

## PAPER

# Remapping the cognitive and neural profiles of children who struggle at school

Duncan E. Astle | Joe Bathelt  | The CALM Team | Joni Holmes

MRC Cognition and Brain Sciences  
Unit, University of Cambridge, Cambridge,  
UK

## Correspondence

Duncan E. Astle, MRC Cognition and Brain  
Sciences Unit, University of Cambridge,  
Cambridge, UK.  
Email: duncan.astle@mrc-cbu.cam.ac.uk

## Funding information

Medical Research Council, Grant/Award  
Number: 5PQ00, 5PQ20 and 5PQ40

## Abstract

Our understanding of learning difficulties largely comes from children with specific diagnoses or individuals selected from community/clinical samples according to strict inclusion criteria. Applying strict exclusionary criteria overemphasizes within group homogeneity and between group differences, and fails to capture comorbidity. Here, we identify cognitive profiles in a large heterogeneous sample of struggling learners, using unsupervised machine learning in the form of an artificial neural network. Children were referred to the Centre for Attention Learning and Memory (CALM) by health and education professionals, irrespective of diagnosis or comorbidity, for problems in attention, memory, language, or poor school progress ( $n = 530$ ). Children completed a battery of cognitive and learning assessments, underwent a structural MRI scan, and their parents completed behavior questionnaires. Within the network we could identify four groups of children: (a) children with broad cognitive difficulties, and severe reading, spelling and maths problems; (b) children with age-typical cognitive abilities and learning profiles; (c) children with working memory problems; and (d) children with phonological difficulties. Despite their contrasting cognitive profiles, the learning profiles for the latter two groups did not differ: both were around 1 SD below age-expected levels on all learning measures. Importantly a child's cognitive profile was not predicted by diagnosis or referral reason. We also constructed whole-brain structural connectomes for children from these four groupings ( $n = 184$ ), alongside an additional group of typically developing children ( $n = 36$ ), and identified distinct patterns of brain organization for each group. This study represents a novel move toward identifying data-driven neurocognitive dimensions underlying learning-related difficulties in a representative sample of poor learners.

## KEYWORDS

cognitive development, education, learning difficulties, machine learning

## 1 | INTRODUCTION

Prevalence rates of developmental disorders linked with learning difficulties, including attention deficit hyperactivity disorder (ADHD), dyslexia, dyscalculia, and specific language impairment (SLI), range

from 3% to 8% (American Psychiatric Association, 2013; Norbury et al., 2016; Polanczyk, Willcutt, Salum, Kieling, & Rohde, 2014; Shalev & Gross-Tsur, 2001). However, the number of children who struggle at school is far higher. In the UK for example, around 30% of the school population fail to meet expected targets in reading or

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. *Developmental Science* Published by John Wiley & Sons Ltd.



maths at age 11 (Department for Education United Kingdom, 2017). The long-term outcomes for children who struggle at school include continued educational underachievement, poor mental health (Roeser & Eccles, 2000), and underemployment (Bynner & Parsons, 2005; de Beer, Engels, Heerkens, & van der Klink, 2014).

Our understanding of the causes of learning difficulties comes largely from studying children with a specific diagnosis (e.g., ADHD or SLI) or those selected from community or clinical samples on the basis of strict inclusion criteria (e.g., children with poor reading skills, but age-typical IQ and maths abilities). Most studies recruit children with “pure” problems (e.g., children with ADHD without comorbid dyslexia, or children with maths problems in the absence of reading problems or low IQ). There are practical advantages to this approach: it outlines clear criteria to inform practitioner decision-making about primary areas of weakness that can be used to identify intervention options.

However, this approach can fail to accommodate the high rates of comorbidity within developmental disorders (Coghill & Sonuga-Barke, 2012; Kotov et al., 2017) and learning-related difficulties (e.g., Angold, Costello, & Erkanli, 1999). Over 80% of children with ADHD meet criteria for at least one additional diagnosis (e.g., Faraone & Biederman, 1998; Willcutt & Pennington, 2000) and 15%–45% have co-occurring reading difficulties (e.g., Biederman et al., 1991; Faraone et al., 1993; Semrud-Clikeman et al., 1992). Reading difficulties also co-occur 50% of the time with maths (Moll et al., 2014) or language problems (McArthur et al., 2000).

Using strict exclusionary criteria also overemphasizes similarities within groups, and the distinctiveness between groups (Coghill & Sonuga-Barke, 2012; Kotov et al., 2017). It is widely documented that symptoms vary between children with the same diagnosis. For example, performance on cognitive tasks within ADHD groups is notoriously variable (Castellanos et al., 2005; Nigg, Willcutt, Doyle, & Sonuga-Barke, 2005). Symptoms also co-occur across groups. For example, symptoms of inattention are common in children with poor literacy and maths skills (Hart et al., 2010; Loe & Feldman, 2007; Zentall, 2007), ADHD, autism spectrum disorder (ASD; Rommelse, Geurts, Franke, Buitelaar & Hartman, 2011), SLI (Duinmeijer, Jong & de Scheper, 2012), and dyslexia (Germano, Gagliano, & Curatolo, 2010; Willcutt & Pennington, 2000). Finally, this approach of selectively grouping children does not capture the majority of struggling learners—they often do not receive a diagnosis or are characterized by complex and comorbid difficulties that would rule them out of studies with strict inclusion criteria.

For these reasons a number of researchers have advocated empirically based quantitative classification systems (Archibald, Cardy, Joannis, & Ansari, 2013; Coghill & Sonuga-Barke, 2012; Ramus, Marshall, Rosen, & van der Lely, 2013; Sonuga-Barke & Coghill, 2014), although few studies have done this. The aim of this approach is to move away from identifying highly selective discrete groups and instead focus on identifying continuous dimensions that distinguish individuals and can be used as potential targets for intervention. Dimensions are derived through data-driven explorations of the data, with no *a priori* assumptions about group membership. For example, factor analysis, a statistical method that groups variables based on shared variance,

## RESEARCH HIGHLIGHTS

- First study to apply machine learning to understand heterogeneity in struggling learners.
- Large sample of struggling learners that includes children with multiple difficulties.
- Rich phenotyping with detailed behavioral, cognitive, and neuroimaging assessments.

is used most commonly to derive underlying dimensions from sets of symptoms or measures (e.g., Kotov et al., 2017). This technique has been used to identify dimensions of phonological and nonphonological skills in children with diagnosed SLI and dyslexia (Ramus et al., 2013) and separate latent constructs for inattention and hyperactivity in children with ADHD (Martel, Von Eye, & Nigg, 2010). An alternative approach, as yet rarely used, is to cluster children together according to shared profiles based on empirical data. In turn this can be used to inform classification systems, and consequently treatment approaches. Clustering algorithms have been used to identify groups of children with distinct learning (Archibald et al., 2013) and behavioral profiles (Bathelt, Holmes, the CALM Team, & Astle, *in press*).

In this study, we use a different data-driven approach—machine learning. Machine learning methods have rarely been applied to understanding developmental disorders (e.g., Fair, Bathula, Nikolas, & Nigg, 2012). Typical applications use supervised machine learning (Peng, Lin, Zhang, & Wang, 2013) in which the algorithm attempts to learn about predefined categories of children. Here, we use an unsupervised learning approach whereby the algorithm attempts to learn about the structure of the data itself rather than which data correspond to predefined groups. Specifically, we used Self Organising Maps (SOMs; Kohonen, 1989), a type of artificial neural network. Due to their efficacy in visualizing multidimensional data, SOMs have been successfully applied to a variety of tasks including textual information retrieval (Lin, Soergel, & Marchionini, 1991), the interpretation of gene expression data (Tamayo et al., 1999), and ecological community modeling (Giraudel & Lek, 2001). SOMs use an algorithm that projects the original data from a multidimensional input space onto a two-dimensional grid of nodes called a “map”, while preserving topographical information. This produces an intervariable representational space, wherein the geometric distance between nodes corresponds to the degree of similarity in the input data. Within the current context, input data are individual children from our sample. The map will represent the cognitive profiles of the children; the closer the children are represented within the map, the more similar their cognitive profiles. In this way, SOMs enable us to map the multidimensional space of our sample—the map will represent how different children group together because of their similar profiles, and in doing so it also learns about the dimensions that most reliably distinguish children.

We applied this technique to a large heterogeneous sample of struggling learners. Children were referred to a research clinic, the

Centre for Attention Learning and Memory (CALM), by health and education professionals for ongoing problems in attention, memory, language, or poor school progress in reading and/or maths. Recruitment was deliberately broad to capture the wide range of poor learners in the school population. Children were accepted into the study irrespective of diagnosis or comorbidity: only non-native English speakers and those with uncorrected sight or hearing problems were excluded. Our first aim was to test whether the multidimensional structure learnt by the map reflects in different sample characteristics, such as the primary reason for referral to the research clinic (e.g., problems in attention, learning, memory, or language).

A second aim of the current study was to use the information from the SOM to identify data-driven groups within the sample. Even though it is likely that the dimensions that distinguish children are continuous, there may be important reasons to need to group children according to their shared cognitive profile: (a) to identify shared etiological mechanisms, which will be easier with data-driven homogenous groups; and (b) to identify groups for a particular intervention. To do this the SOM was combined with another form of machine learning, k-means clustering (Lloyd, 1982). This combination identified groups of children with similar cognitive profiles. Having grouped the children with the cognitive data, we then explored the learning and behavioral profiles of these groups. We also explored differences in white-matter connectivity between the data-driven groups. White matter maturation is a crucial process of brain development that extends into the third decade of life (Lebel, Treit, & Beaulieu, 2017) and relates closely to cognitive development (Clayden et al., 2012; Stevens, Skudlarski, Pearlson, & Calhoun, 2009). The brain can be modeled as a network of brain regions connected by white matter, commonly referred to as a connectome (Hagmann et al., 2008). We derived whole-brain connectomes and compared them across the groups produced by the machine learning. In short, our second aim was to use machine learning to identify groups of children with shared cognitive profiles, and then test whether these groups differ on learning and behavioral measures, and in terms of brain organization.

This mapping process is intentionally exploratory, and given this novel application of the analytical approach alongside a unique sample, it is difficult to make clear predictions about what the algorithm will learn. The children attending the clinic completed assessments of the cognitive skills known to be impaired in children with learning-related problems including measures of phonological processing, short-term and working memory, attention and fluid reasoning (nonverbal IQ). Children with deficits in reading or language, or associated diagnoses of dyslexia or SLI often have phonological processing problems (Bishop & Snowling, 2004; Joanisse, Manis, Keating, & Seidenberg, 2000; Ramus et al., 2010; Vellutino, Fletcher, Snowling, & Scanlon, 2004). In contrast, those with specific problems in maths or diagnosed dyscalculia are typically characterized by more severe deficits in spatial short-term and working memory (Geary, 2004; Holmes, Adams, & Hamilton, 2008; McKenzie, Bull, & Gray, 2003; McLean & Hitch,

1999; Rasmussen & Bisanz, 2005; Simmons, Singleton, & Horne, 2008; Swanson & Sachse-Lee, 2001) and broader executive functions (Bull, Espy, & Wiebe, 2008; Bull, Espy, Wiebe, Sheffield, & Nelson, 2011; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013; Van der Ven, Kroesbergen, Boom, & Leseman, 2012). So, a reasonable prediction is that our large sample of struggling learners will include subgroups of children with either phonological problems or spatial short-term/working memory difficulties, and that these children will predominantly struggle with reading or maths respectively. Below average nonverbal reasoning is common among individuals with reading (Duranovic, Tinjak, & Turbic-Hadzagic, 2014; Gathercole et al., 2016; Pointus, 1981; Winner et al., 2001) and maths problems (Gathercole et al., 2016; Swanson & Beebe-Frankenberger, 2004; Fuchs et al., 2005, 2006; Cirino et al., 2015), as well as those with ADHD (Holmes et al., 2013). So, another reasonable prediction is that our sample of struggling learners will include a subgroup of children with low fluid reasoning skills, and this will be associated with problems in both reading and maths.

