Neuroscience and Biobehavioral Reviews xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Neuroscience and Biobehavioral Reviews

journal homepage: www.elsevier.com/locate/neubiorev

Review article

Neuroanatomy of developmental dyslexia: Pitfalls and promise

Franck Ramus[a,*], Irene Altarelli[b], Katarzyna Jednoróg[c], Jingjing Zhao[d], Lou Scotto di Covella[a]

[a] Laboratoire de sciences cognitives et psycholinguistique (CNRS, ENS, EHESS, PSL Research University), Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris, France
[b] Brain and Learning Lab, Campus Biotech, University of Geneva, 9 Chemin des Mines, 1205 Geneva, Switzerland
[c] Laboratory of Psychophysiology, Department of Neurophysiology, Nencki Institute of Experimental Biology, Polish Academy of Sciences, 3 Pasteur St, 02-093 Warsaw, Poland
[d] School of Psychology, Shaanxi Normal University and Key Laboratory for Behavior and Cognitive Neuroscience of Shaanxi Province, Xi'an, Shaanxi, 710062, China

ARTICLE INFO

Keywords:
Dyslexia
Neuroanatomy
Voxel-based morphometry
Diffusion imaging
Brain volume
Planum temporale

ABSTRACT

Investigations into the neuroanatomical bases of developmental dyslexia have now spanned more than 40 years, starting with the post-mortem examination of a few individual brains in the 60s and 70s, and exploding in the 90s with the widespread use of MRI. The time is now ripe to reappraise the considerable amount of data gathered with MRI using different types of sequences (T1, diffusion, spectroscopy) and analysed using different methods (manual, voxel-based or surface-based morphometry, fractional anisotropy and tractography, multivariate analyses…). While selective reviews of mostly small-scale studies seem to provide a coherent view of the brain disruptions that are typical of dyslexia, involving left perisylvian and occipito-temporal regions, we argue that this view may be deceptive and that meta-analyses and large-scale studies rather highlight many inconsistencies and limitations. We discuss problems inherent to small sample size as well as methodological difficulties that still undermine the discovery of reliable neuroanatomical bases of dyslexia, and we outline some recommendations to further improve this research area.

1. Introduction

Developmental dyslexia is a specific impairment in the acquisition of reading skills that is not solely accounted for by mental age, visual acuity problems, or inadequate schooling (World Health Organization, 2011). Cases with dyslexia have been described for more than a century, and the hypothesis of a specific brain dysfunction was present right from the first one (Morgan, 1896). Investigations into the neuroanatomical bases of developmental dyslexia began with a post-mortem case description by Drake (1968), followed by another series of post-mortem cases by Galaburda and colleagues (Galaburda et al., 1994, 1985; Galaburda and Eidelberg, 1982; Galaburda and Kemper, 1979; Galaburda and Livingstone, 1993; Humphreys et al., 1990). Neuroimaging was rapidly used to broaden the investigations to live individuals and therefore to significantly increase the number of cases studied. The first brain scans of dyslexic individuals may be attributed to Hier et al. (1978) and exploded from the 90s with the advent of MRI, starting with Rumsey et al. (1986).

Post-mortem dissections had the exceptional advantage of offering both macro- and microscopic views of the dyslexic brain, using for the latter standard histological techniques, thus giving some insight into cellular and architectonic phenomena. Disruptions observed in such studies included ectopic neurons, heterotopias, dysplasias, converging towards the hypothesis of a disturbance of neuronal migration (Drake, 1968; Galaburda et al., 1985; Humphreys et al., 1990). Furthermore, disruptions were predominantly localised in left perisylvian regions, suggesting a straightforward link with language and reading networks. While this hypothesis is extremely seductive, it has suffered firstly from its reliance on a small number of cases (a total of 5 male and 3 female dyslexic individuals, compared to 13 male and 3 female controls), and secondly from the lack of additional direct confirmation since those initial observations. Indeed, whereas magnetic resonance imaging (MRI) has allowed researchers to re-assess macroscopic observations on a much larger scale, it has until now lacked the contrast and resolution that would be necessary to observe minute disruptions of neuronal migration. Modern neuroimaging has therefore left that initial hypothesis largely untouched, but has multiplied the opportunities of observing macroscopic alterations of brain structure, using different types of image acquisition sequences, and a growing variety of image analysis techniques.

The present paper attempts to provide a comprehensive view of neuroanatomical differences between dyslexic and control individuals uncovered during 30 years of magnetic resonance imaging (studies focusing on correlations between brain properties and literacy measures

* Corresponding author.
E-mail address: franck.ramus@ens.fr (F. Ramus).

http://dx.doi.org/10.1016/j.neubiorev.2017.08.001
Received 25 November 2016; Received in revised form 28 July 2017; Accepted 2 August 2017

without a clearly identified dyslexic group will not be included). One leading thread will be the gap between a broadly accepted discourse providing an idealised view of the neuroanatomy of dyslexia, and what emerges from a careful, systematic analysis of the data. Indeed, almost all papers on the neural basis of dyslexia start with a paragraph stating that dozens of studies have well established the existence of structural disruptions localised mostly in left perisylvian and occipito-temporal cortex, and in the white matter connecting these areas. Depending on their focus, some papers will additionally mention disruptions in the cerebellum and/or in the thalamus. Literature reviews tend to give a similar picture. We would love to believe that story, and indeed, in our own papers, we have also contributed to this discourse. However, we would like to raise the possibility that this apparent consensus relies on selective reviews of the evidence, with a marked preference for results that seem convergent and that can easily be weaved into a plausible cognitive theory of dyslexia, at the cost of disregarding inconsistencies and findings in regions that lack an obvious cognitive interpretation. We will further discuss the methodological differences that may limit convergence between studies, as well as the impact of small-scale studies on the emergence of a potentially false consensus.

2. Global brain measures

Although this has seldom been a focus of major interest, total brain volume has been reported in many studies of dyslexia, mostly as a control variable for more local measures. Many studies have reported lower brain volumes in dyslexic compared to control individuals. However, the differences reported vary greatly between studies, with effect sizes ranging from d = −1 (Frye et al., 2010) to 1.53 (Eckert et al., 2005). Furthermore, studies vary in the populations studied (children vs. adults) and in the precise nature of the volume reported: total grey and white matter volumes (GMV and WMV), or total brain volume (TBV, adding GMV and WMV, usually including subcortical grey matter but excluding the cerebellum), or total intracranial volume (TIV), including grey matter, white matter, meninges and cerebrospinal fluid. In order to have a clearer picture of global differences in brain volume between dyslexic and control individuals, we carried out a systematic literature search and a meta-analysis of TIV, TBV, GMV and WMV across 18 studies including 1164 participants. This meta-analysis confirmed that dyslexic individuals showed lower total brain volume than controls in all the measures considered (see Table 1 and full results in Supplements 2 and 3). We did not find evidence that this pattern differed between children and adults.

Because of the well-established correlation between total brain volume and IQ (r = 0.3; McDaniel, 2005), and because dyslexic and control groups are often matched in non-verbal IQ, but seldom on verbal and on full-scale IQ (FSIQ), one might wonder whether such a difference in brain volume might be due to differences in IQ. In order to test this hypothesis, we carried out an additional analysis on 81 French and 81 Polish children from Jednoróg et al. (2015), for whom we had both global brain measures and estimated FSIQ scores (based on Wechsler's Blocks and Similarities). We found that, even after controlling for IQ, group differences in TIV, TBV, GMV and WMV remained significant and with an effect size of 0.42 (see Supplement 1). Thus, group differences in total brain volume do not seem to be accounted for by differences in FSIQ. In our data, these effect sizes translate into 53% to 59% of dyslexic individuals having a GMV (respectively WMV) below the lower 95% confidence interval of the control group.

Table 1
Summary of meta-analyses of global brain measures, compared between dyslexic and control individuals.

Measure	N studies	N controls	N dyslexics	Cohen's d	95% CI
Total intracranial volume	5	130	153	0.34	[−0.19; 0.87]
Total brain volume	13	243	259	0.58	[0.32; 0.85]
Total grey matter volume	14	412	458	0.40	[0.12; 0.69]
Total white matter volume	14	331	376	0.40	[0.20; 0.61]

Another potential confounding factor would be body size, which covaries with head size: if dyslexic individuals were shorter than controls, this might explain differences in brain volume. Unfortunately, we have not found any study addressing that question satisfactorily (Lysiak-Seichter et al., 2006 reported differences but with comorbidity and socioeconomic confounds).

Alternative measurements to total brain volume exist, such as total cortical thickness and surface area, with surface showing much higher correlation (0.81) with volume than thickness (0.15) (Panizzon et al., 2009; Winkler et al., 2010). Very few studies have compared whole-brain cortical thickness and surface area across dyslexic and control participants, and no clear picture has emerged. In brief, Frye et al. (2010) found greater total surface area in dyslexic adults compared to controls, while the opposite pattern emerged in the study of Altarelli et al. (2014). The result by Frye et al., based on 16 dyslexic and 16 control adults, is consistent with their finding of greater total GMV in dyslexics (d = −1). However, this result was the most extreme outlier in the opposite direction from the general trend in our meta-analysis of total grey matter volume (d = 0.4). It therefore seems reasonable to disregard it, and to expect more generally that dyslexic individuals show lower total cortical surface than controls, consistently with their lower total brain volume.

Another, different measure of global brain anatomy concerns its global asymmetry. Asymmetries in width and volume of the frontal and occipital poles have been documented in the general population, with the right frontal pole typically being larger than the left, while the opposite applies to the occipital pole (Galaburda et al., 1978; Kertesz et al., 1990; LeMay, 1976; Watkins et al., 2001). Some differences between dyslexic and control individuals have been reported, especially in the occipital pole, found to be either rightward asymmetrical or symmetrical in at least a subsample of the dyslexic population (Duara et al., 1991; Haslam et al., 1981; Hier et al., 1978). These observations have shown mixed replications (Dole et al., 2013; Hynd et al., 1990; Zadina et al., 2006), but overall too few studies have focused on these aspects. As a result, no definitive answer has been reached yet regarding macroscopic asymmetries in dyslexia. Interestingly, a link might exist between global patterns of asymmetry and more localised ones, including the asymmetry of language areas within the Sylvian fissure (Barrick et al., 2005). Further investigations of global asymmetries as well as of their relations with local asymmetries in dyslexia would seem necessary.

Finally, cerebral cortex folding can also be evaluated at the whole brain level. The study of gyrification through automated techniques is relatively young, and thus few studies have applied it in the quest for anatomical bases of developmental dyslexia. While the pioneering work of Casanova et al. (2004) provided evidence for reduced whole-brain gyrification in developmental dyslexia, we have been unable to replicate these observations in our own investigations (Altarelli, 2013; Altarelli et al., submitted).

In sum, the observation of reduced total brain volume in developmental dyslexia seems to be the most robust when compared to other global measures, as these observations have been replicated across a substantial number of separate studies and emerge clearly from our *meta*-analysis. It is important to note that, as far as a link with cognitive function is concerned, very little is currently known. As we have previously mentioned, reduced brain volume cannot be attributed to lower IQ. We are not aware of any hypothesis involving lower brain volume as a specific aetiological factor for dyslexia, although it is of course possible that lower brain volume might be a general susceptibility factor

for many developmental disorders, in conjunction with other, more specific factors. Alternatively, perhaps reduced brain volume is a secondary consequence of other disruptions of brain development. Reductions in brain weight and cortical volume have also been reported in animals with induced cortical anomalies similar to the ones associated with developmental dyslexia (Peiffer et al., 2002); in these animal models, the timing of early cortical insult appeared to be crucial (Threlkeld et al., 2006). However, knocking down the expression of either *Dyx1c1* or *Kiaa0319*, two candidate dyslexia susceptibility gene homologues, although producing a variety of brain defects, did not result in reduced total cerebral volume (Szalkowski et al., 2013, 2012). Thus, whether brain volume reduction participates in the aetiology of dyslexia or whether it is a mere consequence of early disruption of brain development remains to be clarified. At any rate, group differences in total brain volume cannot be ignored, as they may affect many other local brain differences.

3. Voxel-based morphometry

Voxel-based morphometry (VBM) is an automated technique for assessing structural changes in the brain (Ashburner et al., 1998). Since it is quick and relatively easy to use, it has largely grown in popularity since its introduction. The technique typically uses T1-weighted volumetric MRI scans and performs voxel-wise statistical tests across the whole brain to identify structural differences between groups. It has until now been the most popular technique to analyse neuroanatomical differences in dyslexia and other disorders. Around twenty VBM studies have been published[1] with the aim of examining structural differences between dyslexic and non-dyslexic adults and children (Brambati et al., 2004; Brown et al., 2001; Dole et al., 2013; Eckert et al., 2016b, 2005; Evans et al., 2014; Hoeft et al., 2007; Jednoróg et al., 2015, 2014; Krafnick et al., 2014; Kronbichler et al., 2008; Menghini et al., 2008; Pernet et al., 2009a,b; Silani et al., 2005; Siok et al., 2008; Steinbrink et al., 2008; Tamboer et al., 2015; Vinckenbosch et al., 2005; Xia et al., 2016; Yang et al., 2016).

Areas most frequently reported include: left posterior temporal and temporo-parietal regions, where both increased and decreased grey matter volume (GMV) was reported in dyslexia (Brambati et al., 2004; Eckert et al., 2016b; Hoeft et al., 2007; Silani et al., 2005; Steinbrink and Klatte, 2008); left inferior frontal gyrus with reduced GMV in dyslexia (Brown et al., 2001; Eckert et al., 2016b), occipito-temporal regions bilaterally (Brambati et al., 2004; Eckert et al., 2005; Kronbichler et al., 2008) and cerebellum, also showing GMV reduction (Brambati et al., 2004; Brown et al., 2001; Eckert et al., 2005; Kronbichler et al., 2008). Finally, there are studies where no significant differences in GMV between dyslexic and control groups were revealed (Eckert et al., 2016b; Pernet et al., 2009a,b; Tamboer et al., 2015). Qualitative literature reviews have offered relatively optimistic syntheses of these results, suggesting converging evidence for decreased grey matter volumes across a distributed set of regions, predominantly in the left hemisphere (Eckert, 2004; Giraud and Ramus, 2013; Kershner, 2015; Norton et al., 2015; Pugh et al., 2001; Richardson and Price, 2009). However, three meta-analyses have revealed limited consistency of these findings across studies (Eckert et al., 2016b; Linkersdörfer et al., 2012; Richlan et al., 2013). The earlier two meta-analyses included nine VBM studies, of which seven in common, with a total of 277 and 266 participants respectively. Linkersdörfer et al. (2012), using an activation likelihood estimation method, found significantly lower GMV in dyslexics in bilateral supramarginal gyri, cerebellum, right superior temporal gyrus, left fusiform and inferior temporal gyri. Richlan et al. (2013), using effect size signed differential mapping, reported differences only in bilateral superior temporal areas, at $p < 0.005$ uncorrected. Indeed, in that meta-analysis no group difference emerged at a false discovery rate threshold of 0.05. The most recent meta-analysis by Eckert et al. (2016b) included 11 studies involving a total of 462 participants, and reported lower GMV in dyslexics in left orbitofrontal cortex, left posterior STS, and right cerebellum, using similar methods and threshold as Richlan et al. (2013). Obviously, the results of meta-analyses are quite sensitive to the studies included and methods used.

Inconsistencies between studies may have different sources. Although VBM has the advantage of being automatic, objective and user-independent, it has also some drawbacks. It is rarely implemented consistently across studies, and changing user-specified options can alter the results, often in a way similar to the biological differences under investigation (Henley et al., 2009). For instance, modulation (taking into account absolute GMV from the original images) has not always been used (Brown et al., 2001; Vinckenbosch et al., 2005), but results have been interpreted similarly whether this step was included or not. It has been already shown in studies of Williams Syndrome that the choice to modulate images during pre-processing is a likely explanation for discrepant VBM findings (Eckert et al., 2006). Similarly the registration algorithm itself, being either diffeomorphic flow-based algorithm (DARTEL) with group specific template or constrained warp with an a priori template, can introduce differences, with more focal results obtained with the former approach (Peelle et al., 2012). Finally, the choice of smoothing kernel size is also known to affect the results, since the extent of the findings increases with kernel size (Henley et al., 2010). With the improvement of registration methods, the choice of smaller kernels is generally recommended (Shen and Sterr, 2013), however they range from 4 to 12 mm in previous dyslexia studies.

At the level of statistical analyses, it is recommended to include an adjustment for some measure of head size and other nuisance variables in the VBM design matrix, especially when affine and non-linear modulation is used in a standard VBM pre-processing pipeline.[2] Often the lack of a statistically significant group difference in TIV, age or sex is wrongly assumed to imply that these factors do not affect the results (Barnes et al., 2010) and they are therefore seldom included. Previous VBM studies in dyslexia either did not use adjustment for head size (Brown et al., 2001; Evans et al., 2014) or used different indices − TIV (e.g., Eckert et al., 2005; Jednoróg et al., 2014), total GMV (e.g., Eckert et al., 2016b; Hoeft et al., 2007; Kronbichler et al., 2008) or total GMV and WMV (e.g., Menghini et al., 2008). Authors tend to use these global measures indifferently, yet they are not interchangeable. For instance, total GMV, WMV and TIV show different longitudinal trends, both in childhood and in ageing. This may be of importance when comparing different age groups or including a wide age range. In Huntington's disease, it has been shown that adjusting for total GMV removes some disease-related effects (presumably because of global grey matter degeneration), while adjusting for TIV does not (Henley et al., 2010). In dyslexia, where differences in TIV, GMV and WMV of similar effect sizes are observed (see previous section), there may not be a single best global covariate. The most logical, for the purpose of distinguishing local from global differences, would be to use a global covariate of the

[1] For this section and the associated Fig. 1, we carried out a systematic bibliographic search in Pubmed, with the following keyword combination: "dyslexia"[MeSH Terms] OR dyslex*[All Fields] OR reading disab*[All Fields] AND ("brain"[MeSH Terms] OR "brain"[All Fields]) AND (volume*[All Fields] OR voxel*[All Fields]). This resulted in 141 articles. After an initial screening based on relevance of the title and abstract, there remained 59 potentially relevant studies. For Fig. 1, we selected studies that met the following criteria:

- Reporting clusters of grey matter showing significant differences between a group of dyslexic and a group of control individuals, in a whole-brain analysis.
- Reporting sample sizes. This resulted in 20 studies. When several statistical analyses with different thresholds were reported, we chose the numbers corresponding to the most stringent analysis.

