DCDC2, KIAA0319 and CMIP Are Associated with Reading-Related Traits

Tom S. Scerri, Andrew P. Morris, Lyn-Louise Buckingham, Dianne F. Newbury, Laura L. Miller, Anthony P. Monaco, Dorothy V.M. Bishop, and Silvia Paracchini

Background: Several susceptibility genes have been proposed for dyslexia (reading disability; RD) and specific language impairment (SLI). RD and SLI show comorbidity, but it is unclear whether a common genetic component is shared.

Methods: We have investigated whether candidate genes for RD and SLI affect specific cognitive traits or have broad effect on cognition. We have analyzed common risk variants within RD (*MRPL19/C2ORF3*, *KIAA0319*, and *DCDC2*) and language impairment (*CMIP* and *ATP2C2*) candidate loci in the Avon Longitudinal Study of Parents and Children cohort (n = 3725), representing children born in southwest England in the early 1990s.

Results: We detected associations between reading skills and *KIAA0319*, *DCDC2*, and *CMIP*. We show that *DCDC2* is specifically associated with RD, whereas variants in *CMIP* and *KIAA0319* are associated with reading skills across the ability range. The strongest associations were restricted to single-word reading and spelling measures, suggesting that these genes do not extend their effect to other reading and language-related skills. Inclusion of individuals with comorbidity tends to strengthen these associations. Our data do not support *MRPL19/C20RF3* as a locus involved in reading abilities nor *CMIP/ATP2C2* as genes regulating language skills.

Conclusions: We provide further support for the role of *KIAA0319* and *DCDC2* in contributing to reading abilities and novel evidence that the language-disorder candidate gene *CMIP* is also implicated in reading processes. Additionally, we present novel data to evaluate the prevalence and comorbidity of RD and SLI, and we recommend not excluding individuals with comorbid RD and SLI when designing genetic association studies for RD.

Key Words: ALSPAC, association study, dyslexia, language, reading abilities, specific language impairment (SLI)

yslexia (or reading disability, RD) and SLI are common childhood disorders. RD is a specific deficit in learning to read, whereas SLI refers to an impairment in the acquisition of oral language (1). The biological cause of RD and SLI remains poorly understood, but it is clear that their manifestation is the result of multiple interacting factors, many of which have a genetic origin. Family studies have reported that, for both disorders, first-degree relatives of affected individuals also have a 30% to 50% chance of being affected, whereas the general population prevalence is approximately 5% (2,3). Comorbidity between RD and SLI has been consistently reported. Estimates indicate that 43% of children with SLI are later diagnosed with RD (4), and up to 55% of children with RD meet criteria for SLI (5). These figures have led to the hypothesis that SLI and RD may be manifestations of the same underlying deficit or may share etiologic factors, such as genetic determinants (1). Both SLI and RD show increased comorbidity with attentiondeficit/hyperactivity disorder (ADHD), another common neurodevelopmental disorder, affecting 3% to 5% children (6). It is estimated that 25% to 40% of children with RD manifest symptoms of ADHD as well (7), and children with language disorder are at higher risk of developing ADHD (8).

Several genes have been proposed as susceptibility candidates for either RD or language-related skills and have been extensively

Address correspondence to Silvia Paracchini, D.Phil., Wellcome Trust Centre for Human Genetics, OX3 7BN Oxford, United Kingdom; E-mail: silviap@ well.ox.ac.uk.

Received Oct 26, 2010; revised Jan 18, 2011; accepted Feb 5, 2011.

reviewed (9–11). The RD candidates include the *MRPL19/C2ORF3* locus, *ROBO1*, *KIAA0319*, *DCDC2*, and *DYX1C1*. With the exception of *ROBO1* (12), these genes are supported by genetic associations with common single nucleotide polymorphisms (SNPs).

DYX1C1 was the first RD candidate to be identified following breakpoint mapping of a translocation cosegregating with RD (13). Association analysis in a cohort with RD implicated two putative coding variants: the -3A (rs3743205) and the 1249T (rs57809907) variants. A large number of replication studies have not reached consensus in supporting *DYX1C1* RD susceptibility variants (14–22).

More consistent observations have been reported for the KIAA0319 and DCDC2 genes located at the chromosome 6 locus. Most of the associations with KIAA0319 cluster around the 5' end of this gene and generally show the same allelic trend across independent studies (23-27). Functional studies showed that one particular RD-risk haplotype, effectively tagged by the minor allele of the rs2143340 SNP, is associated with reduced expression of KIAA0319 (28). This haplotype also harbors the minor allele of rs9461045, creating a binding site for a nuclear protein, which could explain the reduced gene expression and provides a functional mechanism underlying the genetic associations (29). Other studies did not find associations within KIAA0319 but identified the nearby DCDC2 gene as an RD candidate (30,31). Replication studies in samples selected for RD provided further support for DCDC2 but with modest associations (14,23,32,33). Rare variants located between these two genes have been found to be associated with speech perception in children with dyslexia (34).

The *MRPL19* and *C2ORF3* genes, which appear to be coregulated, are supported by single intergenic SNPs and overlapping haplotypes yielding significant associations in two independent samples of Finnish and German origin (35).

The candidate genes for language include *CMIP*, *ATP2C2*, and *CNTNAP2*. *CMIP* and *ATP2C2* have been associated with nonword repetition, which is regarded as a measure of phonologic short-

From the Wellcome Trust Centre for Human Genetics, University of Oxford, UK (TSS, APM, L-LB, DFN, APM, SP); School of Social and Community Medicine, University of Bristol, UK (LLM); Department of Experimental Psychology, University of Oxford, UK (DVMB).

term memory, in samples of individuals with language impairment (36). Both genes were identified following high-density mapping at the chromosome 16 locus for SLI (37). The associations, originally identified in a cohort of individuals with SLI, were also seen in a subgroup of individuals selected on the basis of low language skills from the Avon Longitudinal Study of Parents and Children (AL-SPAC) cohort (38), albeit with an opposite direction of trend for *CMIP*. Instead, no associations were detected with language traits in the entire ALSPAC cohort, which is representative of the general population. This suggests that the two genes have an effect on nonword repetition on a background of language impairment.

CNTNAP2 genetic variants were found to be associated with language-related phenotypes and a task of verbal short-term memory in the same language-impaired cohort used for *CMIP* and *ATP2C2* (39). *CNTNAP2* is a target of FOXP2 (39), which is implicated in severe and rare forms of language impairment (40).