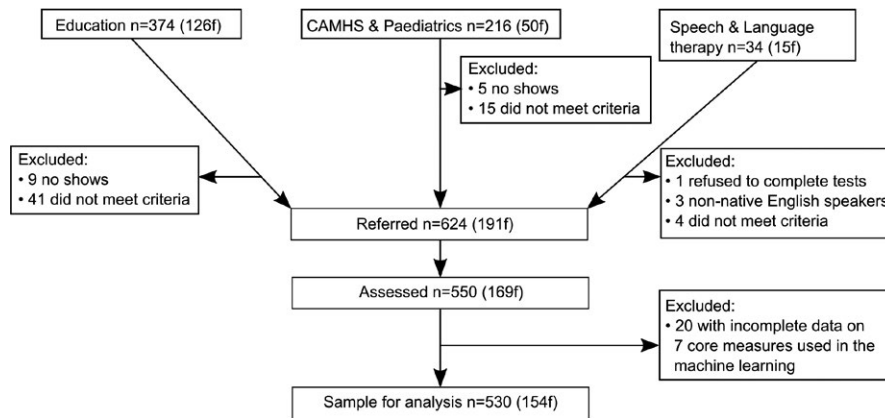
## 2 | METHOD

### 2.1 | Participants

Children were referred by practitioners working in educational or clinical services to the Centre for Attention Learning and Memory (CALM), a research clinic at the MRC Cognition and Brain Sciences Unit, University of Cambridge. Referrers were asked to identify the primary reason for referral, which could include ongoing problems in "attention", "learning", "memory", or "poor school progress". The only exclusion criteria were uncorrected problems in vision or hearing and English as a second language.

The initial sample consisted of 550 children. Twenty children (3.6%) were subsequently removed because of missing data on any one of the seven tasks used for the machine learning. All subsequent details refer to the remaining 530 children (see Figure 1 for recruitment). Thirty three percent were referred for problems with attention, 11% for language difficulties, 10% for memory problems, and 43% for problems with poor school progress (for 3% of children referrers did not provide a primary referral reason). The final sample (mean age = 111 months, range = 65–215 months) contained 366 boys (69%). A high proportion of boys is consistent with prevalence estimates for different developmental disorders within cohort studies (e.g., Russell, Rodgers, Ukomunne, & Ford, 2014).

Children were recruited with single, multiple or no diagnosis. The majority did not have a diagnosis (340, 64%). The prevalence of diagnoses were: ASD = 6%; dyslexia = 6%; obsessive compulsive disorder (OCD) = 2%. Twenty-two percent of the sample had a diagnosis of ADD or ADHD, and further 11% were under assessment for ADHD (on an ADHD clinic waiting list for a likely diagnosis of ADD or ADHD). Finally, 19% of the sample had received support from a Speech and Language Therapist (SLT) within the past 2 years, but did not typically have a diagnosis of SLI.



**FIGURE 1** CONSORT diagram showing recruitment avenues and exclusions

Families attended the CALM clinic for the children's cognitive and learning assessments. Testing lasted approximately 3 hr and was completed over multiple sessions where necessary. Parents/carers were invited to complete multiple questionnaires assessing the child's behavior and all children were invited for a subsequent magnetic resonance imaging (MRI) scan. Ethical approval was granted by the local NHS research ethics committee (Reference: 13/EE/0157). Written parental consent was obtained and children provided verbal assent.

## 2.2 | Measures

### 2.2.1 | Cognitive

A large battery of cognitive, learning, and behavioral measures are administered in the CALM clinic (full protocol: <http://calm.mrc-cbu.cam.ac.uk/protocol/>). Seven cognitive tasks meeting the following criteria were used for the machine learning: (a) data were available for all 530 children; (b) accuracy was the outcome variable; and (c) age standardized norms were available. For all measures, age standardized scores were converted Z scores using the mean and standard deviation from the respective normative samples to put all measures on a common scale (original age norms were a mix of scaled, t, and standard scores). The following measures of fluid and crystallized reasoning were included: Matrix Reasoning, a measure of fluid intelligence (Wechsler Abbreviated Scale of Intelligence [WASI]; Wechsler, 2011); Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007). Phonological processing was assessed using the Alliteration subtest of the Phonological Awareness Battery (PhAB; Frederickson, Reason, & Firth, 1997). Verbal and visuo-spatial short-term and working memory were measured using Digit Recall, Dot Matrix, Backward Digit Recall, and Mr X subtests from the Automated Working Memory Assessment (AWMA; Alloway, 2007).

### 2.2.2 | Learning

Spelling, reading (Word Reading), and maths (Numerical Operations) measures were taken from the Wechsler Individual Achievement Test (WIAT; Wechsler, 2001). Educational assessments were available

for 98% of the sample. The maths fluency subtest from Woodcock Johnson III Test of Achievement (WJ; Woodcock, Schrank, & Mather, 2007) was administered to the first 68 children attending the CALM clinic, instead of the numerical operations measure. We had initially chosen this over the numerical operations subtest from the WIAT because the fluency measure is timed, and therefore quicker to administer. However, when the maths fluency scores were very low we switched to the WIAT—we wondered whether the timed nature of the maths fluency measure was underestimating children's maths ability in this sample. Scores were slightly better with the WIAT, but not significantly so. A two sample Kolmogorov-Smirnoff test indicated that the maths fluency scores from the first 68 participants and the numerical operations scores from the next 68 participants are drawn from the same continuous distribution ( $[D = 0.2144, p = 0.0874]$ ). Age standardized scores were converted to z scores using the normative sample mean and standard deviation for all learning measures.

### 2.2.3 | Behavior

Parents/carers completed the Behavioural Rating Inventory of Executive Function (BRIEF; Gioia, Isquith, Guy, & Kenworthy, 2000). This is designed to assess behavioral skills associated with executive function on eight scales, including planning, working memory, inhibition, impulse control, and emotional regulation. Complete data were available for 99% of our 530 children.

The Children's Communication Checklist (CCC-2; Bishop, 2003) was also administered. This consists of eight scales assessing a child's structural language (e.g., speech, syntax, semantics), pragmatic communication skills (e.g., turn taking, initiation, and use of context), and two additional scales to assess ASD-related dimensions (social relations and interests). Complete CCC-2 data were available for 99% of the sample.

## 2.3 | Statistical methods

A SOM consists of a predefined number of nodes laid out on a two-dimensional grid plane; each node corresponds to a "node-weight vector" with the same dimensionality as the input data. In our case,

each node will have seven weights associated with it (one for each cognitive task). A rule of thumb for determining map size, is to use a number of nodes equal to around 5 times the square root of the number of observations (Tian, Azarian, & Pecht, 2014). In this case, we used a 10 by 10 grid of nodes.

### 2.3.1 | Training the map

SOMs were trained using the neural network toolbox in MATLAB v2017a (MathWorks Inc., Natick, MA). SOMs consist of a predefined number of nodes laid out on a two-dimensional grid plane. Each node corresponds to a weight vector with the same dimensionality as the input data. We initialized the node weight vectors using linear combinations of the first two principal components of the input data. SOMs were then trained using a batch implementation, in which each node  $i$  is associated with a model  $m_i$  and a “buffer memory”. One cycle of the batch algorithm can be broken down into the following: Each input vector  $x^{(t)}$  is mapped onto the node with which it shares the least Euclidean distance at time  $t$ . This node is known as its Best Matching Unit (BMU). Each buffer sums the values of all input vectors  $x^{(t)}$  in the neighborhood set belonging to node  $i$  and divides this by the total number of these input vectors to derive a mean value. All  $m_i$  are then updated concurrently according to these values. In this way, neighboring nodes become more similar to one another. This cycle is repeated, clearing all the buffers on each cycle and distributing new copies of the input vectors into them. The neighborhood size (ND) decreases as a function of  $t$  over  $n$  steps in an “ordering” phase, from the initial neighborhood size (INS) down to 1 (Equation 1). In the “fine tuning” phase the neighborhood size is fixed at  $<1$ , meaning that the node weights are updated according only to the input vectors for which they are the BMU. This node adjustment process is the mechanism by which the SOM learns about the input data. In the current training process, we used 5 “ordering” runs and a single final fine tuning run.

$$ND = 1 + \left[ (INS) * \left( 1 - \left( \frac{t}{n} \right) \right) \right] \quad (1)$$

At the end of the training process: (a) the weight vector for each individual node reflects the scores of the children for whom that node was the BMU; (b) neighboring nodes have similar weights, such that children with similar cognitive profiles are allocated to nodes that are near each other. In essence, the machine learning process generates a model of the multidimensional cognitive data set on which the SOM was trained.

### 2.3.2 | Exploring the distributions of different groups of children

Once the map had been trained we tested whether different groups of children cluster together. For example, if a child's diagnosis predicts their cognitive profile, then children with the same diagnosis

ought to cluster together. That is, they ought to sit on nodes that are near one another. However, if there is no systematic relationship between this characteristic and a child's cognitive profile then this group will be randomly scattered across the map. We tested this both for diagnosis (ASD, dyslexia and ADHD) and the referrer's primary reason for sending the child to the CALM clinic (problems with attention, language, memory, or poor school progress).

To do this, the BMU was tested for each different group. The topographical distribution of this was tested statistically using a version of the Kolmogorov–Smirnov test adapted for 2-dimensional data from two samples (Peacock, 1983). The statistic ( $D$ ) tests whether the two samples are drawn from the same or different 2-dimensional distributions. In each case we compared the distribution of members of a particular category (e.g., referred for language problems) with that of nonmembers (e.g., those *not* referred for language problems). A significant statistic indicates that the two distributions are not drawn from the same underlying population—i.e., that this particular way of categorizing children is significantly predictive of the cognitive profile that they have. Conversely a nonsignificant result indicates that the category's members are equally likely to appear anywhere within the map.

### 2.3.3 | Data driven subgrouping

The artificial neural network maps cognitive profiles in a continuous 2D plane of nodes, where space indicates similarity. We carved our map into sections and grouped the children who fell within that section, thereby clustering children with similar cognitive profiles. Clustering children who sit close together ought to yield groups with relatively homogenous cognitive profiles that are necessarily distinct from children in other clusters.