[2] In Christian Gaser's VBM8 toolbox the modulated images can be optionally saved by correcting for non-linear warping only. By ignoring the affine scaling, the modulation is equivalent to a global scaling with the (inverse) scaling factor of the affine transformation and there is no need to additionally correct for individual brain size.

same nature as the local measure of interest: total GMV for local GMV, total WMV for local WMV, total cortical surface for local surface, whole-brain mean FA for local FA, etc. For instance, in their analysis of regions of interest, Eckert et al. (2016a) reported that group differences in local GMV disappeared once total GMV was controlled, suggesting that they were mere consequences of total GMV differences. Finally, other considerations may also be taken into account in choosing a nuisance covariate, such as collinearity and bias amplification.

Last but not least, very different results can be obtained by varying the level and type of statistical correction for multiple tests, and these differ a lot in previous VBM studies of dyslexia. Because VBM is performed voxel-wise, and because there are many voxels in the brain, the problem of multiple testing is particularly acute with this method. Yet, reporting uncorrected or insufficiently corrected results is not uncommon in dyslexia research, which increases the risk of reporting false positive results.

A related problem is the relatively small sample size of most neuroimaging studies of dyslexia, which typically involve 10–20 subjects per group. Historically, the first anatomical studies had to rely on a very limited number of available post-mortem samples. Because these studies revealed what seemed to be large differences, researchers using neuroimaging could hope to observe such differences with relatively small sample sizes, consistent with the standards of that time. In hindsight, however, it is now evident that neuroimaging studies with 10–20 participants per group are statistically underpowered to reveal group differences at a fully corrected level, hence the widespread use of uncorrected or partly-corrected statistics. In theory, the larger the sample size, the larger the number of significant clusters that should be reported. Yet, it is troubling that the largest four published VBM studies of dyslexia reported either one region of significant GMV difference (Jednoróg et al., 2015; N = 236 across 3 countries) or none at all (Eckert et al., 2016b, N = 255 across 6 sites; Pernet et al., 2009a, N = 77; Tamboer et al., 2015, N = 94) when fully correcting for multiple comparisons. Furthermore, these large studies were not well represented in meta-analyses: The Pernet et al. (2009a) study was included only in the meta-analysis by Linkersdörfer et al. (2012); The Tamboer et al. (2015) only in Eckert et al.'s (2016b) meta-analysis; The most recent and largest two studies (Eckert et al., 2016b; Jednoróg et al., 2015) were not included in any meta-analysis. It will be interesting to see whether any group difference remains significant once they are.

Dyslexia research does not seem to be isolated in this respect. Fusar-Poli et al. (2014) have noted that the number of clusters reported in VBM studies does not seem to increase with sample size as much as would be expected from theoretical statistical power. We have carried out a similar analysis of VBM studies of dyslexia (Fig. 1[3]), and it is pretty obvious that the trend is not in the expected direction, small-scale studies (N < 20 per group) contributing disproportionately to the differences reported in the literature. This suggests that the number of foci reported in small sample VBM studies in general, and in meta-analyses with few studies, is inflated, through various factors including publication bias in favour of the studies that report differences, lax statistical thresholds, and so-called "p-hacking" (multiplying analysis variants until one yields a significant result). This in turn suggests that many of the reported differences in small-scale studies may be false positive results. Thus, at this stage, it is not clear whether any GMV difference previously reported in dyslexia research will stand the test of time.

[3] Obviously this Figure has required a number of debatable choices. We have chosen to classify and represent statistical corrections into three categories. Other factors (discussed in the text) that may affect the reporting of significant clusters include modulation, smoothing, head size adjustment, and the use of other covariates, but it was impossible to take all of them into account. We include as Supplementary data an Excel sheet used to produce the figure and summarising the main characteristics of each study (Supplement 1).

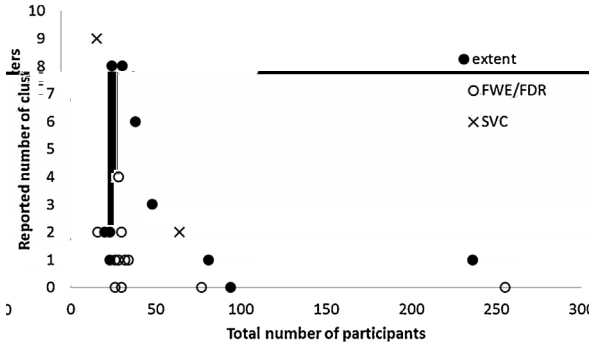

Fig. 1. Number of grey matter clusters reported to differ between control and dyslexic individuals, as a function of sample size, across 20 published whole-brain VBM analyses obtained from a systematic bibliographic search. Different marks indicate the type of correction for multiple comparisons applied. When several levels of correction were reported in a given study, the results of the most stringent one are displayed here. Black discs: Combination of height and extent thresholds. White discs: family-wise error (FWE) or false discovery rate (FDR) correction. Crosses: small volume correction (SVC).

4. Surface-based morphometry

As described in the previous section, the vast majority of structural MRI studies of developmental dyslexia have been exploring grey matter volume differences between groups, applying voxel-based morphometry. This approach however is not ideal, for at least two reasons. First, the required volume-based normalisation step appears to be less effective in matching brains than other procedures (Fischl et al., 1999; Ghosh et al., 2010), while making VBM relatively insensitive to individual anatomical features. Second, it only allows measuring differences in volume, which are sometimes difficult to interpret, given that they could stem either from the thickness of the cortical sheet or from its folding. Indeed, some authors have argued that volume estimation, as a composite measure of cortical thickness and surface area, might not be the most appropriate morphological feature to be investigated, in particular in the study of genetically-influenced disorders (Panizzon et al., 2009; Winkler et al., 2010). Given these shortcomings, a few studies have relied on alternative techniques to explore cortical surface and thickness (Shaw et al., 2012; Wallace et al., 2013), including in the study of developmental dyslexia (Altarelli et al., 2013; Altarelli et al. submitted; Clark et al., 2014; Frye et al., 2010; Ma et al., 2015).

Frye et al. (2010) used an ROI-based approach to compare the cortical thickness and surface area of four brain regions (i.e., inferior frontal, supramarginal, inferior parietal and fusiform) in a population of dyslexic and control adults. Their study revealed greater surface area in controls in the inferior frontal and fusiform regions, as well as differences in the lateralization of cortical thickness in the supramarginal region. The authors also reported inter-hemispheric differences in cortical thickness and surface area across the whole sample. Though yielding novel results, the analysis is rather limited, as it is based on a relatively small population of adults (16 per group) and as it takes under consideration only a restricted set of brain regions, which may or may not have been selected prior to a whole-brain analysis.

In an attempt to replicate these observations in children, we adopted the exact same ROI-based approach, relying on the same segmentation software and on the same brain atlas as Frye and collaborators. Although including more than twice as many subjects as Frye et al., none of their findings on developmental dyslexia or on inter-hemispheric differences were replicated (Altarelli, 2013; Altarelli et al., submitted).

We also went one step further and ran whole-brain vertex-by-vertex analyses of cortical thickness and surface area in two independent

datasets comprising age-matched and sex-matched dyslexic and control children (87 participants in total). Analyses of the first dataset uncovered greater cortical thickness in a left superior frontal cluster in dyslexic children. They also revealed reduced surface area in a precentral region in dyslexic boys compared to control boys, while the opposite pattern was reported for girls. Both results were significant using a Monte-Carlo cluster-wise correction, but not an FDR correction. The second dataset yielded only partially matching findings, namely a marginally significant difference in cortical thickness in the same left superior frontal cluster as in the first cohort. Overall, these results suggest that reproducibility of observations should not be taken for granted in morphometric studies of developmental dyslexia. Quite typically, these null results did not find their way into the published paper, which focused on a functionally-defined region of interest (Altarelli et al., 2013).

Ma and colleagues (Ma et al., 2015) also performed vertex-wise analyses within a pre-established cortical mask spanning temporo-parietal and occipito-temporal regions, in a cohort of dyslexic and control children (total N = 64). They found greater cortical thickness in dyslexic children in the left fusiform gyrus and in the right temporo-parietal region, again with no overlap with the aforementioned work of Frye et al. (2010), and with that of Altarelli et al. (submitted).

It should be noted that, while inter-brain registration is considered to be more successful via a spherical template compared to volume-based normalisation (Fischl et al., 1999; Ghosh et al., 2010), little is known as regards the alignment and matching of functionally relevant regions across participants. In fact, the location of functionally defined regions is known to vary from one subject to another, despite accurate macro-anatomical alignment (Frost and Goebel, 2012). Building on these observations, Altarelli and colleagues (2013) focused on functionally defined regions of interest in the ventral occipito-temporal cortex. Using fMRI, they localised the individual peaks of response to houses, faces and words and investigated cortical thickness differences between dyslexic and control children around those peaks. They found a specific reduction in cortical thickness around the peak of response to words, in the left hemisphere, in dyslexic girls compared to control girls. This result was found to be consistent in a second independent age-matched dataset, as well as in a comparison of dyslexic children with reading-matched controls.

In summary, surface-based morphometry is a technique that is promising, as it has many advantages over voxel-based morphometry, but it has been seldom used in dyslexia research so far, and at present there is not a single result that has been replicated across independent studies.

5. Morphometry of the planum temporale

Given the aforementioned challenges of automated methods, manual delineation of regions of interest remains a valuable strategy, which does not require inter-subject alignment and registration. It is also particularly useful considering the complexity and inter-individual variability inherent to some key anatomical areas. An example in developmental dyslexia is the Sylvian fissure, and more precisely the planum temporale (PT), an associative auditory area located on the upper surface of the superior temporal gyrus, posterior to Heschl's gyrus. The surface area of this region is larger in the left than in the right hemisphere in about 65% of the general population (Geschwind and Levitsky, 1968). The idea that the asymmetry of planum temporale was altered in developmental dyslexia emerged as early as in the 1980s, following post mortem observations on 7 dyslexic brains (Galaburda et al., 1985; Humphreys et al., 1990 Humphreys et al., 1990). Many subsequent (in vivo) MRI studies have however failed to provide clear evidence for or against PT anomalies in dyslexia (Best and Demb, 1999; Bloom et al., 2013; Heiervang et al., 2000; Larsen et al., 1990; Robichon et al., 2000b; Semrud-Clikeman et al., 1996). Among the 20 MRI studies on the topic published thus far, only five reported symmetry or rightward asymmetry of the PT in developmental dyslexia (Altarelli et al., 2014; Bloom et al., 2013; Hynd et al., 1990; Larsen et al., 1990; Semrud-Clikeman et al., 1996).

Several factors may explain these inconsistencies. First of all, the anatomical criteria applied to delineate the planum temporale region are of crucial importance, yet they differ greatly between MRI studies of developmental dyslexia. Indeed, we have verified that the criteria used for both the anterior and the posterior anatomical boundaries of the planum have an impact on the estimated asymmetry index, and influence the observed group differences (see below; Altarelli et al., 2014).

A second methodological factor of importance concerns the reconstruction techniques applied to the MRI images and the measurements used. As an example, some studies compared the length of the PT between dyslexic and control participants as measured from a limited number of sagittal slices (Kibby et al., 2004; Leonard et al., 1993; Semrud-Clikeman et al., 1991), while others analysed its surface area, sometimes taking convolutions into account (Hugdahl et al., 1998; Schultz et al., 1994). Underestimations of the PT can occur when considering its length on a subset of sagittal sections, as highlighted by work on the topic (Kulynych et al., 1993; Loftus et al., 1993), probably because of underestimations of the full lateral extent of Heschl's gyrus and sulcus and of the planum.

Finally, a third relevant factor regards the limited number of participants typically studied (15 per group on average), with heterogeneous criteria for inclusion in the dyslexic group (concerning, for instance, the presence or absence of comorbidities, the evaluation of IQ, the types of reading tests used and cut-off criteria applied). In addition, groups have not always been matched for relevant variables such as age, non-verbal IQ, and handedness.

Taking all of the above into account, in their recent study on the topic, Altarelli et al. (2014) manually delineated and measured PT surface area in 81 children (a population more than twice as large as the one typically studied in previous reports, with groups carefully matched for age, sex, non-verbal IQ and handedness), comparing different anatomical criteria. Closely following Galaburda's (1985) delineation criteria, they provided partial confirmation of Galaburda et al.'s initial hypothesis, in the sense that they found that dyslexic boys tended to have rightward asymmetrical PT surface, when control boys showed the expected leftward asymmetry. Furthermore, Altarelli et al. found that the set of anatomical criteria used is crucial and is likely to explain at least some of the previous discordant findings. In brief, they found that a group difference in PT asymmetry is observed only when the anterior boundary is defined by Heschl's sulcus, and when the posterior boundary is defined by a change in slope or orientation of the superior temporal plane, thereby excluding any posterior ramus of the sylvian fissure. Posterior duplications of Heschl's gyrus on the right hemisphere, which were more frequent in dyslexic boys, played an important role in the observed group differences. Discrepancies in the application of such anatomical criteria may explain most of the inconsistencies between Altarelli et al.'s results and previous studies restricted to dyslexic and control males (Green et al., 1999; Heiervang et al., 2000; Robichon et al., 2000a; Rumsey et al., 1997).

Moreover, another key factor to be taken into consideration is sex, as it was found to interact with diagnosis of dyslexia, such that PT surface area asymmetry was deviant only in dyslexic boys, but not in dyslexic girls (which may also partly explain previous discordant findings) (Altarelli et al., 2014). More recently, a new study using the very same anatomical criteria reported a deviant PT asymmetry pattern in children at family risk of dyslexia

(Vanderauwera et al., in press). No group-sex interaction was reported in that study, but statistical power may have been too limited for that test.[4]

Thus, a deviant asymmetry pattern of the planum temporale may be a reliable finding after all, at least in some dyslexic boys. Further replications of this result, using the very same anatomical criteria, would be desirable. At any rate, it is also worth underlining that planum temporale asymmetry is not an all-or-none marker that can differentiate all dyslexic boys from controls. In our own data (Altarelli et al., 2014), there is considerable overlap of the PT asymmetry index between the two groups, and the effect size of the group difference (in boys) is 0.78. PT asymmetry is therefore best viewed as one neural marker among many others that may increase the risk of dyslexia.

Functional interpretations of the potential impact of a deviant PT asymmetry remain uncertain, especially given controversies surrounding language lateralisation (Bishop, 2013). In this regard, we suggest that it may prove fruitful to connect neuroanatomical studies with those on auditory cortical oscillations, which attribute distinct roles for speech processing to left and right auditory cortices (including the PT) (Giraud and Poeppel, 2012; Goswami, 2011; Poeppel, 2003) and suggest differences in some oscillatory responses (and in their asymmetry) in dyslexic individuals (Hamalainen et al., 2012; Lehongre et al., 2013, 2011; Lizarazu et al., 2015; Poelmans et al., 2012).

6. White matter and diffusion imaging

Diffusion imaging is a non-invasive neuroimaging technique that measures neural tissue microstructure from the diffusion-weighted magnetic resonance (MR) signal. It relies on a quantification of the preferential directions of diffusion of water molecules in brain tissue. The most frequently used method used to process diffusion images is the diffusion tensor imaging model (DTI), which leads to the computation of fractional anisotropy (FA), an index of the extent to which water molecules' motion is constrained in one particular direction. Together with tract-reconstruction algorithms, this can be used to provide an estimation of the structural connectivity between brain regions.

Like in voxel-based morphometry (VBM), it is possible to carry out whole-brain, voxel-based analyses (VBA) of the diffusion tensor. This approach has the same advantages and limitations as typical VBM analysis. Advantages include that 1) it is a hypothesis-free whole-brain exploration; and 2) the analysis is simple to apply and fully automated. Limitations include 1) the difficulty of unambiguously associating statistical differences in voxel clusters to a clearly identified white matter tract; 2) issues of alignment inaccuracies across subjects, such that certain voxels may actually not correspond to the same tract, or even to white matter in all subjects (Simon et al., 2005); 3) in some voxels close to grey matter, estimated differences in FA may rather reflect the relative proportion of different tissue types; 4) the large number of voxels tested requires stringent correction for multiple comparisons that are seldom fully applied.

In order to partly circumvent these problems, "skeleton-based" approaches focus on a more limited set of white matter voxels (the skeleton) at the intersection of white matter volumes across all participants of a study. The most influential of these approaches is "tract-based spatial statistics" (TBSS) (Smith et al., 2007, 2006). A number of studies have used TBSS to compare local FA between dyslexic and control participants (Carter et al., 2009; Deutsch et al., 2005; Keller and Just, 2009; Klingberg et al., 2000; Odegard et al., 2009; Richards et al., 2008; Rimrodt et al., 2010; Steinbrink et al., 2008). According to a meta-analysis by Vandermosten et al. (2012b), these studies converge in suggesting decreased FA in dyslexic participants around the left temporo-parietal junction. These results are also partly consistent with studies showing correlations between FA in similar regions and reading or phonological ability in normal-reading populations (e.g., Beaulieu et al., 2005; Dougherty et al., 2007; Feldman et al., 2012; Nagy et al., 2004; Qiu et al., 2008; Steinbrink et al., 2008; Yeatman et al., 2011). Yet, in unpublished analyses of our own data in 61 participants, TBSS has not revealed any reliable differences between dyslexic and control groups. It is not impossible that more null results remain unpublished.

An alternative to VBA is to use FA maps and tractography algorithms to try and reconstruct specific white matter pathways in each subject's native space, thus circumventing the problem of aligning the brains in a common space. Based on VBA studies and on a priori hypotheses, the arcuate fasciculus (AF) is certainly the white matter tract that has attracted the greatest interest in dyslexia research (Carter et al., 2009; Hoeft et al., 2011; Niogi and McCandliss, 2006; Vanderauwera et al., 2015; Vandermosten et al., 2012a; Zhao et al., 2016). Group differences between dyslexic and control participants have most often been reported in the long segment of the left AF, linking Broca's and Wernicke's regions (Vandermosten et al., 2012b).

However, use of the DTI model for tract reconstruction has limitations. Tract reconstruction does not always succeed in all participants. This has proven a particular problem for the right arcuate fasciculus, whose reconstruction fails in a substantial number of participants (Catani et al., 2007; Vandermosten et al., 2012a; Yeatman et al., 2011; Zhao et al., 2016), even though there is no doubt that they all have this tract (indeed, it can be reconstructed in those same participants using spherical deconvolution: Zhao et al., 2016). One reason may be that the AF crosses other prominent fibre tracts (corona radiata and corpus callosum), and that the DTI model does not handle fibre-crossing regions well, which is a more general problem. Indeed, one key limitation of the standard diffusion tensor model is that it can only recover a single fibre orientation in each voxel.

