility in guiding clinical decisions about appropriate support and whether to prescribe medication. Thus, until
alternative approaches are sufficiently researched and established there is still a strong case to retain the practice of distinguishing conditions despite its substantial flaws. Nonetheless, the current *transdiagnostic revolution* does invite researchers and practitioners to be more critical and aware of the arbitrary nature of diagnostic labels, which should not be seen as fundamental truths and should not be the sole means for determining support/treatment plans (Thapar et al., 2017).

A crucial first step would be a move towards more child-centred rather than diagnosis-centred approaches. How can this be achieved? Most researchers and clinicians will recognise that neurodevelopmental difficulties cannot be fully explained by a set of diagnostic criteria, but often funders, policies, and national guidelines are not as flexible (Thapar et al., 2017). Service provision is too often allocated based on categorical evaluations of whether a child meets diagnostic criteria; a decision that already involves a lengthy and expensive process of assessments and consultations. Therefore, extending assessments beyond diagnostic items to include other relevant information such as social context and the child’s additional strengths and weaknesses may seem like a daunting and expensive ordeal. However, such change in systemic agendas is arguably what is required to make space for complex phenotypic patterns and potential accompanying areas of difficulties to be considered. A more child-centred service provision, focusing on a child’s level of functioning, needs, and concerns rather than diagnostic labels can help identify the relevant areas of support, as well as the appropriate tools to target them.

A potential step in this direction would be the development of diversified feasible transdiagnostic measurement toolkits to be shared by practitioners and researchers. Emerging initiatives such as the harmonised transdiagnostic assessment protocol for neurodevelopment recently outlined by Boulton et al. (2021) take this approach forward. The proposed protocol provides one of the first attempts to incorporate the transdiagnostic framework within services
and clinics. It was formulated in consultation with a large group of clinicians, researchers, and community groups across Australia. A national research committee surveyed stakeholders and organised national summits to reach a consensus on a set of diverse and feasible measurement areas to include in the harmonised protocol. The identified assessment instruments span several domains of needs, such as child functioning and wellbeing, child mental health, caregiver mental health, and family background. The goal of the initiative is to achieve impact at both the clinical and research level: firstly, to enable transdiagnostic clinical assessments for children with neurodevelopmental needs; and secondly, to facilitate harmonised large-scale data collection that has the potential to inform research, policy, and practice. The outputs of this transdiagnostic protocol will enable one of the first empirical evaluations of whether the examination of transdiagnostic needs beyond primary diagnoses could allow the delivery of more integrative care and improve child outcomes. In the long-term, this approach could facilitate the departure from categorical approaches and could subsequently feed into the design of interventions in which enrolment is based on combinations of needs and strengths across domains, as opposed to primary diagnosis (Finlay-Jones et al., 2019). This pragmatic goodness-of-fit matching of interventions to processes relevant to the individual child could enable a flexible treatment that could be applied across a range of presentations including complex, comorbid, and subthreshold difficulties.

Overall, while the categorisation of neurodevelopmental difficulties into disorders has some practical utility there is an argument to be made that a more flexible approach that takes into consideration the high level of heterogeneity and overlap in neurodevelopmental disorders may benefit clinical practice. The first attempts to integrate the neurodevelopmental transdiagnostic framework into clinical settings are just emerging and in due course will facilitate important insights about efficacy and feasibility. Future goals should be to include similar applications in educational settings.
3.2. Theoretical Implications

Following decades of research shaped by the diagnostic categories specified in the DSM, the field is currently undergoing a transitional phase in which exploratory research into how children and their characteristics are clustered is taking priority (Astle et al., 2021). Such research is key for testing, and where necessary accordingly modifying, predominant assumptions about diagnostic boundaries or common factors that explain the associations between symptoms and/or tasks (Astle & Fletcher-Watson, 2020). In some cases, these investigations are informed by relevant theories such as multiple deficit models, (neuro)constructivism, the network theory of psychopathology, compensatory accounts (i.e., difficulty in one domain would lead to compensatory reorganisation in other areas), and mutualism, but most efforts so far have largely been data-driven and hypothesis-free. The studies presented in this thesis are an example of this. Generally, across the three empirical chapters, findings did not align with core deficit/one-to-one mapping accounts and were instead more consistent with (neuro)constructivist, mutualistic, and/or network theories. While theoretical perspectives guided the study design, research questions, and interpretations of the empirical work, none of the investigations explicitly tested or falsified a given theory, and therefore their theoretical implications should be considered with caution.

Arguably, the theoretical contribution of this work manifests through its consolidation with the insights generated by the broader transdiagnostic effort. The consolidation of exploratory findings from this thesis and across the literature can inform subsequent confirmatory hypothesis-driven research studies, which would, in turn, contribute towards refining current and/or building novel theoretical perspectives. Advances in our scientific understanding of neurodevelopment will almost certainly require theories that integrate findings across disciplines and levels of analysis because no single level of analysis could account for the variance in any functional outcome (Beauchaine et al., 2019). Therefore, future
theories would need to move away from poorly supported notions of unitary causes and would instead need to incorporate more complex possibilities such as dynamic emergence. Generative and computational models will certainly play a role in the formulation and testing of such theories, and the challenge will be achieving a balance between parsimony and complexity (Johnson et al., 2021). In the long-term, a heterogeneous mix of different types of theories would likely benefit the field and further data and time will select those that best reflect reality (Johnson, 2020).

4. Limitations and Future Directions

Having considered the potential implications of the transdiagnostic work presented in the thesis, next I turn to its limitations and outline important future goals that will ensure that tangible impact can be achieved.

4.1. Understanding Age Effects

The CALM cohort has many strengths, including transdiagnostic sampling and a diverse set of measurement domains, which allowed me to gain a deeper understanding of the various manifestations of neurodevelopmental needs. However, there were some limitations regarding its age-heterogeneous and cross-sectional nature. The wide age range (5 to 18) in the context of atypical development made questions about developmental periods and age-related change difficult to address. Therefore, I chose to control for age across all analyses, but I consider that age effects should be more comprehensively investigated in the future. This should include empirical evaluations of hypotheses about age differentiation, which are prominent in both the clinical and cognitive development literature. In the clinical realm, the age differentiation hypothesis refers to the proposition that childhood comorbidity can be explained by the differentiation of difficulties with increasing age (e.g., Sterba et al., 2007; Sterba et al., 2010). Similarly, in the field of cognitive development, it is proposed that cognitive faculties integrate from infancy into childhood and then differentiate to become more
specialised and less correlated from childhood into adulthood (e.g., Tideman & Gustafsson, 2004). So far, evidence for these phenomena (integration/differentiation) in both the clinical and cognitive literature has been notoriously mixed (Groen et al., 2019; McElroy et al., 2018; Simpson-Kent et al., 2020).

