

# Genetic and Environmental Influences on Orthographic and Phonological Skills in Children With Reading Disabilities

Javier Gayán

*Wellcome Trust Centre for Human Genetics  
University of Oxford*

Richard K. Olson

*Department of Psychology  
University of Colorado at Boulder*

Data from identical and fraternal twins were analyzed to estimate the proportions of genetic and environmental influences on group deficits in accuracy and, when available, speed for printed word recognition and for related skills in phonological decoding (PD), orthographic coding (OC), and phoneme awareness (PA). In addition, bivariate genetic analyses were employed to estimate the degree of common genetic influence on group deficits across these different reading and language skills. About half of the group deficits in each of the skills were due to genetic influences, and the genetic origins were largely shared among the measures ( $r_g = .53 - .99$ ), except for those between OC and PA ( $r_g = .28 - .39$ ). Implications of the results are discussed for models of reading disability and remediation.

Children with a school history of reading problems tend to have particular difficulty reading words, both in isolation and in context. Their deficits in printed word recognition accuracy, fluency, or both place major constraints on the ultimate goal of reading to comprehend written texts (Perfetti, 1985). The importance of word-reading deficits in reading disability (RD) has led researchers to explore possible component processes that may be etiologically significant. For example, researchers have noted that deficits in printed word recognition are usually related to similar or even greater deficits in phonological decoding (PD), phoneme awareness (PA), and orthographic coding (OC; see definitions following; Olson, Forsberg, & Wise, 1994; Olson, Wise, Conners, Rack, & Fulker, 1989; Rack, Snowling, &

Olson, 1992). This study addresses the genetic and environmental etiology of group deficits in these component processes and in isolated word reading. Data from expanded samples of identical and fraternal twins recruited from Colorado schools were analyzed to estimate the proportional influences from genes, shared environment, and nonshared environment on group deficits in each skill, both in accuracy and, when available, speed of processing. In addition, bivariate genetic analyses of the twin data were used to learn if the same or different genes are responsible for genetic influences on group deficits in word reading and the different component processes, and on speed and accuracy in those processes.

The reading of isolated English words depends on two types of knowledge, PD knowledge of common grapheme–phoneme correspondences and word-specific orthographic knowledge (Baron, 1979; Olson, Kliegl, Davidson, & Foltz, 1985). PD is operationally defined here as the process of “sounding out” a novel letter string in a way that is consistent with the most common correspondences between the graphemes and phonemes. Knowledge for some of the most consistent correspondences is often profitably associated with explicit instruction in the early stages of reading development (Chall, 1967). Additional implicit learning of grapheme–phoneme correspondences may accrue automatically as children’s word-reading vocabulary expands (Plaut, McClelland, Seidenberg, & Patterson, 1996). Several standardized tests of English reading skill recognize the importance of PD by including tests of oral nonword reading accuracy. This study added an experimental measure of oral nonword reading speed because limitations in speed may place an additional constraint, beyond accuracy, on the contribution of PD to reading development. A measure of silent nonword reading accuracy and speed was also included in the study to control for any oral production constraints on participants’ PD.

Unfortunately for beginning readers of English, many grapheme–phoneme correspondences in common English words are not very consistent compared to the higher level of consistency found in other languages such as German, Spanish, and Italian. This fact places an additional constraint on the English reader’s explicit or implicit extraction of the most frequent grapheme–phoneme correspondences (Frith, Wimmer, & Landerl, 1998; Hutzler & Wimmer, 2001). OC refers here to word-specific knowledge that cannot be derived solely from sublexical PD processes and subsequent associations with oral vocabulary. Word-specific orthographic knowledge was tested in this study by asking participants to find the word among two phonologically identical letter strings (e.g., rain/rane) or to select the appropriate homophone to fit a sentence context (e.g., bare/bear). In both tasks, the two choices would sound the same through sublexical PD processes. Selection of the correct response depended on knowledge of the target word’s specific orthographic pattern as well as for the target word’s meaning in the homophone-choice task. OC skills are particularly important in English because of its frequent homophones and “exception” or “strange” words that require word-specific knowledge of grapheme–phoneme correspondences and meaning.

From a developmental perspective, knowledge of common grapheme–phoneme correspondences (PD) is likely to play an important role in children’s learning about word-specific spellings for exception words and homophones (OC) for two reasons (Ehri & Wilce, 1980). First, most exception words have elements that are consistent with common grapheme–phoneme correspondences. Second, knowledge of the most typical correspondences for the exception parts of words may ultimately help children orient to and learn about the exceptional grapheme–phoneme correspondences in those words. Thus, one would expect that growth in PD and OC skills would be correlated. On the other hand, some significant differences are possible in the etiology of phonological and orthographic knowledge. For example, word-specific orthographic knowledge ultimately depends on exposure to those specific words in print, along with contextual or external feedback to derive the correct pronunciation and meaning. In contrast, sublexical PD skills can be based on general knowledge of the statistical regularities of grapheme–phoneme correspondences that are not word specific.