These limitations of the DTI model have motivated the development of a variety of techniques for resolving multiple fibre orientations and capturing complex tract configurations. One such method is spherical deconvolution (SD), which relies on advanced diffusion sequences with high angular diffusion imaging (requiring a diffusion weighting $b > 1300$ s/mm^2). SD tractography is based on multi-tensor models, thus can estimate fibre orientations in regions with fibres from multiple directions (Anderson and Ding, 2002; Tournier et al., 2004). Spherical deconvolution yields anisotropy estimates equivalent to FA, known as hindrance-modulated orientational anisotropy (HMOA; Dell'Acqua et al., 2013). The main difference is that, in any given voxel, there is one HMOA value for each fibre orientation (if more than one), thus providing richer information than a single FA value in fibre-crossing regions (Dell'Acqua et al., 2013; Dubois et al., 2014). A comparison of whole-brain tractography using DTI vs. SD in the same brain is shown in Fig. 2. The difference in the density of the reconstructed fibres is so striking that it really begs the question to what extent the DTI model can be relied upon for meaningful tractography analyses.

Two recent studies have applied both DTI and SD tractography to comparisons between dyslexic and control (or with and without family risk) children (Vanderauwera et al., 2015; Zhao et al., 2016). Both have shown the superiority of SD over DTI, in particular concerning the reconstruction of the arcuate fasciculus, and the ability to distinguish it from both neighbouring and intersecting fibre tracts. Disturbingly, Vanderauwera et al. (2015) did not find any difference in the FA of the left AF between children with and without family risk of dyslexia, whether using DTI or SD (but they found correlations between FA and phonological awareness). We partly replicated the observation of FA differences between control and dyslexic children in the left AF when

[4] In our analysis of Altarelli et al. (2014), we had 81 participants, a 2 × 2 × 2 design with 2 covariates, and a partial eta square = 0.061 (so f = 0.255) for the group x sex interaction. Using GPower (Faul et al., 2007), we calculate that the observed power for this interaction was 0.62. In Vanderauwera et al. (in press), they had 54 pre-reader participants, a 2 × 2 design with 2 covariates. Assuming the same effect size, their power for the group x sex interaction was 0.45. But the expected effect size is probably lower, given that no more than half of the pre-readers should grow up to be dyslexic. For the analysis on 28 adolescents, power was 0.25.

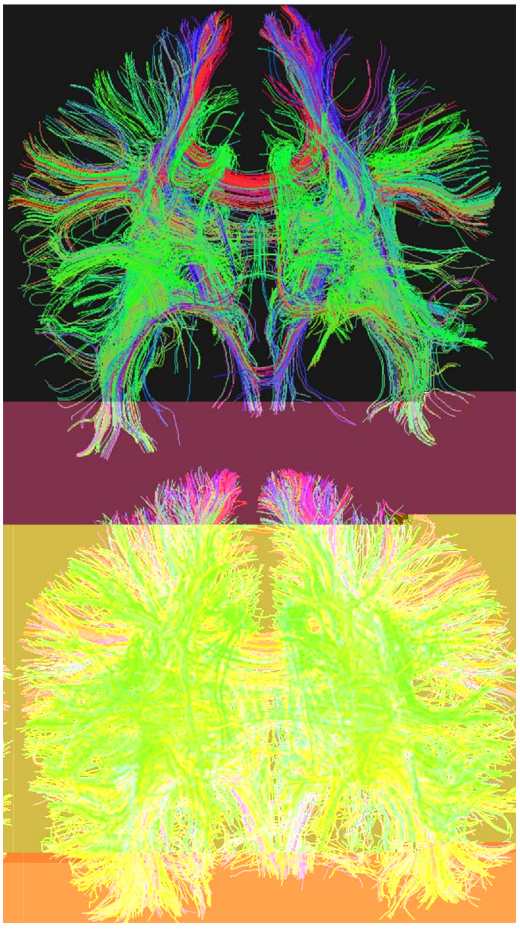

Fig. 2. Examples of whole brain tractography of the same brain, using two different methods. Upper panel: diffusion tensor imaging. Lower panel: spherical deconvolution. Sequence parameters: 32 channels, 1.7 × 1.7 × 1.7 mm resolution, 60 directions, b = 1400 s/mm2, repetition time = 14,000 msec, echo time = 91 msec, total acquisition time = 18 min in three bouts. Data from Zhao et al. (2016).

using the DTI model (across the three segments of the left arcuate fasciculus, rather than specifically the long segment), but this result was not confirmed when using SD. We also observed differences in the asymmetry of the anterior and posterior segments of the AF, and of the neighbouring superior longitudinal fasciculus (SLF II: middle segment), again using the DTI but not the SD model (Zhao et al., 2016). The known limitations of the DTI model suggest at least two possible interpretations for the divergence between results obtained using DTI and SD. Firstly, it may be that group differences attributed to the AF in DTI analyses actually reflect differences in fibres that intersect the AF (such as the corpus callosum and the corona radiata, Ben-Shachar et al., 2007). Secondly, it may be that group differences attributed to the AF in DTI analyses actually reflect differences in neighbouring tracts such as the superior longitudinal fasciculus, which is not distinguished from the AF in DTI studies.[5] More studies using both DTI and spherical deconvolution tractography will be needed to fully resolve this point.

In recent studies, a new trend has emerged to analyse FA, not only averaged across an entire tract (as reported above), but all along the tract length (Colby et al., 2012; Langer et al., 2015; Wang et al., 2016; Yeatman et al., 2012). This approach has allowed investigators to report group differences in FA in sometimes very limited portions of a given tract, even when mean FA over the entire tract did not significantly differ between groups (such as in Wang et al., 2016). There may be good motivations for that, such as focusing the analysis of a tract on portions that are less likely to be affected by crossing fibres (Yeatman et al., 2011), or using variations in FA along a tract as proxies of variations in tract shape/curvature or as reflections of local disruptions. Yet, one question raised by this method is whether it is legitimate to rescue what are primarily null results, by carrying out a larger number of additional comparisons, thus aggravating the risk of reporting false positive results. Standard statistical procedure would suggest that the prior finding of a significant group x location interaction would be necessary to authorise comparing groups at every location. And such comparisons should be stringently corrected for multiple tests. The second issue is one of interpretation: if FA group differences are reliably localised in a very specific portion of a white matter tract, do they reflect differences in the structure of this particular tract, or are they more likely to reflect properties of another tract that crosses the target one at this particular location? Because this approach is a relatively new development, very few studies have used this method, so it remains to be seen to what extent these results will be replicated. In the meantime, we suggest that analyses of tract profiles should be used and interpreted with caution, either focusing on a priori portions of interest selected using anatomical knowledge (e.g., Yeatman et al., 2012), or followed up by an analysis of intersecting fibre tracts.

Overall, our review of the diffusion imaging literature suggests that this technique is still in its infancy, in the sense that acquisition and analysis methods are constantly evolving, with no common standard adopted yet. As a result, different studies typically differ substantially in sequence parameters, in the dependent measure (FA or HMOA), in whether tracts are reconstructed, in tract reconstruction algorithms (when applicable), in the specific fibre bundle that is designated by a certain tract name, and in statistical procedure. Thus at the moment rigorous meta-analyses integrating sufficiently comparable data are not possible and it is too early to say whether any result reported in this literature will stand the test of time.

7. Gyrification patterns

Beyond total brain volume, brain gyrification is the most visible source of anatomical differences between individuals. Furthermore it occurs during early brain development and is pretty stable after birth (Cachia et al., 2016; Mangin et al., 2004; White et al., 2010), making it a prime target for the study of early markers of cognitive function and impairment. Finally, sulci are reliable (although certainly imperfect) landmarks commonly used to identify cytoarchitectonic brain regions, in the absence of direct histological evidence. Despite their great potential relevance, until recently only a few case studies had investigated sulci configuration in relation to dyslexia (Chiarello et al., 2006; Craggs et al., 2006; Leonard et al., 1993), while this approach has been more popular for schizophrenia (Cachia et al., 2008; Gay et al., 2013; Plaze et al., 2009).

More recently, Im et al. (2015) have used a complex graph-theoretical analysis to investigate global sulcal patterns across the temporo-parietal and occipito-temporal regions, in a population of fifty-nine children (28 readers, with about half dyslexic, and 31 pre- or beginning readers, with about half with a family history of dyslexia). They reported atypical sulcal patterns both in the dyslexic and in the at-risk pre-reader groups, characterised by more sulcal basins of smaller size than in the control groups. These results were correlated with reading measures, suggesting a link between neuroanatomy and brain function.

We adopted an entirely different approach, focusing on quantitative and qualitative properties of three semi-manually labelled sulci of interest: the central sulcus, the sylvian fissure and the superior temporal sulcus, in a large population of 142 dyslexic and 106 control children across 3 countries (Scotto di Covella et al., submitted). One result stood out concerning the mean depth of the central sulcus, in the form of a

[5] According to the criteria defined by Catani et al. (2005) and Thiebaut de Schotten et al. (2011), the anterior (fronto-parietal) segment of the AF represents a sub-portion of the 3rd segment of the SLF (SLF III), restricted to language areas. All the other segments are entirely distinct.

triple interaction between hemisphere, sex and group. We found the central sulcus to be deeper in the right hemisphere in control girls, but symmetrical in control boys. However, it was deeper in the left hemisphere in dyslexic girls and deeper in the right in dyslexic boys. No other difference was found in any of these sulci.

Finally, one additional relevant study is worth mentioning, reporting that normal-reading children with a discontinuous left occipitotemporal sulcus show better word reading skills than those with a continuous one (Borst et al., 2016). However, the sample size (10 and 6 respectively) is such that replication is essential. Furthermore, it remains to be assessed whether the frequency of the two configurations differs between groups of diagnosed dyslexic and control children.

Each of these studies is quite unique and would deserve to be replicated before conclusions are drawn. Furthermore, the functional interpretation of the observed gyrification differences remains entirely open. Nonetheless, we think that the study of cortical sulci is a promising research avenue that would deserve much more attention than it has received until now.

8. Metabolites in cerebral tissue

The concentration of a variety of neurometabolites, like N-acetylaspartate (NAA), choline, creatine, GABA, glutamate as well as phospholipids can be evaluated in vivo via the application of Magnetic Resonance Spectroscopy (MRS). The comparison of dyslexic and control populations has revealed some differences in these measures. Richardson et al. (1997) reported anomalies in the cerebral concentration of phosphomonoesters, potentially suggesting an abnormal incorporation of phospholipids into cellular membranes.

Localized alterations have also been reported. In an early study by Rae et al. (1998), lower choline/NAA ratios were observed in the left temporo-parietal region and in the right cerebellum of dyslexic individuals as compared to controls, while no differences were seen in the frontal lobes. These observations were interpreted as emerging from decreased choline levels in these brain regions. Yet others (Laycock et al., 2008) later reported higher Cho/NAA ratio in the right cerebellum of dyslexic participants, potentially related to higher choline levels in this region. Finally, two more recent studies have examined neurometabolites concentration in adults (Bruno et al., 2013) and children (Pugh et al., 2014) with a wide range of reading skills, reporting negative correlations between phonological skills and choline concentration levels in a left temporo-parietal region, and negative correlations between phonological skills and choline concentration levels as well as glutamate levels in an occipital medial region, respectively.

At present, no straightforward interpretation of these metabolic results is available. Alterations in neurometabolite levels may reveal underlying cytoarchitectonic differences, as for instance choline concentration could be related to the relative presence of large neurons in the cerebral area under study, as hypothesised by Rae et al. (1998). Another possibility, put forward by other authors, is that abnormal myelination may influence measured concentrations of choline (Laycock et al., 2008; Pugh et al., 2014). Finally, it has been hypothesised that elevated glutamate levels may be related to enhanced neuronal excitability (Pugh et al., 2014).

Generally speaking, the possibility of assessing in vivo the concentration of neurometabolites in specific brain regions seems extremely promising. However, at the moment MRS suffers from a number of technical limitations (poor spatial resolution, difficulty to distinguish certain molecules, long acquisition times) that will need to be overcome before it can decisively inform the neurobiology of dyslexia and connect with other pieces of the puzzle.

9. Multivariate approaches

To date, most neuroanatomical studies of dyslexia have focused on between-group comparisons in which different neuroanatomical measures were generally explored in isolation. However, such an approach disregards potential interactions between regions and measures. Multivariate classification techniques offer a method for decoding group membership based on a multiplicity of anatomical measures. These techniques have only been recently applied to the study of neuroanatomical markers of developmental dyslexia (Cui et al., 2016; Płoński et al., 2017; Tamboer et al., 2016), although they have been utilized in the past in other developmental disorders like autism (Ecker et al., 2010) and ADHD (Peng et al., 2013).

Leonard and colleagues pioneered multivariate approaches by manually developing an anatomical risk index integrating the asymmetry, size and surface area of several brain structures, including the planum temporale, the anterior lobe of the cerebellum and Heschl's gyri. This risk index was adjusted in such a way as to tease apart dyslexic readers, normal readers, and poor readers with low verbal ability (Leonard et al., 2002a,b). This finding was then replicated in a subsequent study: children with positive risk indices (N = 8, with larger, asymmetrical brain structures) had reading impairment, whereas children with negative risk indices (N = 14, with relatively smaller and symmetrical brain structures) had severe comprehension impairments more typical of specific language impairment (Leonard et al., 2006).

Next, Pernet et al. (2009b) used confidence intervals (CI) to estimate group differences on a larger sample of control (N = 39) and dyslexic adults (N = 38). Specifically, control subjects' brains were used to build 95% CI using a bootstrap procedure, and subsequently each dyslexic brain was classified voxelwise as being within or outside the 95% CI. The internal sensitivity and specificity of the classification results were tested using a 3-fold cross-validation. Two areas for which 100% of dyslexic subjects were out of the normal range were 2 small clusters (6 and 7 voxels) in the right cerebellar declive and the right lentiform nucleus with 98% sensitivity and 85% specificity.

Recently, three studies applied multivariate machine learning techniques to classify dyslexic vs. control subjects based on various sets of neuroanatomical measures: 5 white matter indices (Cui et al., 2016), grey matter volume (Tamboer et al., 2016) and 5 cortical indices: volume, surface area, thickness, curvature and folding (Płoński et al., 2017). The first two studies were performed on groups of 20–30 participants per group (children and adults respectively), using linear support vector machine (LSVM) classifier and leave-one-out cross validation (LOOCV). The classifier based on white matter measures achieved high accuracy (83.6%) with many different features distinguishing between dyslexic and control children, such as the superior longitudinal fasciculus, inferior fronto-occipital fasciculus, thalamocortical projections, corpus callosum, cingulum, fornix, the cerebellar peduncle, corona radiata, and corticospinal tract (Cui et al., 2016). The classifier based on GMV achieved 80% accuracy and the voxels discriminating between dyslexic and control students were situated in the left occipital fusiform gyrus (LOFG), in the right occipital fusiform gyrus (ROFG), and in the left inferior parietal lobule (LIPL). However, when this classifier was tested on an large independent sample of young adults, the classification performance dropped to 59%, since a large percentage of false alarms was found. Finally, we tested three different feature selection and classification algorithms on a big multisite sample of children (106 controls and 130 dyslexics) while the generalizability of classification was assessed with both 10-fold and leave-one-out cross validation (Płoński et al., 2017a). Classification achieved moderate, but above chance accuracy of 65%, and features that discriminated between dyslexic and control children were exclusively situated in the left hemisphere, including superior and middle temporal gyri, sub-parietal sulcus and prefrontal areas. They-were related to geometric properties of the cortex, with generally higher mean curvature and a greater folding index characterizing the dyslexic group.

The large discrepancy between the classification accuracies obtained by the three studies suggests that liberal LOOCV together with relatively small groups may have inflated decoding accuracies. Indeed,

in a very large-scale analysis of autistic and control brains, Haar et al. (2014) failed to obtain above chance accuracy of classification, and at the same time showed that smaller group sizes (N around 20) yield wider decoding accuracy distributions, regardless of the anatomical measures used to perform the classification. Since there are usually multiple feature sets that can be chosen, the use of small groups and the choice of an arbitrary feature set holds the potential for greatly inflating decoding accuracies (30% of analyses on small samples will spuriously yield greater than 60% classification accuracy). As for VBM, the greatest classification rates are generally obtained in small-scale studies, while large-scale studies yield more modest ones (Woo et al., 2017).

Thus, multivariate analysis techniques are a promising way to go beyond single brain features, obtain a more global view of the dyslexic brain, and optimise classification algorithms. However, they should be used with caution, given that the richer the MRI data and the smaller the sample sizes, the greater the risk of overfitting the data with random combinations of features, and the greater the risk of spuriously inflating classification rates.

10. Discussion and perspectives

As we have illustrated in this review, the study of the neuroanatomical bases of dyslexia using brain imaging is rife with difficulties. In this discussion, we will reflect on the following issues: the heterogeneity of dyslexia, sex differences, effect sizes and sample sizes, causes and consequences, and manual vs. automated image processing.

10.1. Heterogeneity of dyslexia: subtypes, cultures and languages

It is commonplace to state that developmental dyslexia is a heterogeneous disorder, and that there are several subtypes of dyslexia. If this is true, then it is of utmost importance to take those subtypes into account in any investigation of the neuroanatomical bases of dyslexia. Indeed, different cognitive deficits imply different neural bases. Furthermore, the existence of distinct subtypes also has (potentially worrying) implications on the size of groups to be compared and thus on the statistical power of studies.

Yet, subtypes have typically not been taken into account (with a few exceptions, such as Jednoróg et al., 2014; Leonard et al., 2006). One obvious reason is that there is still no widespread agreement with respect to subtypes of dyslexia. There has been Boder's (1973) classical distinction between dysphonetic and dyseidetic dyslexias, Castles and Coltheart's (1993) phonological vs. surface dyslexias, Leonard et al.'s (2002a) distinction between cases with and without an oral language deficit, Bosse et al's (2007) distinction between phonological and visual attention span deficits, as well as more bottom-up approaches based on cluster analyses (Jednoróg et al., 2014; Morris et al., 1998). It is fair to say that none of these typologies has been sufficiently broadly accepted to provide general guidance for the study of the neuroanatomy of dyslexia.