There is no clear theoretical rationale for how many clusters the map should be carved into. By definition, the map is fully continuous without clear boundaries. One way to validate the clusters is to test whether they generalize to data not included in the initial machine learning—this could be other cognitive data, learning measures, behavioral questionnaires, or brain data. For example, if clusters cannot be distinguished with unseen data then it suggests that the machine learning is over-fitting the data and/or the number of clusters is too high. In this case, the maps would need to be trained with fewer repetitions, a reduced set of nodes, or most likely a reduced number of clusters. To foreshadow our results, in the current sample we can identify four clusters of children. This is the maximum number of clusters that yield generalizable unique profiles. The Supplementary Materials includes a five cluster solution, which replicates the clusters from the four cluster solution, and a statistical comparison between the two. The Supplementary Materials also includes an alternative means of grouping children that is not reliant on machine learning—community detection via a network analysis (e.g., Bathelt et al. 2018).

To identify data-driven clusters the node weight values from the SOM were submitted to k-means clustering. Once the nodes were grouped according to the similarity of their weights, we identified



children assigned to each group of nodes. This provided us with clusters of children based on nodes they were assigned to in the original mapping. This process was repeated 1,000 times, with the map re-trained on every iteration and the k-means clustering recalculated, to check that the clusters were robust. Inevitably some children sit on the arbitrary cluster boundary within the map and thus fall inconsistently into multiple different clusters on each iteration. Across the 1,000 iterations we were able to identify the children's modal cluster, which was used for subsequent analyses. There was a clear modal cluster for 529 children (chi-squared test,  $ps < 0.05$ ). To check the clustering, each cluster distribution was plotted on the original map. If the process had worked then all cluster members ought to sit on neighboring nodes within the original map.

The cognitive profiles of the clusters were compared to identify the ways in which they differ (it is necessarily the case that they will differ). Importantly the groups were then compared on other measures not included in the machine learning, namely learning and behavioral assessments and in terms of brain organization. For all of our assessments we corrected for multiple comparisons using a Bonferroni Correction within each data type (i.e., cognition, learning, and behavioral measures).

## 2.4 | Neuroimaging

### 2.4.1 | MRI participant sample

254 children participated in the MRI part of the study. 64 scans were not useable due to excessive motion ( $>3$  mm movement during the diffusion sequence estimated through FSL eddy or visual inspection of T1-weighted images). The final sample for MRI analysis consisted of 184 children (123 male, Age [months]:  $M = 117.62$ ,  $SE = 1.938$ ). The ratios of SOM-defined groups did not differ from the behavioral sample (Cluster 1:  $n = 48$ , Cluster 2:  $n = 44$ , Cluster 3:  $n = 51$ , Cluster 4:  $n = 41$ ,  $\chi^2 = 0.01$ ,  $p > 0.999$ ). There were no significant differences between the groups in residual movement (see Table 1). For an additional comparison with a typically developing sample, we selected children from a concurrent study about risk and resilience in education that shared many of the same cognitive assessments and used an identical neuroimaging protocol (Ethical approval number: Pre.2015.11). This sample included children attending mainstream school in the UK with normal or corrected-to-normal vision or hearing and no history of brain injury who were recruited via local schools and through advertisements in public places (childcare and community centers, libraries). Children were selected to use in the current analysis if they fell within the right age-bracket, had useable MRI data (i.e., good quality T1, movement during the diffusion sequence  $<3$  mm and 69 diffusion-weighted volumes) and had cognitive scores within the normal range—scores above the 40th percentile for their age on assessments of fluid reasoning, vocabulary, verbal, and visuospatial short-term and working memory were selected. This additional comparison sample consisted of 36 children (18 male, Age [months]:  $M = 117.79$ ,  $SE = 3.129$ , range: 83.02–150.05, mean fluid IQ = 53 [age expected =  $50 \pm 10$  SD]).

**TABLE 1** Comparison of residual movement during the diffusion sequence between groups. The upper triangle of the table shows the  $p$ -value of an independent sample  $t$ -test. The lower triangle shows the corresponding  $t$ -value

	C1	C2	C3	C4	C0
C1		0.119	0.668	0.401	0.208
C2	1.57		0.247	0.291	0.847
C3	0.43	-1.17		0.225	0.309
C4	-0.84	-1.69	-1.22		0.100
C0	1.27	0.19	1.02	1.66	

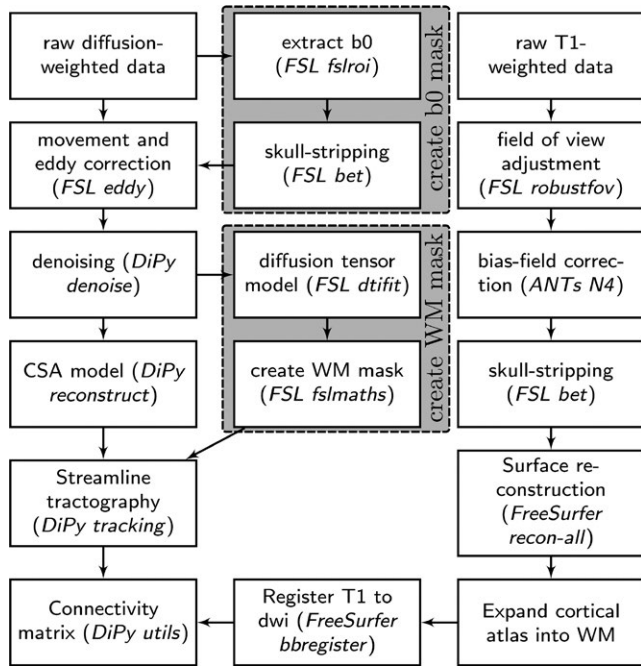
### 2.4.2 | MRI data acquisition

Magnetic resonance imaging data were acquired at the MRC Cognition and Brain Sciences Unit in Cambridge, on the Siemens 3 T Tim Trio system (Siemens Healthcare) 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 2.98 ms, and repetition time was 2,250 ms. Diffusion scans were acquired using echo-planar diffusion-weighted images with an isotropic set of 60 noncollinear directions, using a weighting factor of  $b = 1,000 \text{ s} \times \text{mm}^{-2}$ , interleaved with a T2-weighted ( $b = 0$ ) volume. Whole brain coverage was obtained with 60 contiguous axial slices and isometric image resolution of 2 mm. Echo time was 90 ms and repetition time was 8,400 ms.

### 2.4.3 | Structural connectome construction and comparison

First, MRI scans were converted from the native DICOM to compressed NIfTI-1 format. Next, correction for motion, eddy currents, and field inhomogeneities was applied using FSL eddy (see Figure 2 for an overview of processing steps). Furthermore, we submitted the images to nonlocal means de-noising (Manjon, Coupe, Martí-Bonmati, Collins, & Robles, 2009) using DiPy v0.11 (Garyfallidis et al., 2014) to boost signal-to-noise ratio. The diffusion tensor model was fitted to derive maps of fractional anisotropy (FA) using dtfit in FSL v.5.0.6 (Behrens et al., 2003). A constant solid angle (CSA) model was fitted to the 60-gradient-direction diffusion-weighted images using a maximum harmonic order of 8 using DiPy. Whole-brain probabilistic tractography was performed with 8 seeds in any voxel with a General FA value higher than 0.1. The step size was set to 0.5 and the maximum number of crossing fibers per voxel to 2.

For ROI definition, T1-weighted images were submitted to nonlocal means denoising in DiPy, robust brain extraction using ANTs v1.9 (Avants et al., 2011), and reconstruction in FreeSurfer v5.3 (<http://surfer.nmr.mgh.harvard.edu>). Regions of interests (ROIs) were based on the Desikan-Killiany parcellation of the MNI template (Desikan et al., 2006) with 34 cortical ROIs per hemisphere and 17 subcortical ROIs. The cortical parcellation was expanded by 2 mm into the



**FIGURE 2** Overview of processing steps to reconstruct a white matter connectome from diffusion-weighted and T1-weighted MRI data

subcortical white matter. The parcellation was moved to diffusion space using FreeSurfer tools.

For each pairwise combination of ROIs, the number of streamlines intersecting both ROIs was calculated. A symmetric intersection was used, i.e., streamlines starting and ending in each ROI were averaged. The weight of the connection matrices represented the  $\log_{10}$ -transformed number of streamlines between the ROIs.

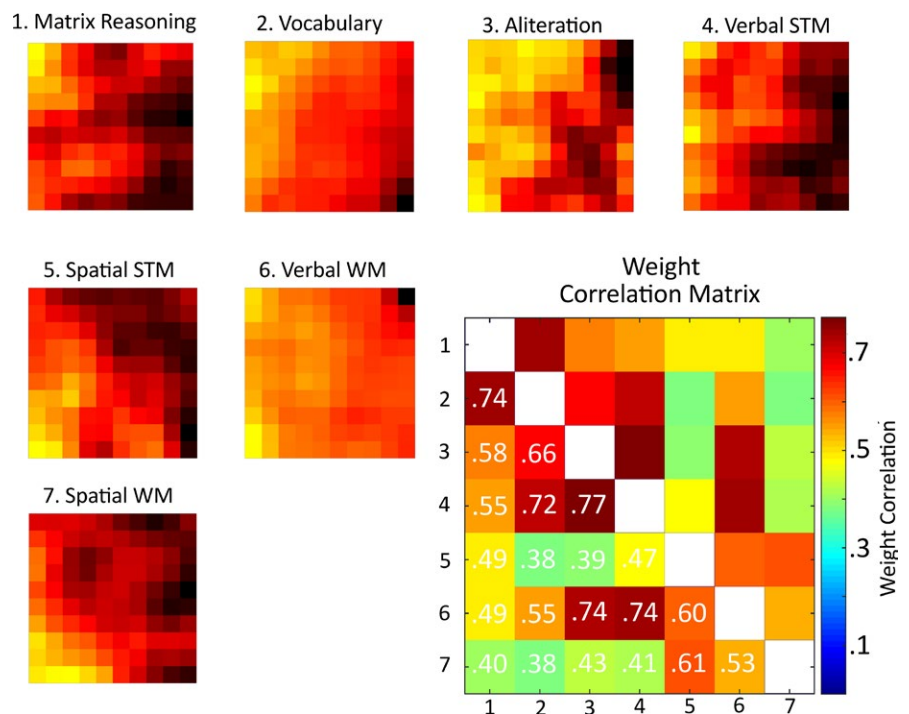
To investigate regional differences, we calculated the sum of all connections per region within the connectome. Regions that showed a significant difference between a deficit group (C1, C2, C4) and an age-appropriate performance group (C3) were selected ( $t$ -test:  $p_{\text{uncorrected}} < 0.05$ ) and further tested against the external comparison group (method adapted from Shen et al., 2017). Only regions that displayed a significant difference relative to the external comparison sample were included (FDR-corrected  $p < 0.05$ ).

### 3 | RESULTS

#### 3.1 | Comparison of the weight matrices

A good way to demonstrate how the SOM represents the cognitive data is to plot the values for each weight vector (i.e., the weights that correspond to each individual task) across the grid of nodes. This can be seen in Figure 3.