The use of epidemiologic cohorts has proved to be a valid approach to investigate further genetic associations with some of these genes. The *KIAA0319* RD-associated haplotype (26) was significantly associated with reading skills in both ALSPAC (41) and in a twin-based Australian sample (42), but with an opposite trend in the latter. The same Australian sample was also used to investigate *DCDC2* (43) and *DYX1C1* (44).

Phenotype definition is a key component when investigating the genetics of language and reading disorders. Tests of singleword reading are the most commonly used measures in genetic association studies of RD (45). Nonword repetition is a good marker for heritable SLI (46). However, an important issue is how far language problems should be identified solely by psychometric tests, which may miss key features of communication difficulties. Parental reports can be highly effective in identifying heritable communication problems (47,48), but they typically identify a different subset of children than those identified on direct language testing (49).

Bishop and Snowling (50) noted that RD and SLI were for a long time regarded as distinct disorders but in recent years have been reconceptualized as points on a continuum. This is an oversimplification, because different components of language and reading skills can fractionate, but it is possible that the same genetic components could contribute to both disorders and explain, at least partially, the observed comorbidity. Previous studies exploring the role of shared genes in contributing to both RD and SLI indicate KIAA0319 as a possible common risk factor supported by associations with language-related measures in samples selected for language impairment (51,52). One of these studies also showed that CMIP was associated with both reading and language-related measures in the same sample selected for language impairment (52). No association with language measures was reported for an investigation of DCDC2 and DYX1C1 in a sample of families ascertained for dyslexia (53). A genetic overlap has been suggested for RD and ADHD by a linkage study (54), and DCDC2 has been suggested to contribute to both RD and ADHD (55). ATP2C2 has also been found associated with ADHD (56).

Here, we investigated in the ALSPAC cohort genetic associations reported in the literature from samples selected for either RD or SLI. We conducted association analysis to 1) replicate associations with reading and nonword repetition measures, 2) dissect the phenotypic components of such associations by testing different but related quantitative phenotypes to pinpoint the underlying cognitive deficit(s), and 3) test for pleiotropic effects across reading- and language-related measures. We identified association between reading abilities and the *DCDC2*, *KIAA0319*, and *CMIP* genes and have shown that these associations follow different patterns; whereas *DCDC2* is associated more specifically with dyslexia, *CMIP*

and *KIAA0319* are associated with reading abilities in the normal range. In addition, we show that these genes have a specific effect on a test of single-word reading rather than a more generalized impact. Lastly, we have evaluated the effect of individuals with comorbid RD and SLI in association analysis. Our results suggest that the inclusion of these individuals may increase power in genetic association studies for dyslexia.

Methods and Materials

We genotyped the ALSPAC children cohort ($n \sim 11,000$) using either Sequenom iPLEX assays (San Diego, California) or the KBiosciences (Herts, United Kingdom) service using their in-house technology. Nineteen SNPs passed the quality control criteria of a call rate greater than 90%, error rate less than 1.5% (estimate derived from approximately 3% of samples blindly distributed in duplicates), minor allele frequency greater than .05, and genotype frequencies in Hardy–Weinberg equilibrium (p > .05). We had good quality data for SNPs within *MRPL19/C2ORF3*, *KIAA0319*, *DCDC2*, *ATP2C2*, and *CMIP* but not for *DYX1C1* and *CNTNAP2*, which therefore were not included in this analysis.

Both quantitative and case—control analyses were performed within PLINK (57) testing for additive effects.

We based our initial analysis on the F1 sample (Figure 1; Methods in Supplement 1), which includes all available individuals after filtering for missing data, ethnicity, IQ, and autistic traits. Individuals were then assigned to the groups of RD, SLI, ADHD, any of the four comorbid combinations of these three disorders, or unaffected (Figure 1, Table 1). See Methods in Supplement 1 for full description of sample subgroups.

Results

Observed Disorder Prevalence

The ALSPAC children were assigned to one of eight affection status subgroups (Methods in Supplement 1); unaffected, RD, SLI, ADHD, or one of the four comorbid combinations (Figure 2, Table 1). From the initial sample, we filtered sequentially for missing data, ethnicity, performance IQ, and signs of autism, selecting 3725 individuals to calculate disorder prevalences. The prevalence of RD (6.04%) and SLI (6.44%) in this subgroup of the ALSPAC cohort is comparable to other studies (2,58), and the prevalence of ADHD (1.05%) is lower than previous reports of approximately 5% (59). The low prevalence of ADHD is explained primarily by a conservative assignment criterion but a specific dropout of children with ADHD from the ALSPAC study has also been suggested (60). Levels of comorbidity (Table 1) were comparable to other studies (1,9), but our conservative criterion for ADHD would have an impact on the rates of comorbidity with ADHD in this sample.

The quantitative measures selected for either ascertainment criteria or association analysis (Table 2 and Table S1 in Supplement 1) show various degrees of correlation (Table S2 in Supplement 1). A strong correlation (.514 $\leq r \leq$.814) was observed across the read-ing-related measures (excluding MEMSPAN). Low correlation was observed across the language-related measures (.099 $\leq r \leq$.197), and NW_REPT showed higher correlation with the reading measures. This is consistent with observations in our cohort of families with SLI (36). This also fits with the notion that the different language tests measure distinct language components and identify different groups of impaired individuals (49).

Quantitative Genetic Analysis

We analyzed 19 SNPs for association with READ and NW_REPT to replicate previous findings with RD and SLI, respectively (Table S3 in

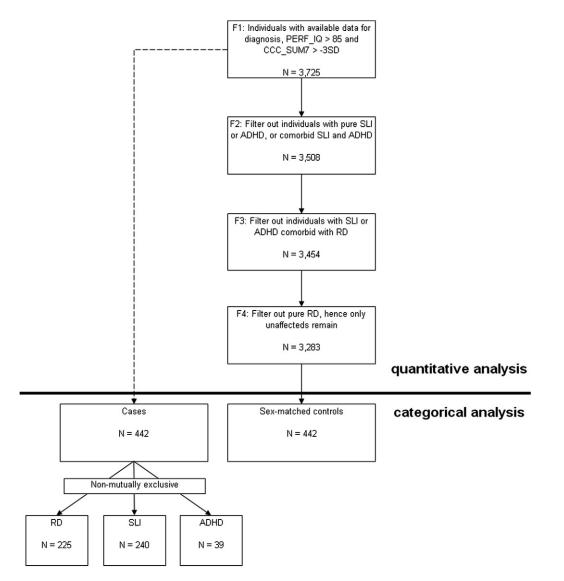


Figure 1. Diagram illustrating how phenotypic subgroups were identified. The subgroups above the black horizontal lines were used for quantitative analysis while the ones below were used for case-control analysis. The extent of co-morbidity (hence the non-mutually exclusive definition of cases), can be seen in Figure 2. ADHD, attention-deficit/hyperactivity disorder; CCC_SUM7, sum of first seven scales from the Children's Communication Checklist; PERF_IQ, performance IQ; RD, reading disability; SLI, specific language impairment.