Another, more optimistic outlook is that, compared with other disorders (such as autism or schizophrenia), dyslexia is relatively homogeneous after all. One reason is that the central diagnostic phenotype (reading) is defined pretty narrowly (as opposed to, say, social functioning). Another reason is that it turns out that, at the cognitive level, a majority of dyslexic individuals seem to share a similar kind of deficit, that is, in the phonological domain (even though profiles of phonological impairment may of course vary) (Ramus, 2003; Saksida et al., 2016; White et al., 2006). Thus, while there remain debates about the aetiology of the phonological deficit, only a fraction of the population with dyslexia (who do not have a phonological deficit) may contribute substantial heterogeneity at the cognitive level. This may concern individuals with a primary visual, or visual-attentional, rather than phonological impairment. Within this non-phonological subtype, there is space for several types of deficits, and hypotheses remain largely open (Bosse et al., 2007; Gori and Facoetti, 2015; Saksida et al., 2016; Vidyasagar and Pammer, 2010). Nevertheless, to the extent that only a minority of dyslexic individuals belong to a non-phonological subtype, any study that includes a sufficiently large random sample of dyslexic individuals should produce results that tend to reflect the majority phonological subtype.[6] Minority subtypes will add noise to the data, may diminish statistical power, may slightly skew the results, but should not render the whole enterprise impossible. The observed effects may also sometimes be driven by outliers, whether individuals with atypical anatomy or with an unrecognized syndrome (such as the case of bilateral perisylvian syndrome described by Eckert et al., 2016a). Under that view, only small-scale studies are at significant risk of including a sample whose distribution of subtypes may be strongly distorted, and therefore of producing results that cannot be generalised. Nevertheless, it remains an important goal for the future to achieve a better understanding of the typology of dyslexia subtypes, and consequently to take them into account in neuroimaging studies. More generally, it would seem desirable for any neuroimaging study of dyslexia to explicitly state its underlying assumptions with respect to dyslexia subtypes, and to discuss the consequences on the interpretation of the results.

An additional layer of complexity comes from the fact that dyslexia is expressed to some extent differently in different languages (Hadzibeganovic et al., 2010). There is indeed some evidence that structural brain differences between dyslexic and control groups may differ between alphabetic and morphographic languages (e.g., Liu et al., 2013; Siok et al., 2008; Su et al., submitted). Whether they may also differ to some extent between alphabetic languages of different transparencies remains to be established. Some of these brain differences may of course be the consequences of different cognitive demands and learning paths. But it is not impossible that the set of individuals who become dyslexic in a given language would not be entirely the same if they had to learn to read in a different language. There are suggestions, for instance, that visual deficits, as well as deficits in morphological awareness, may play a greater role in dyslexia in Chinese than in alphabetic languages (e.g., Shu et al., 2006; Siok et al., 2009). This implies that certain types of brain disruptions may be more represented among dyslexic individuals in certain languages than in others. Add to that cultural differences (in particular, the education system and methods to teach reading) that may also have an impact on the manifestations of dyslexia, and this renders cross-linguistic and cross-national comparisons of neuroimaging studies particularly risky, especially since many methodological differences (scanner, sequences, analysis methods…) are typically confounded with linguistic and cultural differences. One caveat, though, is that these considerations should not lead one to dismiss all inconsistencies between studies as a mere consequence of linguistic or cultural differences. Our analysis of VBM studies shows that inconsistencies are as large between studies carried out in English-speaking countries as between different languages. There are good reasons, other than language and culture, for such inconsistencies, which we address thereafter.

10.2. Sex: an important and overlooked source of heterogeneity

Although the uneven sex ratio in dyslexia has been known for a long time, sex has never been a factor of much interest in dyslexia research, whether at the cognitive or at the neural level. Early neuroimaging studies either focused exclusively on males (e.g., Brown et al., 2001; Paulesu et al., 2001), or included too few male and female participants to afford splitting groups by sex. Even larger studies often neglected to

[6] Of course, this majority phonological subtype may itself be characterised by multiple brain differences in various locations, with individual differences in the specific spatial pattern of disruption, much in the same way as different dyslexic individuals show diverse profiles on phonological awareness, verbal short-term memory and rapid automatized naming tasks.

do so or still had too few participants of one sex (e.g., C. R. Pernet et al., 2009a; Tamboer et al., 2015). Yet, the few studies that had enough participants of each sex and that deliberately entered sex as a factor in their analyses have often found that the brain differences between dyslexic and control participants depend on sex. This has been reported in a VBM study (Evans et al., 2014), in our study on the cortical thickness of the visual word-form area (Altarelli et al., 2013), in our study of the planum temporale (Altarelli et al., 2014), as well as in our study of cortical sulci (Scotto di Covella et al., submitted). It is also worth recalling that Galaburda et al's initial post-mortem observations also differed to some extent between male and female brains (Galaburda et al., 1985; Humphreys et al., 1990 Humphreys et al., 1990).

Thus, there are indications that the neural basis of developmental dyslexia may to some extent differ between males and females. The reasons are not fully known, although a role for foetal sex hormones may seem plausible, as hypothesised by Geschwind and Galaburda (1985). According to one scenario, female hormones would play a protective role with respect to various disruptions of brain development. Thus females may require more severe brain disruptions to show an equal level of cognitive disruption as males (Ramus, 2006), a prediction consistent with animal studies (Fitch et al., 1997; Hall et al., 1991; Roof et al., 1994), with post-mortem work on dyslexia (Galaburda et al., 1985; Humphreys et al., 1990 Humphreys et al., 1990), and with the study showing a group difference in the visual word form area exclusively in females (Altarelli et al., 2013). Consistent with the idea of greater female resilience to brain disruptions, there is also evidence that boys' reading ability suffers more from early left hemisphere damage than that of girls (Frith and Vargha-Khadem, 2001), which may suggest that boys are more strongly left-lateralized for language and may therefore show more limited plasticity following unilateral brain disruption (Chen et al., 2007; Spironelli et al., 2010). However, some neuroimaging studies also reported certain group differences in males only (Altarelli et al., 2014), or different patterns in males and females (Evans et al., 2014; Scotto di Covella et al., submitted), thus suggesting that the picture is more complex than a simple severity gradient, and that the neural basis of dyslexia may partly differ between males and females.

The study of sex differences in dyslexia is still in its infancy. However, it deserves more attention than it has received so far, and it should be investigated more systematically in all dyslexia studies that have enough participants of each sex. Indeed, disruptions of brain development operate against a genetic and hormonal background that differs between the sexes, and that sometimes lead to basic sex differences in the neuroanatomical properties of interest. For instance, in our studies on the neuroanatomy of dyslexia, we have serendipitously discovered sex differences in the asymmetry of Heschl's gyrus surface (Altarelli et al., 2014), and of central sulcus depth (Scotto di Covella et al., submitted). More generally, the importance of paying attention to sex is not restricted to dyslexia research, but has recently been recognised for all of neuroscience (Cahill, 2006) and medical research (Clayton and Collins, 2014).

10.3. What effect sizes should we expect and what are the consequences?

In all discussions of the replication crisis (Button et al., 2013; Ioannidis, 2011; Szucs and Ioannidis, 2017), the notion of effect size is central. The smaller an effect one seeks to detect, the larger the sample size needed to reliably detect it (if true). What effect size can reasonably be expected to be found between the dyslexic and the control brain? It may be worth noting that, compared to almost all other neurodevelopmental or neurodegenerative disorders, developmental dyslexia is a very, very mild one, characterised by primary symptoms in a very narrowly defined cognitive domain, and having an impact on everyday functioning only in proportion to the social importance accorded to reading in certain societies at a certain time in history. Whereas it may be plausible to expect major brain disruptions in autism, schizophrenia, or Alzheimer's disease, this is certainly not the case for dyslexia (although even for autism, major brain disruptions are not confirmed; see Haar et al., 2014). Whatever brain difference there is to be found between dyslexic and control participants has to be small, otherwise dyslexic individuals would not only be dyslexic but would also show much more severe problems in various domains. This consideration applies even more to most neuroimaging studies of dyslexia, which, rather than selecting particularly severe cases of dyslexia in order to increase effect size, use relatively liberal inclusion criteria, such as 1.25 or even 1 standard deviation below the norm on one reading test (sometimes out of 2 or 3 taken), thereby mechanically diminishing the expected effect size. Furthermore, it is likely that there are multiple neural risk factors for dyslexia, and unlikely that any single of them might play a role in all cases of dyslexia.

Unfortunately, neuroimaging studies of dyslexia almost never report the effect sizes of group differences. They also seldom report the values and standard deviations of the measures that are claimed to differ (grey or white matter volume, cortical thickness or surface, white matter anisotropy). Meta-analyses have to infer effect sizes from t values and

studies, reporting moderate effect sizes, unless they have taken very special steps to reduce heterogeneity and/or to increase severity. Of course, it is also important to acknowledge and take into account the challenges that arise with large-scale studies: pooling data across multiple sites with different sampling strategies and technical characteristics, as well as broadening recruitment efforts, are all likely to increase population heterogeneity, and therefore increase the risk of false negative findings. This certainly does not mean that large-scale studies are hopeless or less reliable than small-scale ones: rather, this implies that their statistical power does not rise as steeply as the raw numbers would suggest, and that all steps that can be taken to reduce heterogeneity across sites need to be considered.

10.4. Disentangling causes and consequences

A major issue in all of dyslexia research is, whenever a group difference is found, to what extent it reflects an actual cause of dyslexia, and to what extent it may merely reflect a consequence of less exposure to writing, of a deviant learning path, or of compensatory mechanisms. This issue probably applies as much to neural as to cognitive group differences, given that most brain measures that are studied in dyslexia research (such as grey and white matter volumes, cortical thickness and surface, white matter anisotropy) have been shown to change with learning and experience (Draganski et al., 2004; Engvig et al., 2010; Haier et al., 2009; Scholz et al., 2009).

In the context of dyslexia and more generally of developmental research, several approaches have been used to overcome this problem. One such approach is to match groups for the level of performance that may be the cause of the group differences of interest. For instance, showing that dyslexic children had weak phonological skills, even compared to control children matched in reading ability (therefore younger), was a particularly compelling piece of evidence in favour of the hypothesis that the phonological deficit is a cause, and not only a consequence, of reading disability (e.g., Snowling et al., 1986). Most neuroimaging studies have studied groups matched in chronological age. Nevertheless, a few have also used reading-matched control groups, suggesting that some neuroanatomical differences in dyslexia may not be entirely due to altered reading experience (Altarelli et al., 2013; Hoeft et al., 2007; Krafnick et al., 2014; Xia et al., 2016). However, a reading-matched control group is not a definitive proof of a brain-to-reading causation. Firstly, groups may be matched on one reading measure, but not on others; Secondly, reading performance may be equated, but reading experience is not: dyslexic readers typically have much more reading experience than younger children with the same reading ability, and their reading experience may be qualitatively different too; Finally, the fact that reading-matched control children are younger (typically at least 2 years younger) is a particular problem for neuroimaging studies, in which the brain measures of interest (volume, surface, thickness…) change significantly and non-linearly with age during childhood (and sometimes depending on sex), such that group differences may to some extent be confounded with growth effects.

The alternative approach, to identify actual causes of dyslexia, is of course to acquire MRI images before the onset of reading acquisition. Because only 5% of children grow up to be dyslexic, if dyslexia is the outcome of interest, then the initial population must be enriched with children at risk of becoming dyslexic, which is usually achieved by over-recruiting children with dyslexic parents or siblings, or children with an early language delay. Even before reading acquisition, it is then possible to conduct comparisons between children at family risk and a control group without risk. A number of such studies have been conducted and have suggested neuroanatomical differences (Black et al., 2012; Hosseini et al., 2013; Im et al., 2015; Kraft et al., 2015; Langer et al., 2015; Raschle et al., 2017, 2011; Vandermosten et al., 2015; Wang et al., 2016). One drawback of such comparisons is that the at-risk group only includes about 40–50% of children who will really end up dyslexic. This is bound to dilute the effects of any predictor of dyslexia and further decrease statistical power, in studies that often already have a limited sample size. Thus such group comparisons may reflect, if anything, neural correlates of the family risk itself more than of dyslexia. Indeed one study suggested that group differences may even reflect the parent conferring the dyslexia risk (Black et al., 2012).

The real gold standard, and indeed the holy grail of research on the neural basis of dyslexia, is to continue such studies of children at risk of dyslexia longitudinally, until those children learn to read, and eventually obtain (or not) a dyslexia diagnosis. Only then is it possible to directly test the hypothesis that structural brain properties may predict dyslexia ahead of time, and may therefore play a genuine causal role in the aetiology of dyslexia. As far as we are aware, only one study published so far has achieved that aim (Kraft et al., 2016), and another one partly so (Clark et al., 2014), scanning the children just slightly too late (grade 1) to entirely escape the cause-consequence problem. Nevertheless, several other studies of at-risk children cited above should reach that aim soon. One study also used this longitudinal approach to identify neural predictors of reading ability in normal-reading populations, but then was not in a position to identify predictors of dyslexia per se (Myers et al., 2014).

It must be emphasised that these holy-grail longitudinal studies are particularly difficult to carry out. On top of the usual challenges of MRI research on children (e.g., recruitment, compliance, head movement), they face specific challenges including: 1) the children must be young enough not to have learned to read (at most 6 years old in most countries); 2) a substantial group of children must carry a dyslexia risk, thereby severely constraining recruitment; 3) they must be followed for several years (at least 2 or 3) in order to receive a reliable diagnosis of dyslexia, thereby facing the risk of a substantial attrition rate. Thus everything conspires to make it difficult to gather a large sample size. Yet, these studies have a particularly acute need of large sample sizes, given an ultimate and largely under-appreciated challenge: that because of the time span between brain measures (time 1) and the dyslexia diagnosis (time 2), they are looking for smaller effect sizes than cross-sectional studies (see Fig. 3). Indeed, even the same measure made twice at two time points can only imperfectly correlate with itself. The most reliable measures such as IQ scores correlate 0.8–0.9 with themselves at several weeks interval (Wechsler, 2005). Reading scores in children correlate 0.6–0.85 at 6 months interval (Caravolas et al., 2013). Such correlations can only decrease with time, as many factors intervene. The best known early cognitive predictor of reading development and dyslexia, phonemic awareness, correlates 0.43 with later word decoding skills, according to a meta-analysis (Melby-Lervåg et al., 2012). In a longitudinal study of 2000 children from USA, Australia,

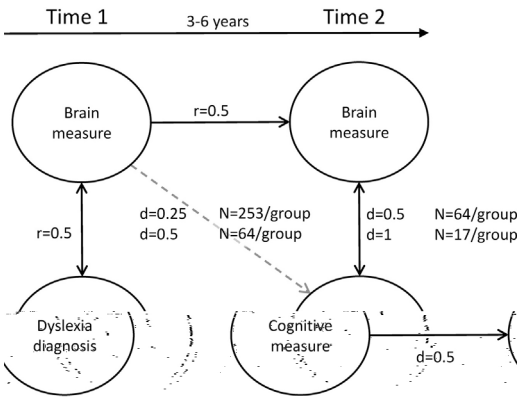

Fig. 3. The challenge of prospective longitudinal studies of the structural brain predictors of dyslexia. Illustration of the highest plausible effect sizes for group comparisons of cognitive and brain measures, and for correlations between cognitive and brain measures, either cross-sectionally or longitudinally (from Time 1 to Time 2). Effect sizes shown are Cohen's d for group comparisons, Pearson's r for correlations. For each value of d, sample sizes required for 80% statistical power are indicated.

Norway and Sweden, a collection of 17 preschool language measures predicted together about 20% of the variance in reading skills in grades 1 and 2 (again equivalent to r = 0.45) (Furnes and Samuelsson, 2010). On the neural side, it is difficult to obtain an estimation of the longitudinal stability of brain measures. This may in fact depend on the type of brain measure, as discussed below. The most popular measures (grey/white matter volume, cortical thickness, fractional anisotropy) change significantly with age (Giedd and Rapoport, 2010) and are also known to change with experience and training, which makes it unlikely that they should show much greater longitudinal stability than cognitive measures (see Dubois and Adolphs, 2016, for a more general discussion of the test-retest reliability of MRI measures).

Now, the holy-grail studies attempt to bridge two gaps in one leap: the brain/cognition gap, and the time gap. Estimating the composite effect size over successive paths is not mathematically tractable without actual data, nevertheless it can be stated with confidence that the effects that these studies are trying to discover are necessarily smaller than those of the mediating paths. Fig. 3 illustrates this notion and the consequences on the necessary sample sizes. In a nutshell, if one accepts our estimate of d = 0.5 for likely structural differences between dyslexic and control children, and if a time lag of 3 years reduces this expected effect size to, say, d = 0.25, then holy-grail studies would need 253 participants per group in order to have 80% power to detect the effect. Even in the wildly optimistic scenario of a 0.5 effect size for the longitudinal structure-diagnosis prediction, they would still need 64 participants per group. Which, considering the fact that only about 40% of children at family risk become dyslexic (Pennington and Lefly, 2001; Puolakanaho et al., 2007; Snowling et al., 2007), would require 160 children in the at-risk group at the start (neglecting attrition rate). None of the longitudinal studies cited above come close, with sample sizes of the at-risk group ranging from 10 to 36. For instance, in the largest such study with 36 children at risk (and 35 controls), Vandermosten et al. (2015) had 55% chance of detecting an effect with the optimistic size of 0.5. If 15 of those 36 children turn out to be dyslexic and none drop out, the power of the ultimate group comparison will drop to 40% for d = 0.5 and to 14% for d = 0.25. This is not meant to be a criticism of such studies. As we have said, they are the most important studies to understand the neuroanatomical causes of dyslexia, and they are particularly difficult to conduct. The conundrum is that they are the studies that look for the smallest effects, and are therefore in need of the largest sample sizes, yet they are the ones for which it is most difficult to gather large samples. They are therefore at high risk of finding no predictor of dyslexia, or if they find one, of finding a false positive one.

These difficulties are well illustrated by the two published longitudinal prediction studies. Kraft et al. (2016) started with 71 kindergarten children, with about half carrying a family risk for dyslexia. 59 were scanned between 4 and 6 years of age, and 53 had usable images. In the end only 35 of those could have their reading skills tested at the end of 1st or 2nd grade, yielding a final comparison between 12 dyslexic and 23 control children, and therefore a power ranging from 0.1 to 0.28 to detect an effect size of 0.25–0.5 respectively. The authors report testing the longitudinal prediction of group membership by two brain measures: The T1 intensity of a cluster attributed to the anterior segment of the left AF is reported to significantly predict group membership (p = 0.039), but the cortical thickness of the left supramarginal gyrus does not. The size of the group difference in T1 intensity is impossible to assess given the data reported, but has to be much larger than 0.5 to be significant given the low power. How plausible is that? Another issue is correction for multiple tests: the authors report testing only two brain measures (including a nonstandard one, T1 intensity, that cannot be compared with any other study), but how many have they really tested (given that they have full FA and cortical thickness maps and 4 reconstructed tracts)? How many tests should actually have been taken into account in the correction for multiple testing? In another words, how confident can we be that the reported difference in T1 intensity of the AF is not a false positive result obtained after thorough data mining?