4.2. Studying Developmental Change

To address these hypotheses with sufficient rigour longitudinal data from the same children followed over time is essential. At the time of conducting the empirical work of the thesis, the longitudinal data from the CALM cohort was not yet available: data-collection is currently in progress. In addition to understanding age effects, longitudinal follow-up data across relevant timeframes would also enable the formal investigation of other key questions about the stability of individual profiles, developmental cascades, bidirectional cross-domain couplings, and sensitive periods. Studying patterns of desistence and persistence could help identify potentially modifiable factors that can promote positive outcomes. Relatedly, shedding light on how interactions across domains vary in strength and kind across developmental time may be crucial for informing the timing of interventions (Kievit, 2020). Such insights could enable researchers to detect developmental challenges early and intervene before they self-reinforce and/or spill over to other domains.

4.3. Focus on Mechanisms Across Levels

Data-driven descriptive studies documenting strengths and weaknesses and their associations across areas of functioning are useful and have their place. However, to achieve greater impact the insights gained from the current approach need to be supplemented by an enhanced understanding of the mechanisms of developmental change. As discussed in Chapter I, the word mechanism is used in a broad sense here to refer to the entire multifaceted, multi-step dynamic process that leads from an initial state to an outcome state (Benton, 2021). Crucially, candidate mechanisms may involve low-level processes that may not concurrently
be related to the outcome domain and their influence may instead unfold over time. Therefore, the endeavour to study mechanisms should be dynamic and longitudinal and would ideally be approached from very early infancy onwards.

4.4. A Life-Course Perspective

Consequently, an important future direction for transdiagnostic research is the study of early neurodevelopment. Investigations focusing on early infancy will be especially informative as they may elucidate the mechanistic pathways that lead to the phenotypic differences observed in school-aged children. Studying early signs of neurodevelopmental difficulties could potentially identify key liabilities before they become compounded by years of atypical interactions with the environment (Johnson, 2017). Emerging transdiagnostic efforts pursuing this direction follow the younger siblings of children with diagnosed neurodevelopmental difficulties longitudinally. In this case, recruitment is independent of diagnostic status but targeted towards a group of children at increased neurodevelopmental risk (i.e., in the form of a potential genetic predisposition). Examples include the British Autism Study of Infant Siblings (BASIS) and the Study of Autism and ADHD Risk in Siblings (STAARS; Green et al., 2013; Gui et al., 2020), which aim to identify early neural and cognitive precursors of difficulties manifested in later years. Relatedly, The Roots of Autism and ADHD Twin Study in Sweden (RATSS) aims to disentangle how genetic and environmental factors shape neurodevelopmental risk by focusing on monozygotic twin pairs in which only one of the siblings has been diagnosed with a neurodevelopmental condition (Bölte et al., 2014). Such studies can help answer questions about the degree to which later poor functional outcomes are predicted by early atypicalities. It can also help identify cases where specific features arise as a compensatory consequence of early liabilities (Tesink et al., 2009). Large longitudinal nationally representative cohort studies, such as the Adolescent Brain Cognitive Development or Millennium Cohort Study (Garavan et al., 2018; Patalay et al., 2017), which span from
childhood to adulthood can complement these findings with insights about maturational change. A life-course dynamic perspective would be most informative, but this can be difficult to implement and requires substantial resources and commitment from children and families.

4.5. Measurement Practices

One substantial challenge to the life-course developmentally informative research agenda is the difficulty in selecting measurement sets that are appropriate across the lifespan. Performance-based cognitive tasks administered in infancy would not be suitable in early childhood, and those used in early childhood may not tap the same constructs in adolescents. Carefully designed neurocognitive batteries, such as the one used in CALM, offer an intermediate solution because they have been developed to ensure good developmental continuity and cover long developmental periods. In terms of mental health or behavioural ratings, informants typically change from parents to self-report in later life. Latent variable models aiming to capture change at the latent level are usually employed in these cases, but as discussed in Chapter I they are not without drawbacks and cannot be seen as a panacea. Another important measurement challenge worth critical consideration is the vast array of tasks and tools labelled as assessing the same construct, which often show poor correspondence with one another and do not seem to be tapping the same target domain (Dockrell et al., 2017). This poses a large problem for the broader field because reliance on ill-specified measurement constructs can lead to very different results across studies and hamper our ability to consolidate findings and draw conclusions. Generally, while diverse measurement sets are needed to study complexity, their selection remains a challenging task and their suitability and validity should be empirically evaluated.

4.6. Consider Context

Finally, the approach taken in the thesis was largely focused on child-level characteristics and did not sufficiently consider contextual/environmental factors. As discussed
in Chapter I, these are important sources of variance which need to be integrated into future investigations. For example, findings that institutionally-reared children exposed to early severe deprivation display elevated levels of attention deficits and overactivity provide support for the notion that disorder features can arise in atypical social contexts, and stress the need to consider why social risk matters (e.g., Kreppner et al., 2001). Additionally, factors such as socioeconomic status and school environment are known to influence many of the relations studied in this thesis. For instance, the relationship between socioeconomic disadvantage and internalising emotional difficulties is reported to be stronger for children with weaker verbal skills (Flouri et al., 2014). Similarly, the negative influence of low prosociality on internalising and externalising problems is exacerbated in children attending low-performing schools or living in socially disadvantaged neighbourhoods (Flouri & Sarmadi, 2016). Some of these effects could be mediated by the influence of early life stressors, which are known to increase the risk for both internalising and externalising difficulties (March-Llanes et al., 2017). Therefore, low childhood socioeconomic status and early life stress are increasingly being recognised as transdiagnostic risk factors associated with a host of adverse neurodevelopmental outcomes (Lynch et al., 2021). Longitudinal observational designs are important for understanding how such factors operate, but where ethical and feasible, quasi-experimental setups and/or randomised control trails will also be needed to assess causality.

4.7. Summary

Overall, mechanistic research spanning multiple levels of description and following children across development is necessary to address many of the challenges discussed here. This research agenda requires substantial resources and therefore increased data-sharing and collaborative efforts will be crucial for its success. Large-scale longitudinal observations are arguably most relevant for many of the pressing questions within the neurodevelopmental literature, but in due time the insights afforded by this approach may be formalised into
mathematical models and further studied through other methods such as generative or computational models (Akarca et al., 2021; Thomas et al., 2019). Carefully controlled lab experiments and intervention studies will also have a crucial role to play. To sum up this section, I outline three broad themes of questions relevant for advancing our understanding of neurodevelopment:

◇ What drives neurodevelopmental change? How do different developmental domains interact? How does context (e.g., socioeconomic) contribute to neurodevelopmental risk? Can we identify modifiable factors (child-level or environmental) that can support the positive developmental cascades and prevent the negative?