Supplement 1). This analysis was conducted in the F1 sample (Figure 1). *DCDC2*, *KIAA0319*, and *CMIP* showed associations with READ (Table 3 and Table S4 in Supplement 1). The association with rs2143340 (*KIAA0319*) was statistically significant (p < .0023) and in

Affection Status	Frequency	%
Unaffected	3283	88.13
RD	171	4.59
SLI	186	4.99
ADHD	26	.70
RD and SLI	46	1.23
RD and ADHD	5	.13
SLI and ADHD	5	.13
RD, SLI, and ADHD	3	.08

ADHD, attention-deficit hyperactivity disorder; RD, reading disability; SLI, specific language impairment.

the same direction as previously reported. The only signal observed for NW_REPT was with the *DCDC2* rs793862 marker (p = .03).

To follow up the results observed with READ and NW_REPT, we analyzed the SNPs showing *p* values < .05 with the other available reading and language-related measures (Table S5 in Supplement 1). We detected the strongest associations with SPELL (*DCDC2*, *KIAA0319*, and *CMIP*; Table 4) and other weak signals with NW_READ (*DCDC2* and *KIAA0319*; minimum p = .01) and MEMSPAN (*CMIP*; minimum p = .03). *DCDC2* yielded slightly stronger associations with SPELL than READ, consistent with previous findings where *DCDC2* was originally identified in a sample of individuals with spelling impairments (31). This analysis suggests that the *KIAA0319*, *DCDC2*, and *CMIP* genes contribute specifically to reading abilities and in particular to single-word reading and single-word spelling tests.

We then tested whether the associations with READ and SPELL in this population cohort were driven by the inclusion of impaired individuals (Figure 1; Methods in Supplement 1). First we tested a

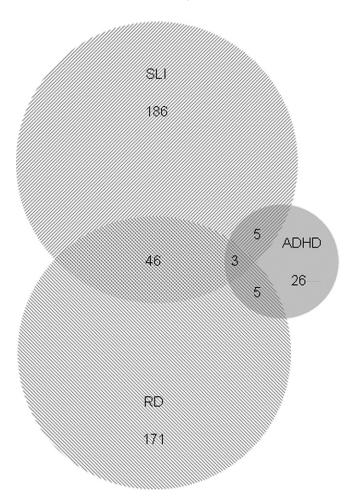


Figure 2. A Venn diagram illustrating the distributions of the cases identified for reading disabilities (RD), specific language impairment (SLI), and attention-deficit/hyperactivity disorder (ADHD) from Sample F1. Circle size is proportional to sample size, and circle overlaps represent comorbidity.

sample that retained the unaffected and RD cases but excluded any cases with pure SLI and/or pure ADHD (F2). Then we removed the cases of RD that had comorbidity with SLI, ADHD, or both (F3).

for association in the unaffected group only (F4).
aled different patterns of association underlying
ed in F1 (Table 4 and Table S4 and S6 in Supple-
lly, the data show that the <i>DCDC2</i> associations are
the small proportion of individuals with RD. For
52 is associated with READ ($p = .004$) and SPELL
subgroup including all cases with RD (F2) and
signal strength to F1 ($p = .006$, READ; $p = .003$,

This analysis revea the results detected ment 1). Specifical indeed driven by example, rs793862 (p = .003) in the showing similar sig SPELL). The associations become progressively weaker when removing the approximately 50 RD cases comorbid with SLI or ADHD $(p \sim .01; F3)$ and in the unaffected group (F4; minimum p = .03). Conversely, the SNPs that showed the strongest associations at the KIAA0319 locus (rs2143340) and in CMIP (rs6564903) had similar effect sizes in the different subgroups with little variation from F1 to F4. Two other SNPs in CMIP (rs12927866 and rs16955705) showed the same pattern. However, two of the other four SNPs tested in KIAA0319 (rs6935076 and rs9461045) showed modest associations with a pattern similar to DCDC2, where association disappeared in the unaffected subgroup (F4).

Finally, we tested

The associations with DCDC2 show the same allelic trends as previously reported (Table 3 and Table S3 in Supplement 1). This is also the case for KIAA0319, with the exception of rs6935076, which showed the opposite trend from the original report in a UK sample of individuals with RD (24). This is surprising because the major allele of rs6935076, which we found to be associated with poor reading, is in high linkage disequilibrium with all the other associated alleles at this locus in populations of European descendent (ALSPAC and our cohort of dyslexic individuals). Regarding CMIP, the original study was conducted in two samples: a cohort of individuals with SLI and a subgroup derived from ALSPAC for being language impaired. The associations showed opposite trend between these two samples (36). Our present associations show a trend consistent with the original report in that ASLPAC subgroup.

In summary, our data suggest that DCDC2 has a specific effect on RD, while CMIP and one variant at the KIAA0319 locus (rs2143340) are significantly associated with general reading abilities. The actual location of rs2143340 is within the gene TTRAP, but it is in linkage disequilibrium with KIAA0319 variants and is tagging the risk haplotype that originally refined the association to KIAA0319 (26). The other two KIAA0319 markers, showing an association pattern suggestive of a more specific role in RD, are instead located in the first intron (rs6935076) or regulatory sequences (rs9461045) of KIAA0319.

Measure	Assignment/ Phenotype ^a	Summary Description	Target Age	Reference
READ ^b	A/P	Single-word reading accuracy	7.5 year	67
READ@9	Α	Single-word reading accuracy	9.5 year	68
SPELL	Р	Single-word spelling accuracy	7.5 year	68
PHONEME	Р	Phoneme awareness	7.5 year	69
NW-READ	Р	Single-non-word reading accuracy	9.5 year	68
MEMSPAN	Р	Working memory	10.5 year	70
WOLD	A/P	Listening and comprehension test	8.5 year	71
NW-REPT ^c	A/P	Phonological short-term memory test	8.5 year	72
CCC-SUM7	A/P	Sum of first seven scales from Children's Communication Checklist	7.5 year	73
Speech/language therapy	А	Child has ever had speech/language therapy	7.6 year	
DAWBA DSM-IV	А	Attention-deficit hyperactivity disorder diagnosis	7.6 year – 8.5 year	74
PERF_IQ	А	Performance IQ	8.5 year	75

See Table S1 in Supplement 1 for more details.