Similarly, Clark et al. (2014) scanned Norwegian children on 3 successive occasions, at age 6 (1st grade), 8, and 12. Dyslexia diagnosis was done at age 12. They reported that cortical thickness measured at age 6 differed in 5 regions between children who were diagnosed as dyslexic at 12 and those who were not. However, their sample size for this analysis was extremely small: 7 dyslexic and 10 control children (as pointed out by Kraft et al., 2015). How could 5 significant differences be found with such low statistical power? Besides choosing a lenient statistical threshold, the fact is that the effect sizes of the differences observed were huge: d > 2, as revealed by their Fig. 1 and by the authors in their reply (Clark et al., 2015). Given that plausible cross-sectional structural correlates of dyslexia have an effect size no greater than 1, it seems very unlikely that group differences greater than 2 might exist, with 6 years interval between the MRI and the diagnostic measure. Thus these group differences bear all the hallmarks of false positive results. Indeed, the one group difference that remained at age 12 (thinner cortex in left Heschl's gyrus in dyslexics), and that could therefore be potentially compared with cross-sectional studies of similarly aged children, was not observed in the two studies that measured it, one with 81 children (Altarelli et al., 2014) and the other with 64 (Ma et al., 2015).

Thus, studies of the early neuroanatomical predictors of dyslexia are both the most exciting and the most handicapped in terms of statistical power. The characteristics of these studies imposing tough limits on the size of samples that can be collected, the only viable solution to avoid the now familiar combination of low statistical power, failure to correct for multiple statistical tests, flexibility in data analysis and other forms of *p*-hacking, and publication of false-positive results, would be to require replication of results across independent longitudinal studies.

Finally, let us point out a last, and largely overlooked possibility to overcome the problem of disentangling causes and consequences of dyslexia. It is that not all brain properties are equally plastic. To the extent that one can measure brain properties that are fixed early on (before or shortly after birth) and that do not change much afterwards, targeting these specific brain properties might prove a fruitful approach for the discovery of true neuroanatomical causes of dyslexia, and does not necessarily require longitudinal studies. The next section elaborates on this idea.

10.5. The limits of automated voxel-based, template-based morphometry

Understandably, the most popular methods for the analysis of neuroanatomy are entirely automated ones: voxel-based morphometry, voxel-based fractional anisotropy analyses (such as TBSS), vertex-based analyses (of cortical thickness and surface area) and analyses focusing on automatically-defined regions of interest (see Lerch et al., 2017 for a recent review). Unfortunately, these methods suffer from a number of limitations:

- To the extent that they are applied to the entire brain (and they usually are, given that this comes at no additional cost), they yield a very large number of dependent measures (volume, surface, thickness or anisotropy at every single voxel or vertex), and thus statistical tests, which then require drastic corrections that are rarely fully applied, as we have discussed in the case of VBM studies. Together with the many options available in the analysis pipelines further increasing the number of tests, these methods are prone to encourage the reporting of false positive results. The number of tests may of course be reduced by targeting regions of interest. And indeed, there is ample scope for more studies carefully investigating theoretically and empirically-defined regions of interest. However, this is a solution only to the extent that such regions of interest are really determined a priori and set once and for all (using for instance preregistration, see Table 2), rather than chosen after a first whole-brain analysis of group differences, as is all too often the case. In a

Table 2
Recommendations for more reliable research on the neuroanatomy of dyslexia. Some of these recommendations are very general and inspired from other sources (e.g., Button et al., 2013).

Recommendation	Explanation
Pre-register studies ahead of data collection	Pre-registration should include experimental protocol, a priori hypotheses, and complete analysis plan for these hypotheses (including pre-processing stages, regions of interest, covariates, etc.). This is the only solution for reported hypotheses to really be a priori hypotheses, and therefore for p values to be meaningful. All analyses not declared in pre-registration can still be carried out and reported, as exploratory analyses. Online platforms such as the Open Science Framework can be used for this purpose (https://osf.io/).
Power analyses	Estimate the likely effect sizes that are being sought, and carry out power analyses at the time of designing experiments, in order to estimate the necessary sample size.
Larger sample sizes	Realistic estimation of effect sizes and power analyses should logically lead to a substantial increase in sample sizes.
Use stringent criteria for dyslexia	This should increase the expected effect sizes, hence statistical power (but to be weighed against the reduction in sample size that may follow).
Study reliable brain properties in well-defined brain regions	Avoid template-based matching and probabilistic atlases that offer low anatomical precision. Make measurements in each individual's native space, informed by knowledge of cytoarchitectonic, genetic or functional architecture.
Study less plastic brain properties	Brain properties that are fixed early in development and that change little with age and learning (e.g., gyral or sulcal landmarks, some asymmetries) may escape the cause-consequence problems typical of more plastic measures such as grey/white matter volumes and fractional anisotropy.
Systematically use relevant covariates	These typically should include age, sex, and a global measure related to the local measure of interest (e.g., total cortical surface when local cortical surface is investigated), in order to differentiate local from global effects. IQ (usually non-verbal IQ in dyslexia research) may also be relevant when it is suspected to covary with the measures of interest. Linear effects should be checked rather than assumed, and proper fitting should take place when nonlinear effects are found.
Full disclosure	Fully disclose analysis variants that were tried before arriving at the variant that produced the reported result. This is essential for readers to evaluate the likelihood of a false positive result.
Open data	Make data available at the time of publication, for other researchers to be able to reproduce analyses and produce new ones.
Collaboration for independent replications	When obtaining a new result, ask a collaborator or use a public database (such as http://www.dyslexiadata.org/) to test whether the result holds in another dataset.

sense, powerful software has made it easier to perform automated whole-brain analyses than to think about appropriate regions of interest and to study them carefully.

- Statistical comparison across participants relies on voxel-wise or vertex-wise mapping individual anatomical measures onto a reference template. This implies blurring and deformation of the images with consequent loss of detail and imprecisions; in a number of cases the mapping may be such that different brain regions of different participants are treated as the same (and the opposite as well). For major neurological disorders, such imprecision may just add a little noise to the analysis; however, for such a mild neurodevelopmental disorder as dyslexia, this probably means making errors of the same order of magnitude as the differences of interest.
- When specific regions of the brain are targeted (such as the planum temporale, or a particular fibre tract), probabilistic atlases defined on reference templates are used in order to automatically identify the region of interest in each individual brain. However, these atlases are typically established on the basis of too few individuals to accurately represent human anatomical variation, and mapping between each individual brain and the atlas can only have the precision of the mapping to the reference template. As a consequence, the reliability of such atlases is probably quite low. To give an example, an entirely automated study of planum temporale asymmetry in 2000+ brains using the HO Cortical structure atlas (http://www.cma.mgh.harvard.edu/fsl_ atlas.html) in SPM8 yielded planum asymmetry index (L-R/L + R) values ranging from 0 to 0.25 (Guadalupe et al., 2015). This implies that almost all participants had greater left than right planum, and that the range of asymmetry was very limited. For a comparison, the values (adjusted to the same scale) reported in our own study based on manual delineation ranged from −0.5 to 0.5 (Altarelli et al., 2014), consistent with Geschwind and Levitsky (1968). Thus, although absolute values may not have mattered much for the specific purposes of Guadalupe and colleagues (genetic association), it is clear that estimations of planum temporale surface or volume provided by the HO atlas combined with SPM are not comparable to those obtained by expert anatomists. The parcellation provided by the Destrieux atlas (Destrieux et al., 2010) used in Freesurfer is known to be inadequate for other reasons, implementing incorrect criteria for the posterior boundary of the planum (Guadalupe et al., 2015). For anybody interested in the actual dimensions of the left and right planum temporale, such methods are therefore to be avoided.
- As a result of the previous two points, automated methods can only detect relatively large differences, probably larger than many of the real differences of interest (e.g., disruptions of neuronal migration). Furthermore, the typically reported group differences in grey/white matter volume or fractional anisotropy are not interpretable in themselves. For instance, volume differences may reflect a variety of phenomena, including differences in cortical thickness, or surface, or differences in sulcal patterns, or focal concentrations of disruptions of neuronal migration, or a different alignment quality with the common template. In a nutshell, finding such differences should be the beginning of much more fine-grained investigations, rather than an end in itself.
- The most popular measures yielded by these methods (grey/white matter volume, fractional anisotropy) are known to change with experience and training, which makes them less than ideal to solve the cause-consequence problem discussed above.

The alternative approach to entirely automated measurements, template-based mapping, and probabilistic atlases is to manually delineate brain regions in each individual brain, in its native space. In principle, the "ground truth" of functional brain architecture is cytoarchitectonics, as described in the work of Brodmann, von Economo, and their modern followers (Amunts et al., 2013). However, current brain imaging techniques do not provide direct in vivo access to such information. An equally or even more interesting, but not more accessible, type of information would be individual gene expression patterns across brain tissues. A more feasible alternative is to use fMRI to define regions with a given functional role in each individual brain, in the same native space as structural images (e.g., Altarelli et al., 2013). Finally, one possible compromise is to focus on well-described brain regions or landmarks that can be reliably identified from their morphology, such as various gyri and sulci. Indeed there is a relationship between macroanatomical landmarks and cytoarchitectonic regions, albeit an imperfect one (Morosan et al., 2001). For instance, we have

found that delineating the planum temporale may provide insights into the neuroanatomy of dyslexia, if (and only if) it is delineated by using cytoarchitectonic knowledge from post-mortem studies (Altarelli et al., 2014). Similarly, the dimensions and shape of sulci may also provide some insights into cognitive differences and disorders (Borst et al., 2016; Cachia et al., 2014; Plaze et al., 2009; e.g., Scotto di Covella et al., submitted). Even if such regions or sulci do not show a perfect, systematic relationship with cytoarchitectonic areas, they still are more meaningful than coordinates in a reference template, and may provide other advantages, including that:

- They can be reliably identified and delineated in each individual brain (in most cases), in their native space.
- Early disruptions of cortical development and organisation may disrupt their dimensions and/or shape, such that, even if the dimensions and/or shape of these landmarks do not play a direct causal role in the phenotype of interest, they may be strong correlates thereof.
- They seem to be fixed early on in development and to show little plasticity, making them better candidates for early brain markers of dyslexia than measures of cortical volume, thickness, or fractional anisotropy.

Of course, there are more choices than the two extremes discussed here. There is room for a judicious combination of automated and manual methods (Eckert et al., 2008, 2005). For instance, automated surface-based morphometry is known to yield tissue segmentation errors: correcting them manually is important in order to improve estimates of cortical thickness in several regions (McCarthy et al., 2015; Popescu et al., 2016). To the extent that this is done, and that one measures cortical thickness in the native space in well-defined cortical regions, rather than on atlas-based regions in a common template, automated surface-based processing is a tool that should not be neglected. Furthermore, manual segmentation techniques can be aided by algorithms. For instance, although cortical sulci can be defined entirely manually, this process can be considerably facilitated by relying on an initial automatic labelling of sulci, such as that provided by Brainvisa (Mangin et al., 2004). Thus, the message is not so much to ban automated techniques, but to use the expert human eye to review and correct their outputs for anatomical plausibility, rather than using them blindly. Most importantly, making tangible measurements in each individual's native space circumvents the many problems introduced by voxel-based or vertex-based template-matching and probabilistic atlases.

11. Conclusion

In conclusion, across all the techniques available to investigate neuroanatomical differences, and across several dozen studies, we find that there are relatively few results that stand out as robust, well-replicated neuroanatomical differences between control and dyslexic individuals. In fact, the only one that can be asserted with a high degree of confidence is the difference in total brain size (whether measured by TIV, TBV, GMV, WMV or total brain surface). Yet it is certainly not the theoretically most exciting difference, and it is associated with a modest effect size (0.4). We would like to believe that the difference in planum asymmetry in boys is a reliable one, yet it still needs to be replicated again using similar methods to increase our degree of confidence. The technique that has been used in the greatest number of studies, VBM, has been the most disappointing: both large-scale studies and meta-analyses suggest that most published results may be false positive ones, and until now not a single difference in local GMV can be claimed to have been reliably proven. Diffusion imaging studies may seem to converge on the left arcuate fasciculus, but methodological limitations and differences across studies do not allow for a proper evaluation of the nature and the consistency of that result.

In this review, we have emphasised the limitations of many of the past studies on the neuroanatomy of dyslexia. With the aim of improving the reliability of future research, we also propose a number of recommendations (Table 2), and what seem to be contradictory injunctions. On the one hand, in order to increase the reliability of the findings and in particular statistical power, we advocate a substantial increase in sample sizes. On the other hand, we also recommend selecting more severe cases with dyslexia in order to increase expected effect sizes, rather than diluting the potential findings with mildly impaired readers. Unfortunately, severe cases are fewer than mild ones, so it may be particularly challenging to recruit large samples of severe cases. We have also advocated scepticism toward the most popular image processing techniques based on entirely automated processing and mapping onto common templates, as they are probably too imprecise for the purpose of uncovering the neuroanatomical bases of dyslexia. Yet, the more careful, manual methods that we favour are obviously much more time-consuming, and therefore more difficult and expensive to apply to large datasets.

Nevertheless, we think that our recommendations are not totally incompatible. But they come at a cost. They imply that any study attempting to discover neuroanatomical bases of dyslexia has to be a much larger and more expensive project than the usual 10–20 poor vs. 10–20 good readers. Much larger samples have to be gathered, and samples from different laboratories have to be pooled together for joint or meta-analyses. Furthermore, personnel have to be trained to skilfully and reliably identify the neuroanatomical markers of interest, implying that the cost of image analysis becomes somewhat proportional to sample size. The price of a publishable study will incur a steep rise as a consequence. Cognitive neuroscientists who find such a prospect disheartening may find moderate comfort in learning that they are not alone: geneticists have been through that phase already (Ioannidis et al., 2001). They have learned the hard way to distrust whole sectors of their literature, whose results could not be replicated. As a result, the standards for publication have been dramatically upgraded. Necessary sample sizes for whole genome association studies now range in the thousands. Leading journals also require any genetic association to be observed in two independent populations before publication can be considered (e.g., Editorial, 2005). Besides the increase in reliability of published results, a positive side-effect has been an increase in collaborations and in consortia. It is now customary, once a result is obtained in one's own data, to ask collaborators to replicate the same analysis, and to publish the results jointly. Such a strategy could also certainly bring benefits to the cognitive neuroscience community in general, and to the study of the neuroanatomy of dyslexia in particular. For this purpose, the recently created Dyslexia data consortium (http://www.dyslexiadata.org/) might be a promising platform to share data and run replication studies.

Acknowledgements

This work was supported by Agence Nationale de la Recherche [grant numbers ANR-11-BSV4-014-01, ANR-10-LABX-0087 IEC, ANR-10-IDEX-0001-02 PSL*]; European Union's Horizon 2020 research and innovation programme [Marie Sklodowska-Curie grant agreement No 661667]; Polish Ministry of Science and Higher Education [grant numbers IP2010 015170, IP2011 020271, IP2012 025272]; National Science Centre grant [grant number 2014/14/A/HS6/00294]. The present paper is an expansion on Ramus et al. (in press). We thank Jessica Dubois and Michel Thiebaut de Schotten for their input on the diffusion imaging section, and Cyril Pernet and an anonymous reviewer for their thoughtful comments on a previous version of this paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.neubiorev.2017.08.001.

References

Altarelli, I., Monzalvo, K., Iannuzzi, S., Fluss, J., Billard, C., Ramus, F., Dehaene-Lambertz, G., 2013. A functionally guided approach to the morphometry of occipito-temporal regions in developmental dyslexia: evidence for differential effects in boys and girls. J. Neurosci. 33, 11296–11301.

Altarelli, I., Leroy, F., Monzalvo, K., Fluss, J., Billard, C., Dehaene-Lambertz, G., Galaburda, A.M., Ramus, F., 2014. Planum temporale asymmetry in developmental dyslexia: revisiting an old question. Hum. Brain Mapp. 35, 5717–5735.

Altarelli, I., Monzalvo, K., Billard, C., Fluss, J., Dehaene-Lambertz, G., Ramus, F., submitted. Cortical thickness, surface area and gyrification in children with developmental dyslexia: atlas and vertex-based approaches.

Altarelli, I., 2013. Neuroanatomical Correlates of Developmental Dyslexia. Université Pierre et Marie Curie, Paris (Ph.D Dissertation).

Amunts, K., Lepage, C., Borgeat, L., Mohlberg, H., Dickscheid, T., Rousseau, M.E., Bludau, S., Bazin, P.L., Lewis, L.B., Oros-Peusquens, A.M., Shah, N.J., Lippert, T., Zilles, K., Evans, A.C., 2013. BigBrain: an ultrahigh-resolution 3D human brain model. Science 340, 1472–1475. http://dx.doi.org/10.1126/science.1235381.

Anderson, A., Ding, Z., 2002. Sub-voxel measurement of fiber orientation using high angular resolution diffusion tensor imaging. Book of Tenth Annual Meeting of the International Society for Magnetic Resonance in Medicine. ISMRM, Berkeley, CA (p. 440).

Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Price, C., Friston, K., 1998. Identifying global anatomical differences: deformation-based morphometry. Hum. Brain Mapp. 6, 348–357. http://dx.doi.org/10.1002/(Sici)1097-0193(1998)6:5/6<348:Aid-Hbm4>3.0.Co;2-P.

Barnes, J., Ridgway, G.R., Bartlett, J., Henley, S.M., Lehmann, M., Hobbs, N., Clarkson, M.J., MacManus, D.G., Ourselin, S., Fox, N.C., 2010. Head size, age and gender adjustment in MRI studies: a necessary nuisance? Neuroimage 53, 1244–1255.

Barrick, T.R., Mackay, C.E., Prima, S., Maes, F., Vandermeulen, D., Crow, T.J., Roberts, N., 2005. Automatic analysis of cerebral asymmetry: an exploratory study of the relationship between brain torque and planum temporale asymmetry. Neuroimage 24, 678–691. http://dx.doi.org/10.1016/j.neuroimage.2004.09.003.

Beaulieu, C., Plewes, C., Paulson, L.A., Roy, D., Snook, L., Concha, L., Phillips, L., 2005. Imaging brain connectivity in children with diverse reading ability. Neuroimage 25, 1266–1271.

Ben-Shachar, M., Dougherty, R.F., Wandell, B.A., 2007. White matter pathways in reading. Curr. Opin. Neurobiol. 17, 258–270.

Best, M., Demb, J.B., 1999. Normal planum temporale asymmetry in dyslexics with a magnocellular pathway deficit. Neuroreport 10, 607–612.

Bishop, D.V., 2013. Cerebral asymmetry and language development: cause, correlate, or consequence? Science 340, 1230531. http://dx.doi.org/10.1126/science.1230531.

Black, J.M., Tanaka, H., Stanley, L., Nagamine, M., Zakerani, N., Thurston, A., Kesler, S., Hulme, C., Lyytinen, H., Glover, G.H., Serrone, C., Raman, M.M., Reiss, A.L., Hoeft, F., 2012. Maternal history of reading difficulty is associated with reduced language-related gray matter in beginning readers. Neuroimage 59, 3021–3032.