PD and OC skills may also differ in their dependence on a basic analytic language skill called PA, which in this study is defined as the ability to isolate and manipulate abstract subsyllabic phonemes in speech. It is measured here by assessing children’s ability to play a Pig Latin game and to say a word or nonword that is left after deleting a spoken phoneme from a spoken word or nonword. Many researchers have documented the close association between growth in PA in language and PD in reading (Liberman, Shankweiler, Fischer, & Carter, 1974; Wagner, Torgesen, & Rashotte, 1994). The causal nature of this association appears to be reciprocal: Learning to read promotes growth in PA (Morais, Cary, Alegria, & Bertelson, 1979), and preschoolers’ level of PA predicts later reading development, particularly in PD (Lundberg, Frost, & Peterson, 1988; Wagner et al., 1994). Relatively little study has been done on the developmental relation between PA and OC, but cross-sectional data have suggested that this relation is not as strong. Olson, Forsberg, and Wise (1994) found that measures of PA and PD loaded on a common factor, whereas the OC measures loaded on a separate correlated factor that had a stronger loading for a measure of print exposure. Converging evidence on the partial independence of OC has come from comparisons of older disabled and younger normal readers matched on standard measures of word recognition. The older disabled groups were lower in both PD and PA, but not in OC (Olson, Wise, Conners, Rack, & Fulker, 1989; Rack, Snowling, & Olson, 1992). Taken together, these results suggest that the development of OC may have a partly independent etiology from that for PD and PA. In this study, we explore the shared and independent etiology of group deficits in the different skills through behavioral genetic analyses.

The exploration of possible genetic influences on reading problems began with observations that these problems tend to run in families (Thomas, 1905; Fisher, 1905; Hallgren, 1950). However, familial transmission is necessary but not sufficient evidence for genetic etiology because family members typically share both

their genes and their environment (DeFries, 1985). More recently, studies of identical and fraternal twins reared together have been used to separate the relative contributions of genetic and environmental factors to RDs. Identical and fraternal twin pairs both share their home and school environments (see qualifications in the discussion herein), but they differ in their average genetic similarity: Identical or monozygotic (one egg, MZ) twins share all their genes, whereas fraternal or dizygotic (two egg, DZ) twins share half of their segregating genes on average. Thus, if MZ twins were more likely to share a reading disorder, this would provide evidence for some degree of genetic influence on the disorder.

Early behavior–genetic twin studies of reading disability usually classified individual twins dichotomously as “dyslexic” or “normal.” Then researchers compared the degree to which both members of a twin pair were “concordant” (i.e., they shared the disorder) or “discordant” (i.e., only one twin was “dyslexic”) depending on their zygosity (MZ or DZ). The vast majority of these studies reported significantly higher concordance rates for MZ twins, suggesting some genetic etiology (Bakwin, 1973; Halgren, 1950; Hermann, 1959; Norrie, 1939; Zerbin-Rudin, 1967). However, experts now generally recognize that reading ability is normally distributed in the population (Rodgers, 1983; Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). Therefore, dichotomous classification and comparison of concordance rates is not optimal for behavioral genetic studies of reading disability. (We do not use the term *dyslexia* herein for reasons described in the participants section.)

DeFries and Fulker (1985) developed a more appropriate regression procedure for analyzing twin data wherein at least one member of each pair (the “proband”) was from the low tail of the normal distribution. The average proportion of genetic influence on the proband group’s deficit was determined by comparing the average regression to the population mean for the MZ and DZ cotwins of the selected probands (see the analysis section). DeFries, Fulker, and LaBuda (1987) first applied this analysis to a small group of MZ and DZ twins below approximately the 10th percentile in a composite measure of word recognition, reading comprehension, and spelling from the Peabody Individual Achievement Test (PIAT; Dunn & Markwardt, 1970). They found significant evidence for genetic influence on the group deficit: On average, DZ cotwins regressed significantly more than MZ cotwins toward the population mean for the composite reading measure. A recent analysis with a much larger twin sample yielded a heritability estimate of .58, indicating that about half of the group deficit was due to genetic influence (Wadsworth, Olson, Pennington, & DeFries, 2000).

Olson et al. (1989) first explored separately the genetic etiology of group deficits in word recognition, PD, PA, and OC. All but OC showed significant evidence for genetic etiology, although no significant contrast was found among the different levels of genetic influence across the measures in this small initial study. Subsequent analyses (Gayán & Olson, 1999; Olson, Forsberg, & Wise, 1994) with a larger sam-

ple revealed significant heritabilities for group deficits in all the measures (word recognition = .45, PD = .61, OC = .58, and PA = .56). Olson, Forsberg, and Wise (1994) also conducted the first bivariate genetic analyses that showed at least some common genetic etiology for group deficits in accuracy on all the measures.

The previous analyses (Gayán & Olson, 1999; Olson, Forsberg, & Wise, 1994; Olson et al., 1989) were limited by the relatively small sample sizes and single measures of each skill domain. This sample is nearly twice as large as the sample Olson and colleagues (1989; Olson, Forsberg, & Wise, 1994) analyzed, allowing a more stringent criterion for group-deficit membership, and results are reported now for at least two measures of each skill domain. Another important extension beyond our previous analyses is the inclusion herein of measures of processing speed in each skill domain. Experts increasingly recognize that deficits in automaticity or fluency in component word-reading skills may sometimes place important additional constraints on general reading ability beyond those imposed by deficits in accuracy alone (Wolf, 1997). Therefore, we also explored the genetic etiology of processing speed deficits for accurate responses in this analyses. Finally, we estimated the genetic correlations between group deficits across the various tasks.