 Table 2.
 Description of Phenotypic Measures

^aSpecifies whether the measure was used for assignment (A) of case status or as phenotype (P) for quantitative analysis.

^bCore measure for RD.

^cCore measure for specific language impairment.

Table 3. Associations Results of the 19 SNPs Tested in F1 with READ and NW_REPT

				REA	D						
Chr.	Gene Locus	SNP	n	β	SE	р	n	β	SE	р	Risk Allele
2	MRPL19/C2ORF3	rs1000585	3,050	.00	.03	.972	3,048	.00	.03	.928	
2	MRPL19/C2ORF3	rs917235	3,165	.00	.03	.949	3,163	02	.03	.353	
2	MRPL19/C2ORF3	rs714939	3,041	.02	.03	.427	3,039	.01	.03	.646	
6	DCDC2	rs793862	3,117	08	.03	.006	3,115	06	.03	.031	A (minor)
6	DCDC2	rs807701	3,193	05	.03	.033	3,191	03	.03	.185	G (minor)
6	DCDC2	rs807724	3,085	07	.03	.015	3,083	03	.03	.257	C (minor)
6	DCDC2	rs1087266	3,198	03	.03	.219	3,196	.00	.03	.915	
6	KIAA0319	rs761100	3,190	03	.03	.211	3,188	01	.03	.603	
6	KIAA0319	rs6935076	3,006	.07	.03	.011	3,004	.02	.03	.482	G (major) ^a
6	KIAA0319	rs2038137	3,053	02	.03	.374	3,051	02	.03	.544	
6	KIAA0319	rs9461045	3,126	08	.03	.024	3,124	03	.03	.368	T (minor)
6	KIAA0319 ^b	rs2143340	3,042	11	.04	.001	3,040	04	.04	.242	G (minor)
16	CMIP	rs12927866	3,055	07	.03	.005	3,053	04	.03	.136	T (minor) ^a
16	CMIP	rs6564903	3,157	08	.02	.002	3,155	02	.02	.360	T (minor) ^a
16	CMIP	rs4265801	3,052	.02	.03	.449	3,050	.03	.03	.289	
16	CMIP	rs16955705	3,050	06	.03	.029	3,048	02	.03	.482	C (minor) ^a
16	ATP2C2	rs16973771	3,009	.01	.03	.691	3,007	.02	.03	.493	
16	ATP2C2	rs2875891	3,049	.00	.03	.950	3,047	.02	.03	.458	
16	ATP2C2	rs8045507	3,046	.00	.03	.979	3,044	.01	.03	.588	

Only one *p* value was statistically significant (< .0023; Methods in Supplement 1) and is highlighted in bold; β (beta) values are standardized and relative to the minor allele (as defined in Table S3 in Supplement 1). Risk allele is reported only for markers showing *p* values < .05.

SNP, single nucleotide polymorphism. ^aOpposite trend compared with original reports (24,36).

^bWithin *TTRAP*.

As well as positive findings, we also report lack of replications. We could not detect associations between NW_REPT and the language candidates *CMIP* and *ATP2C2* in the general population. This is consistent with our previous study, which found associations with nonword repetition for *CMIP* and *ATP2C2* only in a subgroup of individuals with language impairment (36). Our data do not support the role of the *MRPL19/C20RF3* locus in influencing reading abili-

ties. *MRPL19/C2ORF3* was tested using both single markers and haplotypes according to previous reports (35), but none showed any associations (haplotype analysis not shown).

Case-Control Analysis

To test directly for association between the candidate genes and RD or SLI, we analyzed the 19 SNPs in a case—control setting. We

Table 4.	Summar	of Results Showing	Association (p <	.05) with C	Quantitative Measures
----------	--------	--------------------	------------------	-------------	-----------------------

				F1				F2				F3			F	4: Unaf	fected	k	
	Gene									RE	AD								
Chr.	Locus	SNP	Ν	β	SE	Р	Ν	β	SE	Р	Ν	β	SE	Р	Ν	β	SE	Р	Risk Allele
6	DCDC2	rs793862	3,117	08	.03	.006	2,936	09	.03	.004	2,890	08	.03	.010	2,740	06	.03	.042	A (minor)
6	DCDC2	rs807701	3,193	05	.03	.033	3,003	04	.03	.090	2,954	03	.03	.276	2,803	02	.02	.376	G (minor)
6	DCDC2	rs807724	3,085	07	.03	.015	2,898	07	.03	.018	2,850	05	.03	.091	2,700	02	.03	.422	C (minor)
6	KIAA0319	rs6935076	3,006	.07	.03	.011	2,831	.08	.03	.003	2,784	.07	.03	.006	2,646	.05	.02	.028	G (major) ^a
6	KIAA0319	rs9461045	3,126	08	.03	.024	2,947	08	.03	.026	2,901	08	.03	.022	2,752	05	.03	.162	T (minor)
6	KIAA0319 ^b	rs2143340	3,042	11	.04	.001	2,864	12	.04	.001	2,817	12	.04	.001	2,677	11	.03	.001	G (minor)
16	CMIP	rs12927866	3,055	07	.03	.005	2,874	08	.03	.004	2,829	07	.03	.005	2,690	07	.02	.005	T (minor) ^a
16	CMIP	rs6564903	3,157	08	.02	.002	2,966	08	.03	.002	2,919	08	.03	.002	2,768	07	.02	.002	T (minor) ^a
16	CMIP	rs16955705	3,050	06	.03	.029	2,869	06	.03	.022	2,824	06	.03	.019	2,684	05	.02	.027	C (minor) ^a
										SP	ELL								
6	DCDC2	rs793862	3,094	09	.03	.003	2,913	09	.03	.003	2,871	08	.03	.009	2,729	06	.03	.030	A (minor)
6	DCDC2	rs807724	3,065	08	.03	.007	2,878	08	.03	.011	2,834	06	.03	.050	2,691	04	.03	.204	C (minor)
6	KIAA0319 ^b	rs2143340	3,023	10	.04	.004	2,845	10	.04	.005	2,802	11	.04	.004	2,669	10	.04	.006	G (minor)
16	CMIP	rs12927866	3,036	06	.03	.014	2,855	07	.03	.009	2,814	07	.03	.011	2,682	06	.03	.014	T (minor) ^a
16	CMIP	rs6564903	3,136	07	.02	.008	2,945	07	.03	.003	2,901	07	.03	.004	2,758	07	.02	.008	T (minor) ^a
16	CMIP	rs16955705	3,030	06	.03	.026	2,849	06	.03	.019	2,808	06	.03	.017	2,675	06	.03	.027	C (minor) ^a

Only SNPs showing *p* values < .05 in any group tested are reported; *p* values statistically significant (< .0023; Methods in Supplement 1) are in bold; β (beta) values are standardized and relative to the minor allele (as defined in Table S3 in Supplement 1).