Bloom, J.S., Garcia-Barrera, M.A., Miller, C.J., Miller, S.R., Hynd, G.W., 2013. Planum temporale morphology in children with developmental dyslexia. Neuropsychologia 51, 1684–1692.

Boder, E., 1973. Developmental dyslexia: a diagnostic approach based on three atypical reading-spelling patterns. Dev. Med. Child Neurol. 15, 663–687.

Borst, G., Cachia, A., Tissier, C., Ahr, E., Simon, G., Houdé, O., 2016. Early Cerebral Constraints on Reading Skills in School-Age Children: An MRI Study. Mind Brain Educ. 10, 47–54. http://dx.doi.org/10.1111/mbe.12098.

Bosse, M.L., Tainturier, M.J., Valdois, S., 2007. Developmental dyslexia: the visual attention span deficit hypothesis. Cognition 104, 198–230.

Brambati, S.M., Termine, C., Ruffino, M., Stella, G., Fazio, F., Cappa, S., Perani, D., 2004. Regional reductions of gray matter volume in familial dyslexia. Neurology 63, 742–745.

Brown, W.E., Eliez, S., Menon, V., Rumsey, J.M., White, C.D., Reiss, A.L., 2001. Preliminary evidence of widespread morphological variations of the brain in dyslexia. Neurology 56, 781–783.

Bruno, J.L., Lu, Z.-L., Manis, F.R., 2013. Phonological processing is uniquely associated with neuro-metabolic concentration. Neuroimage 67, 175–181. http://dx.doi.org/10.1016/j.neuroimage.2012.10.092.

Button, K.S., Ioannidis, J.P.A., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafò, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. Nat. Rev. Neurosci. 14, 365–376. http://dx.doi.org/10.1038/nrn3475.

Cachia, A., Paillère-Martinot, M.-L., Galinowski, A., Januel, D., de Beaurepaire, R., Bellivier, F., Artiges, E., Andoh, J., Bartrés-Faz, D., Duchesnay, E., 2008. Cortical folding abnormalities in schizophrenia patients with resistant auditory hallucinations. Neuroimage 39, 927–935.

Cachia, A., Borst, G., Vidal, J., Fischer, C., Pineau, A., Mangin, J.-F., Houdé, O., 2014. The shape of the ACC contributes to cognitive control efficiency in preschoolers. J. Cogn. Neurosci. 26, 96–106.

Cachia, A., Borst, G., Tissier, C., Fisher, C., Plaze, M., Gay, O., Rivière, D., Gogtay, N., Giedd, J., Mangin, J.-F., Houdé, O., Raznahan, A., 2016. Longitudinal stability of the folding pattern of the anterior cingulate cortex during development. Dev. Cogn. Neurosci. 19, 122–127. http://dx.doi.org/10.1016/j.dcn.2016.02.011.

Cahill, L., 2006. Why sex matters for neuroscience. Nat. Rev. Neurosci. 7, 477–484.

Caravolas, M., Lervåg, A., Defior, S., Seidlová Málková, G., Hulme, C., 2013. Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. Psychol. Sci. http://dx.doi.org/10.1177/0956797612473122.

Carter, J.C., Lanham, D.C., Cutting, L.E., Clements-Stephens, A.M., Chen, X., Hadzipasic, M., Kim, J., Denckla, M.B., Kaufmann, W.E., 2009. A dual DTI approach to analyzing white matter in children with dyslexia. Psychiatry Res. Neuroimaging 172, 215–219. http://dx.doi.org/10.1016/j.pscychresns.2008.09.005.

Casanova, M.F., Araque, J., Giedd, J., Rumsey, J.M., 2004. Reduced brain size and gyrification in the brains of dyslexic patients. J. Child Neurol. 19, 275–281.

Castles, A., Coltheart, M., 1993. Varieties of developmental dyslexia. Cognition 47, 149–180.

Catani, M., Jones, D.K., ffytche, D.H., 2005. Perisylvian language networks of the human brain. Ann. Neurol. 57, 8–16. http://dx.doi.org/10.1002/ana.20319.

Catani, M., Allin, M.P., Husain, M., Pugliese, L., Mesulam, M.M., Murray, R.M., Jones, D.K., 2007. Symmetries in human brain language pathways correlate with verbal recall. Proc. Natl. Acad. Sci. 104, 17163–17168.

Chen, C., Xue, G., Dong, Q., Jin, Z., Li, T., Xue, F., Zhao, L., Guo, Y., 2007. Sex determines the neurofunctional predictors of visual word learning. Neuropsychologia 45, 741–747. http://dx.doi.org/10.1016/j.neuropsychologia.2006.08.018.

Chiarello, C., Lombardino, L.J., Kacinik, N.A., Otto, R., Leonard, C.M., 2006. Neuroanatomical and behavioral asymmetry in an adult compensated dyslexic. Brain Lang. 98, 169–181.

Clark, K.A., Helland, T., Specht, K., Narr, K.L., Manis, F.R., Toga, A.W., Hugdahl, K., 2014. Neuroanatomical precursors of dyslexia identified from pre-reading through to age 11. Brain. http://dx.doi.org/10.1093/brain/awu229.

Clark, K.A., Helland, T., Specht, K., Narr, K.L., Manis, F.R., Toga, A.W., Hugdahl, K., 2015. Reply: cortical differences in preliterate children at familiar risk of dyslexia are similar to those observed in dyslexic readers. Brain 138http://dx.doi.org/10.1093/brain/awv037. e379–e379.

Clayton, J.A., Collins, F.S., 2014. NIH to balance sex in cell and animal studies. Nature 509, 282–283.

Colby, J.B., Soderberg, L., Lebel, C., Dinov, I.D., Thompson, P.M., Sowell, E.R., 2012. Along-tract statistics allow for enhanced tractography analysis. Neuroimage 59, 3227–3242. http://dx.doi.org/10.1016/j.neuroimage.2011.11.004.

Craggs, J.G., Sanchez, J., Kibby, M.Y., Gilger, J.W., Hynd, G.W., 2006. Brain morphology and neuropsychological profiles in a family displaying dyslexia and superior nonverbal intelligence. Cortex 42, 1107–1118.

Cui, Z., Xia, Z., Su, M., Shu, H., Gong, G., 2016. Disrupted white matter connectivity underlying developmental dyslexia: a machine learning approach. Hum. Brain Mapp. 37, 1443–1458. http://dx.doi.org/10.1002/hbm.23112.

Dell'Acqua, F., Simmons, A., Williams, S.C.R., Catani, M., 2013. Can spherical deconvolution provide more information than fiber orientations? Hindrance modulated orientational anisotropy, a true-tract specific index to characterize white matter diffusion. Hum. Brain Mapp. 34, 2464–2483. http://dx.doi.org/10.1002/hbm.22080.

Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. Neuroimage 53, 1–15.

Deutsch, G.K., Dougherty, R.F., Bammer, R., Siok, W.T., Gabrieli, J.D.E., Wandell, B., 2005. Children's reading performance is correlated with white matter structure measured by diffusion tensor imaging. Cortex 41, 354–363.

Dole, M., Meunier, F., Hoen, M., 2013. Gray and white matter distribution in dyslexia: a VBM study of superior temporal gyrus asymmetry. PLoS One 8, e76823. http://dx.doi.org/10.1371/journal.pone.0076823.

Dougherty, R.F., Ben-Shachar, M., Deutsch, G.K., Hernandez, A., Fox, G.R., Wandell, B.A., 2007. Temporal-callosal pathway diffusivity predicts phonological skills in children. Proc. Natl. Acad. Sci. U. S. A. 104, 8556–8561.

Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., May, A., 2004. Neuroplasticity: changes in grey matter induced by training. Nature 427, 311–312.

Drake, W.E., 1968. Clinical and pathological findings in a child with a developmental learning disability. J. Learn. Disabil. 1, 486–502.

Duara, R., Kushch, A., Gross-Glenn, K., Barker, W.W., Jallad, B., Pascal, S., Loewenstein, D.A., Sheldon, J., Rabin, M., Levin, B., 1991. Neuroanatomic differences between dyslexic and normal readers on magnetic resonance imaging scans. Arch. Neurol. 48, 410–416.

Dubois, J., Adolphs, R., 2016. Building a science of individual differences from fMRI. Trends Cogn. Sci. 20, 425–443. http://dx.doi.org/10.1016/j.tics.2016.03.014.

Dubois, J., Dehaene-Lambertz, G., Kulikova, S., Poupon, C., Hüppi, P.S., Hertz-Pannier, L., 2014. The early development of brain white matter: a review of imaging studies in fetuses, newborns and infants. Neuroscience: Secrets CNS White Matter 276, 48–71. http://dx.doi.org/10.1016/j.neuroscience.2013.12.044.

Ecker, C., Marquand, A., Mourao-Miranda, J., Johnston, P., Daly, E.M., Brammer, M.J., Maltezos, S., Murphy, C.M., Robertson, D., Williams, S.C., Murphy, D.G., 2010. Describing the brain in autism in five dimensions–magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. J. Neurosci. 30, 10612–10623.

Eckert, M.A., Leonard, C.M., Wilke, M., Eckert, M., Richards, T., Richards, A., Berninger, V., 2005. Anatomical signatures of dyslexia in children: unique information from manual and voxel based morphometry brain measures. Cortex 41, 304–315.

Eckert, M.A., Tenforde, A., Galaburda, A.M., Bellugi, U., Korenberg, J.R., Mills, D., Reiss, A.L., 2006. To modulate or not to modulate: differing results in uniquely shaped Williams syndrome brains. Neuroimage 32, 1001–1007. http://dx.doi.org/10.1016/j.neuroimage.2006.05.014.

Eckert, M.A., Lombardino, L.J., Walczak, A.R., Bonihla, L., Leonard, C.M., Binder, J.R., 2008. Manual and automated measures of superior temporal gyrus asymmetry: concordant structural predictors of verbal ability in children. Neuroimage 41, 813–822.

Eckert, M.A., Berninger, V.W., Hoeft, F., Vaden Jr., K.I., 2016a. A case of Bilateral Perisylvian Syndrome with reading disability. Cortex 76, 121–124. http://dx.doi.org/10.1016/j.cortex.2016.01.004.

Eckert, M.A., Berninger, V.W., Vaden, K.I., Gebregziabher, M., Tsu, L., 2016b. Gray

matter features of reading disability: a combined meta-analytic and direct analysis approach. eneuro 3 (1). http://dx.doi.org/10.1523/eneuro.0103-15.2015. ENEURO. 0103-15.2015.

Eckert, M., 2004. Neuroanatomical markers for dyslexia: a review of dyslexia structural imaging studies. Neuroscientist 10, 362–371.

Editorial, 2005. Framework for a fully powered risk engine. Nat. Genet. 37http://dx.doi.org/10.1038/ng1105-1153. (1153–1153).

Engvig, A., Fjell, A.M., Westlye, L.T., Moberget, T., Sundseth, Ø., Larsen, V.A., Walhovd, K.B., 2010. Effects of memory training on cortical thickness in the elderly. Neuroimage 52, 1667–1676. http://dx.doi.org/10.1016/j.neuroimage.2010.05.041.

Evans, T.M., Flowers, D.L., Napoliello, E.M., Eden, G.F., 2014. Sex-specific gray matter volume differences in females with developmental dyslexia. Brain Struct. Funct. 219, 1041–1054.

Faul, F., Erdfelder, E., Lang, A.-G., Buchner, A., 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav. Res. Methods 39, 175–191. http://dx.doi.org/10.3758/BF03193146.

Feldman, H.M., Lee, E.S., Yeatman, J.D., Yeom, K.W., 2012. Language and reading skills in school-aged children and adolescents born preterm are associated with white matter properties on diffusion tensor imaging. Neuropsychologia 50, 3348–3362. http://dx.doi.org/10.1016/j.neuropsychologia.2012.10.014.

Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., 1999. High-resolution intersubject averaging and a coordinate system for the cortical surface. Hum. Brain Mapp. 8, 272–284.

Fitch, R.H., Brown, C.P., Tallal, P., Rosen, G.D., 1997. Effects of sex and MK-801 on auditory-processing deficits associated with developmental microgyric lesions in rats. Behav. Neurosci. 111, 404–412.

Frith, U., Vargha-Khadem, F., 2001. Are there sex differences in the brain basis of literacy related skills?: Evidence from reading and spelling impairments after early unilateral brain damage. Neuropsychologia 39, 1485–1488.

Frost, M.A., Goebel, R., 2012. Measuring structural-functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment. Neuroimage 59, 1369–1381. http://dx.doi.org/10.1016/j.neuroimage.2011.08.035.

Frye, R.E., Liederman, J., Malmberg, B., McLean, J., Strickland, D., Beauchamp, M.S., 2010. Surface area accounts for the relation of gray matter volume to reading-related skills and history of dyslexia. Cereb. Cortex 20, 2625–2635.

Furnes, B., Samuelsson, S., 2010. Predicting reading and spelling difficulties in transparent and opaque orthographies: a comparison between scandinavian and US/Australian children. Dyslexia 16, 119–142.

Fusar-Poli, P., Radua, J., Frascarelli, M., Mechelli, A., Borgwardt, S., Di Fabio, F., Biondi, M., Ioannidis, J.P.A., David, S.P., 2014. Evidence of reporting biases in voxel-based morphometry (VBM) studies of psychiatric and neurological disorders. Hum. Brain Mapp. 35, 3052–3065. http://dx.doi.org/10.1002/hbm.22384.

Galaburda, A.M., Eidelberg, D., 1982. Symmetry and asymmetry in the human posterior thalamus: II. Thalamic lesions in a case of developmental dyslexia. Arch. Neurol. 39, 333–336.

Galaburda, A.M., Kemper, T.L., 1979. Cytoarchitectonic abnormalities in developmental dyslexia: a case study. Ann. Neurol. 6, 94–100.

Galaburda, A.M., Livingstone, M., 1993. Evidence for a magnocellular defect in developmental dyslexia. Ann. N. Y. Acad. Sci. 682, 70–82.

Galaburda, A.M., LeMay, M., Kemper, T.L., Geschwind, N., 1978. Right-left asymmetrics in the brain. Science 199, 852–856.

Galaburda, A.M., Sherman, G.F., Rosen, G.D., Aboitiz, F., Geschwind, N., 1985. Developmental dyslexia: four consecutive patients with cortical anomalies. Ann. Neurol. 18, 222–233.

Galaburda, A.M., Menard, M.T., Rosen, G.D., 1994. Evidence for aberrant auditory anatomy in developmental dyslexia. Proc. Natl. Acad. Sci. U. S. A. 91, 8010–8013.

Gay, O., Plaze, M., Oppenheim, C., Mouchet-Mages, S., Gaillard, R., Olié, J.-P., Krebs, M.-O., Cachia, A., 2013. Cortex morphology in first-episode psychosis patients with neurological soft signs. Schizophr. Bull. 39, 820–829. http://dx.doi.org/10.1093/schbul/sbs083.

Geschwind, N., Galaburda, A.M., 1985. Cerebral lateralization. Biological mechanisms associations, and pathology: I. A hypothesis and a program for research. Arch. Neurol. 42, 428–459.

Geschwind, N., Levitsky, W., 1968. Human brain: left-right asymmetries in temporal speech region. Science 161, 186–187.

Ghosh, S.S., Kakunoori, S., Augustinack, J., Nieto-Castanon, A., Kovelman, I., Gaab, N., Christodoulou, J.A., Triantafyllou, C., Gabrieli, J.D.E., Fischl, B., 2010. Evaluating the validity of volume-based and surface-based brain image registration for developmental cognitive neuroscience studies in children 4–11 years of age. Neuroimage 53, 85–93. http://dx.doi.org/10.1016/j.neuroimage.2010.05.075.

Giedd, J.N., Rapoport, J.L., 2010. Structural MRI of pediatric brain development: what have we learned and where are we going? Neuron 67, 728–734. http://dx.doi.org/10.1016/j.neuron.2010.08.040.

Giraud, A.L., Poeppel, D., 2012. Cortical oscillations and speech processing: emerging computational principles and operations. Nat. Neurosci. 15, 511–517.

Giraud, A.L., Ramus, F., 2013. Neurogenetics and auditory processing in developmental dyslexia. Curr. Opin. Neurobiol. 23, 37–42.

Gori, S., Facoetti, A., 2015. How the visual aspects can be crucial in reading acquisition? The intriguing case of crowding and developmental dyslexia. J. Vis. 15, 8.

Goswami, U., 2011. A temporal sampling framework for developmental dyslexia. Trends Cogn. Sci. 15, 3–10.

Green, R.L., Hutsler, J.J., Loftus, W.C., Tramo, M.J., Thomas, C.E., Silberfarb, A.W., Nordgren, R.E., Nordgren, R.A., Gazzaniga, M.S., 1999. The caudal infrasylvian surface in dyslexia Novel magnetic resonance imaging–based findings. Neurology 53, 974. http://dx.doi.org/10.1212/WNL.53.5.974.

Guadalupe, T., Zwiers, M.P., Wittfeld, K., Teumer, A., Vasquez, A.A., Hoogman, M., Hagoort, P., Fernandez, G., Buitelaar, J., van Bokhoven, H., 2015. Asymmetry within and around the human planum temporale is sexually dimorphic and influenced by genes involved in steroid hormone receptor activity. Cortex 62, 41–55.

Haar, S., Berman, S., Behrmann, M., Dinstein, I., 2014. Anatomical abnormalities in autism? Cereb. Cortex. http://dx.doi.org/10.1093/cercor/bhu242.

Hadzibeganovic, T., van den Noort, M., Bosch, P., Perc, M., van Kralingen, R., Mondt, K., Coltheart, M., 2010. Cross-linguistic neuroimaging and dyslexia: a critical view. Cortex 46, 1312–1316. http://dx.doi.org/10.1016/j.cortex.2010.06.011.

Haier, R.J., Karama, S., Leyba, L., Jung, R.E., 2009. MRI assessment of cortical thickness and functional activity changes in adolescent girls following three months of practice on a visual-spatial task. BMC Res. Notes 2, 174. http://dx.doi.org/10.1186/1756-0500-2-174.

Hall, E.D., Pazara, K.E., Linseman, K.L., 1991. Sex differences in postischemic neuronal necrosis in gerbils. J. Cereb. Blood Flow Metab. 11, 292–298.

Hamalainen, J.A., Rupp, A., Soltesz, F., Szucs, D., Goswami, U., 2012. Reduced phase locking to slow amplitude modulation in adults with dyslexia: an MEG study. Neuroimage 59, 2952–2961. http://dx.doi.org/10.1016/j.neuroimage.2011.09.075.