◇ Can we identify neurodevelopmental subgroups that transcend diagnostic boundaries and predict functional outcomes? Would these profiles be clinically and scientifically useful? Should we search for subgroupings based on biological characteristics (genotype, neural profile), phenotype (behavioural, cognitive), response to treatment, environmental experiences/exposures, or a combination?

◇ Would harmonised transdiagnostic assessment protocols improve service provision? Could they subsequently be used to devise transdiagnostic treatments/support plans? Would transdiagnostic interventions prove to be more effective relative to disorder-centred approaches?

4.8. Conclusion

The work conducted within this thesis demonstrates the benefits of a multi-level, multifactorial transdiagnostic framework over the prevailing categorical approach. I have argued that to make progress in understanding the causes and consequences of neurodevelopmental difficulties we need to relax our reliance on outdated assumptions and embrace novel methodological and conceptual directions. This will require continued effort to study diverse populations and to focus not only on the isolated features but also on the dynamic
processes which shape individual outcomes irrespective of diagnostic status. This research should be conducted from early infancy onwards, simultaneously considering multiple levels and their two-way chains of interaction. Enabling evidence-based strategies for alleviating distress and adverse outcomes in later life should be one of the central motivations for studying neurodevelopment and the transdiagnostic approach can help us achieve this. If our samples are more representative of the children we aim to understand, and our methods better suited to capture the dynamics of development, then our theories, although undoubtedly more complex, will likely have greater practical significance. Overall, this thesis contributes to the wider transdiagnostic research effort, which I am optimistic in due time would help many children who struggle neurodevelopmentally to thrive and fulfil their potential.
Appendix I

1. Alternative Network Estimation

The network model presented in the chapter was based on all possible pairwise correlations. To evaluate the robustness of this method I compared the extent to which symptom interrelations based on pairwise correlations are associated with those derived from the listwise correlations \((N = 668)\) by correlating the regularised partial correlation adjacency matrices estimated using each procedure (Figure AI.1). The network presented in Chapter II was derived using sample-centred scores. To check the robustness of this model, estimates derived from networks based on raw scores (without centring) and norm-referenced scores both with and without transformation were compared. Finally, the adjacency matrices derived from networks estimated without thresholding and without regularisation were compared. All estimation methods produced similar solutions as indicated by the high correlations across the derived adjacency matrices (Figure AI.1, \(r\) range 1 to 0.89).
Correlation across the adjacency matrices derived from networks estimated with norm-referenced (nrm.), raw, and sample-centred standardised scores (std.) with (trns.) /without (non-trns.) nonparanormal transformation based on listwise (list)/pairwise (pair) correlations. The final two matrices are derived from networks estimated without thresholding (no-thr.) and without regularisation (unrg), respectively. The network presented in the Chapter II is based on sample-centred scores, applying nonparanormal transformation and thresholding, and using pairwise correlations (std.trns.pair).
2. Centrality Measures Derived from Alternative Estimation Methods

The differences in centrality indices derived from alternative network estimation methods were scrutinised. Some differences across methods were observed (Figure A1.2). Nonetheless, across all models working memory and language coherence were among the most central nodes in line with the results presented in Chapter II.
Figure A1.2

Centrality measures strength and expected influence across the subscales of Conners-3, BRIEF (Behaviour Rating Inventory of Executive Function), and CCC-2 (Children’s Communication Checklist-2) estimated from unregularised, regularised thresholded, and regularised non-thresholded networks.

Notes. Conners-3: Inatt = Inattention; Hyp = Hyperactivity/Impulsivity; Learn = Learning Problems; Agg = Aggression; Peer = Peer Relationships; BRIEF: Inhib = Inhibition; Shift=Shifting; Emot = Emotional Control; Initi = Initiation; WM = Working memory; Plan = Planning/Organisation; Org = Organisation of Materials; Monit = Monitoring; CCC-2: Synt = Syntax; Seman = Semantics; Coher = Coherence; Inap. Initi = Inappropriate Initiation; Stereo = Stereotyped Language; Context = Use of Context; Nonver = Nonverbal Communication; Social = Social Relations; Interest = Interests.
3. Robustness Analyses

To test the accuracy of the regularised partial correlation model the edge weights were bootstrapped ($N$ boots = 2000). Edge weight confidence intervals (CIs) are useful for indicating the accuracy of edge weights and for comparing their strength. The bootstrapped CIs had acceptable widths, suggesting sufficient accuracy (Figure AI.3, separated into three panels to improve visibility). Notably, due to the network estimation method, the CIs are not to be interpreted as significance tests to zero (Epskamp et al., 2018). Instead, sample-derived edge weights and the proportion of time edges were set as different from zero across the 2000 bootstraps are presented in Figure II.3 in Chapter II. The CIs derived from the times the parameter was not set to zero are presented in Figure AI.4 (separated into three panels to improve visibility).
Figure A1.3

The estimated cross-symptom edge weights are represented by the red dots and the means of the bootstrapped edge weights are represented by the black dots. The corresponding bootstrap confidence intervals indicate the edge weight accuracy.
Notes. Conners-3: Inatt = Inattention; Hyp = Hyperactivity/Impulsivity; Learn = Learning Problems; Agg = Aggression; Peer = Peer Relationships; BRIEF: Inhib = Inhibition; Shift=Shifting; Emot = Emotional Control; Initi = Initiation; WM = Working memory; Plan = Planning/Organisation; Org = Organisation of Materials; Monit = Monitoring; CCC-2: Synt = Syntax; Seman = Semantics; Coher = Coherence; Inap. Initi = Inappropriate Initiation; Stereo = Stereotyped Language; Context = Use of Context; Nonver = Nonverbal Communication; Social = Social Relations; Interest = Interests.
Figure A1.4