^aOpposite trend compared with original reports (24,36).

^bWithin TTRAP.

SNP, single nucleotide polymorphism.

Table 5. Summary of the Results of the Case-Control Analysis

	Gene Locus	SNP			SLI Onl	у	SLI and Comorbid Cases			RD Only			RD and Comorbid Cases			
Chr.			No. of Controls	n	p	Odds Ratio	n	р	Odds Ratio	n	р	Odds Ratio	n	р	Odds Ratio	Risk Allele
2	MRPL19/ C2ORF3	rs917235	375	162	.033	1.33	211	.103	1.22	148	.610	1.07	197	.672	1.06	G (minor)
6	DCDC2	rs793862	375	155	.418	1.13	201	.101	1.26	150	.021	1.42	196	.005	1.47	a (minor)
6	DCDC2	rs807701	379	161	.173	1.21	210	.016	1.36	151	.173	1.21	200	.018	1.36	G (minor)
6	DCDC2	rs807724	371	158	.754	1.05	206	.146	1.24	150	.035	1.40	198	.003	1.52	C (minor)
6	KIAA0319	rs6935076	363	149	.993	1.00	196	.661	.95	138	.026	.72	185	.011	.71	G (major) ^a
6	KIAA0319	rs9461045	375	153	.692	1.08	199	.561	1.10	149	.026	1.47	195	.035	1.40	T (minor)

RD, reading disability; SLI, specific language impairment; SNP, single nucleotide polymorphism.

Only SNPs showing p values < .05 in any of group tested are reported.

^aOpposite trend compared with original report (24).

used four subgroups of cases against a unique control group (see Methods and Materials; Figure 1). The different subgroups of cases included individuals with SLI only, RD only, SLI including cases showing comorbidity for RD and/or ADHD, and RD including cases showing comorbidity for SLI and/or ADHD. The strongest associations were observed for *DCDC2* (Table 5 and Table S7 in Supplement 1) in the cases selected for RD and including individuals with comorbidity with SLI and ADHD (minimum p = .003). Other association signals were observed for *KIAA0319* in the RD cases regardless of comorbidity with SLI and ADHD.

These results complement the findings observed in the quantitative analysis and support the idea that DCDC2 is associated with RD. The SNPs rs793862 and rs807724 consistently showed the strongest associations for DCDC2 in both the quantitative and the case-control analysis. This also agrees with our recent case-control analysis of these candidate genes in samples of individuals with RD where the DCDC2 rs807724 marker showed the strongest association (rs793862 was not tested) (52). The case – control analysis of KIAA0319 also agrees with the quantitative analysis, but the associations were of modest size. Associations in the case-control analysis were detected for rs6935076 and rs9461045 in the RD samples; interestingly, both these SNPs showed an association pattern in the quantitative analysis suggestive of a specific effect on RD. Conversely, the rs2143340 marker, which showed the strongest signal in the quantitative analysis and was associated with variation in the normal range, was not associated with RD in the case-control analysis.

The only other observed signal was for the *MRPL19/C2ORF3* locus showing a weak association with SLI. Our analysis found no role of *MRPL19/C2ORF3* in contributing to RD, nor was there any evidence that *CMIP* or *ATP2C2* influenced SLI.

In summary, our case—control analysis provides support for *DCDC2* and suggestive evidence for *KIAA0319* as candidate genes for RD. Inclusion of cases showing comorbidity between RD and SLI or ADHD contributed to the association signals.

Discussion

We have described a genetic association analysis of candidate genes for RD and SLI based on the ALSPAC children cohort. The large sample size made it possible to conduct the analysis in different sample subgroups to answer specific questions. First, we sought to replicate associations reported in clinical samples, and then we tested whether these associations are detectable with specific or multiple measures to understand whether shared genetic effects contribute to the comorbidity observed between RD and SLI. Our findings support association between *DCDC2*, *KIAA0319*, and *CMIP* specifically with reading measures, but not for associations of *MRPL19/C2ORF3* with RD nor of *ATP2C2* or *CMIP* with language measures. We did not detect any pleiotropic effect, which could partly explain the comorbidity between RD and SLI, although *CMIP*, selected as a candidate for language disorder, showed association with reading.

Our strategy tested whether associations were driven by the most severe individuals. We assessed the contribution to associations of individuals that meet criteria for disorder diagnosis by removing them from the quantitative analysis (Table 4) or evaluating them directly in case-control tests (Table 5). To the best of our knowledge this is the first study that used such a strategy. The association signals we detected were supported by complementary results obtained in the two types of analysis. We show that rs2143340, the most strongly associated marker at the KIAA0319 locus, and CMIP variants are significantly associated with reading and spelling skills regardless of the inclusion of the RD individuals. Consistently, these SNPs did not show associations in the casecontrol analysis, supporting the hypothesis that these variants contribute to reading ability variation in the normal range. Conversely, the associations detected for DCDC2 are driven by the most impaired individuals; the associations disappear from the quantitative analysis when RD cases are removed and DCDC2 showed the strongest associations in case-control analysis. These findings suggest that DCDC2 is associated specifically with RD. A pattern similar to DCDC2 is observed for two KIAA0319 markers. Interestingly, rs9461045, one of these two markers, has a functional effect on the expression of KIAA0319 (29). One could speculate that different genetic variants at the KIAA0319 locus have different effects with some variants involved in the general reading processes and other directly involved in RD.

Association between reading abilities in the general population and *KIAA0319* and *DCDC2* have been reported in previous studies (42,43), including our own analysis of *KIAA0319* in ALSPAC (41). It would be interesting to see whether similar patterns will be observed in the Australian sample (42,43) when removing the most severely impaired individuals.