## METHOD

### Participants

This study is based on an analysis of data from the twin sample of the Colorado Learning Disabilities Research Center (CLDRC; DeFries et al., 1997). This sample of twins and their families has been ascertained since 1982 from 27 school districts in Colorado. First, school records are used to identify all twin pairs in a school. Then, those twin pairs in which at least one of the twins has a school history of reading problems (i.e., low reading test scores or assignment to remedial reading classes) are invited to the laboratory at the University of Colorado at Boulder to undergo an extensive battery of standard psychometric tests and experimental measures. The initial sample of twin pairs with some school history, prior to the exclusions described later, included 382 MZ twin pairs and 288 same-sex DZ twin pairs between age 7.98 and 20.99 (90 percent were younger than age 16). A comparison group of twin pairs in which neither member of the twin pair had a school history of reading problems was selected as controls, although some exhibited reading difficulties when tested in the laboratory. The control group was matched to the proband group in age, gender, and school district. It included 232 MZ pairs and 147 same-sex DZ pairs. The school-history and control-twin pairs were given the same battery of tests.

Zygosity of same-sex twin pairs was determined by administering selected items of the Nichols and Bilbro (1966) questionnaire. Those twin pairs whose zygosity re-

mained doubtful after this test were resolved by analysis of blood or buccal samples. For this study, opposite-sex DZ twins were excluded from analyses.

Twins with evidence of serious neurological, emotional, or uncorrected sensory deficits were excluded from the analyses. Twin pairs for whom English is a second language were not included in the initial sample. Finally, for these analyses, twins were selected to have a Wechsler (Wechsler, 1974, 1981) verbal or performance IQ of at least 90. This is an arbitrary but commonly used minimum IQ criterion for specific RD. It tends to yield slightly higher heritability estimates for group deficits in reading and related skills than when children whose IQs are lower than 90 are included in the sample (Knopik, 2000; Olson, Datta, Gayán, & DeFries, 1999; Wadsworth et al., 2000).

We did not use a minimum discrepancy between IQ and reading in this study. A discrepancy criterion has often been used to define “dyslexic” readers whose reading is substantially lower than expected from their IQ, but increasingly apparent is that a discrepancy criterion is not related to basic phonological processing deficits in poor readers (Stanovich & Siegel, 1994) or their response to remediation efforts (Vellutino, Scanlon, & Lyon, 2000).

After these exclusions, the twin sample with a positive school history of RD in at least one member of each pair included a total of 515 twin pairs (288 MZ and 227 DZ pairs). In addition, the control group of twins comprised 215 MZ and 133 DZ twin pairs. The overall mean age of the twins at the time of testing was 11.63 years, ranging from 8.02 to 20.24 years. Ninety percent of the twins were between age 8 and 15.5.

## Measures

The twins were administered an extensive battery of psychometric tests including the Wechsler Intelligence Scale for Children–Revised (WISC–R; Wechsler, 1974) or the Wechsler Adult Intelligence Scale–Revised (WAIS–R; Wechsler, 1981), the PIAT (Dunn & Markwardt, 1970), and several experimental tests especially designed to assess important reading and language skill areas. The experimental measures included time–limited word recognition, OC, PD, and PA. The standard psychometric measures were administered in an approximately 2-hr session (including breaks), and the experimental measures were administered in another 2-hr session on the same day. The morning versus afternoon order of standardized and experimental test sessions was random across twin pairs, but was the same for both members within a twin pair.

*Word recognition.* The standardized measure of word recognition was from the PIAT Word Recognition (PWR; Dunn & Markwardt, 1970). This test required

participants to read across rows of increasingly difficult, unrelated words until they reached an ending criterion of 5 errors in the last 7 words. The final score was based on the number of items correct. As in most standardized measures of word recognition, there was no time constraint. The published test–retest reliability is .89 (Dunn & Markwardt, 1970).

The experimental timed word recognition (TWR) test (Olson, Forsberg, Wise, & Rack, 1994; Olson et al., 1989) assessed word recognition accuracy and processing speed when a difficulty-ordered list of up to 182 isolated words was presented individually in lowercase letters on the computer screen. Participants were initially placed at different difficulty starting levels in the list depending on their performance in a 14-item screening list. To be scored as correct for continued progression through the main test list, the participant's correct response had to be initiated within 2 sec after appearance of the word on the screen, as indicated by a voice key. Participants were instructed to "... read the words as quickly as you can without making mistakes," and to "... sound it out or give it a good guess" for unknown words. List presentation was terminated when the participant failed to read 10 of the last 20 words correctly within the 2-sec limit for response initiation. A test–retest correlation of .93 was obtained with an independent sample of 123 third- through sixth-grade poor readers across 4 months (Olson, Forsberg, Wise, & Rack, 1994).

A composite measure of word recognition (CWR) was created by averaging  $z$  scores from the TWR and PWR tasks.

*Orthographic coding.* Three specific measures were administered to assess OC skills. Orthographic choice (OCH), required the recognition of a target word versus a phonologically identical background foil that was not a word (i.e., rain/rane; sammon/salmon; Olson et al., 1985, 1989). The task included 8 practice trials, and then a total of 80 trials in two 40-trial blocks. The first block consisted of relatively easy targets (i.e., rain). The second block's trials had more difficult lower frequency words (i.e., salmon). The items were balanced to ensure that the distracter item was a plausible string. Thus, the pair certain/sertain was balanced by the trial serpent/serpent: school/scool was balanced by scoop/schoop. This means that participants must make use of word-specific knowledge rather than general information about orthographic structure. Participants were instructed to select the word target by pressing a right or left button. They were told that

A number will appear on the screen if you answered correctly. The faster you answer, the smaller the number will be. If you were wrong, "error" will appear on the screen. Try to answer as quickly as you can while making as few errors as possible.