The estimated cross-symptom edge weights are represented by the red dots and the means of the bootstrapped edge weights are represented by the black dots. The width of the lines corresponds to 95% confidence intervals only for the times the parameter was not set to zero. The transparency of the intervals shows how often an edge was included. Lighter lines indicate that the edge was frequently set to zero.
Notes. Conners-3: Inatt = Inattention; Hyp = Hyperactivity/Impulsivity; Learn = Learning Problems; Agg = Aggression; Peer = Peer Relationships; BRIEF: Inhib = Inhibition; Shift = Shifting; Emot = Emotional Control; Initi = Initiation; WM = Working memory; Plan = Planning/Organisation; Org = Organisation of Materials; Monit = Monitoring; CCC-2: Synt = Syntax; Seman = Semantics; Coher = Coherence; Inap. Initi = Inappropriate Initiation; Stereo = Stereotyped Language; Context = Use of Context; Nonver = Nonverbal Communication; Social = Social Relations; Interest = Interests.
Subcortical and limbic network sub-regions, which showed a significant reduction in connection strength relative to the non-referred sample in only one of the data-driven subgroups. Means represent age, gender, and motion-corrected residuals and are 10% trimmed to account for outliers.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>Non-referred $M_{trim}$</th>
<th>Non-referred $M_{trim}$</th>
<th>$p_{adj}$ Effect size Low 95% CI Up 95% CI</th>
<th>Gyrus</th>
<th>Hemisphere</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Subgroup 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ventromedial putamen</td>
<td>-1.77</td>
<td>2.12</td>
<td>0.03</td>
<td>0.27</td>
<td>0.17</td>
</tr>
<tr>
<td>intermediate lateral area 20</td>
<td>-3.28</td>
<td>-1.64</td>
<td>0.04</td>
<td>0.24</td>
<td>0.12</td>
</tr>
<tr>
<td><strong>Subgroup 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral amygdala</td>
<td>-1.48</td>
<td>0.00</td>
<td>0.02</td>
<td>0.27</td>
<td>0.19</td>
</tr>
<tr>
<td>caudoventral of area 20</td>
<td>-4.59</td>
<td>2.62</td>
<td>0.004</td>
<td>0.36</td>
<td>0.18</td>
</tr>
<tr>
<td>dorsal caudate area TI (temporal agranular insular cortex)</td>
<td>1.12</td>
<td>-3.25</td>
<td>0.04</td>
<td>0.23</td>
<td>0.15</td>
</tr>
<tr>
<td>medial area 38</td>
<td>-5.63</td>
<td>3.92</td>
<td>0.04</td>
<td>0.22</td>
<td>0.13</td>
</tr>
<tr>
<td><strong>Subgroup 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dorsolateral putamen</td>
<td>-6.42</td>
<td>1.29</td>
<td>0.03</td>
<td>0.26</td>
<td>0.10</td>
</tr>
<tr>
<td>rostroventral area 24</td>
<td>-1.53</td>
<td>1.03</td>
<td>0.02</td>
<td>0.27</td>
<td>0.15</td>
</tr>
<tr>
<td>caudal temporal thalamus</td>
<td>-1.25</td>
<td>0.54</td>
<td>0.01</td>
<td>0.30</td>
<td>0.12</td>
</tr>
</tbody>
</table>

*Note.* Subgroup 1: most severe structural language difficulties; Subgroup 2: most severe cool executive difficulties; Subgroup 3: most severe difficulties with hot executive skills & pragmatic communication.
Table AII.2
Subcortical and limbic network sub-regions, which showed a significant reduction in connection strength relative to the non-referred sample in multiple data-driven subgroups. Means represent age, gender, and motion-corrected residuals and are 10% trimmed to account for outliers.