This is the first study reporting an effect of *CMIP* on the reading abilities of the general population. We previously analyzed *CMIP* and *ATP2C2* in the ALSPAC sample and reported an association with nonword repetition for both genes but only in a specific subgroup of language impaired individuals (36). Both quantitative and case—control analyses were carried out within that specific subgroup, the latter by comparing the two tails of the phenotypic distribution.

In that study, we also failed to detect any effect of *CMIP* and *ATP2C2* on language skills in the entire ALSPAC cohort. Therefore, it is possible that the associations between nonword repetition and *CMIP* and *ATP2C2* can only be detected on a background of language impairment. This is consistent with our current findings suggesting that these two genes cannot be considered as general susceptibility factors for SLI. The association between *CMIP* and reading instead represents a direct replication of our recent findings showing that *CMIP* is associated with reading measures in the same SLI cohort where it was originally found associated with nonword repetition (52).

It has been shown that SLI and RD share a common high heritability if the child had poor nonword repetition abilities (50,61). Therefore, we might expect to see evidence of overlapping genetic associations for RD and SLI that might only become apparent in samples with specific deficits. It might be possible that the same CMIP variants have an effect on both reading and language problems depending on the presence of other risk factors. Our data do not support a pleiotropic effect of KIAA0319 on reading and language-related measures as reported previously (51,52). One possible explanation is that the previously reported associations between KIAA0319 and language skills were confined to individuals selected as language-impaired. Another explanation could be the use of psychometric tests not available in ALSPAC. These included the Omnibus language test (62) reported by Rice et al. (51) and measures of expressive and receptive language based on the scales of the Clinical Evaluation of Language Fundamentals (CELF-R) (63) reported by Newbury et al. (52).

It is striking how the associations detected in this study are specific to the single-word reading and spelling tests and not to other reading or language-related measures, despite the correlation across measures. This observation does not exclude that KIAA0319, DCDC2, and CMIP affect additional cognitive functions not tested here. Nevertheless this is interesting in relation to the biological function proposed for some of these genes. KIAA0319 and DCDC2 have been shown to play a role during the development of the cerebral cortex by regulating neuronal migration, a critical step of cortex development (64). Defects in neuronal migration lead to several human syndromes with various degrees of symptoms from epilepsy to mental retardation (65). It is therefore notable that genes involved in such a general process can lead to specific disorders rather than have a broad impact on cognition or behavior. Subtle neuronal migration defects have been suggested to be causative of RD (66). With the data reported here, we reinforce the idea that KIAA0319 and DCDC2, with proven roles in neuronal migration, affect specific phenotypes.

Another important observation stems from our ability to test the effect of comorbidity on association analysis. We were able to show for the first time that inclusion of individuals with comorbid RD and SLI or ADHD do not weaken the association but rather can strengthen it, as in the case-control analysis of DCDC2 (Table 5). This may result from an increase in sample size by including comorbid cases. Given previous reports of associations between DCDC2 and ADHD (55), it is also possible that the associations we observe here for this gene are the combined effects of this gene on RD and ADHD separately. We could not test this hypothesis here because our ADHD sample was small (n = 39). In either case, these findings have an important implication. It is common practice to exclude individuals with SLI and ADHD when designing RD genetic studies to obtain samples as homogeneous as possible and to avoid confounding effects. Our data suggest that the same genes contribute to reading impairment even in the background of different disorders. This would imply also that the same cognitive deficit is at the

basis of reading problems regardless of other clinical diagnoses. Providing that our observations are valid for other RD susceptibility genes, we suggest that individuals with RD comorbid for SLI or ADHD should not be excluded when designing genetic studies of RD, and their inclusion could improve sample power.

This study was funded by the United Kingdom Medical Research Council (Grant No. G0800523/86473) and a Wellcome Trust core grant (Grant No. 075491/Z/04). For the Avon Longitudinal Study of Parents and Children (ALSPAC) study, we thank the midwives for their help in recruiting the families, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists, and nurses. The UK Medical Research Council (Grant No. 74882), the Wellcome Trust (Grant no. 076467) and the University of Bristol provided core support for ALSPAC. This publication is the work of the authors, and Silvia Paracchini will serve as guarantor for the contents of this article.

All authors reported no biomedical financial interests or potential conflicts of interest.

Supplementary material cited in this article is available online.

- Pennington BF, Bishop DV (2009): Relations among speech, language, and reading disorders. Annu Rev Psychol 60:283–306.
- Barry JG, Yasin I, Bishop DV (2007): Heritable risk factors associated with language impairments. *Genes Brain Behav* 6:66–76.
- Fisher SE, DeFries JC (2002): Developmental dyslexia: Genetic dissection of a complex cognitive trait. Nat Rev Neurosci 3:767–780.
- Snowling M, Bishop DV, Stothard SE (2000): Is preschool language impairment a risk factor for dyslexia in adolescence? J Child Psychol Psychiatry 41:587–600.
- McArthur GM, Hogben JH, Edwards VT, Heath SM, Mengler ED (2000): On the "specifics" of specific reading disability and specific language impairment. J Child Psychol Psychiatry 41:869–874.
- Burd L, Klug MG, Coumbe MJ, Kerbeshian J (2003): Children and adolescents with attention deficit-hyperactivity disorder: 1. Prevalence and cost of care. J Child Neurol 18:555–561.
- Pennington BF (2006): From single to multiple deficit models of developmental disorders. *Cognition* 101:385–413.
- Beitchman JH, Nair R, Clegg M, Ferguson B, Patel PG (1986): Prevalence of Psychiatric Disorders in Children with Speech and Language Disorders. J Am Academy Child Psychiatry 25:528–535.
- 9. Scerri TS, Schulte-Korne G (2010): Genetics of developmental dyslexia. *Eur Child Adolesc Psychiatry* 19:179–197.
- Newbury DF, Fisher SE, Monaco AP (2010): Recent advances in the genetics of language impairment. *Genome Med* 2:6.
- Paracchini S, Scerri T, Monaco AP (2007): The genetic lexicon of dyslexia. Annu Rev Genomics Hum Genet 8:57–79.
- Hannula-Jouppi K, Kaminen-Ahola N, Taipale M, Eklund R, Nopola-Hemmi J, Kaariainen H, Kere J (2005): The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. *PLoS Genet* 1:e50.
- Taipale M, Kaminen N, Nopola-Hemmi J, Haltia T, Myllyluoma B, Lyytinen H, et al. (2003): A candidate gene for developmental dyslexia encodes a nuclear tetratricopeptide repeat domain protein dynamically regulated in brain. Proc Natl Acad Sci U S A 100:11553–11558.
- Brkanac Z, Chapman NH, Matsushita MM, Chun L, Nielsen K, Cochrane E, et al. (2007): Evaluation of candidate genes for DYX1 and DYX2 in families with dyslexia. Am J Med Genet B Neuropsychiatr Genet 144B:556–560.
- Dahdouh F, Anthoni H, Tapia-Paez I, Peyrard-Janvid M, Schulte-Korne G, Warnke A, et al. (2009): Further evidence for DYX1C1 as a susceptibility factor for dyslexia. *Psychiatr Genet* 19:59–63.
- Scerri TS, Fisher SE, Francks C, MacPhie IL, Paracchini S, Richardson AJ, et al. (2004): Putative functional alleles of DYX1C1 are not associated with dyslexia susceptibility in a large sample of sibling pairs from the UK. J Med Genet 41:853–857.
- Wigg KG, Couto JM, Feng Y, Anderson B, Cate-Carter TD, Macciardi F, et al. (2004): Support for EKN1 as the susceptibility locus for dyslexia on 15q21. Mol Psychiatry 9:1111–1121.