If tasks discriminate children in similar ways they should have similar node weight topographies. This was quantified by correlating the weight vectors. The resulting correlation matrix can be seen in the bottom right corner of Figure 2. There are some noteworthy relationships. For example, the two measures traditionally combined to produce a full-scale IQ score, the Matrix Reasoning and PPVT vocabulary measure, have very highly correlated weights. Tasks that share a phonological component have highly correlated weight matrices: alliteration, verbal STM, and verbal WM measures. Finally, spatial STM and WM measures are somewhat distinct from other measures, with weight matrices that are only moderately correlated with the other tasks.



**FIGURE 3** Weight distributions from the self-organizing map, split by task. For each task the map depicts high weights (i.e., good performance) as yellow squares and low weights (i.e., poor performance) as black squares. The Pearson correlation between the weight distributions can be seen in the bottom-right matrix



### 3.2 | Exploring distributions of different categories of children

To explore whether different sample characteristics (diagnostic status, referral reason) are reflected within the map, the best matching node for different groups of children was selected. If category membership significantly predicts a child's cognitive profile then these children should sit together in the map. Conversely, if membership is not predictive then the distribution of members should not differ significantly from that of nonmembers. Figure 4 shows the distribution of all children within our network, then for each category of primary referral reason and then each of the major diagnoses. The statistics are shown under each topography. None are significant. That is, children are evenly scattered regardless of the primary reason for referral or diagnosis; each of these characteristics provides no information about a child's cognitive profile on our measures.

### 3.3 | Common cognitive profiles

To identify children with common cognitive profiles, the map was carved into four sections by applying k-means clustering to the node weights of the SOM. Each cluster has a distinct spatial distribution within the map (Figure 5), as expected, and this is reflected in the distribution statistic.

Each group necessarily has a distinct cognitive profile. The first cluster includes children with broad and severe cognitive difficulties—these children are around a standard deviation or more below the age-expected level on all cognitive measures. The third cluster includes children with age-typical cognitive abilities, performing close to age-expected levels on all tasks. These two clusters are subsequently referred to as the “Broad Cognitive Deficits” and the “Age Appropriate” groups respectively. The remaining two clusters have intermediate profiles. They have similar moderate difficulties with Matrix Reasoning, but distinct profiles on the remaining measures. The second cluster has difficulties on the spatial STM, and verbal and spatial WM measures. This group is called the “Working Memory Deficits” group. The fourth cluster has difficulties tasks with a verbal component: vocabulary, phonological awareness, verbal STM, and verbal WM. This cluster is called the “Phonological Deficits” group.

The profiles of the four clusters can be seen in Figure 5, with scores and group comparisons presented in Table 2. All measures differed significantly across groups (all  $p$ s < 0.001). Post hoc Tukey tests were used to identify the underlying pairwise comparisons that produce these significant effects. For Matrix Reasoning all post hoc tests were significant at  $p$  < 0.001, except between clusters 2 and 4 (Working Memory vs. Phonological Deficits groups). The Working Memory Deficits and Phonological Deficits groups have equivalent Matrix Reasoning scores ( $p$  = 0.86). For Vocabulary, all post hocs were significant at  $p$  < 0.001. For the Phonological Awareness task, all post hocs were significant at  $p$  < 0.007. For Verbal STM, all post hocs were significant at  $p$  < 0.001. For Spatial STM, all post hocs were significant at  $p$  < 0.001, except between clusters 1 and 2 (Broad Deficits vs. Working Memory deficits groups). The Broad

Cognitive Deficits and Working Memory Deficits groups have equivalent Spatial STM scores ( $p$  = 0.51). For Verbal WM, all post hocs were significant at  $p$  < 0.001, except for between cluster 2 and 4 (Working Memory vs. Phonological Deficits groups). The Working Memory Deficits and Phonological Deficits groups had equivalent Verbal WM scores ( $p$  = 0.68). And finally, for Spatial WM all post hocs were significant at  $p$  < 0.001, except for between clusters 3 and 4 (Age Appropriate vs. Phonological Deficits groups). The Age Appropriate and Phonological Deficits group had equivalent Spatial WM performance ( $p$  = 0.13).

The four groups are roughly equivalent in size, and although the children in the Phonological Deficits group tend to be younger, there are no significant age differences. The Broad Cognitive Deficit group contains a disproportionate number of girls, relative to the rest of the sample ( $\chi^2$  = 6.12,  $p$  = 0.0133). Conversely the Age Appropriate group contains more boys than expected ( $\chi^2$  = 6.80,  $p$  = 0.009). The Working Memory deficit and Phonological Deficit groups contain the proportions of boys and girls expected ( $\chi^2$  = 0.01,  $p$  = 0.91; and  $\chi^2$  = 0.01,  $p$  = 0.92 respectively).

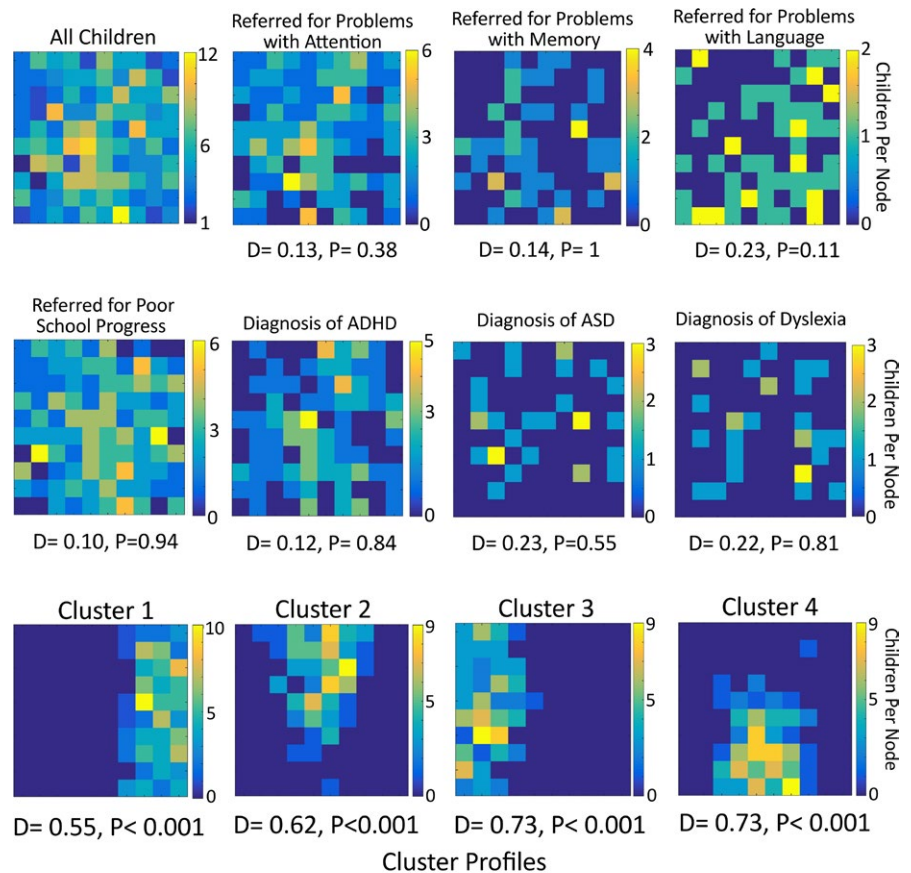
Children referred primarily for problems with attention, poor learning, or memory were equally likely to be assigned to each group. Similarly, a diagnosis did not predict group membership. The only category predictive of group membership was whether the child was under the care of an SLT. These children were disproportionately likely to be members of either the Broad Cognitive Deficits or Phonological Deficits groups. All of these statistics can be found in Table 2.

### 3.4 | Learning and behavioral profiles of the data-driven groups

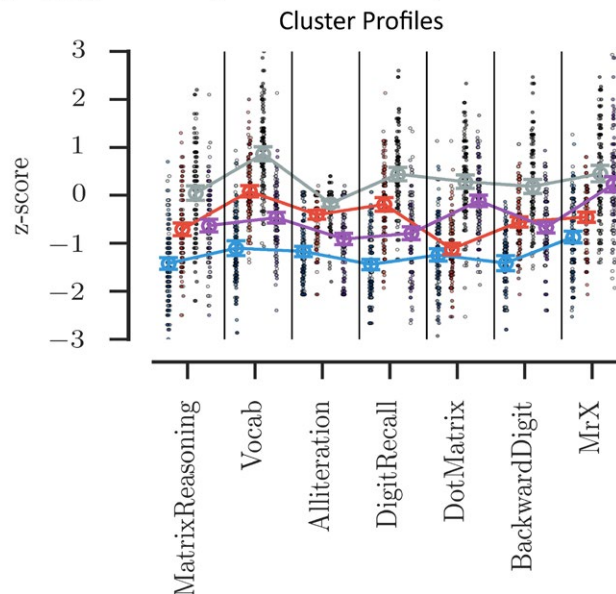
The four clusters also have important differences on other measures not included in the machine learning, which are also shown in Table 2. Age Appropriate children had age appropriate learning skills across spelling, reading, and maths. Children in the Broad Cognitive Deficits group had severe problems on all learning outcomes, being more than 1.5 standard deviations below the age expected levels. The other two groups, despite their highly contrasting cognitive profiles did not differ in their learning profiles—moderate phonological problems or working memory difficulties were associated with very similar learning profiles. This is reflected in the statistics—all measures show a significant group difference (all  $p$ s < 0.001), and all the post hoc tests are significant at  $p$  < 0.001, except between the Phonological and Working Memory deficit groups (spelling,  $p$  = 0.22; reading,  $p$  = 0.41; maths,  $p$  = 0.79).

The subscale scores for both BRIEF and CCC-2 questionnaires, split by group, can be seen in Table 2. Correlation matrices for both the BRIEF and CCC-2 can be found in Tables S1 and S2. Before comparing the groups a PCA was conducted separately for the subscales of each questionnaire to reduce the number of comparisons. These analyses identified two factors in the BRIEF, which together explained 76.1% of the variance. The rotated factor solution and scale loadings can be found in Table S3. The first factor captured the





**FIGURE 4** The distributions of children's best matching unit (BMU) within the map. This is first shown for all children and then for children categorized by referral reason and diagnosis. Beneath each plot the statistic indicates whether the BMUs are evenly scattered or grouped



**FIGURE 5** The top panel shows the distributions of children assigned to each of the four clusters. Beneath each map the statistic indicates that all four clusters occupy a nonrandom set of nodes within the map. Beneath the maps the cognitive profile of each cluster is shown, ordered by cluster number. The scale indicates performance as a z score relative to age expected levels. The dots indicate individual children with the shade indicating the child's consistency within that cluster over the 1,000 iterations—the darker the shade the more consistent the child

working memory, initiate, planning, organization, and monitor subscales. The emotional control, shift, inhibit, and monitor subscales loaded most highly on the second factor. The first factor therefore corresponds to “Cold” executive functions associated with behavioral regulation, while the second corresponds more closely to “Hot” cognitive aspects of executive function. Factor scores were saved and compared across groups: there were no significant differences in behavior across the clusters (all  $p$ s > 0.05).