<table>
<thead>
<tr>
<th>Sub-region</th>
<th>M trimmed</th>
<th>Subgroup</th>
<th>Nref. M trimmed</th>
<th>p</th>
<th>Effect size</th>
<th>Low 95% CI</th>
<th>Up 95% CI</th>
<th>Gyrus</th>
<th>Hemi.</th>
</tr>
</thead>
<tbody>
<tr>
<td>caudal hippocampus</td>
<td>-0.77</td>
<td>1</td>
<td>10.31</td>
<td>0.03</td>
<td>0.26</td>
<td>0.19</td>
<td>0.42</td>
<td>Hippocampus</td>
<td>L</td>
</tr>
<tr>
<td>caudal hippocampus</td>
<td>-5.61</td>
<td>2</td>
<td>10.31</td>
<td>0.00</td>
<td>0.42</td>
<td>0.27</td>
<td>0.49</td>
<td>Hippocampus</td>
<td>L</td>
</tr>
<tr>
<td>caudal hippocampus caudoventral of area 20</td>
<td>-5.97</td>
<td>3</td>
<td>10.31</td>
<td>0.01</td>
<td>0.33</td>
<td>0.14</td>
<td>0.52</td>
<td>Inferior Temporal Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>caudoventral of area 20</td>
<td>2.64</td>
<td>1</td>
<td>3.75</td>
<td>0.03</td>
<td>0.28</td>
<td>0.09</td>
<td>0.43</td>
<td>Inferior Temporal Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>globus pallidus</td>
<td>0.03</td>
<td>1</td>
<td>3.15</td>
<td>0.04</td>
<td>0.23</td>
<td>0.12</td>
<td>0.39</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>globus pallidus</td>
<td>-1.47</td>
<td>2</td>
<td>3.15</td>
<td>0.02</td>
<td>0.28</td>
<td>0.19</td>
<td>0.37</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>globus pallidus</td>
<td>-4.85</td>
<td>2</td>
<td>4.00</td>
<td>0.02</td>
<td>0.28</td>
<td>0.11</td>
<td>0.30</td>
<td>Basal Ganglia</td>
<td>R</td>
</tr>
<tr>
<td>globus pallidus</td>
<td>-4.34</td>
<td>3</td>
<td>3.15</td>
<td>0.00</td>
<td>0.47</td>
<td>0.28</td>
<td>0.63</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>globus pallidus intermediate ventral area 20</td>
<td>-6.10</td>
<td>3</td>
<td>4.00</td>
<td>0.01</td>
<td>0.31</td>
<td>0.10</td>
<td>0.45</td>
<td>Basal Ganglia</td>
<td>R</td>
</tr>
<tr>
<td>intermediate ventral area 20</td>
<td>-1.05</td>
<td>1</td>
<td>1.55</td>
<td>0.04</td>
<td>0.22</td>
<td>0.13</td>
<td>0.28</td>
<td>Inferior Temporal Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>intermediate ventral area 20</td>
<td>-1.46</td>
<td>2</td>
<td>1.55</td>
<td>0.04</td>
<td>0.21</td>
<td>0.04</td>
<td>0.35</td>
<td>Inferior Temporal Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>intermediate ventral area 20</td>
<td>-1.22</td>
<td>3</td>
<td>1.55</td>
<td>0.02</td>
<td>0.27</td>
<td>0.15</td>
<td>0.39</td>
<td>Inferior Temporal Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>medial area 11</td>
<td>-4.25</td>
<td>2</td>
<td>4.97</td>
<td>0.04</td>
<td>0.24</td>
<td>0.11</td>
<td>0.31</td>
<td>Orbital Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>medial area 11</td>
<td>-6.02</td>
<td>3</td>
<td>4.97</td>
<td>0.01</td>
<td>0.34</td>
<td>0.20</td>
<td>0.57</td>
<td>Orbital Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>thalamus</td>
<td>-1.72</td>
<td>1</td>
<td>2.94</td>
<td>0.04</td>
<td>0.22</td>
<td>0.08</td>
<td>0.38</td>
<td>Thalamus</td>
<td>L</td>
</tr>
<tr>
<td>medial pre-frontal thalamus</td>
<td>-0.87</td>
<td>2</td>
<td>2.94</td>
<td>0.04</td>
<td>0.23</td>
<td>0.13</td>
<td>0.36</td>
<td>Thalamus</td>
<td>L</td>
</tr>
<tr>
<td>thalamus</td>
<td>-2.39</td>
<td>3</td>
<td>2.94</td>
<td>0.03</td>
<td>0.26</td>
<td>0.19</td>
<td>0.44</td>
<td>Thalamus</td>
<td>L</td>
</tr>
<tr>
<td>nucleus accumbens</td>
<td>1.23</td>
<td>1</td>
<td>4.93</td>
<td>0.04</td>
<td>0.26</td>
<td>0.16</td>
<td>0.35</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>nucleus accumbens</td>
<td>-2.18</td>
<td>2</td>
<td>4.93</td>
<td>0.01</td>
<td>0.32</td>
<td>0.22</td>
<td>0.39</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>nucleus accumbens</td>
<td>-5.52</td>
<td>3</td>
<td>4.93</td>
<td>0.01</td>
<td>0.32</td>
<td>0.21</td>
<td>0.47</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>rostroventral area 20</td>
<td>-1.73</td>
<td>1</td>
<td>9.17</td>
<td>0.04</td>
<td>0.22</td>
<td>0.10</td>
<td>0.34</td>
<td>Fusiform Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>rostroventral area 20</td>
<td>-7.99</td>
<td>2</td>
<td>9.17</td>
<td>0.01</td>
<td>0.34</td>
<td>0.14</td>
<td>0.49</td>
<td>Fusiform Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>rostroventral area 20</td>
<td>-4.33</td>
<td>3</td>
<td>9.17</td>
<td>0.01</td>
<td>0.31</td>
<td>0.15</td>
<td>0.47</td>
<td>Fusiform Gyrus</td>
<td>L</td>
</tr>
<tr>
<td>ventral caudate</td>
<td>-4.80</td>
<td>1</td>
<td>5.61</td>
<td>0.03</td>
<td>0.32</td>
<td>0.29</td>
<td>0.42</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>ventral caudate</td>
<td>-2.64</td>
<td>1</td>
<td>2.23</td>
<td>0.04</td>
<td>0.24</td>
<td>0.06</td>
<td>0.26</td>
<td>Basal Ganglia</td>
<td>R</td>
</tr>
<tr>
<td>ventral caudate</td>
<td>-1.19</td>
<td>2</td>
<td>5.61</td>
<td>0.04</td>
<td>0.20</td>
<td>0.03</td>
<td>0.26</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
<tr>
<td>ventral caudate</td>
<td>-2.32</td>
<td>2</td>
<td>2.23</td>
<td>0.04</td>
<td>0.21</td>
<td>0.09</td>
<td>0.34</td>
<td>Basal Ganglia</td>
<td>R</td>
</tr>
<tr>
<td>ventral caudate</td>
<td>-3.85</td>
<td>3</td>
<td>5.61</td>
<td>0.01</td>
<td>0.39</td>
<td>0.20</td>
<td>0.50</td>
<td>Basal Ganglia</td>
<td>L</td>
</tr>
</tbody>
</table>

Note: Subgroup 1: most severe structural language difficulties; Subgroup 2: most severe cool executive difficulties; Subgroup 3: most severe difficulties with hot executive skills & pragmatic communication. Nref. = non-referred.
Figure AII.1

Subnetworks identified as significantly weaker in Subgroup 3 relative to the non-referred comparison sample when a more liberal family-wise error correction (p<.05) was applied.

Note. Subgroup 3: most severe difficulties with hot executive skills & pragmatic communication.
Subnetworks identified as significantly different from the non-referred CALM200 sample in Subgroup 1, Subgroup 2, and Subgroup 3 across several t-value thresholds. The figure demonstrates how the size of the identified subnetworks reduced with increasing t-value thresholds.

Note. Subgroup 1: most severe structural language difficulties; Subgroup 2: most severe cool executive difficulties; Subgroup 3: most severe difficulties with hot executive skills & pragmatic communication.
Appendix III

1. Software and Library Versions

All analyses were performed in R (R Core Team, 2020) using the packages bootnet version 1.4.2 (Epskamp et al., 2018), factoextra version 1.0.7 (Kassambara & Mundt, 2020), igraph version 1.2.5 (Csardi & Nepusz, 2006), matchit version 3.0.2 (Ho et al., 2011), mgm version 1.2.9 (Haslbeck & Waldorp, 2015), mice version 3.9.0 (van Buuren & Groothuis-Oudshoorn, 2011), NetworkComparisonTest version 2.2.1 (Van Borkulo et al., 2016), qgraph version 1.6.5 (Epskamp et al., 2012), and psych version 1.9.12.31 (Revelle, 2019).
2. Results Following Missing Data Imputation with MICE

The overall proportion of missing data was similar across cohorts (NKI-RS: 8.17%; CALM: 8.74%). There is no consensus approach for handling missing data in network analysis. Therefore, for robustness, two approaches were simultaneously estimated and compared: pairwise associations and multivariate imputation by chained equations (MICE) via R package mice (van Buuren & Groothuis-Oudshoorn, 2011). The MICE procedure was used to generate 100 imputed datasets with 50 iterations per cohort, using predictive mean matching algorithms (van Buuren & Groothuis-Oudshoorn, 2011). The 100 imputed datasets per cohort were used to estimate 100 networks per cohort following the same method as described in Chapter IV. The estimated task interrelations were then averaged to generate one average network model per cohort. Spearman correlation coefficients of task-interrelation estimates were calculated to examine how similar these average models were to the models presented in Chapter IV (estimated with pairwise associations). The correspondence across methods was excellent: CALM: $r_s = .98, p < .001$; NKI-RS: $r_s = .97, p < .001$. Furthermore, Spearman coefficients comparing the strength centrality estimates generated based on the two approaches for handling missing data also suggested excellent agreement across methods (CALM: $r_s = 1, p < .001$; NKI-RS: $r_s = .98, p < .001$).