- Marino C, Citterio A, Giorda R, Facoetti A, Menozzi G, Vanzin L, et al. (2007): Association of short-term memory with a variant within DYX1C1 in developmental dyslexia. *Genes Brain Behav* 6:640–646.
- Cope NA, Hill G, van den Bree M, Harold D, Moskvina V, Green EK, et al. (2005): No support for association between dyslexia susceptibility 1 candidate 1 and developmental dyslexia. *Mol Psychiatry* 10:237–238.
- Bellini G, Bravaccio C, Calamoneri F, Donatella Cocuzza M, Fiorillo P, Gagliano A, et al. (2005): No evidence for association between dyslexia and DYX1C1 functional variants in a group of children and adolescents from southern Italy. J Mol Neurosci 27:311–314.
- Meng H, Hager K, Held M, Page GP, Olson RK, Pennington BF, et al. (2005): TDT-association analysis of EKN1 and dyslexia in a Colorado twin cohort. Hum Genet 118:87–90.
- Paracchini S, Ang QW, Stanley FJ, Monaco AP, Pennell CE, Whitehouse AJ (2011): Analysis of dyslexia candidate genes in the Raine cohort representing the general Australian population. *Genes Brain Behav* 10: 158–165.
- Harold D, Paracchini S, Scerri T, Dennis M, Cope N, Hill G, et al. (2006): Further evidence that the KIAA0319 gene confers susceptibility to developmental dyslexia. *Mol Psychiatry* 11:1085–1091:1061.
- Cope N, Harold D, Hill G, Moskvina V, Stevenson J, Holmans P, et al. (2005): Strong evidence that KIAA0319 on chromosome 6p Is a Susceptibility Gene for Developmental Dyslexia. Am J Hum Genet 76:581–591.
- Deffenbacher KE, Kenyon JB, Hoover DM, Olson RK, Pennington BF, DeFries JC, Smith SD (2004): Refinement of the 6p21.3 quantitative trait locus influencing dyslexia: Linkage and association analyses. *Hum Genet* 115:128–138.
- Francks C, Paracchini S, Smith SD, Richardson AJ, Scerri TS, Cardon LR, et al. (2004): A 77-kilobase region of chromosome 6p22.2 is associated with dyslexia in families from the United Kingdom and from the United States. Am J Hum Genet 75:1046–1058.
- 27. Kaplan DE, Gayan J, Ahn J, Won TW, Pauls D, Olson RK, *et al.* (2002): Evidence for linkage and association with reading disability on 6p21.3-22. *Am J Hum Genet* 70:1287–1298.
- Paracchini S, Thomas A, Castro S, Lai C, Paramasivam M, Wang Y, et al. (2006): The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Hum Mol Genet* 15:1659–1666.
- Dennis MY, Paracchini S, Scerri TS, Prokunina-Olsson L, Knight JC, Wade-Martins R, et al. (2009): A common variant associated with dyslexia reduces expression of the KIAA0319 gene. PLoS Genet 5 e1000436.
- Meng H, Smith SD, Hager K, Held M, Liu J, Olson RK, et al. (2005): DCDC2 is associated with reading disability and modulates neuronal development in the brain. Proceedings of the Natl Academy Sci U S A 102:17053– 17058.
- Schumacher J, Anthoni H, Dahdouh F, Konig IR, Hillmer AM, Kluck N, et al. (2006): Strong genetic evidence of DCDC2 as a susceptibility gene for dyslexia. Am J Hum Genet 78:52–62.
- Wilcke A, Weissfuss J, Kirsten H, Wolfram G, Boltze J, Ahnert P (2009): The role of gene DCDC2 in German dyslexics. Ann Dyslexia 59:1–11.
- Ludwig KU, Schumacher J, Schulte-Korne G, Konig IR, Warnke A, Plume E, et al. (2008): Investigation of the DCDC2 intron 2 deletion/compound short tandem repeat polymorphism in a large German dyslexia sample. *Psychiatr Genet* 18:310–312.
- Czamara D, Bruder J, Becker J, Bartling J, Hoffmann P, Ludwig KU, et al. (2011): Association of a rare variant with mismatch negativity in a region between KIAA0319 and DCDC2 in dyslexia. *Behav Genet* 41:110–119.
- Anthoni H, Zucchelli M, Matsson H, Muller-Myhsok B, Fransson I, Schumacher J, et al. (2007): A locus on 2p12 containing the co-regulated MRPL19 and C2ORF3 genes is associated to dyslexia. Hum Mol Genet 16:667–677.
- Newbury DF, Winchester L, Addis L, Paracchini S, Buckingham LL, Clark A, et al. (2009): CMIP and ATP2C2 modulate phonological short-term memory in language impairment. Am J Hum Genet 85:264–272.
- SLI Consortium (SLIC) (2004): Highly significant linkage to the SLI1 locus in an expanded sample of individuals affected by specific language impairment. Am J Hum Genet 74:1225–1238.
- Golding J, Pembrey M, Jones R (2001): ALSPAC—The Avon Longitudinal Study of Parents and Children. I. Study methodology. *Paediatr Perinat Epidemiol* 15:74–87.
- Vernes SC, Newbury DF, Abrahams BS, Winchester L, Nicod J, Groszer M, et al. (2008): A functional genetic link between distinct developmental language disorders. N Engl J Med 359:2337–2345.

- Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP (2001): A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 413:519–523.
- Paracchini S, Steer CD, Buckingham LL, Morris AP, Ring S, Scerri T, et al. (2008): Association of the KIAA0319 dyslexia susceptibility gene with reading skills in the general population. Am J Psychiatry 165:1576–1584.
- Luciano M, Lind PA, Duffy DL, Castles A, Wright MJ, Montgomery GW, et al. (2007): A haplotype spanning KIAA0319 and TTRAP is associated with normal variation in reading and spelling ability. *Biol Psychiatry* 62:811– 817.
- Lind PA, Luciano M, Wright MJ, Montgomery GW, Martin NG, Bates TC (2010): Dyslexia and DCDC2: Normal variation in reading and spelling is associated with DCDC2 polymorphisms in an Australian population sample. *Eur J Hum Genet* 18:668–673.
- Bates TC, Lind PA, Luciano M, Montgomery GW, Martin NG, Wright MJ (2009): Dyslexia and DYX1C1: Deficits in reading and spelling associated with a missense mutation. *Mol Psychiatry* 15:1190–1196.
- 45. Shaywitz SE (1998): Dyslexia. N Engl J Med 338:307-312.
- Bishop DV, North T, Donlan C (1996): Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. J Child Psychol Psychiatry 37:391–403.
- Bishop DV, Laws G, Adams C, Norbury CF (2006): High heritability of speech and language impairments in 6-year-old twins demonstrated using parent and teacher report. *Behav Genet* 36:173–184.
- Bishop DV, Hayiou-Thomas ME (2008): Heritability of specific language impairment depends on diagnostic criteria. *Genes Brain Behav* 7:365– 372.
- Bishop DV, McDonald D (2009): Identifying language impairment in children: Combining language test scores with parental report. Int J Lang Commun Disord 44:600–615.
- Bishop DV, Snowling MJ (2004): Developmental dyslexia and specific language impairment: Same or different? *Psychol Bull* 130:858–886.
- Rice ML, Smith SD, Gayan J (2009): Convergent genetic linkage and associations to language, speech and reading measures in families of probands with specific language Impairment. J Neurodev Disord 1:264– 282.
- Newbury DF, Paracchini S, Scerri TS, Winchester L, Addis L, Walter J, et al. (2011): Investigation of dyslexia and SLI risk variants in reading- and language-impaired subjects. *Behav Genet* 41:90–104.
- Marino C, Mascheretti S, Riva V, Cattaneo F, Rigoletto C, Rusconi M, et al. (2011): Pleiotropic effects of DCDC2 and DYX1C1 genes on language and mathematics traits in nuclear families of developmental dyslexia. Behav Genet 41:67–76.
- Willcutt EG, Pennington BF, Smith SD, Cardon LR, Gayan J, Knopik VS, et al. (2002): Quantitative trait locus for reading disability on chromosome 6p is pleiotropic for attention-deficit/hyperactivity disorder. Am J Med Genet 114:260–268.
- Couto JM, Gomez L, Wigg K, Ickowicz A, Pathare T, Malone M, et al. (2009): Association of attention-deficit/hyperactivity disorder with a candidate region for reading disabilities on chromosome 6p. *Biol Psychi*atry 66:368–375.
- Lesch KP, Timmesfeld N, Renner TJ, Halperin R, Roser C, Nguyen TT, et al. (2008): Molecular genetics of adult ADHD: Converging evidence from genome-wide association and extended pedigree linkage studies. J Neural Transm 115:1573–1585.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. (2007): PLINK: A tool set for whole-genome association and populationbased linkage analyses. *Am J Hum Genet* 81:559–575.
- Snowling MJ, Muter V, Carroll J (2007): Children at family risk of dyslexia: A follow-up in early adolescence. J Child Psychol Psychiatry 48:609–618.
- Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA (2007): The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *Am J Psychiatry* 164:942–948.
- Wolke D, Waylen A, Samara M, Steer C, Goodman R, Ford T, Lamberts K (2009): Selective drop-out in longitudinal studies and non-biased prediction of behaviour disorders. *Br J Psychiatry* 195:249–256.
- Bishop DV (2001): Genetic influences on language impairment and literacy problems in children: Same or different? J Child Psychol Psychiatry 42:189–198.
- 62. Zimmerman IL, Steiner VG, Pond RE (1992): *Preschool Language Scale*—3. San Antonio, TX: Psychological Corporation.
- 63. Semel Em WEH, Secord W (1992): Clinical Evaluation of Language Fundamentals—Revised. San Antonio, TX: Psychological Corporation.

- 64. Galaburda AM, LoTurco J, Ramus F, Fitch RH, Rosen GD (2006): From genes to behavior in developmental dyslexia. *Nat Neurosci* 9:1213–1217.
- 65. McManus MF, Golden JA (2005): Neuronal migration in developmental disorders. J Child Neurol 20:280–286.
- Galaburda AM, Sherman GF, Rosen GD, Aboitiz F, Geschwind N (1985): Developmental dyslexia: Four consecutive patients with cortical anomalies. *Ann Neurol* 18:222–233.
- 67. Rust J, Golombok S, Trickey G (1993): WORD: Wechsler Objective Reading Dimensional Manual. Sidcup, UK: Psychological Corporation.
- Nunes T, Bryant P, Olsson J (2003): Learning morphological and phonological spelling rules: An intervention study. *Sci Stud Reading* 7:298–307.
- Rosner J, Simon DP (1971): The auditory analysis test: An initial report. J Learn Disabilities 4:40 – 48.
- Case R, Kurland DM, Goldberg J (1982): Operational efficiency and the growth of short-term memory span. J Exp Child Psychol 33:386–404.

- 71. Rust J (1996): WOLD Wechsler Objective Language Dimensions Manual London UK: Psychological Corporation.
- Gathercole SE, Willis CS, Baddeley AD, Emslie H (1994): The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory* 2:103–127.
- Bishop DV (1998): Development of the Children's Communication Checklist CCC: A method for assessing qualitative aspects of communicative impairment in children. J Child Psychol Psychiatry 39:879– 891.
- 74. Goodman R, Ford T, Richards H, Gatward R, Meltzer H (2000): The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *J Child Psychol Psychiatry* 41:645–655.
- 75. Wechsler D, Golombok S, Rust J (1992): WISC-IIIUK: Wechsler Intelligence Scale for Children. Sidcup, UK: Psycological Corporation.