Participants' final orthographic choice accuracy (OCHA) scores and median orthographic choice latency (OCHL) scores for correct trials were based on their average age-adjusted  $z$  scores across the two trial blocks.

Homonym choice (HCH) required that participants listen to recorded sentences such as “Which is a fruit?” and subsequently choose the appropriate word from a pair of homophones on the computer screen (i.e., pair/pear) by pressing right or left buttons corresponding to the spatial position of the target word. Latency and accuracy feedback was presented after each response following the same procedure as in the OCH task. The HCH task included 5 practice trials, and then 65 experimental trials of randomly ordered difficulty level (Olson, Forsberg, & Wise, 1994). Barker, Wagner, and Torgesen (1992) and Stanovich and West (1989) have used similar tasks. Both homonym choice accuracy (HCHA) scores and homonym choice latency (HCHL) scores on correct responses were obtained for this task.

A third measure that Olson, Forsberg, and Wise (1994) previously found to load more on an orthographic than a phonological factor was the 84 trial PIAT spelling (PSP) subtest (Dunn & Markwardt, 1970), which required the untimed forced choice of a target word (e.g., cloudy), which had been presented orally, from four orthographically and often phonologically similar alternatives printed on a card (e.g., clowdy/clowdy/cloudey/cloudy). This task ended if the participant answered incorrectly in five out of the last seven trials in the difficulty-ordered list. Participants’ PSP final score was based on the number of trials answered minus incorrect answers. The published test–retest reliability is .64 (Dunn & Markwardt, 1970).

*Phonological decoding.* PD is typically measured in standardized tests through the oral reading of pronounceable nonwords (Woodcock, 1987). The present experimental oral phonological decoding (OPD) task included a block of 45 1-syllable nonwords (i.e., ter, strale), and then a block of 40 2-syllable nonwords (i.e., lobsel, vogger), presented one at a time on a computer screen. Each block was preceded by four practice trials. For the first block, the nonword target was removed from the screen as soon as an initial vocalization was detected by the voice key. In the second block, the target stayed on the screen for 2 sec after the voice key was triggered. No accuracy or numerical latency feedback was provided. This task was scored both for oral phonological decoding accuracy (OPDA) and median OPD latency (OPDL) on correct trials, averaged across the two trial blocks. A composite score (OPD) was also created by combining *z* scores for accuracy and median correct reaction time. Oral responses were tape recorded and scored for accuracy off-line.

A silent phonological decoding (SPD) task was presented in which participants choose a printed nonword that would sound like a common word if read aloud. Each target was accompanied by two nonword foils that would not sound like a known word (e.g., coam/baim/goam). Participants indicated their choice by pressing one of three buttons corresponding to the horizontal spatial position of the target nonword. Participants were told that “The computer will underline the item you pick. After that the computer will highlight the correct answer. We’ll start with five practices. Try to answer quickly, but also as accurately as you can” (Olson,



Forsberg, Wise, & Rack, 1994; Olson et al., 1989). Silent phonological decoding accuracy (SPDA) scores and silent phonological decoding latency (SPDL) scores on correct responses were obtained from this task. The age-adjusted correlation with OPD was .80.

*Phoneme Awareness.* Three measures of PA were included in the test battery. The phoneme transposition (PTP) task was a Pig Latin game in which participants were required to take the first sound from the front of a word, put it at the end, and add the sound /ay/. For example, “rope” would become “ope-ray”. Five initial practice examples were discussed with the children, 9 practice trials where they were given the correct answer if wrong, and 45 experimental trials with no feedback. All words were within the listening vocabulary of elementary-school aged children. The final score was based on percent correct for the 45 experimental trials.

The phoneme deletion (PDL) task (Olson, Forsberg, & Wise, 1994) was based on tasks Bruce (1964) and Rosner and Simon (1971) developed. In Part I of the test, participants were presented with recorded spoken nonwords, which they were asked to repeat. They were then asked to remove a specified phoneme from the nonword and if done correctly, the result was a target word (e.g., “say prot,” “now say prot without the /r/ sound”—“pot”). There were 6 practice trials with feedback, and 40 experimental trials without feedback. In Part II of the test, participants were presented with recorded spoken words, which they were asked to repeat. Then they were asked to remove a specified phoneme from the word and pronounce the resulting nonword. There were 5 practice and 28 experimental trials. In both parts of the task, 2 sec were allowed for participants to repeat the initial stimulus, and 6 sec were allowed for the participants to give their response to the deletion instructions. The final PDL score was based on the average age-adjusted  $z$  scores from Part I and Part II of the test. The age-adjusted correlation with PTP was .78.

The Lindamood auditory conceptualization (LAC) test (Lindamood & Lindamood, 1979) used colored blocks to represent phonemes and participants were required to move blocks to reflect changes in sequences of sounds spoken by the tester. The final score was based on a complex weighting scheme for correct responses across 34 trials. The age-adjusted correlation of LAC with PDL was .67.