Finally, all cohort comparisons were repeated. The results were in agreement with those based on pairwise correlations. In terms of the similarity in task interrelationships between CALM and NKI-RS, the Spearman coefficient based on the average of the 100 imputed datasets per cohort was similar to the one presented in Chapter IV ($r_s = .62, p < .001$). A permutation test with 1000 iterations comparing the 100 pairs of imputed datasets, provided similar conclusions to those presented in Chapter IV. The average observed difference in global strength was 0.10 (min = 0.003; max = 0.74), with 99% of comparisons suggesting no differences in global strength ($\alpha = 0.05$). For structure invariance, permutations based on each of the 100 pairs of imputed datasets suggested a significant difference in global structure ($\alpha = 0.05$). Looking at the edges across the 100 comparisons, the four edges flagged as significantly different between cohorts in Chapter IV were the ones that most often fell below the significance threshold following a false discovery rate correction (see Table AIII.1).
**Table AIII.1**

*Mean (M), Minimum (Min), and Maximum (Max) observed edge weights (ew) for CALM and NKI-RS across network estimations based on 100 imputed datasets per cohort. Mean p-values and percentage of comparisons in which edge weights were flagged as significantly different following permutation tests (adjusted for false discovery rate).*

<table>
<thead>
<tr>
<th>Edge</th>
<th>CALM</th>
<th>NKI-RS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M_{ew}</td>
<td>Min_{ew}</td>
</tr>
<tr>
<td>Math - MxReas</td>
<td>0.33</td>
<td>0.28</td>
</tr>
<tr>
<td>Lit - Math</td>
<td>0.20</td>
<td>0.14</td>
</tr>
<tr>
<td>MxReas - Switch</td>
<td>0.10</td>
<td>0.04</td>
</tr>
<tr>
<td>BDR - Math</td>
<td>0.32</td>
<td>0.27</td>
</tr>
</tbody>
</table>

*Note.* Only edges flagged as significantly different in at least half of the comparisons are displayed in the table. BDR = WISC-R/AWMA Backward digit recall; Lit = Literacy (WIAT-II Single word reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix reasoning; Switch = D-KEFS Trails number-letter sequencing task. Sign. = significantly
3. Homogeneity of Variance

Fligner-Killeen’s non-parametric test of homogeneity of variance indicated that there was a significant difference in variance across the two cohorts (Table AIII.2). These differences were not systematic (i.e., for some measures the variance was greater in CALM, for others the opposite was true). Differences in the variance across groups can alter the strength of the connections in the networks (Terluin et al., 2016). Therefore, to test the robustness of the results obtained from the raw data, all key analyses were repeated following a data transformation that decreased the difference in variances. The data transformation was conducted in three steps: first, the data from both cohorts was pooled together; second, participant percentage-ranks ranging from zero to one were separately estimated for each task; and finally, the resulting ranks were then mapped onto a standard normal distribution using a normal quantile function. This procedure limits the influence of univariate outliers by altering the distances between scores to resemble a normal distribution while maintaining individual ranks across tasks (Bignardi et al., 2021; Gregory, 2014). Following the transformation, the differences in variance between groups were no longer significant (Table AIII.2). The transformed data was then analysed following the same procedures as those used with the raw data (i.e., as described in Chapter IV). The networks estimated with the transformed data retained good correspondence with the networks based on the raw data (Spearman correlation coefficients of task-interrelation estimates: CALM: $r_s=0.96, p<.0001$; NKI-RS: $r_s = 0.88, p < .001$). The strength centrality estimates generated based on the transformed data were also similar to the ones derived from the raw data (CALM: $r_s = 0.93, p < .001$; NKI-RS: $r_s =0.82, p = .01$). In terms of cohort comparisons, the task interrelation similarity between CALM and NKI-RS after data transformation was again moderate ($r_s = 0.54, p <.001$). A permutation test with 1000 iterations comparing the cohort networks, estimated following the transformation, provided similar conclusions to those suggested by the raw data. Consistent with the outcomes from the raw data, the differences in global strength were not significant ($S = 0.19, p = 0.629$), and the network structures were not invariant ($M = 0.31, p <0.001$). Following a false discovery rate correction, three of the edges flagged as significantly different based on the analysis of the raw data were again below the significance threshold (Matrix reasoning – Maths: $r_{\text{CALM}} = 0.36$, $r_{\text{NKI-RS}} = 0.08$, $p < .001$; Literacy – Maths: $r_{\text{CALM}} = 0.25$, $r_{\text{NKI-RS}} = 0.46$, $p = .045$; Matrix reasoning – Switching: $r_{\text{CALM}} = 0.08$, $r_{\text{NKI-RS}} = 0.29$)
The edge between Backward digit span and Maths, which was significantly different between cohorts when the raw data was analysed was not significantly different following the data transformation: \( r_{\text{CALM}} = 0.20, r_{\text{NKI-RS}} = 0.06, p = .27 \).

Table AIII.2

The output of Fligner-Killeen’s non-parametric test of homogeneity of variance across CALM and NKI-RS before and after data transformation.

<table>
<thead>
<tr>
<th></th>
<th>Raw data</th>
<th>Transformed data</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CALM Variance</td>
<td>NKI-RS Variance</td>
<td>( \chi^2 )</td>
<td>( p )</td>
<td>CALM Variance</td>
<td>NKI-RS Variance</td>
</tr>
<tr>
<td>Backward Digit Span</td>
<td>0.72</td>
<td>1.62</td>
<td>17.10</td>
<td>&lt;.001</td>
<td>0.57</td>
<td>0.54</td>
</tr>
<tr>
<td>Forward Digit Span</td>
<td>1.03</td>
<td>1.51</td>
<td>16.25</td>
<td>&lt;.001</td>
<td>0.82</td>
<td>1</td>
</tr>
<tr>
<td>Literacy</td>
<td>179.81</td>
<td>116.87</td>
<td>13.72</td>
<td>&lt;.001</td>
<td>0.84</td>
<td>0.80</td>
</tr>
<tr>
<td>Numerical Operations1</td>
<td>78.65</td>
<td>113.60</td>
<td>14.09</td>
<td>&lt;.001</td>
<td>0.79</td>
<td>0.82</td>
</tr>
<tr>
<td>Matrix Reasoning</td>
<td>30.50</td>
<td>21.52</td>
<td>16.35</td>
<td>&lt;.001</td>
<td>0.91</td>
<td>0.74</td>
</tr>
<tr>
<td>Visual Scanning</td>
<td>135.05</td>
<td>91.18</td>
<td>13.05</td>
<td>&lt;.001</td>
<td>0.99</td>
<td>0.93</td>
</tr>
<tr>
<td>Motor Speed</td>
<td>293.35</td>
<td>262.53</td>
<td>7.94</td>
<td>.005</td>
<td>0.86</td>
<td>1</td>
</tr>
<tr>
<td>Number-Letter Switching</td>
<td>3,717.25</td>
<td>3,373.50</td>
<td>9.27</td>
<td>.002</td>
<td>0.76</td>
<td>0.98</td>
</tr>
<tr>
<td>Tower Achievement</td>
<td>14.46</td>
<td>14.10</td>
<td>0.04</td>
<td>0.83</td>
<td>0.98</td>
<td>0.98</td>
</tr>
</tbody>
</table>
4. Investigating Age Effects: Network Moderation Analysis