There were also two factors within the CCC-2, explaining 74% of the variance. The rotated factor solution can be found in

Table S4. Subscales tapping pragmatic aspects of communication load most highly on Factor 1: coherence, inappropriate initiation, stereotyped language, context, nonverbal social skills, and interests subscales. Factor 2 was comprised of scales measuring structural language skills: speech, syntax, semantics, and coherence. These factors were labeled “Pragmatic Communication” and “Structural Language” respectively. There were no significant group differences in Pragmatic Communication factor scores. The groups did, however, differ significantly on Structural Language factor scores ( $p$  < 0.001). Post hoc tests revealed children in the Broad Cognitive Deficits

**TABLE 2** Cognitive, learning, and behavioral measures split by cluster

Label	Cluster 1 Broad deficits	Cluster 2 WM deficits	Cluster 3 Age appropriate	Cluster 4 Phon. deficits		
Descriptives						
N	146	121	132	131		
Male	87	83	105	91		
Female	59	38	27	40	<i>F</i>	<i>p</i> value
Mean age	113	114	112	106	2.136	0.0947
Reason for referral					Chi Sq	<i>p</i> corr
Attention	36	48	52	39	6.755	0.320
Memory	16	11	12	16	0.877	1.000
Language	25	7	12	15	8.321	0.159
Poor school	67	51	51	58	0.938	1.000
Diagnosis					Chi Sq	<i>p</i> corr
ADD/ADHD	29	35	22	28	4.718	0.968
SLT	43	16	15	24	14.931	0.010
Dyslexia	11	9	5	5	3.186	1.000
ASD	9	6	11	6	1.850	1.000
Sus ADHD	15	14	18	10	2.312	1.000
Cognitive measures					<i>F</i>	<i>P</i> corr
Matrix reasoning	-1.42	-0.71	0.04	-0.63	82.40	<0.001
Vocab	-1.10	0.08	0.87	-0.47	157.13	<0.001
Phon. Aware.	-1.17	-0.41	-0.17	-0.91	86.46	<0.001
Verbal STM	-1.44	-0.19	0.44	-0.79	157.35	<0.001
Spatial STM	-1.24	-1.12	0.29	-0.11	139.36	<0.001
Verbal WM	-1.41	-0.55	0.18	-0.66	95.59	<0.001
Spatial WM	-0.87	-0.45	0.45	0.23	73.48	<0.001
Learning measures					<i>F</i>	<i>P</i> corr
Spelling	-1.58	-1.05	-0.47	-1.17	40.613	<0.001
Reading	-1.60	-0.90	0.02	-1.09	65.932	<0.001
Maths	-1.77	-1.02	-0.32	-1.13	55.957	<0.001
Behavioral measures						
Exec. functions					<i>F</i>	<i>P</i> corr
Cold factor	0.05	0.03	0.01	-0.09	0.520	1.000
Hot factor	0.08	-0.01	-0.14	0.07	1.376	0.997
Inhibit	66.3	65.4	64.5	64.9		
Shift	69.6	68.2	66.1	68.5		
Emot. control	65.1	64.4	62.7	65.4		
Initiate	68.1	66.2	65.9	66.6		
WM	75.9	74.2	72.8	72.9		
Planning	72.2	71.3	71.9	71.1		
Organization	58.1	61.4	61.0	59.9		
Monitor	66.5	66.4	64.2	65.1		
Communication					<i>F</i>	<i>p</i> corr
Pragmatic	-0.10	-0.07	0.07	0.11	1.452	0.9076
Structural	-0.53	0.28	0.55	-0.23	38.191	<0.001
Speech	3.6	6.6	7.1	4.6		

(Continues)

**TABLE 2** (Continued)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Label	Broad deficits	WM deficits	Age appropriate	Phon. deficits
Syntax	3.0	6.0	7.3	4.2
Semantics	3.5	5.2	6.2	4.3
Coherence	3.3	4.6	5.3	4.3
Inappro. initiation	5.0	5.8	6.1	5.8
Stereo	4.1	5.5	6.6	5.3
Contex	2.7	4.0	5.4	3.7
Nonverbal	4.0	4.7	5.3	4.5
Social	4.1	4.7	5.2	5.2
Interest	5.3	5.5	5.8	5.7

or Phonological Deficits groups were rated as having significantly greater structural language problems than either of the other two groups (all  $p$ s < 0.001). The respective Structural Language difference between the Age Appropriate and Working Memory Deficit groups was marginal ( $p = 0.043$ ), as was that between the Broad Cognitive Deficits and Phonological Deficits groups ( $p = 0.043$ ).

### 3.5 | White matter differences between the data-driven groups

Differences in white matter connections between the SOM-defined groups were investigated to uncover the neurobiological correlates of the grouping. Each of the deficit groups (Clusters 1, 2 and 4) was compared to the Age Appropriate group (Cluster 3) and an independent sample of typically developing children (TD). Statistical comparison of connection strengths by region indicate significantly lower connection strengths for frontal, temporal, parietal, and subcortical connections in Cluster 1 compared to Cluster 3 and TD (see Table 3 and Figure 6). There was no significant difference in regional connection strength between Cluster 2 and Cluster 3 or between Cluster 2 and TD. The comparison of Cluster 4 and Cluster 3 indicated significantly lower strength of parietal connections and the comparison with TD indicated significantly different frontal connections.

Regional comparison indicated a significant reduction for Cluster 1 (Broad Deficits) compared to both comparison groups for the right inferior frontal gyrus (see Figure 6, C1:  $M = 0.59$ ,  $SE = 0.027$ ; C3:  $M = 0.65$ ,  $SE = 0.025$ ; C0:  $M = 0.72$ ,  $SE = 0.028$ ;  $t(82) = -3.19$ ,  $p_{\text{corrected}} = 0.018$ ), the right lateral orbitofrontal gyrus (C1:  $M = 0.66$ ,  $SE = 0.021$ ; C3:  $M = 0.72$ ,  $SE = 0.021$ ; C0:  $M = 0.75$ ,  $SE = 0.025$ ;  $t(82) = -2.91$ ,  $p_{\text{corrected}} = 0.032$ ), the left fusiform gyrus (C1:  $M = 0.69$ ,  $SE = 0.017$ ; C3:  $M = 0.76$ ,  $SE = 0.017$ ; C0:  $M = 0.79$ ,  $SE = 0.021$ ;  $t(82) = -3.69$ ,  $p_{\text{corrected}} = 0.011$ ), and the left precentral gyrus (C1:  $M = 0.95$ ,  $SE = 0.019$ ; C3:  $M = 1.01$ ,  $SE = 0.019$ ; C0:  $M = 1.04$ ,  $SE = 0.016$ ;  $t(82) = -3.42$ ,  $p_{\text{corrected}} = 0.013$ ). The comparison between Cluster 4 (Phonological Deficits) and both comparison groups indicated significantly lower connection strength in the left precentral gyrus (C4:  $M = 0.97$ ,  $SE = 0.016$ ; C3:  $M = 1.01$ ,

$SE = 0.019$ ; C0:  $M = 1.04$ ,  $SE = 0.016$ ; C4 vs. C0:  $t(75) = -3.03$ ,  $p_{\text{corrected}} = 0.013$ ) and left rostral anterior cingulate gyrus (C4:  $M = 0.30$ ,  $SE = 0.009$ ; C3:  $M = 0.33$ ,  $SE = 0.01$ ; C0:  $M = 0.35$ ,  $SE = 0.009$ ; C4 vs. C0:  $t(75) = -3.51$ ,  $p_{\text{corrected}} = 0.006$ ). There were no significant differences between Cluster 2 (Working Memory Deficits) and the comparison groups, once controlling for multiple comparisons.

## 4 | DISCUSSION

We used machine learning to identify the cognitive profiles within a large heterogeneous sample of children with learning-related problems. These profiles were represented as topographical maps. None of the known characteristics of the children (e.g., diagnosis or referral route) were predictive of the cognitive profiles identified by the machine learning. To highlight the cognitive profiles that exist within the dataset, we subsequently carved the topographical maps into four sections. The children that correspond to these four sections will necessarily have distinct cognitive profiles, but they could also be distinguished in terms of learning and behavioral scores, and patterns of brain organization. The four groups cut across any traditional diagnostic groups that existed within the data.

More than half of the sample fell into two extreme groups, one with age-appropriate cognitive abilities and the other with widespread cognitive deficits that were at least one standard deviation below age-typical levels across all tasks. There was no evidence that children with age-expected scores on the cognitive measures had learning difficulties. Their performance was in the age-typical range across all measures of learning and their structural communication skills were rated as normal for their age. But we should be very cautious in regarding these children as typically developing; they have been referred by professionals in children's services, and as a group they have elevated behavioral difficulties. For this reason in our neuroimaging analysis we used an additional external comparison group.

The learning scores of the broad deficit group place them within the bottom 5% of the population on measures of spelling, reading, and maths, and they were rated as having difficulties in both

TABLE 3 Results of regional connection strengths between C3 and the other groups

	C1		C2		C4		C3		C0	
	Median	Mad	Median	Mad	Median	Mad	Median	Mad	Median	Mad
Frontal	0.68	0.053	0.70	0.047	0.69	0.058	0.72	0.069	0.72	0.038
Temporal	0.52	0.037	0.54	0.030	0.54	0.039	0.55	0.033	0.54	0.025
Parietal	0.66	0.052	0.67	0.045	0.67	0.042	0.70	0.067	0.68	0.050
Occipital	0.60	0.073	0.63	0.062	0.63	0.048	0.63	0.093	0.62	0.081
Subcortical	0.55	0.074	0.59	0.055	0.57	0.055	0.58	0.055	0.59	0.046
	C1 vs. C3			C2 vs. C3			C4 vs. C3			
	U	p	p <sub>corr.</sub>	U	p	p <sub>corr.</sub>	U	p	p <sub>corr.</sub>	
Frontal	823	0.003	<b>0.013</b>	979	0.144	0.216	793	0.024	0.060	
Temporal	772	0.001	<b>0.006</b>	1027	0.240	0.277	924	0.171	0.224	
Parietal	758	0.001	<b>0.006</b>	890	0.042	0.090	715	0.005	<b>0.018</b>	
Occipital	996	0.056	0.104	1064	0.334	0.358	889	0.110	0.184	
Subcortical	916	0.016	<b>0.047</b>	1085	0.393	0.393	928	0.179	0.224	
	C1 vs. C0			C2 vs. C0			C4 vs. C0			
	U	p	p <sub>corr.</sub>	U	p	p <sub>corr.</sub>	U	p	p <sub>corr.</sub>	
Frontal	520	0.001	<b>0.010</b>	657	0.196	0.321	505	0.012	<b>0.034</b>	
Temporal	588	0.08	<b>0.032</b>	785	0.344	0.364	726	0.307	0.364	
Parietal	577	0.004	<b>0.032</b>	694	0.327	0.364	567	0.022	0.056	
Occipital	751	0.113	0.212	685	0.214	0.321	735	0.247	0.336	
Subcortical	594	0.007	<b>0.032</b>	755	0.364	0.364	617	0.057	0.122	

Bold text indicates significant effects at corrected  $p < 0.05$ .