## Analysis

Behavior genetic analyses can be used to separate the genetic and environmental influences on a trait by taking advantage of a twin-study design. Environmental influences can be further decomposed into the shared environmental and the nonshared environmental factors. Shared environment includes all the environmental influences that are shared by the twins in a twin pair, whereas nonshared environment encompasses environmental experiences that are unique to the individual. Nonshared environment also includes any variance due to test error.

The DeFries–Fulker (DF) multiple regression analysis of twin data (DeFries & Fulker, 1985) provides a powerful test of the genetic and environmental influences on group deficits in a trait. This multiple regression technique is particularly appropriate for samples that have been selected for deviant scores on a continuous dimension such as reading. Affected individuals (i.e., probands) are selected as those who have a positive school history of reading deficits and fall below a deficit criterion on the continuous dimension; the other member of each twin pair is termed the cotwin. Because MZ twins are genetically identical, whereas DZ twins share only half of their segregating genes on average, any genetic influence would cause DZ cotwins to regress more than MZ cotwins toward the mean of the unselected population. Thus, we can use the differential regression of MZ and DZ cotwins toward the population mean to estimate the heritability of the trait.

Several articles have previously explained the details of the DF regression method (DeFries & Fulker, 1985, 1988; LaBuda, DeFries, & Fulker). The sample to be analyzed comprises MZ and DZ probands (P) and cotwins (C). The coefficient of relation (R) expresses the genetic relatedness of twin pairs and is coded for MZ (R = 1) and DZ (R = 0.5) twins. The regression of the cotwin's score onto the proband's score and the coefficient of relation is known as the basic DF model:

$$C = \beta_0 + \beta_1 P + \beta_2 R \quad (1)$$

In this multiple regression, the partial regression coefficient  $\beta_2$  estimates the differential regression of cotwin scores by zygosity, controlling for proband score. In fact, it estimates twice the difference between the means of MZ and DZ cotwins, adjusting for MZ and DZ proband mean differences. If prior to analysis the data are transformed by expressing each score as a deviation from the mean of the unselected population and dividing each score by the difference between the proband mean and the population mean ( $\bar{P} - \mu$ ), then  $\beta_2$  provides a direct estimate of the heritability of the group deficit ( $h_g^2$ ), an index of the extent to which the deficit of the proband group is caused by genetic factors. The effect of shared environment on the group deficit ( $c_g^2$ ) can be estimated as

$$c_g^2 = (2\bar{C}_{DZ}) - \bar{C}_{MZ} \quad (2)$$

Finally, the effect of nonshared environment ( $e_g^2$ ) can be estimated with transformed data so that  $\bar{P}_{MZ} = 1$  and  $V_{P_{MZ}} = 1$ , as

$$e_g^2 = \bar{P}_{MZ} - \bar{C}_{MZ} = 1 - \bar{C}_{MZ} \quad (3)$$

or could simply be computed as the subtraction of the  $h_g^2$  and  $c_g^2$  estimates from unity because these three effects are designed to be exhaustive.

As a consequence of the truncate selection employed to ascertain proband twins in the CLDRC, double entry of concordant twin pairs was performed prior to analysis (DeFries & Gillis, 1991) and standard errors of parameter estimates were adjusted accordingly.

In addition, a bivariate extension of this basic DF model allows for the regression of a cotwin's score on one variable ( $C_y$ ) onto the proband's score on a second variable ( $P_x$ ) and the coefficient of relationship (Olson, Forsberg, & Wise, 1994):

$$C_y = \beta_0 + \beta_1 P_x + \beta_2 R \tag{4}$$

When this bivariate DF model is applied to transformed data in a sample selected for deviant scores in the  $y$  variable, the partial regression coefficient  $\beta_{2(xy)}$  estimates the bivariate heritability (Light & DeFries, 1995):

$$\beta_{2(xy)} = h_{g(x)} h_{g(y)} r_g \tag{5}$$

an index of the extent to which deficits in the proband variable are due to genetic factors that also influence the cotwin variable. Using the regression coefficients that estimate heritability from the univariate and bivariate regression models, an estimate of the genetic correlation between deficits in two variables can also be computed (Knopik, Alarcón, & DeFries, 1997):

$$r_g = \sqrt{\frac{\beta_{2(xy)} \beta_{2(yx)}}{\beta_{2(x)} \beta_{2(y)}}} \tag{6}$$

Nonetheless, the sample described previously is selected for positive school history of RD, which is moderately correlated with most reading and language measures given in the laboratory. Therefore, this selected sample is already indirectly selected for deviant scores in each of the two variables used in the bivariate analysis, which could bias the bivariate parameter estimates. For this reason, we combined the group selected for school history of RD and the control group of twins to create a relatively unselected sample. This sample was then analyzed with the bivariate regression model by selecting probands in each variable as being below a cut-off criterion, namely 1.5 *SDs* below the mean of the control group.

## RESULTS

### Correlations

We selected twin pairs in which at least one member of the pair had a positive school history of RD and a verbal or performance IQ score higher than 90. For this

selected sample of 515 twin pairs, Table 1 presents Pearson product-moment phenotypic correlations for the variables described in the methods section. For those variables with several versions, correlations are presented for only one version: Accuracy scores were selected for OCH (i.e., OCHA), HCH (i.e., HCHA), and SPD (i.e., SPDA), and the composite score was selected for OPD. All phenotypic correlations among reading and language tasks are positive and significant, ranging from small (.19) to very large (.96). These correlations are lower than those obtained in the unselected population due to the restricted range of this selected sample. In addition, the use of twins as individuals could bias these estimates. For the computation of these correlations, we randomly chose one twin from each pair and consider it an independent observation. Correlations based on double entry of all twin pairs are similar but slightly larger.