To evaluate whether age moderated the strength of network interrelations in either cohort, a regularised Moderated Network Model (MNM) was estimated for each cohort, with age specified as a continuous moderator (Haslbeck, 2020). As discussed in Chapter IV, this analysis was not sufficiently powered to detect the small age-moderation effects typically observed in the literature and should be interpreted with caution. MNMs extend the regularised pairwise Gaussian Graphical Model to include moderation effects. The procedure was implemented in the R package mgm (Haslbeck & Waldorp, 2015) as follows: first, a regularised MNM with age specified as a continuous moderator was fitted to each of the hundred imputed datasets per cohort; second, the stability of the MNM results was assessed by bootstrapping (N boots = 100) each imputed dataset to obtain confidence intervals around the obtained estimates; finally, the bootstrapped estimates from all 100 imputed datasets were averaged for each cohort and are shown in Figures AIII.1 and AIII.2. Across the 100 imputed datasets per cohort, age did not significantly moderate (p > .05) any of the edges in either cohort, which could likely be due to low power to detect the moderation effects typically reported in the literature.
Figure AIII.1

The bootstrapped age moderation coefficients averaged across the 100 imputed CALM datasets are represented by the location of the dot. The width of the lines corresponds to the 95% confidence intervals. The value inside the dot corresponds to the average proportion of bootstraps in which the estimate was set as different from zero.
Figure AIII.2

The bootstrapped age moderation coefficients averaged across the 100 imputed NKI datasets are represented by the location of the dot. The width of the lines corresponds to the 95% confidence intervals. The value inside the dot corresponds to the average proportion of bootstraps in which the estimate was set as different from zero.
5. Network Stability

The bootstrapped 95% confidence intervals (CIs) around the estimated edge weights/task interrelations (N boots = 1000) are presented in Figure AIII.3 (CALM) and Figure AIII.4 (NKI-RS). The CIs derived from the times the parameter was not set to zero, together with the proportion of bootstrapped samples in which edges were set as different from zero across the 1000 iterations are presented in Figure AIII.5 (CALM) and Figure AIII.6 (NKI-RS). These results together with edge weight estimates are again summarised in Figure 5 in Chapter IV. As shown in Figure AIII.7 and Figure AIII.8, the smallest edge weights in each cohort had overlapping CIs and did not differ significantly from other edges, indicating that their magnitude and rank-ordering should be interpreted with caution.
Figure AIII.3

The estimated task interrelations/edge weights for CALM are represented by the red dots and the means of the bootstrapped edge weights are represented by the black dots. The widths of the corresponding bootstrapped 95% confidence intervals indicate the edge weight accuracy.
Figure AIII.4

The estimated task interrelations/edge weights for NKI-RS are represented by the red dots and the means of the bootstrapped edge weights are represented by the black dots. The widths of the corresponding bootstrapped 95% confidence intervals indicate the edge weight accuracy.
The estimated edge weights for CALM are represented in red and the means of the bootstrapped edge weights are in black. The width of the lines corresponds to 95% confidence intervals for the times the parameter was not set to zero. Transparency indicates how often an edge was included, with lighter lines showing that it was frequently set to zero.
Figure AIII.6

The estimated edge weights for NKI-RS are represented in red and the means of the bootstrapped edge weights are in black. The width of the lines corresponds to 95% confidence intervals for the times the parameter was not set to zero. Transparency indicates how often an edge was included, with lighter lines showing that it was frequently set to zero.

Figure AIII.7

Bootstrapped edge difference tests ($\alpha = 0.05$) across non-zero task interrelations in CALM. The colours on the diagonal indicate the magnitude of task interrelations, with darker colours reflecting stronger
associations. The grey boxes suggest no significant difference across task interrelations, whereas the black boxes indicate significantly different task interrelations (i.e., the Math-MxReas interrelation is significantly stronger than most other edges).

Note. BDR = WISC-R/AWMA Backward digit recall; FDR = WISC-R/AWMA Forward digit recall; Lit = Literacy (WIAT-II Single word reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
Figure AIII.8
Bootstrapped edge difference tests ($\alpha = 0.05$) across non-zero task interrelations in NKI-RS. The colours on the diagonal indicate the magnitude of task interrelations, with darker colours reflecting stronger associations. The grey boxes suggest no significant difference across task interrelations, whereas the black boxes indicate significantly different task interrelations (i.e., the Lit-Math interrelation is significantly stronger than most others).

Note. Note. BDR = WISC-R/AWMA Backward digit recall; FDR = WISC-R/AWMA Forward digit recall; Lit = Literacy (WIAT-II Single word reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
6. Network Comparison via Confirmatory Fit

The robustness of the results was examined through a different analysis pipeline. For this analysis, the NKI-RS was used as a training cohort to identify the best-fitting Gaussian graphical model via the R package psychonetrics, version 0.7.2 (Epskamp, 2020). The fit of the identified model was then tested in CALM. In the training cohort (NKI-RS), the prune function was used to automatically and reclusively remove parameters that do not meet the $\alpha = 0.05$ level. This was done in combination with the stepup function, which adds edges according to modification indices at $\alpha = 0.05$ level. This approach resulted in a sparser network compared to the regularisation procedure presented in Chapter IV (see the resulting network in Figure AIII.9). Subsequently, the model structure identified in NKI-RS (i.e., the edges included in the final NKI-RS model) was then fitted to CALM. Fit indices suggested poor fit (see Table AIII.3). Modification indices were then examined to flag potential differences in the task interrelation structure. In agreement with the results of the permutation testing reported in Chapter IV, the top modification index pointed at one of the edges, which was identified as significantly different based on the permutation test: Maths - Matrix Reasoning. Allowing this parameter to be included in the CALM model significantly improved the fit (see Table AIII.3). For the remaining estimates, the largest differences in edge weights based on this analysis pipeline were consistent with those reported in Chapter IV (Maths – Literacy, Maths – Backward digit recall, Matrix reasoning – Switching).
Figure AIII.9

Displayed on the right are model parameters for the model identified based on NKI-RS data. Shown on the left are parameters estimated following a confirmatory fit of the same model structure to the CALM data with one modification (Maths – Matrix Reasoning).