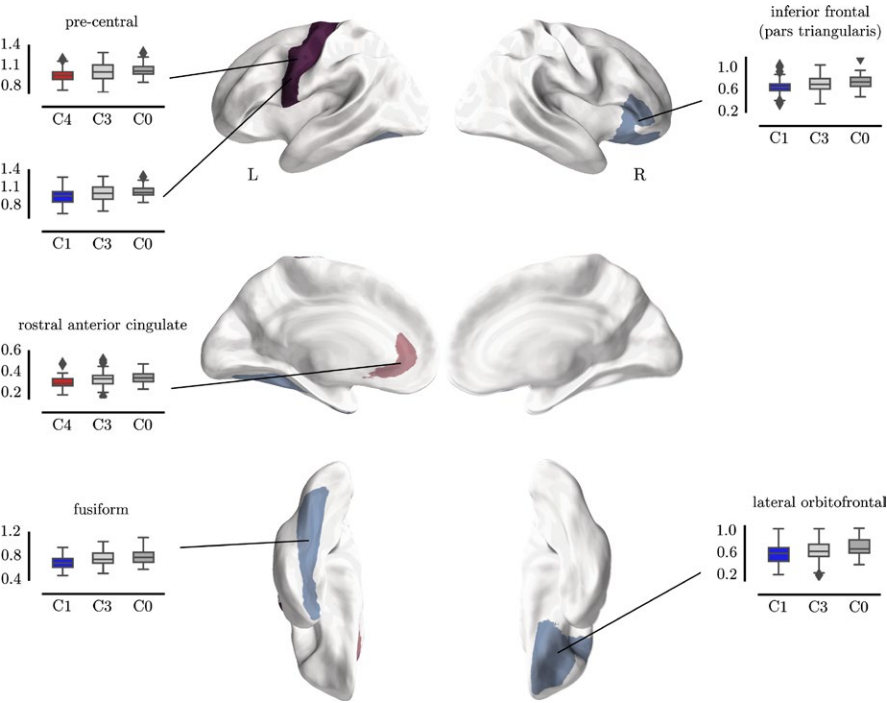


FIGURE 6 Regions with consistent significant differences in node degree between Cluster 1 and the control groups (blue) and Cluster 4 and the control groups (red)

structural and pragmatic aspects of communication. Generalized cognitive deficits therefore appear to constrain multiple aspects of learning. They also had behavioral problems related to executive function, although this was true for all four groups. Relative to both

comparison groups, this group also had reduced structural connectivity in the left precentral gyrus, right inferior frontal gyrus, right lateral occipital cortex, and the left fusiform. These areas have been previously identified as playing a key role in multiple higher order



cognitive skills. For example, the right inferior frontal gyrus is implicated in multiple different executive functions, most commonly measures of inhibitory control (Aron, Robbins, & Poldrack, 2014); the lateral occipital cortex has been found to be modulated by visual attention (Sprague & Serences, 2013); left premotor areas have been linked to language-related difficulties in both children and adults (Mayes, Reilly, & Morgan, 2015; Scott, McGettigan, & Eisner, 2009); and the fusiform gyrus has been suggested as a locus of immature processing of word forms in dyslexia (Tamboer, Vorst, Ghebrea, & Scholte, 2016). These general struggling learners are rarely studied, but our data suggest that they are common amongst those coming to the attention of children's specialist services. Their relative underrepresentation in studies of learning-related problems means that we have little understanding of the key underlying deficits, mechanisms or potential routes to effective intervention. It is also interesting to note that girls were disproportionately common in this group, relative to the sample as a whole or indeed relative to most studies of learning difficulties. Conversely very few girls appeared in the age-appropriate cognitive profile group. In short, the girls referred to the study tended to have more severe cognitive and learning difficulties. One possibility is that there is a gender bias in the reason for children coming to the attention of children's specialist services, with boys being identified more commonly for behavioral difficulties (which may be less closely tied to cognitive and learning profiles), whereas more severe cognitive or learning difficulties are needed for girls to come to the attention of specialists.

Two intermediate groups, both with fluid reasoning scores in the low-average range, were also identified. One intermediate group was characterized by problems on tasks requiring phonological processing, with performance around three quarters of a standard deviation below age-expected levels on measures of phonological awareness, and verbal short-term and working memory. These children had significant problems with structural aspects of communication, mirroring the well-documented link between phonological processing difficulties and specific difficulties with language (Bishop & Norbury, 2002; Bishop & Snowling, 2004; Ramus et al., 2013). However, the learning profile demonstrates equivalent and large deficits across measures of reading, spelling, and mathematics. Poor phonological processing is associated with both poor reading (Carroll & Snowling, 2004; Snowling, 1995; Wagner & Torgesen, 1987) and mathematical development (De Smedt, Taylor, Archibald, & Ansari, 2010; Hecht, Torgesen, Wagner, & Rashotte, 2001; Swanson & Sachse-Lee, 2001). A consistent finding within the field of learning difficulties is that phonological problems are linked selectively with reading. The majority of these findings come from studies that select poor readers, but this is not the same as demonstrating that phonological impairments will always result in selective reading difficulties. Our data suggest that children selected on the basis of phonological difficulties will actually have more widespread learning problems. Membership of the phonological deficit group was associated with reduced structural connectivity in the left precentral gyrus and rostral anterior cingulate, relative to both comparison groups. The precentral gyrus has been implicated in language processing and is thought to be involved

in speech production and also decoding via articulatory simulation (Scott et al., 2009). This area has also been implicated in selective language impairment (Mayes et al., 2015). Furthermore, tracts of the perisylvian language network that connect temporal and frontal language areas deficits are passing the precentral gyrus and may be substantially contributing the connectomics differences. Differences in white matter properties of these tracts have been repeatedly implicated in language deficits (Rimrodt, Peterson, Denckla, Kaufmann, & Cutting, 2010; Roberts et al., 2014). This would also mirror the structural communication difficulties that these children demonstrate. Indeed, this is the only behavioral measure that aligns well with the cognitive profiles—children who perform poorly on phonological tasks are also rated as having significant structural language problems by their parents. Other behavioral measures of executive control do not align well with cognitive profiles.

The fourth group had a somewhat contrasting profile of cognitive deficit to the phonological deficit group. They were characterized by similar fluid IQ scores but had more pronounced difficulties in working memory. Their spatial short-term memory scores were over a standard deviation below age-expected levels, and half a standard deviation down on the verbal and spatial working memory measures. Their phonological abilities were less impaired, they were not rated as having the structural language difficulties reported for the phonological deficit group, and their neural profile was less homogenous. One possibility is that multiple different etiological routes can result in this profile of difficulties.

Despite contrasting cognitive and neural profiles, the learning profiles of the working memory and phonological deficit groups were nearly identical. This diverges strongly from a preceding literature that emphasizes a marked association between phonological difficulties and problems with literacy (Lyytinen et al., 2004; Snowling, Bishop, & Stothard, 2000; Tanaka et al., 2011), and an emerging literature that suggests strong associations between spatial short-term and working memory problems and numeracy difficulties (Bull et al., 2008; Raghubar, Barnes, & Hecht, 2010; Szucs et al., 2013). These previous studies all recruit on the basis of highly selective learning profiles (e.g., maths problems in the absence of reading difficulties) or diagnostic group, which will have overestimated the distinctiveness of these impairments within the general population of struggling learners.

Despite their utility, machine learning approaches to exploring cognitive profiles have limitations. The current combination of a multidimensional mapping method with a data-driven clustering algorithm suffers from the drawback that the number of groups within the data is underspecified. The mapping process is continuous, with no obvious boundaries, which makes it difficult to have a clear rationale about the formation of groups. Inevitably some children will sit close to a group boundary within the map. Our approach was to add clusters until the clusters did not differ on measures not included in the machine learning. This is how we arrived at four clusters. This is a relatively conservative approach, since different cognitive profiles could exist that genuinely have identical learning, behavioral, and neural correlates. Furthermore, we suspect that datasets with





higher dimensionality, stemming from a more widespread battery of measures, could have greater success in identifying more widely differing cognitive profiles.

An alternative to machine learning is to use a network analysis with a community detection algorithm (e.g., Bathelt et al., 2018; Fair et al., 2012). An example of this approach applied to our data can be found in our Supplementary Materials section. This represents the children as nodes and the correlation between their profiles as edges. It is possible to use this approach to identify communities of clusters that maximize the correlation within cluster and the distinctiveness across clusters. This iterative process includes a quality of separation metric (Q) which the clustering algorithm is designed maximize. A major advantage of this approach is that no a priori assumptions about the number of clusters need to be made. However, there are also drawbacks to this alternative. The primary limitation is that a network analysis clusters children on the basis of a correlation matrix. As such it is blind to overall severity. The current sample contains a large number of children with relatively consistent poor scores across all cognitive measures and many children with stable age-appropriate scores. A network analysis would not be able to distinguish these two groups because the two profiles are highly correlated (this is indeed the case, see Supplementary Materials). The SOM uses Euclidean Distance as its primary means of locating children within the 2D topographical space, and as such is able represent both selective cognitive impairments and overall differences in severity. A further limitation is sample size. Whilst we included 530 children in the topographical mapping process, only 220 children were used in the structural neuroimaging comparison. This likely means that we only have sufficient power to detect the largest and most consistent group differences. More diffuse but equally important differences in whole brain connectome organization might exist, but a larger sample would be needed to identify them.

In summary, we used a machine learning approach that represents high-dimensional data as a 2D topography, to map the profiles of struggling learners. We combined this with a clustering algorithm to identify particular cognitive profiles represented within the map. Specifically, four profiles could be identified that comprise children with: (a) general and severe deficits, (b) age-appropriate performance, (c) working memory deficits, (d) phonological deficits. Furthermore, these data-driven groups are likely to align closely with underlying etiological mechanisms, as evidenced by differences in brain organization across two of the deficit groups, and provide the opportunity to devise interventions that more specifically target the cognitive difficulties faced by individuals with particular profiles.

## ACKNOWLEDGMENTS

The Centre for Attention Learning and Memory (CALM) research clinic is based at and supported by funding from the MRC Cognition and Brain Sciences Unit, University of Cambridge. The Principal Investigators are Joni Holmes (Head of CALM), Susan Gathercole (Chair of CALM Management Committee), Duncan Astle, Tom Manly and Rogier Kievit. Data collection is assisted by a team of

researchers and PhD students at the CBSU. This currently includes: Sarah Bishop, Annie Bryant, Sally Butterfield, Fanchea Daily, Laura Forde, Erin Hawkins, Sinead O'Brien, Clíodhna O'Leary, Joseph Rennie, and Mengya Zhang. The authors 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.