In general, correlations among the word recognition tasks are very large, and they exhibit moderate to large correlations with the other reading and language measures. Orthographic tasks correlate moderately among themselves, and slightly lower with the PD and PA measures. The OPD task correlates moderately with the SPD task. These PD tasks also correlate moderately with PA tasks. Finally, the PA tasks correlate in the moderate to large range among themselves.

### Univariate Analysis

Within this selected sample of twin pairs, probands (i.e., affected individuals) were identified as those with positive school history of reading deficits and scores 1.5

TABLE 1  
Phenotypic Correlations Between the Reading and Language Tasks for the RD School History Selected Sample<sup>a</sup>

	<i>CWR</i>	<i>TWR</i>	<i>PWR</i>	<i>OCHA</i>	<i>HCHA</i>	<i>PSP</i>	<i>OPD</i>	<i>SPDA</i>	<i>PTP</i>	<i>PDL</i>
<i>TWR</i>	.95									
<i>PWR</i>	.96	.83								
<i>OCHA</i>	.49	.44	.49							
<i>HCHA</i>	.53	.53	.50	.56						
<i>PSP</i>	.64	.62	.61	.47	.51					
<i>OPD</i>	.78	.75	.74	.44	.39	.48				
<i>SPDA</i>	.65	.60	.65	.52	.43	.40	.58			
<i>PTP</i>	.45	.39	.47	.37	.28	.28	.47	.55		
<i>PDL</i>	.55	.50	.55	.37	.19	.29	.59	.57	.79	
<i>LAC</i>	.48	.42	.50	.34	.25	.26	.38	.47	.58	.61

*Note.* CWR = composite word recognition; TWR = timed word recognition; PWR = PIAT word recognition; OCHA = orthographic choice accuracy; HCHA = homonym choice accuracy; PSP = PIAT spelling; OPD = oral phonological decoding; SPDA = silent phonological decoding; PTP = phoneme transposition; PDL = phoneme deletion; LAC = Lindamood auditory conceptualization.

<sup>a</sup>All correlations are statistically significant.

*SDs* below the mean of the control group in the relevant measure. Table 2 presents MZ and DZ proband and cotwin means (expressed as deviations from the unselected population), as well as double-entry and single-entry sample sizes. It is important to note that some sample size differences are found across variables due to the  $-1.5$  *SD* selection, as well as some participants having missing data for some variables. Also, sample size differences exist across the composite, accuracy, and latency versions of two of the OC tasks (OCH and HCH) and of the PD tasks (OPD and SPD). Differences between the accuracy and composite versions of each task are probably due to the  $-1.5$  *SD* selection. In addition, latency exhibited smaller sample sizes than accuracy, and these sample size differences were due to a distributional effect. Latency scores for these variables were closer to the control mean than accuracy scores. Therefore, the  $-1.5$  *SD* selection yielded fewer probands in the latency measures.

The selected samples of probands and their cotwins on each variable were then analyzed with the DF multiple regression method to obtain estimates of the genetic and environmental effects on the group deficits (see Table 3). In this fashion, we estimated the heritability of the group deficit in CWR to be about 54% of the total variance, whereas shared environmental effects account for 39%, and the remaining 6% is due to nonshared environmental influences and test error. This CWR variable is a composite of PWR and TWR. The partition of variance for PWR and TWR is very similar to that of the composite CWR measure.

In addition to the word recognition measures, Table 3 also provides estimates of the heritability, shared environment, and nonshared environment of group deficits of related reading and language measures. The sample size for the SPDL was so small that very little confidence can be placed on the estimates for this measure; therefore, we excluded it from the discussion of the estimates. In general, group deficits in OC, PD, and PA tasks are significantly heritable, genetic effects accounting for about 46–72% of the group deficits. The influence of shared environment on these variables is sometimes significant, but of smaller magnitude than the genetic influence (10–39%). Finally, nonshared environment influences, including test error, range from .06 to .37.

Because of the large age range of the twin pairs analyzed, we also tested for the possibility that the genetic etiology of group deficits in these skills changed across the age range. However, the interaction between age and the genetic effect was small and nonsignificant for all skills.

## Bivariate Analysis

Bivariate regression analyses provided estimates of bivariate heritability across several reading and language measures. Genetic correlations were then computed from the univariate and bivariate regression coefficients. Analysis of all possible

TABLE 2  
 Monozygotic and Dizygotic Proband and Cotwin Means Expressed as Deviations From the Unselected Population, and Double Entry and Single Entry Sample Sizes for Each Task<sup>a</sup>