Haslam, R., Dalby, J., Johns, R., Rademaker, A., 1981. Cerebral asymmetry in developmental dyslexia. Arch. Neurol. 38, 679–682. http://dx.doi.org/10.1001/archneur.1981.00510110039003.

Heiervang, E., Hugdahl, K., Steinmetz, H., Inge Smievoll, A., Stevenson, J., Lund, A., Ersland, L., Lundervold, A., 2000. Planum temporale, planum parietale and dichotic listening in dyslexia. Neuropsychologia 38, 1704–1713.

Henley, S.M., Ridgway, G.R., Scahill, R.I., Kloppel, S., Tabrizi, S.J., Fox, N.C., Kassubek, J., 2010. Pitfalls in the use of voxel-based morphometry as a biomarker: examples from huntington disease. AJNR Am. J. Neuroradiol. 31, 711–719. http://dx.doi.org/10.3174/ajnr.A1939.

Hier, D.B., LeMay, M., Rosenberger, P.B., Perlo, V.P., 1978. Developmental dyslexia: evidence for a subgroup with a reversal of cerebral asymmetry. Arch. Neurol. 35, 90–92.

Hoeft, F., Meyler, A., Hernandez, A., Juel, C., Taylor-Hill, H., Martindale, J.L., McMillon, G., Kolchugina, G., Black, J.M., Faizi, A., Deutsch, G.K., Siok, W.T., Reiss, A.L., Whitfield-Gabrieli, S., Gabrieli, J.D., 2007. Functional and morphometric brain dissociation between dyslexia and reading ability. Proc. Natl. Acad. Sci. U. S. A. 104, 4234–4239.

Hoeft, F., McCandliss, B.D., Black, J.M., Gantman, A., Zakerani, N., Hulme, C., Lyytinen, H., Whitfield-Gabrieli, S., Glover, G.H., Reiss, A.L., Gabrieli, J.D., 2011. Neural systems predicting long-term outcome in dyslexia. Proc. Natl. Acad. Sci. U. S. A. 108, 361–366.

Hosseini, S.M.H., Black, J.M., Soriano, T., Bugescu, N., Martinez, R., Raman, M.M., Kesler, S.R., Hoeft, F., 2013. Topological properties of large-scale structural brain networks in children with familial risk for reading difficulties. Neuroimage 71, 260–274. http://dx.doi.org/10.1016/j.neuroimage.2013.01.013.

Hugdahl, K., Heiervang, E., Nordby, H., Smievoll, A.I., Steinmetz, H., Stevenson, J., Lund, A., 1998. Central auditory processing, MRI morphometry and brain laterality: applications to dyslexia. Scand. Audiol. 27, 26–34. http://dx.doi.org/10.1080/010503998420621.

Humphreys, P., Kaufmann, W.E., Galaburda, A.M., 1990. Developmental dyslexia in women: neuropathological findings in three patients. Ann. Neurol. 28, 727–738.

Hynd, G.W., Semrud-Clikeman, M., Lorys, A.R., Novey, E.S., Eliopulos, D., 1990. Brain morphology in developmental dyslexia and attention deficit disorder/hyperactivity. Arch. Neurol. 47, 919–926.

Im, K., Raschle, N.M., Smith, S.A., Ellen Grant, P., Gaab, N., 2015. Atypical sulcal pattern in children with developmental dyslexia and at-risk kindergarteners. Cereb. Cortex. http://dx.doi.org/10.1093/cercor/bhu305.

Ioannidis, J.P.A., Ntzani, E.E., Trikalinos, T.A., Contopoulos-Ioannidis, D.G., 2001. Replication validity of genetic association studies. Nat. Gen. 29 (3), 306–309. http://dx.doi.org/10.1038/ng749.

Ioannidis, J.P.A., 2011. Excess significance bias in the literature on brain volume abnormalities. Arch. Gen. Psychiatry Archgenpsychiatry (2011. 28 v1).

Jednoróg, K., Gawron, N., Marchewka, A., Heim, S., Grabowska, A., 2014. Cognitive subtypes of dyslexia are characterized by distinct patterns of grey matter volume. Brain Struct. Funct. 219, 1697–1707.

Jednoróg, K., Marchewka, A., Altarelli, I., Monzalvo, K., van Ermingen-Marbach, M., Grande, M., Grabowska, A., Heim, S., Ramus, F., 2015. How reliable are grey matter disruptions in specific reading disability across multiple countries and languages? Insights from a large-scale voxel-based morphometry study. Hum. Brain Mapp. 36, 1741–1754.

Keller, T.A., Just, M.A., 2009. Altering cortical connectivity: remediation-induced changes in the white matter of poor readers. Neuron 64, 624–631.

Kershner, J.R., 2015. A mini-review: toward a comprehensive theory of dyslexia. J. Neurol. Neurosci.

Kertesz, A., Polk, M., Black, S.E., Howell, J., 1990. Sex, handedness, and the morphometry of cerebral asymmetries on magnetic resonance imaging. Brain Res. 530, 40–48.

Kibby, M.Y., Kroese, J.M., Morgan, A.E., Hiemenz, J.R., Cohen, M.J., Hynd, G.W., 2004. The relationship between perisylvian morphology and verbal short-term memory functioning in children with neurodevelopmental disorders. Brain Lang. 89, 122–135.

Klingberg, T., Hedehus, M., Temple, E., Salz, T., Gabrieli, J.D., Moseley, M.E., Poldrack, R.A., 2000. Microstructure of temporo-parietal white matter as a basis for reading ability: evidence from diffusion tensor magnetic resonance imaging. Neuron 25, 493–500.

Krafnick, A.J., Flowers, D.L., Luetje, M.M., Napoliello, E.M., Eden, G.F., 2014. An investigation into the origin of anatomical differences in dyslexia. J. Neurosci. 34, 901–908. http://dx.doi.org/10.1523/jneurosci.2092-13.2013.

Kraft, I., Cafiero, R., Schaadt, G., Brauer, J., Neef, N.E., Müller, B., Kirsten, H., Wilcke, A.,

Boltze, J., Friederici, A.D., 2015. Cortical differences in preliterate children at familiar risk of dyslexia are similar to those observed in dyslexic readers. Brain 138, e378.

Kraft, I., Schreiber, J., Cafiero, R., Metere, R., Schaadt, G., Brauer, J., Neef, N.E., Müller, B., Kirsten, H., Wilcke, A., Boltze, J., Friederici, A.D., Skeide, M.A., 2016. Predicting early signs of dyslexia at a preliterate age by combining behavioral assessment with structural MRI. Neuroimage 143, 378–386. http://dx.doi.org/10.1016/j.neuroimage.2016.09.004.

Kronbichler, M., Wimmer, H., Staffen, W., Hutzler, F., Mair, A., Ladurner, G., 2008. Developmental dyslexia: gray matter abnormalities in the occipitotemporal cortex. Hum. Brain Mapp. 29, 613–625.

Kulynych, J.J., Vladar, K., Jones, D.W., Weinberger, D.R., 1993. Three-dimensional surface rendering in MRI morphometry: a study of the planum temporale. J. Comput. Assist. Tomogr. 17, 529–535.

Langer, N., Peysakhovich, B., Zuk, J., Drottar, M., Sliva, D.D., Smith, S., Becker, B.L.C., Grant, P.E., Gaab, N., 2015. White matter alterations in infants at risk for developmental dyslexia. Cereb. Cortex. http://dx.doi.org/10.1093/cercor/bhv281.

Larsen, J.P., Høien, T., Lundberg, I., Ødegaard, H., 1990. MRI evaluation of the size and symmetry of the planum temporale in adolescents with developmental dyslexia. Brain Lang. 39, 289–301.

Laycock, S.K., Wilkinson, I.D., Wallis, L.I., Darwent, G., Wonders, S.H., Fawcett, A.J., Griffiths, P.D., Nicolson, R.I., 2008. Cerebellar volume and cerebellar metabolic characteristics in adults with dyslexia. Ann. N. Y. Acad. Sci. 1145, 222–236. http://dx.doi.org/10.1196/annals.1416.002.

LeMay, M., 1976. Morphological cerebral asymmetries of modern man, fossil man, and nonhuman primate. Ann. N. Y. Acad. Sci. 280, 349–366.

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., Giraud, A.L., 2011. Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. Neuron 72, 1080–1090.

Lehongre, K., Morillon, B., Giraud, A.L., Ramus, F., 2013. Impaired auditory sampling in dyslexia: further evidence from combined fMRI and EEG. Front. Hum. Neurosci. 7, 454. http://dx.doi.org/10.3389/fnhum.2013.00454.

Leonard, C.M., Voeller, K.K., Lombardino, L.J., Morris, M.K., Hynd, G.W., Alexander, A.W., Andersen, H.G., Garofalakis, M., Honeyman, J.C., Mao, J., et al., 1993. Anomalous cerebral structure in dyslexia revealed with magnetic resonance imaging. Arch. Neurol. 50, 461–469.

Leonard, C.M., Lombardino, L.J., Walsh, K., Eckert, M.A., Mockler, J.L., Rowe, L.A., Williams, S., DeBose, C.B., 2002a. Anatomical risk factors that distinguish dyslexia from SLI predict reading skill in normal children. J. Commun. Disord. 35, 501–531.

Leonard, C.M., Lombardino, L.J., Walsh, K., Eckert, M.A., Mockler, J.L., Rowe, L.A., Williams, S., DeBose, C.B., 2002b. Anatomical risk factors that distinguish dyslexia from SLI predict reading skill in normal children. J. Commun. Disord. 35, 501–531.

Leonard, C., Eckert, M., Given, B., Berninger, V., Eden, G., 2006. Individual differences in anatomy predict reading and oral language impairments in children. Brain 129, 3329–3342. http://dx.doi.org/10.1093/brain/awl262.

Lerch, J.P., van der Kouwe, A.J.W., Raznahan, A., Paus, T., Johansen-Berg, H., Miller, K.L., Smith, S.M., Fischl, B., Sotiropoulos, S.N., 2017. Studying neuroanatomy using MRI. Nat. Neurosci. 20, 314–326. http://dx.doi.org/10.1038/nn.4501.

Linkersdörfer, J., Lonnemann, J., Lindberg, S., Hasselhorn, M., Fiebach, C.J., 2012. Grey matter alterations co-localize with functional abnormalities in developmental dyslexia: an ALE meta-analysis. PLoS One 7, e43122. http://dx.doi.org/10.1371/journal.pone.0043122.

Liu, L., You, W., Wang, W., Guo, X., Peng, D., Booth, J., 2013. Altered brain structure in Chinese dyslexic children. Neuropsychologia 51, 1169–1176. http://dx.doi.org/10.1016/j.neuropsychologia.2013.03.010.

Lizarazu, M., Lallier, M., Molinaro, N., Bourguignon, M., Paz-Alonso, P.M., Lerma-Usabiaga, G., Carreiras, M., 2015. Developmental evaluation of atypical auditory sampling in dyslexia: Functional and structural evidence. Hum. Brain Mapp. 36 (12), 4986–5002. http://dx.doi.org/10.1002/hbm.22986.

Loftus, W.C., Tramo, M.J., Thomas, C.E., Green, R.L., Nordgren, R.A., Gazzaniga, M.S., 1993. Three-dimensional quantitative analysis of hemispheric asymmetry in the human superior temporal region. Cereb. Cortex N. Y. N 1991 (3), 348–355.

Lysiak-Seichter, M., Gredes, M., Proff, P., Gedrange, T., Splieth, C.H., 2006. Evaluation of disparities in physical development and dental age of two specific groups of boys. Eur. J. Paediatr. Dent. Off. J. Eur. Acad. Paediatr. Dent. 7, 169–173.

Ma, Y., Koyama, M.S., Milham, M.P., Castellanos, F.X., Quinn, B.T., Pardoe, H., Wang, X., Kuzniecky, R., Devinsky, O., Thesen, T., Blackmon, K., 2015. Cortical thickness abnormalities associated with dyslexia, independent of remediation status. NeuroImage Clin. 7, 177–186. http://dx.doi.org/10.1016/j.nicl.2014.11.005.

Mangin, J.F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., Scifo, P., Ochiai, T., Brunelle, F., Regis, J., 2004. A framework to study the cortical folding patterns. Neuroimage 23 (Suppl. 1), S129–38.

McCarthy, C.S., Ramprashad, A., Thompson, C., Botti, J.-A., Coman, I.L., Kates, W.R., 2015. A comparison of FreeSurfer-generated data with and without manual intervention. Brain Imaging Methods 379. http://dx.doi.org/10.3389/fnins.2015.00379.

McDaniel, M.A., 2005. Big-brained people are smarter: a meta-analysis of the relationship between in vivo brain volume and intelligence. Intelligence 33, 337–346.

Melby-Lervåg, M., Lyster, S.-A.H., Hulme, C., 2012. Phonological skills and their role in learning to read: a meta-analytic review. Psychol. Bull. 138, 322.

Menghini, D., Hagberg, G.E., Petrosini, L., Bozzali, M., Macaluso, E., Caltagirone, C., Vicari, S., 2008. Structural correlates of implicit learning deficits in subjects with developmental dyslexia. Ann. N. Y. Acad. Sci. 1145, 212–221.

Morgan, W.P., 1896. A case of congenital word blindness. Br. Med. J. 2, 1378.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., Zilles, K., 2001. Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13, 684–701.

Morris, R.D., Stuebing, K.K., Fletcher, J.M., Shaywitz, S.E., Lyon, G.R., Shankweiler, D.P., Katz, L., Francis, D.J., Shaywitz, B.A., 1998. Subtypes of reading disability: variability around a phonological core. J. Educ. Psychol. 90, 347–373.

Myers, C.A., Vandermosten, M., Farris, E.A., Hancock, R., Gimenez, P., Black, J.M., Casto, B., Drahos, M., Tumber, M., Hendren, R.L., Hulme, C., Hoeft, F., 2014. White matter morphometric changes uniquely predict children's reading acquisition. Psychol. Sci. 25, 1870–1883. http://dx.doi.org/10.1177/0956797614544511.

Nagy, Z., Westerberg, H., Klingberg, T., 2004. Maturation of white matter is associated with the development of cognitive functions during childhood. J. Cogn. Neurosci. 16, 1227–1233.

Niogi, S.N., McCandliss, B.D., 2006. Left lateralized white matter microstructure accounts for individual differences in reading ability and disability. Neuropsychologia 44, 2178–2188.

Norton, E.S., Beach, S.D., Gabrieli, J.D.E., 2015. Neurobiology of dyslexia. Curr. Opin. Neurobiol. 30, 73–78. http://dx.doi.org/10.1016/j.conb.2014.09.007.

Odegard, T.N., Farris, E.A., Ring, J., McColl, R., Black, J., 2009. Brain connectivity in non-reading impaired children and children diagnosed with developmental dyslexia. Neuropsychologia 47, 1972–1977.

Płoński, P., Gradkowski, W., Altarelli, I., Monzalvo, K., van Ermingen-Marbach, M., Grande, M., Heim, S., Marchewka, A., Bogorodzki, P., Ramus, F., Jednoróg, K., 2017. Multi-parameter machine learning approach to the neuroanatomical basis of developmental dyslexia. Hum. Brain Mapp. 38, 900–908. http://dx.doi.org/10.1002/hbm.23426.

Panizzon, M.S., Fennema-Notestine, C., Eyler, L.T., Jernigan, T.L., Prom-Wormley, E., Neale, M., Jacobson, K., Lyons, M.J., Grant, M.D., Franz, C.E., Xian, H., Tsuang, M., Fischl, B., Seidman, L., Dale, A., Kremen, W.S., 2009. Distinct genetic influences on cortical surface area and cortical thickness. Cereb. Cortex 19, 2728–2735.

Paulesu, E., Démonet, J.F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa, S.F., Cossu, G., Habib, M., Frith, C.D., Frith, U., 2001. Dyslexia: cultural diversity and biological unity. Science 291, 2165–2167.

Peelle, J.E., Cusack, R., Henson, R.N., 2012. Adjusting for global effects in voxel-based morphometry: gray matter decline in normal aging. Neuroimage 60, 1503–1516. http://dx.doi.org/10.1016/j.neuroimage.2011.12.086.

Peiffer, A.M., Rosen, G.D., Fitch, R.H., 2002. Rapid auditory processing and MGN morphology in microgyric rats reared in varied acoustic environments. Brain Res. Dev. Brain Res. 138, 187–193.

Peng, X., Lin, P., Zhang, T., Wang, J., 2013. Extreme learning machine-Based classification of ADHD using brain structural MRI data. PLoS One 8, e79476. http://dx.doi.org/10.1371/journal.pone.0079476.

Pennington, B.F., Lefly, D.L., 2001. Early reading development in children at family risk for dyslexia. Child Dev. 72, 816–833.

Pernet, C., Andersson, J., Paulesu, E., Demonet, J.F., 2009a. When all hypotheses are right: a multifocal account of dyslexia. Hum. Brain Mapp. 30, 2278–2292.

Pernet, C.R., Poline, J.B., Demonet, J.F., Rousselet, G.A., 2009b. Brain classification reveals the right cerebellum as the best biomarker of dyslexia. BMC Neurosci. 10, 67.

Plaze, M., Paillère-Martinot, M.-L., Penttilä, J., Januel, D., de Beaurepaire, R., Bellivier, F., Andoh, J., Galinowski, A., Gallarda, T., Artiges, E., Olié, J.-P., Mangin, J.-F., Martinot, J.-L., Cachia, A., 2009. 'Where Do Auditory Hallucinations Come From?'—A Brain Morphometry Study of Schizophrenia Patients With Inner or Outer Space Hallucinations. Schizophr Bull 37 (1), 212–221. http://dx.doi.org/10.1093/schbul/sbp081.

Poelmans, H., Luts, H., Vandermosten, M., Boets, B., Ghesquiere, P., Wouters, J., 2012. Auditory steady state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia. Ear Hear. 33, 134–143. http://dx.doi.org/10.1097/AUD.0b013e31822c26b9.

Poeppel, D., 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as asymmetric sampling in time. Speech Commun. 41, 245–255. http://dx.doi.org/10.1016/S0167-6393(02)00107-3.

Popescu, V., Klaver, R., Versteeg, A., Voorn, P., Twisk, J.W.R., Barkhof, F., Geurts, J.J.G., Vrenken, H., 2016. Postmortem validation of MRI cortical volume measurements in MS. Hum. Brain Mapp. 37, 2223–2233. http://dx.doi.org/10.1002/hbm.23168.

Pugh, K.R., Mencl, W.E., Jenner, A.R., Katz, L., Frost, S.J., Lee, J.R., Shaywitz, S.E., Shaywitz, B.A., 2001. Neurobiological studies of reading and reading disability. J. Commun. Disord. 34, 479–492.