## ORCID

Joe Bathelt  <http://orcid.org/0000-0001-5195-956X>

## REFERENCES

- Alloway, T. (2007). *Automated Working Memory Assessment (AWMA)*. London, UK: Pearson Assessment.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry*, 40, 57–87. <https://doi.org/10.1111/1469-7610.00424>
- Archibald, L. M. D., Cardy, J., Joanisse, M. F., & Ansari, D. (2013). Language, reading, and math learning profiles in an epidemiological sample of school age children. *PLoS ONE*, 8(10), e77463. <https://doi.org/10.1371/journal.pone.0077463>
- Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in Cognitive Sciences*, 18(4), 177–185. <https://doi.org/10.1016/j.tics.2013.12.003>
- Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ANTs similarity metric performance in brain image registration. *NeuroImage*, 54(3), 2033–2044. <https://doi.org/10.1016/j.neuroimage.2010.09.025>
- Bathelt, J., Holmes, C., the CALM team, & Astle, D. E. (2018). Data-Driven Subtyping of Executive Function-Related Behavioral Problems in Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57(4), 252–262.e4. <https://doi.org/10.1016/j.jaac.2018.01.014>
- Behrens, T. E. J., Woolrich, M. W., Jenkinson, M., Johansen-Berg, H., Nunes, R. G., Clare, S., ... Smith, S. M. (2003). Characterization and propagation of uncertainty in diffusion-weighted MR imaging. *Magnetic Resonance in Medicine*, 50(5), 1077–1088. <https://doi.org/10.1002/mrm.10609>
- Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of Attention Deficit Hyperactivity Disorder with conduct, anxiety, and other disorders. *American Journal of Psychiatry*, 148, 564–577.
- Bishop, D. V. M. (2003). *Children's Communication Checklist (CCC-2)*. London, UK: Pearson Assessment.
- Bishop, D. V. M., & Norbury, C. F. (2002). Inferential processing and story recall in children with communication problems: A comparison of specific language impairment, pragmatic language impairment and high-functioning autism. *International Journal of Language & Communication Disorders*, 37(3), 227–251. <https://doi.org/10.1080/13682820210136269>
- Bishop, D. V. M., & Snowling, M. J. (2004). Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*, 130(6), 858–886. <https://doi.org/10.1037/0033-2909.130.6.858>
- Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, 33(3), 205–228. <https://doi.org/10.1080/87565640801982312>



- Bull, R., Espy, K. A., Wiebe, S. A., Sheffield, T. D., & Nelson, J. M. (2011). Using confirmatory factor analysis to understand executive control in preschool children: Sources of variation in emergent mathematical achievement. *Developmental Science*, 14, 679–692. <https://doi.org/10.1111/j.1467-7687.2010.01012.x>
- Bynner, J., & Parsons, S. (2005). *Does numeracy matter more?*. London: National Research and Development Centre for Adult Literacy and Numeracy, Institute of Education.
- Carroll, J. M., & Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry*, 45, 631–640. <https://doi.org/10.1111/j.1469-7610.2004.00252.x>
- Castellanos, F. X., Sonuga-Barke, E. J. S., Scheres, A., Di Martino, A., Hyde, C., & Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biological Psychiatry*, 57(11), 1416–1423. <https://doi.org/10.1016/j.biopsych.2004.12.005>
- Cirino, P. T., Fuchs, L. S., Elias, J. T., Powell, S. R., & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities*, 48(2), 156–175.
- Clayden, J. D., Jentschke, S., Muñoz, M., Cooper, J. M., Chadwick, M. J., Banks, T., ... Vargha-Khadem, F. (2012). Normative development of white matter tracts: Similarities and differences in relation to age, gender, and intelligence. *Cerebral Cortex*, 22(8), 1738–1747. <https://doi.org/10.1093/cercor/bhr243>
- Coghill, D., & Sonuga-Barke, E. J. S. (2012). Annual Research Review: Categories versus dimensions in the classification and conceptualisation of child and adolescent mental disorders – Implications of recent empirical study. *Journal of Child Psychology and Psychiatry*, 53, 469–489. <https://doi.org/10.1111/j.1469-7610.2011.02511.x>
- de Beer, J., Engels, J., Heerkens, Y., & van der Klink, J. (2014). Factors influencing work participation of adults with developmental dyslexia: A systematic review. *BMC Public Health*, 14, 77. <https://doi.org/10.1186/1471-2458-14-77>
- De Smedt, B., Taylor, J., Archibald, L., & Ansari, D. (2010). How is phonological processing related to individual differences in children's arithmetic skills? *Developmental Science*, 13, 508–520. <https://doi.org/10.1111/j.1467-7687.2009.00897.x>
- Department for Education United Kingdom. (2017). National curriculum assessments at key stage 2 in England, 2017. Retrieved from [www.gov.uk](http://www.gov.uk)
- Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., Buckner, R. L., ... Albert, M. S. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31(3), 968–980. <https://doi.org/10.1016/j.neuroimage.2006.01.021>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test*. Minneapolis, MN: Pearson Education.
- Duinmeijer, I., de Jong, J., & Scheper, A. (2012). Narrative abilities, memory and attention in children with a specific language impairment. *International Journal of Language & Communication Disorders*, 47, 542–555. <https://doi.org/10.1111/j.1460-6984.2012.00164.x>
- Duranovic, M., Tinjak, S., & Turbic-Hadzagic, A. (2014). Morphological knowledge in children with dyslexia. *Journal of Psycholinguistic Research*, 43(6), 699–713. <https://doi.org/10.1007/s10936-013-9274-2>
- Fair, D. A., Bathula, D., Nikolas, M. A., & Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proceedings of the National Academy of Sciences of the United States of America*, 109(17), 6769–6774. <https://doi.org/10.1073/pnas.1115365109>
- Faraone, S. V., Biederman, J., Lehman, B. K., Spencer, T., Norman, D., Seidman, L. J., Kraus, I., Perrin, J., Chen, W. J., & Tsuang, M. T. (1993). Intellectual performance and school failure in children with attention deficit hyperactivity disorder and in their siblings. *Journal of Abnormal Psychology*, 102(4), 616–623.
- Faraone, S. V., & Biederman, J. (1998). Neurobiology of attention-deficit hyperactivity disorder. *Biological Psychiatry*, 44(10), 951–958. [https://doi.org/10.1016/S0006-3223\(98\)00240-6](https://doi.org/10.1016/S0006-3223(98)00240-6)
- Frederickson, N., Reason, R., & Firth, U. (1997). *Phonological Awareness Battery (PhAB)*. London, UK: GL Assessment.
- Frederickson, N., Frith, U., & Reason, R. (1997). *Phonological Assessment Battery*. London, UK: GL Assessment.
- Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The Prevention, Identification, and Cognitive Determinants of Math Difficulty. *Journal of Educational Psychology*, 97(3), 493–513. <https://doi.org/10.1037/0022-0663.97.3.493>
- Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., ... Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, 98(1), 29–43. <http://dx.doi.org/10.1037/0022-0663.98.1.29>
- Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., van der Walt, S., Descoteaux, M., ... Contributors, D. (2014). DiPy a library for the analysis of diffusion MRI data. *Frontiers in Neuroinformatics*, 8, 8. <https://doi.org/10.3389/fninf.2014.00008>
- Gathercole, S. E., Woolgar, F., the CALM team, Kievit, R., Astle, D., Manly, T., & Holmes, J. (2016). How common are WM deficits in children with difficulties in reading and mathematics? *Journal of Applied Research in Memory and Cognition*, 5(4), 384–394. <https://doi.org/10.1016/j.jarmac.2016.07.013>
- Geary, D. C. (2004). Mathematics and learning disability. *Journal of Learning Disabilities*, 37(1), 4–15. <https://doi.org/10.1177/00222194040370010201>
- Germano, E., Gagliano, M. D., & Curatolo, P. (2010). Comorbidity of ADHD and dyslexia. *Developmental Neuropsychology*, 35(5), 475–493. <https://doi.org/10.1080/87565641.2010.494748>
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). *Behaviour Rating Inventory of Executive Function – Ages 5–18 (BRIEF)*. Lutz, FL: Psychological Assessment Resources.
- Girard, J. L., & Lek, S. (2001). A comparison of self-organizing map algorithm and some conventional statistical methods for ecological community ordination. *Ecological Modelling*, 146(1–3), 329–339. [https://doi.org/10.1016/S0304-3800\(01\)00324-6](https://doi.org/10.1016/S0304-3800(01)00324-6)
- Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLoS Biology*, 6, e159. <https://doi.org/10.1371/journal.pbio.0060159>
- Hart, S. A., Petrill, S. A., Willcutt, E., Thompson, L. A., Schatschneider, C., Deater-Deckard, K., & Cuttin, L. E. (2010). Exploring how symptoms of Attention-Deficit/Hyperactivity Disorder are related to reading and mathematics performance: General genes, general environments. *Psychological Science*, 21(11), 1708–1715. <https://doi.org/10.1177/0956797610386617>
- Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology*, 79(2), 192–227. <https://doi.org/10.1006/jecp.2000.2586>
- Holmes, J., Adams, J. W., & Hamilton, C. J. (2008). The relationship between visuospatial sketchpad capacity and children's mathematical skills. *European Journal of Cognitive Psychology*, 20(2), 272–289. <https://doi.org/10.1080/09541440701612702>
- Holmes, J., Hilton, K. A., Place, M., Alloway, T. P., Elliott, J. G., & Gathercole, S. E. (2014). Children with low working memory and children with ADHD: same or different? *Frontiers in Human Neuroscience*, <https://doi.org/10.3389/fnhum.2014.00976>
- Joanisse, M. F., Manis, F. R., Keating, P., & Seidenberg, M. S. (2000). Language deficits in dyslexic children: Speech perception. *Phonology*,