<i>Task</i>	<i>Proband Mean</i>		<i>Cotwin Mean</i>		<i>Sample Size</i>			
	<i>MZ</i>	<i>DZ</i>	<i>MZ</i>	<i>DZ</i>	<i>MZ DE</i>	<i>SE</i>	<i>DZ DE</i>	<i>SE</i>
CWR	-2.67	-2.62	-2.50	-1.74	325	215	195	159
TWR	-2.49	-2.47	-2.33	-1.61	301	201	167	141
PWR	-2.63	-2.56	-2.41	-1.59	303	203	192	159
OCH	-2.62	-2.57	-2.30	-1.49	219	150	144	119
OCHA	-2.30	-2.27	-1.93	-1.14	187	132	125	107
OCHL	-2.83	-2.89	-2.07	-1.44	128	97	77	67
HCH	-2.42	-2.55	-1.95	-1.26	93	69	74	65
HCHA	-2.15	-2.26	-1.69	-1.12	82	62	66	57
HCHL	-2.31	-2.48	-1.45	-0.98	64	49	36	32
PSP	-2.25	-2.27	-1.84	-1.23	233	173	140	124
OPD	-2.66	-2.67	-2.35	-1.42	234	161	142	122
OPDA	-2.34	-2.34	-2.06	-1.32	249	176	149	128
OPDL	-2.85	-2.77	-2.07	-1.14	143	109	102	91
SPD	-2.91	-2.91	-2.56	-1.69	156	110	96	85
SPDA	-2.59	-2.60	-2.25	-1.49	152	109	94	84
SPDL	-2.43	-2.44	-0.79	-0.51	31	30	15	14
PTP	-3.22	-2.94	-2.56	-1.33	161	121	95	84
PDL	-2.83	-2.90	-2.46	-1.48	119	88	86	75
LAC	-2.30	-2.19	-1.63	-1.05	92	74	81	72

*Note.* MZ = monozygotic; DZ = dizygotic; DE = double entry; SE = single entry; CWR = composite word recognition; TWR = timed word recognition; PWR = PIAT word recognition; OCH = orthographic choice composite; OCHA = orthographic choice accuracy; OCHL = orthographic choice latency; HCH = homonym choice composite; HCHA = homonym choice accuracy; HCHL = homonym choice latency; PSP = PIAT spelling; OPD = oral phonological decoding composite; OPDA = oral phonological decoding accuracy; OPDL = oral phonological decoding latency; SPD = silent phonological decoding composite; SPDA = silent phonological decoding accuracy; SPDL = silent phonological decoding latency; PTP = phoneme transposition; PDL = phoneme deletion; LAC = Lindamood auditory conceptualization.

<sup>a</sup>Proband selection criterion of positive school history and  $-1.5$  SD.

bivariate combinations of the 19 variables would result in 171 genetic correlations. Thus, to simplify results, we only used several selected variables for these analyses. The CWR was analyzed together with three OC tasks (OCHA, HCHA, and PSP), two PD tasks (OPD composite and SPDA), and three PA tasks (PTP, PDL, and LAC). Then OCHA was selected as the representative OC task and was analyzed with two PD tasks (OPD and SPDA) and two PA variables (PTP and PDL). The genetic correlation between PD and PA was assessed by bivariate regression of OPD with PTP and PDL, and finally, these two PA measures were analyzed together. In addition, latency and accuracy versions of each available task were compared.

For this bivariate analysis the group selected for school history of RD and the control group were combined to create an unselected sample of twins. Importantly, some of the twins in the control group could now be considered probands if their score on the task analyzed fell below the  $-1.5$  SDs criterion used in this bivariate analysis. Table 4 presents bivariate proband and cotwin means where it is evident that when probands are selected on one task, DZ cotwins' scores on a second task regress more toward the population mean than MZ cotwins' scores. Bivariate heritabilities among word recognition (WR), OC, PD, and PA tasks are generally significant (Table 4), providing evidence for common genetic etiologies for deficits across these reading and language skills.

TABLE 3  
Heritability ( $h_g^2$ ), Shared Environment ( $c_g^2$ ), and Nonshared Environment ( $e_g^2$ ) Estimates of Group Deficits in Each Task<sup>a</sup>

<i>Task</i>	$h_g^2$	<i>SE</i>	<i>p</i>	$c_g^2$	<i>SE</i>	<i>p</i>	$e_g^2$	<i>SE</i>
CWR	.54	.08	< .001	.39	.09	< .001	.06	.03
TWR	.57	.08	< .001	.37	.10	< .001	.06	.03
PWR	.59	.08	< .001	.33	.09	< .001	.08	.03
OCH	.60	.11	< .001	.28	.11	.005	.12	.04
OCHA	.67	.12	< .001	.17	.12	.086	.16	.04
OCHL	.47	.17	.004	.27	.15	.036	.27	.07
HCH	.62	.16	< .001	.18	.14	.090	.20	.06
HCHA	.58	.16	< .001	.20	.14	.072	.21	.06
HCHL	.47	.25	.035	.16	.24	.250	.37	.08
PSP	.55	.10	< .001	.26	.10	.004	.18	.04
OPD	.71	.10	< .001	.18	.10	.045	.12	.04
OPDA	.64	.09	< .001	.25	.10	.008	.12	.03
OPDL	.63	.16	< .001	.10	.13	.230	.27	.06
SPD	.60	.13	< .001	.28	.13	.014	.12	.05
SPDA	.59	.13	< .001	.28	.13	.015	.13	.04
SPDL	.23	.29	.222	.10	.29	.370	.68	.09
PTP	.69	.15	< .001	.11	.14	.216	.20	.06
PDL	.72	.14	< .001	.15	.14	.137	.13	.06
LAC	.46	.16	.002	.25	.14	.041	.29	.06

*Note.* CWR = composite word recognition; TWR = timed word recognition; PWR = PIAT word recognition; OCH = orthographic choice composite; OCHA = orthographic choice accuracy; OCHL = orthographic choice latency; HCH = homonym choice composite; HCHA = homonym choice accuracy; HCHL = homonym choice latency; PSP = PIAT spelling; OPD = oral phonological decoding composite; OPDA = oral phonological decoding accuracy; OPDL = oral phonological decoding latency; SPD = silent phonological decoding composite; SPDA = silent phonological decoding accuracy; SPDL = silent phonological decoding latency; PTP = phoneme transposition; PDL = phoneme deletion; LAC = Lindamood auditory conceptualization.