Pugh, K.R., Frost, S.J., Rothman, D.L., Hoeft, F., Tufo, S.N.D., Mason, G.F., Molfese, P.J., Mencl, W.E., Grigorenko, E.L., Landi, N., Preston, J.L., Jacobsen, L., Seidenberg, M.S., Fulbright, R.K., 2014. Glutamate and choline levels predict individual differences in reading ability in emergent readers. J. Neurosci. 34, 4082–4089. http://dx.doi.org/10.1523/JNEUROSCI.3907-13.2014.

Puolakanaho, A., Ahonen, T., Aro, M., Eklund, K., Leppanen, P.H.T., Poikkeus, A.M., Tolvanen, A., Torppa, M., Lyytinen, H., 2007. Very early phonological and language skills: estimating individual risk of reading disability. J. Child Psychol. Psychiatry 48, 923–931.

Qiu, D., Tan, L.-H., Zhou, K., Khong, P.-L., 2008. Diffusion tensor imaging of normal white matter maturation from late childhood to young adulthood: voxel-wise evaluation of mean diffusivity, fractional anisotropy, radial and axial diffusivities, and correlation with reading development. Neuroimage 41, 223–232. http://dx.doi.org/10.1016/j.neuroimage.2008.02.023.

Rae, C., Lee, M.A., Dixon, R.M., Blamire, A.M., Thompson, C.H., Styles, P., Talcott, J., Richardson, A.J., Stein, J.F., 1998. Metabolic abnormalities in developmental dyslexia detected by 1H magnetic resonance spectroscopy. Lancet 351, 1849–1852.

Ramus, F., Altarelli, I., Jednoróg, K., Zhao, J., Scotto di Covella, L., in press Brain asymmetries and sex differences in developmental dyslexia, in: Galaburda, A.M., Gaab, N., Hoeft, F., McCardle, P. (Eds.), Dyslexia and Neuroscience: The Geschwind-Galaburda Hypothesis, 30 Years Later. Brookes; Baltimore, MD.

Ramus, F., 2003. Developmental dyslexia: specific phonological deficit or general

sensorimotor dysfunction? Curr. Opin. Neurobiol. 13, 212–218.
Ramus, F., 2006. A neurological model of dyslexia and other domain-specific developmental disorders with an associated sensorimotor syndrome. In: Rosen, G.D. (Ed.), The Dyslexic Brain: New Pathways in Neuroscience Discovery. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 75–101.
Raschle, N.M., Chang, M., Gaab, N., 2011. Structural brain alterations associated with dyslexia predate reading onset. Neuroimage 57, 742–749.
Raschle, N.M., Becker, B.L.C., Smith, S., Fehlbaum, L.V., Wang, Y., Gaab, N., 2017. Investigating the influences of language delay and/or familial risk for dyslexia on brain structure in 5-year-olds. Cereb. Cortex 27, 764–776. http://dx.doi.org/10.1093/cercor/bhv267.
Richards, T., Stevenson, J., Crouch, J., Johnson, L.C., Maravilla, K., Stock, P., Abbott, R., Berninger, V., 2008. Tract-based spatial statistics of diffusion tensor imaging in adults with dyslexia. Am. J. Neuroradiol. 29, 1134–1139. http://dx.doi.org/10.3174/ajnr.A1007.
Richardson, F.M., Price, C.J., 2009. Structural MRI studies of language function in the undamaged brain. Brain Struct. Funct. 213, 511–523. http://dx.doi.org/10.1007/s00429-009-0211-y.
Richardson, A.J., Cox, I.J., Sargentoni, J., Puri, B.K., 1997. Abnormal cerebral phospholipid metabolism in dyslexia indicated by phosphorus-31 magnetic resonance spectroscopy. NMR Biomed. 10, 309–314. http://dx.doi.org/10.1002/(SICI)1099-1492(199710)10:7<309:AID-NBM484>3.0.CO;2-0.
Richlan, F., Kronbichler, M., Wimmer, H., 2013. Structural abnormalities in the dyslexic brain: a meta-analysis of voxel-based morphometry studies. Hum. Brain Mapp. 34, 3055–3065.
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, 739–749. http://dx.doi.org/10.1016/j.cortex.2009.07.008.
Robichon, F., Bouchard, P., Demonet, J., Habib, M., 2000a. Developmental dyslexia: re-evaluation of the corpus callosum in male adults. Eur. Neurol. 43, 233–237.
Robichon, F., Levrier, O., Farnarier, P., Habib, M., 2000b. Developmental dyslexia: atypical cortical asymmetries and functional significance. Eur. J. Neurol. 7, 35–46.
Roof, R.L., Duvdevani, R., Braswell, L., Stein, D.G., 1994. Progesterone facilitates cognitive recovery and reduces secondary neuronal loss caused by cortical contusion injury in male rats. Exp. Neurol. 129, 64–69.
Rumsey, J.M., Dorwart, R., Vermess, M., Denckla, M.B., Kruesi, M.J., Rapoport, J.L., 1986. Magnetic resonance imaging of brain anatomy in severe developmental dyslexia. Arch. Neurol. 43, 1045–1046.
Rumsey, J.M., Donohue, B.C., Brady, D.R., Nace, K., Giedd, J.N., Andreason, P., 1997. A magnetic resonance imaging study of planum temporale asymmetry in men with developmental dyslexia. Arch. Neurol. 54 (12), 1481–1489.
Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J.F., Bricout, L., Billard, C., N'Guyen-Morel, M.A., Le Heuzey, M.-F., Soares-Boucaud, I., George, F., Ziegler, J.C., Ramus, F., 2016. Phonological skills, visual attention span, and visual stress in developmental dyslexia: insights from a population of French children. Dev. Psychol. 52, 1503–1516.
Scholz, J., Klein, M.C., Behrens, T.E., Johansen-Berg, H., 2009. Training induces changes in white-matter architecture. Nat. Neurosci. 12, 1370–1371.
Schultz, R.T., Cho, N.K., Staib, L.H., Kier, L.E., Fletcher, J.M., Shaywitz, S.E., Shankweiler, D.P., Katz, L., Gore, J.C., Duncan, J.S., Shaywitz, B.A., 1994. Brain morphology in normal and dyslexic children: the influence of sex and age. Ann. Neurol. 35, 732–742. http://dx.doi.org/10.1002/ana.410350615.
Scotto di Covella, L., Leroy, F., Cauvet, E., Girard, F., Couvignou, M., Caron-Guyon, J., Altarelli, I., Monzalvo, K., Dehaene-Lambertz, G., Rivière, D., Marchewka, A., Jednorog, K., Heim, S., Van Ermingen-Marbach, M., Ramus, F., submitted. Morphometry of three sulci in developmental dyslexia.
Semrud-Clikeman, M., Hynd, G.W., Novey, E.S., Eliopulos, D., 1991. Dyslexia and brain morphology: relationships between neuroanatomical variation and neurolinguistic tasks. Learn. Individ. Differ. 3, 225–242. http://dx.doi.org/10.1016/1041-6080(91)90009-P. Special Issue: Integrating Cognitive and Neurodevelopmental Approaches I.
Semrud-Clikeman, M., Hooper, S.R., Hynd, G.W., Hern, K., Presley, R., Watson, T., 1996. Prediction of group membership in developmental dyslexia, attention deficit hyperactivity disorder, and normal controls using brain morphometric analysis of magnetic resonance imaging. Arch. Clin. Neuropsychol. 11, 521–528. http://dx.doi.org/10.1093/arclin/11.6.521.
Shaw, P., Malek, M., Watson, B., Sharp, W., Evans, A., Greenstein, D., 2012. Development of cortical surface area and gyrification in attention-deficit/hyperactivity disorder. Biol. Psychiatry Oxytocin Soc. Bonds Dev. 72, 191–197. http://dx.doi.org/10.1016/j.biopsych.2012.01.031.
Shen, S., Sterr, A., 2013. Is DARTEL-based voxel-based morphometry affected by width of smoothing kernel and group size? A study using simulated atrophy. J. Magn. Reson. Imaging 37, 1468–1475. http://dx.doi.org/10.1002/jmri.23927.
Shu, H., McBride-Chang, C., Wu, S., Liu, H.Y., 2006. Understanding Chinese developmental dyslexia: morphological awareness as a core cognitive construct. J. Educ. Psychol. 98, 122–133.
Silani, G., Frith, U., Demonet, J.F., Fazio, F., Perani, D., Price, C., Frith, C.D., Paulesu, E., 2005. Brain abnormalities underlying altered activation in dyslexia: a voxel based morphometry study. Brain 128, 2453–2461.
Siok, W.T., Niu, Z., Jin, Z., Perfetti, C.A., Tan, L.H., 2008. A structural-functional basis for dyslexia in the cortex of Chinese readers. Proc. Natl. Acad. Sci. U. S. A. 105, 5561–5566.
Siok, W.T., Spinks, J.A., Jin, Z., Tan, L.H., 2009. Developmental dyslexia is characterized by the co-existence of visuospatial and phonological disorders in Chinese children. Curr. Biol. 19, R890–2.
Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.J., 2006. Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. Neuroimage 31, 1487–1505. http://dx.doi.org/10.1016/j.neuroimage.2006.02.024.
Smith, S.M., Johansen-Berg, H., Jenkinson, M., Rueckert, D., Nichols, T.E., Miller, K.L., Robson, M.D., Jones, D.K., Klein, J.C., Bartsch, A.J., Behrens, T.E.J., 2007. Acquisition and voxelwise analysis of multi-subject diffusion data with tract-based spatial statistics. Nat. Protoc. 2, 499–503. http://dx.doi.org/10.1038/nprot.2007.45.
Snowling, M.J., Goulandris, N., Bowlby, M., Howell, P., 1986. Segmentation and speech perception in relation to reading skill: a developmental analysis. J. Exp. Child Psychol. 41, 489–507.
Snowling, M.J., Muter, V., Carroll, J., 2007. Children at family risk of dyslexia: a follow-up in early adolescence. J. Child Psychol. Psychiatry 48, 609–618. http://dx.doi.org/10.1111/j.1469-7610.2006.01725.x.
Spironelli, C., Penolazzi, B., Angrilli, A., 2010. Gender differences in reading in school-aged children: an early ERP study. Dev. Neuropsychol. 35, 357–375. http://dx.doi.org/10.1080/87565641.2010.480913.
Steinbrink, C., Klatte, M., 2008. Phonological working memory in german children with poor reading and spelling abilities. Dyslexia 14, 271–290.
Steinbrink, C., Vogt, K., Kastrup, A., Muller, H.P., Juengling, F.D., Kassubek, J., Riecker, A., 2008. The contribution of white and gray matter differences to developmental dyslexia: insights from DTI and VBM at 3. 0 T. Neuropsychologia 46, 3170–3178.
Su, M., Zhao, J., Thiebaut de Schotten, M., Zhou, W., Gong, G., Ramus, F., Shu, H., submitted. Alterations in white matter pathways underlying phonological and morphological processing in Chinese developmental dyslexia.
Szalkowski, C.E., Fiondella, C.G., Galaburda, A.M., Rosen, G.D., LoTurco, J.J., Fitch, R.H., 2012. Neocortical disruption and behavioral impairments in rats following in utero RNAi of candidate dyslexia risk gene Kiaa0319. Int. J. Dev. Neurosci. 30, 293–302. http://dx.doi.org/10.1016/j.ijdevneu.2012.01.009.
Szalkowski, C.E., Booker, A.B., Truong, D.T., Threlkeld, S.W., Rosen, G.D., Fitch, R.H., 2013. Knockdown of the candidate dyslexia susceptibility gene homolog dyx1c1 in rodents: effects on auditory processing, visual attention, and cortical and thalamic anatomy. Dev. Neurosci. 35 (1), 50–68. http://dx.doi.org/10.1159/000348431.
Szucs, D., Ioannidis, J.P.A., 2017. Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. PLoS Biology 15 (3), e2000797. http://dx.doi.org/10.1371/journal.pbio.2000797.
Tamboer, P., Scholte, H.S., Vorst, H.C., 2015. Dyslexia and voxel-based morphometry: correlations between five behavioural measures of dyslexia and gray and white matter volumes. Ann. Dyslexia 65, 121–141.
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 Clin. 11, 508–514. http://dx.doi.org/10.1016/j.nicl.2016.03.014.
Thiebaut de Schotten, M., Dell'Acqua, F., Forkel, S.J., Simmons, A., Vergani, F., Murphy, D.G.M., Catani, M., 2011. A lateralized brain network for visuospatial attention. Nat. Neurosci. 14, 1245–1246. http://dx.doi.org/10.1038/Nn.2905.
Threlkeld, S.W., McClure, M.M., Rosen, G.D., Fitch, R.H., 2006. Developmental time-frames for induction of microgyria and rapid auditory processing deficits in the rat. Brain Res. 1109, 22–31.
Tournier, J.-D., Calamante, F., Gadian, D.G., Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. Neuroimage 23, 1176–1185. http://dx.doi.org/10.1016/j.neuroimage.2004.07.037.
Vanderauwera, J., Vandermosten, M., Dell'Acqua, F., Wouters, J., Ghesquière, P., 2015. Disentangling the relation between left temporoparietal white matter and reading: a spherical deconvolution tractography study. Hum. Brain Mapp. 36, 3273–3287. http://dx.doi.org/10.1002/hbm.22848.
Vanderauwera, J., Altarelli, I., Vandermosten, M., De Vos, A., Wouters, J., & Ghesquière, P. (in press). Atypical Structural Asymmetry of the Planum Temporale is Related to Family History of Dyslexia. Cerebral Cortex. https://doi.org/10.1093/cercor/bhw348.
Vandermosten, M., Boets, B., Poelmans, H., Sunaert, S., Wouters, J., Ghesquière, P., 2012a. A tractography study in dyslexia: neuroanatomic correlates of orthographic, phonological and speech processing. Brain 135, 935–948. http://dx.doi.org/10.1093/brain/awr363.
Vandermosten, M., Boets, B., Wouters, J., Ghesquière, P., 2012b. A qualitative and quantitative review of diffusion tensor imaging studies in reading and dyslexia. Neurosci. Biobehav. Rev. 36, 1532–1552. http://dx.doi.org/10.1016/j.neubiorev.2012.04.002.
Vandermosten, M., Vanderauwera, J., Theys, C., De Vos, A., Vanvooren, S., Sunaert, S., Wouters, J., Ghesquière, P., 2015. A DTI tractography study in pre-readers at risk for dyslexia. Dev. Cogn. Neurosci. 14, 8–15. http://dx.doi.org/10.1016/j.dcn.2015.05.006.
Vidyasagar, T.R., Pammer, K., 2010. Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. Trends Cogn. Sci. 14, 57–63.
Vinckenbosch, E., Robichon, F., Eliez, S., 2005. Gray matter alteration in dyslexia: converging evidence from volumetric and voxel-by-voxel MRI analyses. Neuropsychologia 43, 324–331.
Wallace, G.L., Robustelli, B., Dankner, N., Kenworthy, L., Giedd, J.N., Martin, A., 2013. Increased gyrification, but comparable surface area in adolescents with autism spectrum disorders. Brain 136, 1956–1967. http://dx.doi.org/10.1093/brain/awt106.
Wang, Y., Mauer, M.V., Raney, T., Peysakhovich, B., Becker, B.L.C., Sliva, D.D., Gaab, N., 2016. Development of tract-specific white matter pathways during early reading development in at-risk children and typical controls. Cereb. Cortex. http://dx.doi.org/10.1093/cercor/bhw095.
Watkins, K.E., Paus, T., Lerch, J.P., Zijdenbos, A., Collins, D.L., Neelin, P., Taylor, J.,

Worsley, K.J., Evans, A.C., 2001. Structural asymmetries in the human brain: a voxel-based statistical analysis of 142 MRI scans. Cereb. Cortex 11, 868–877. http://dx.doi.org/10.1093/cercor/11.9.868.

Wechsler, D., 2005. WISC-IV: Echelle d'Intelligence de Wechsler pour Enfants – Quatrième Edition. Editions du Centre de Psychologie Appliquée, Paris.

White, S., Milne, E., Rosen, S., Hansen, P.C., Swettenham, J., Frith, U., Ramus, F., 2006. The role of sensorimotor impairments in dyslexia: a multiple case study of dyslexic children. Dev. Sci. 9, 237–255.

White, T., Su, S., Schmidt, M., Kao, C.-Y., Sapiro, G., 2010. The development of gyrification in childhood and adolescence. Brain Cogn. Adolesc. Brain Dev.: Curr. Themes Future Dir. 72, 36–45. http://dx.doi.org/10.1016/j.bandc.2009.10.009.

Winkler, A.M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P.T., Duggirala, R., Glahn, D.C., 2010. Cortical thickness or grey matter volume?: The importance of selecting the phenotype for imaging genetics studies. Neuroimage 53, 1135–1146.

Woo, C.-W., Chang, L.J., Lindquist, M.A., Wager, T.D., 2017. Building better biomarkers: brain models in translational neuroimaging. Nat. Neurosci. 20, 365–377. http://dx.doi.org/10.1038/nn.4478.

World Health Organization, 2011. International Statistical Classification of Diseases and Related Health Problems – 10th Revision, 4th ed. World Health Organization, Geneva.

Xia, Z., Hoeft, F., Zhang, L., Shu, H., 2016. Neuroanatomical anomalies of dyslexia: disambiguating the effects of disorder, performance, and maturation. Neuropsychologia 81, 68–78.

Yang, Y.-H., Yang, Y., Chen, B.-G., Zhang, Y.-W., Bi, H.-Y., 2016. Anomalous cerebellar anatomy in chinese children with dyslexia. Front. Psychol. 7. http://dx.doi.org/10.3389/fpsyg.2016.00324.

Yeatman, J.D., Dougherty, R.F., Rykhlevskaia, E., Sherbondy, A.J., Deutsch, G.K., Wandell, B.A., Ben-Shachar, M., 2011. Anatomical properties of the arcuate fasciculus predict phonological and reading skills in children. J. Cogn. Neurosci. 23, 3304–3317. http://dx.doi.org/10.1162/jocn_a_00061.

Yeatman, J.D., Dougherty, R.F., Myall, N.J., Wandell, B.A., Feldman, H.M., 2012. Tract profiles of white matter properties: automating fiber-tract quantification. PLoS One 7, e49790.

Zadina, J.N., Corey, D.M., Casbergue, R.M., Lemen, L.C., Rouse, J.C., Knaus, T.A., Foundas, A.L., 2006. Lobar asymmetries in subtypes of dyslexic and control subjects. J. Child Neurol. 21, 922–931.

Zhao, J., Thiebaut de Schotten, M., Altarelli, I., Dubois, J., Ramus, F., 2016. Altered hemispheric lateralization of white matter tracts in developmental dyslexia: evidence from spherical deconvolution tractography. Cortex 76, 51–62.