- and Morphology. *Journal of Experimental Child Psychology*, 77(1), 30–60. <https://doi.org/10.1006/jecp.1999.2553>
- Kohonen, T. (1989). *Self-organizing feature maps. Self-organization and associative memory*. Springer Series in Information Sciences, vol 8. Berlin, Germany: Springer. <https://doi.org/10.1007/978-3-642-88163-3>
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., & Zimmerman, M. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): A dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126(4), 454–477. <https://doi.org/10.1037/abn0000258>
- Lebel, C., Treit, S., & Beaulieu, C. (2017). A review of diffusion MRI of typical white matter development from early childhood to young adulthood. *NMR in Biomedicine*, 3(1), e3778–23. <https://doi.org/10.1002/nbm.3778>
- Lin, X., Soergel, D., & Marchionini, G. (1991). A self-organizing semantic map for information retrieval (pp. 262–269). SIGIR '91 Proceedings of the 14th annual international ACM SIGIR conference on Research and development in information retrieval.
- Lloyd, S. P. (1982). "Least squares quantization in PCM" (PDF). *IEEE Transactions on Information Theory*, 28(2), 129–137. <https://doi.org/10.1109/TIT.1982.10564s89>
- Loe, I. M., & Feldman, H. M. (2007). Academic and educational outcomes of children with ADHD. *Journal of Pediatric Psychology*, 32(6), 643–654. <https://doi.org/10.1093/jpepsy/jsl054>
- Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T., Kulju, P., Laakso, M. L., ... Richardson, U. (2004). Early development of children at familial risk for dyslexia—follow-up from birth to school age. *Dyslexia*, 10(3), 146–178. <https://doi.org/10.1002/dys.274>
- Manjon, J. V., Coupe, C., Martí-Bonmati, L., Collins, D. L., & Robles, M. (2009). Spatially adaptive non-local MRI denoising. *NeuroImage*, 47(July), S72. [https://doi.org/10.1016/s1053-8119\(09\)70443-4](https://doi.org/10.1016/s1053-8119(09)70443-4)
- Martel, M. M., Von Eye, A., & Nigg, J. T. (2010). Revisiting the latent structure of ADHD: Is there a 'g' factor? *Journal of Child Psychology and Psychiatry*, 51, 905–914. <https://doi.org/10.1111/j.1469-7610.2010.02232.x>
- Mayes, A. K., Reilly, S., & Morgan, A. T. (2015). Neural correlates of childhood language disorder: A systematic review. *Developmental Medicine & Child Neurology*, 57(8), 706–717. <https://doi.org/10.1111/dmcn.12714>
- McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children's arithmetical performance. *Educational and Child Psychology*, 20(3), 93–108.
- McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology*, 74(3), 240–260. <https://doi.org/10.1006/jecp.1999.2516>
- McArthur, G., Hogben, J., Edwards, V., Heath, S., & Mengler, E. (2000). On the "Specifics" of Specific Reading Disability and Specific Language Impairment. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(7), 869–874.
- Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoof, N., ... Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, 29, 65–77. <https://doi.org/10.1016/j.learninstruc.2013.09.003>
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., & Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biological Psychiatry*, 57(11), 1224–1230. <https://doi.org/10.1016/j.biopsych.2004.08.025>
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., ... Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, 57, 1247–1257. <https://doi.org/10.1111/jcpp.12573>
- Peacock, J. A. (1983). Two-dimensional goodness-of-fit testing in astronomy. *Monthly Notices of the Royal Astronomical Society*, 202(3), 615–627. <https://doi.org/10.1093/mnras/202.3.615>
- Peng, X., Lin, P., Zhang, T., & Wang, J. (2013). Extreme learning machine-based classification of ADHD using brain structural MRI data. *PLoS ONE*, 8(11), e79476. <https://doi.org/10.1371/journal.pone.0079476>
- Pointus, A. A. (1981). Geometric figure-rotation task and face representation in dyslexia: Role of spatial relations and orientation. *Perceptual and Motor Skills*, 53(2), 607–614. <https://doi.org/10.2466/pms.1981.53.2.607>
- Polanczyk, G. V., Willcutt, E. G., Salum, G. A., Kieling, C., & Rohde, L. A. (2014). ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis. *International Journal of Epidemiology*, 43(2), 434–442. <https://doi.org/10.1093/ije/dyt261>
- Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, 20(2), 110–122. <https://doi.org/10.1016/j.lindif.2009.10.005>
- Ramus, F., Marshall, C. R., Rosen, S., & van der Lely, H. K. J. (2013). Phonological deficits in specific language impairment and developmental dyslexia: Towards a multidimensional model. *Brain*, 136(2), 630–645. <https://doi.org/10.1093/brain/aws356>
- Ramus, F., Marshall, C.R., Rosen, S., & van der Lely, H.K. (2013). Phonological deficits in specific language impairment and developmental dyslexia: towards a multidimensional model. *Brain*, 136(2), 630–645.
- Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology*, 91(2), 137–157. <https://doi.org/10.1016/j.jecp.2005.01.004>
- Rimrodt, S. L., Peterson, D. J., Denckla, M. B., Kaufmann, W. E., & Cutting, L. E. (2010). White matter microstructural differences linked to left perisylvian language network in children with dyslexia. *Cortex*, 46(6), 739–749. <https://doi.org/10.1016/j.cortex.2009.07.008>
- Roberts, T. P. L., Heiken, K., Zarnow, D., Dell, J., Nagae, L., Blaskey, L., ... Edgar, J. C. (2014). Left hemisphere diffusivity of the arcuate fasciculus: Influences of autism spectrum disorder and language impairment. *AJNR. American Journal of Neuroradiology*, 35(3), 587–592. <https://doi.org/10.3174/ajnr.A3754>
- Roeser, R. W., & Eccles, J. S. (2000). Schooling and mental health. In A. J. Sameroff, M. Lewis & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (pp. 135–156). Boston, MA: Springer.
- Rommelse, N. N. J., Geurts, H. M., Franke, B., Buitelaar, J. K., & Hartman, C. A. (2011). A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neuroscience & Biobehavioral Reviews*, 35(6), 1363–1396. <https://doi.org/10.1016/j.neubiorev.2011.02.015>
- Russell, G., Rodgers, L. R., Ukomunne, O. C., & Ford, T. (2014). Prevalence of parent-reported ASD and ADHD in the UK: Findings from the millennium cohort study. *Journal of Autism and Developmental Disorders*, 44(1), 31–40. <https://doi.org/10.1007/s10803-013-1849-0>
- Russell, G., Rodgers, L. R., Ukoumunne, O. C., & Ford, T. (2014). Prevalence of parent-reported ASD and ADHD in the UK: Findings from the Millennium Cohort Study. *Journal of Autism and Developmental Disorders*, 44(1), 31–40. <https://doi.org/10.1007/s10803-013-1849-0>
- Scott, S. K., McGettigan, C., & Eisner, F. (2009). A little more conversation, a little less action – Candidate roles for the motor cortex in speech perception. *Nature Reviews Neuroscience*, 10(4), 295–302. <https://doi.org/10.1038/nrn2603>
- Sermud-Clíkeman, M., Biederman, J., Sprich-Buckminster, S., Lehman, B.K., Faraone, S.V., & Norman, D. (1992). Comorbidity between ADDH and Learning Disability: A review and report in a clinically



- referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, 31(3), 439–448.
- Shalev, R. S., & Gross-Tsur, V. (2001). Developmental dyscalculia. *Pediatric Neurology*, 24(5), 337–342. [https://doi.org/10.1016/S0887-8994\(00\)00258-7](https://doi.org/10.1016/S0887-8994(00)00258-7)
- Shen, X., Finn, E. S., Scheinost, D., Rosenberg, M. D., Chun, M. M., Papademetris, X., & Constable, R. T. (2017). Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *Nature Protocols*, 12(3), 506–518. <https://doi.org/10.1038/nprot.2016.178>
- Simmons, F., Singleton, C., & Horne, J. (2008). Brief report—Phonological awareness and visual-spatial sketchpad functioning predict early arithmetic attainment: Evidence from a longitudinal study. *European Journal of Cognitive Psychology*, 20(4), 711–722. <https://doi.org/10.1080/09541440701614922>
- Snowling, M. J. (1995). Phonological processing and developmental dyslexia. *Journal of Research in Reading*, 18(2), 132–138. <https://doi.org/10.1111/j.1467-9817.1995.tb00079.x>
- Snowling, M., Bishop, D. V., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 41(5), 587–600. <https://doi.org/10.1111/1469-7610.00651>
- Sonuga-Barke, E. J. S., & Coghill, D. (2014). The foundations of next generation attention-deficit/hyperactivity disorder neuropsychology: Building on progress during the last 30 years. *Journal of Child Psychology and Psychiatry*, 55, e1–e5. <https://doi.org/10.1111/jcpp.12360>
- Sprague, T. C., & Serences, J. T. (2013). Attention modulates spatial priority maps in the human occipital, parietal and frontal cortices. *Nature Neuroscience*, 16(12), 1879–1887. <https://doi.org/10.1038/nn.3574>
- Stevens, M. S., Skudlarski, P., Pearlson, G. D., & Calhoun, V. D. (2009). Age-related cognitive gains are mediated by the effects of white matter development on brain network integration. *NeuroImage*, 48(4), 738–746. <https://doi.org/10.1016/j.neuroimage.2009.06.065>
- Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, 96(3), 471–491. <https://doi.org/10.1037/0022-0663.96.3.471>
- Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology*, 79(3), 294–321. <https://doi.org/10.1006/jecp.2000.2587>
- Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, 49(10), 2674–2688. <https://doi.org/10.1016/j.cortex.2013.06.007>
- Tamayo, P., Slonim, D., Mesirov, J., Zhu, Q., Kitareewan, S., Dmitrovsky, E., ... Golub, T. R. (1999). Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 96, 2907–2912. <https://doi.org/10.1073/pnas.96.6.2907>
- Tamboer, P., Vorst, H. C. M., Ghebreab, S., & Scholte, H. S. (2016). Machine learning and dyslexia: Classification of individual structural neuro-imaging scans of students with and without dyslexia. *NeuroImage Clinical*, 11, 508–514. <https://doi.org/10.1016/j.nicl.2016.03.014>
- Tanaka, H., Black, J. M., Hulme, C., Stanley, L. M., Kesler, S. R., Whitfield-Gabrieli, S., ... Hoeft, F. (2011). The brain basis of the phonological deficit in dyslexia is independent of IQ. *Psychological Science*, 22(11), 1442–1451. <https://doi.org/10.1177/0956797611419521>
- Tian, J., Azarian, M. H., & Pecht, M. (2014). *Anomaly detection using self-organizing maps-based K-nearest neighbor algorithm*. Proceeding of the European Conference of the Prognostics and Health Management Society.
- Van der Ven, S. H. G., Kroesbergen, E. H., Boom, J., & Leseman, P. P. M. (2012). The development of executive functions and early mathematics: A dynamic relationship. *British Journal of Educational Psychology*, 82, 100–119. <https://doi.org/10.1111/j.2044-8279.2011.02035.x>
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45, 2–40. <https://doi.org/10.1046/j.0021-9630.2003.00305.x>
- Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101(2), 192–212. <https://doi.org/10.1037/0033-2909.101.2.192>
- Wechsler, D. (2001). *Wechsler Individual Achievement Test* (2nd ed.). London, UK: Pearson Assessment.
- Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence* (2nd ed.). London, UK: Pearson Assessment.
- Willcutt, E., & Pennington, B. (2000). Psychiatric comorbidity in children and adolescents with reading disability. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(8), 1039–1048. <https://doi.org/10.1111/1469-7610.00691>
- Winner, E., von Karolyi, C., Malinsky, D., French, L., Seliger, C., Ross, E., & Weber, C. (2001). Dyslexia and visual-spatial talents: Compensation vs deficit model. *Brain and Language*, 76(2), 81–110. <https://doi.org/10.1006/brln.2000.2392>
- Woodcock, R. W., Schrank, F. A., & Mather, N. (2007). *Woodcock-Johnson III*. Rolling Meadows, IL: Riverside.
- Zentall, S. S. (2007). Math performance of students with ADHD: Cognitive and behavioral contributors and interventions. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 219–243). Baltimore, MD, US: Paul H Brookes Publishing.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Astle DE, Bathelt J, the CALM team, Holmes J. Remapping the cognitive and neural profiles of children who struggle at school. *Dev Sci*. 2019;22:e12747. <https://doi.org/10.1111/desc.12747>