Note. BDR = WISC-R/AWMA Backward digit recall; FDR = WISC-R/AWMA Forward digit recall; Lit = Literacy (WIAT-II Single word reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
Table AIII.3

Model fit indices following model identification based on NKI-RS data, confirmatory fit of NKI-RS identified structure on CALM data, and a modified model for CALM based on modification indices (i.e., including the edge Maths – Matrix Reasoning).

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>DF</th>
<th>p</th>
<th>BIC</th>
<th>CFI</th>
<th>TLI</th>
<th>RMSEA</th>
<th>$\Delta \chi^2$</th>
<th>$\Delta$ DF</th>
<th>$\Delta$ p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NKI-RS</td>
<td>28.52</td>
<td>23</td>
<td>0.20</td>
<td>20,269</td>
<td>0.99</td>
<td>0.03</td>
<td></td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CALM</td>
<td>116.07</td>
<td>23</td>
<td>0</td>
<td>21,694</td>
<td>0.94</td>
<td>0.88</td>
<td>0.11</td>
<td>60.82</td>
<td>1</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>CALM Mod.</td>
<td>55.25</td>
<td>22</td>
<td>0</td>
<td>21,639</td>
<td>0.98</td>
<td>0.95</td>
<td>0.07</td>
<td>50.07</td>
<td>1</td>
<td>&lt; .001</td>
</tr>
</tbody>
</table>

Note. BIC = Bayesian Information Criterion; CFI = Comparative Fit Index; TLI = Tucker Lewis index; RMSEA = Root Mean Square Error of Approximation.
7. Joint Network Estimation with Fused Graphical Lasso

Joint network estimation via fused graphical lasso (FGL) provides an alternative approach for investigating patterns of similarities and differences across groups (Danaher et al., 2014). The main advantage of FGL over other methods is that it exploits the similarities across groups to improve estimates, without masking differences across groups (Costantini et al., 2019). The method has previously been employed to compare symptom interrelations across individuals with diagnosed psychiatric conditions and community samples (Richetin et al., 2017). Similar to the graphical least absolute shrinkage and selection operator, the FGL uses a tuning parameter to apply a penalty to the sum of absolute values of the concentration matrix elements. In addition, it uses a second tuning parameter to penalise the sum of absolute values of the differences between the corresponding concentration matrix elements across groups. In the current analysis, the values of both tuning parameters were selected based on information criteria (EBIC) for consistency with the analysis presented in Chapter IV. All analyses were implemented in the R package EstimateGroupNetwork, version 0.2.2 (Costantini & Epskamp, 2017). The resulting estimates are displayed in Figure AIII.10. Overall, the results retained substantial similarity to those based on the networks presented in Chapter IV (CALM: $r_s = 1, p < .001$; NKI-RS: $r_s = 0.99, p < .001$). The degree of similarity across cohorts was again moderate – $r_s = .65, p < .001$. The largest differences in cohort edge weights corresponded to those identified through the separate model estimation presented in Chapter IV: Maths – Matrix Reasoning, Maths – Backward Digit Recall, Maths – Literacy, and Maths – Switching.
Figure AIII.10

Edge weights for CALM and NKI-RS following a joint network estimation based on fused graphical lasso.

Note. BDR = WISC-R/AWMA Backward digit recall; FDR = WISC-R/AWMA Forward digit recall; Lit = Literacy (WIAT-II Single word reading and Spelling); Math = WIAT-II Numerical operations; MxReas = WASI-II Matrix reasoning; Scan = D-KEFS Visual scanning; Speed = D-KEFS Motor speed; Switch = D-KEFS Trails number-letter sequencing task; Tower = D-KEFS Tower.
References


https://doi.org/10.1176/appi.ajp.2016.15111435


https://doi.org/10.1016/J.NEUROIMAGE.2016.06.058


Best, W., Hughes, L., Masterson, J., Thomas, M. S. C., Howard, D., Kapikian, A., &


Dockrell, J. E., Lindsay, G., & Palikara, O. (2011). Explaining the academic achievement at school leaving for pupils with a history of language impairment: Previous academic achievement and literacy skills. Child Language Teaching and Therapy, 27(2), 223–237. https://doi.org/10.1177/0265669011398671


Durkin, K., & Conti-Ramsden, G. (2010). Young people with specific language impairment:


Cognition From Childhood to Adolescence. *Developmental Psychology, 43*(6), 1460–1473. https://doi.org/10.1037/0012-1649.43.6.1460


Clinical Psychological Science, 6(3), 335–351.
https://doi.org/10.1177/2167702617745092

https://doi.org/10.1093/biostatistics/kxm045

https://doi.org/10.1016/J.CORTEX.2016.04.023


https://doi.org/10.1002/acp.934


estimating the number of dimensions in psychological research. *PLoS ONE, 12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035


Green, J., Wan, M. W., Guiraud, J., Holsgrove, S., McNally, J., Slonims, V., Elsabbagh, M., Charman, T., Pickles, A., Johnson, M., Baron-Cohen, S., Bedford, R., Bolton, P.,


http://www.jstatsoft.org/


https://doi.org/http://dx.doi.org/10.1037/edu0000644


https://doi.org/10.3389/fnhum.2014.00976


Kassambara, A., & Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses*. https://cran.r-


Kushki, A., Anagnostou, E., Hammill, C., Duez, P., Brian, J., Iaboni, A., Schachar, R.,


C. Thomas, D. L. Segal, & M. Hersen (Eds.), *Comprehensive Handbook of Personality and Psychopathology* (pp. 23–43). John Wiley & Sons Inc.


and Psychiatry and Allied Disciplines, 61(6), 681–688.

https://doi.org/10.1111/jcpp.13150


Salazar, F., Baird, G., Chandler, S., Tseng, E., O’sullivan, T., Howlin, P., Pickles, A., &


Szücs, D. (2016). Subtypes and comorbidity in mathematical learning disabilities:


Van Borkulo, C. D., Epskamp, S., & Maintainer, A. M. (2016). *Statistical Comparison of Two Networks Based on Three Invariance Measures.* https://cran.r-


https://doi.org/10.1111/1467-9604.12027


https://doi.org/10.1111/j.1467-7687.2007.00567.x


https://doi.org/10.1080/02664763.2010.498507


https://doi.org/10.1016/J.LINDIF.2008.01.003