<sup>a</sup>Proband selection criterion of positive school history and  $-1.5$  SD.

TABLE 4

Bivariate Monozygotic and Dizygotic Proband and Cotwin Means, Bivariate Heritability ( $h_g^2_{(xy)}$ ), and Genetic Correlation ( $r_g$ ) Computed as a Function of the Univariate and Bivariate Regression Coefficients

<i>Proband Task</i>	<i>Cotwin Task</i>	<i>Proband MZ</i>	<i>Mean DZ</i>	<i>Cotwin MZ</i>	<i>Mean DZ</i>	$h_g^2_{(xy)}$	<i>SE</i>	$r_g$
CWR	OCHA	-2.56	-2.49	-1.32	-0.98	.24	.09	
OCHA	CWR	-2.26	-2.24	-2.26	-1.59	.58	.14	0.81
CWR	HCHA	-2.60	-2.57	-1.38	-1.01	.28	.11	
HCHA	CWR	-2.10	-2.23	-2.34	-1.69	.71	.19	0.94
CWR	PSP	-2.55	-2.48	-1.62	-1.25	.27	.08	
PSP	CWR	-2.20	-2.18	-2.36	-1.75	.55	.12	0.88
CWR	OPD	-2.56	-2.49	-2.01	-1.45	.41	.09	
OPD	CWR	-2.56	-2.58	-2.46	-1.71	.59	.10	0.99
CWR	SPDA	-2.56	-2.55	-1.93	-1.46	.36	.12	
SPDA	CWR	-2.49	-2.46	-2.27	-1.80	.36	.13	0.97
CWR	PTP	-2.56	-2.47	-1.70	-1.22	.34	.14	
PTP	CWR	-3.07	-2.87	-2.30	-1.61	.38	.10	0.70
CWR	PDL	-2.58	-2.55	-1.69	-1.40	.21	.15	
PDL	CWR	-2.71	-2.75	-2.30	-1.65	.49	.14	0.67
CWR	LAC	-2.60	-2.55	-1.21	-1.04	.11	.11	
LAC	CWR	-2.21	-2.17	-2.09	-1.62	.41	.18	0.53
OCHA	OPD	-2.27	-2.24	-1.91	-1.40	.43	.14	
OPD	OCHA	-2.57	-2.58	-1.40	-0.94	.36	.10	0.73
OCHA	SPDA	-2.33	-2.30	-1.88	-1.49	.32	.17	
SPDA	OCHA	-2.49	-2.46	-1.37	-1.07	.23	.14	0.67
OCHA	PTP	-2.27	-2.23	-1.73	-1.34	.33	.20	
PTP	OCHA	-3.08	-2.89	-1.30	-1.00	.15	.11	0.39
OCHA	PDL	-2.34	-2.30	-1.49	-1.37	.08	.20	
PDL	OCHA	-2.71	-2.75	-1.28	-0.94	.26	.14	0.28
OPD	PTP	-2.54	-2.57	-1.77	-1.25	.42	.16	
PTP	OPD	-3.05	-2.88	-2.00	-1.36	.36	.10	0.64
OPD	PDL	-2.54	-2.76	-1.79	-1.45	.37	.16	
PDL	OPD	-2.71	-2.75	-2.03	-1.52	.39	.14	0.67
PTP	PDL	-3.13	-3.05	-2.23	-1.56	.40	.13	
PDL	PTP	-2.71	-2.74	-2.42	-1.59	.62	.18	0.85
TWR	PWR	-2.41	-2.37	-2.28	-1.71	.45	.10	
PWR	TWR	-2.52	-2.44	-2.18	-1.56	.45	.08	0.98
OCHA	OCHL	-2.26	-2.24	-1.19	-0.89	.26	.17	
OCHL	OCHA	-2.68	-2.71	-1.31	-0.95	.27	.12	0.66
HCHA	HCHL	-2.11	-2.23	-0.96	-0.79	.20	.18	
HCHL	HCHA	-2.31	-2.37	-1.18	-1.15	.06	.18	0.23
OPDA	OPDL	-2.25	-2.25	-1.32	-0.90	.37	.14	
OPDL	OPDA	-2.83	-2.73	-1.69	-1.17	.34	.10	0.71
SPDA	SPDL	-2.43	-2.45	-0.31	-0.11	.16	.14	
SPDL	SPDA	-2.21	-2.20	-1.21	-0.94	.24	.25	0.44

*Note.* MZ = monozygotic; DZ = dizygotic; CWR = composite word recognition; OCHA = orthographic choice accuracy; HCHA = homonym choice accuracy; PSP = PIAT spelling; OPD = oral phonological decoding composite; SPDA = silent phonological decoding accuracy; PTP = phoneme transposition; PDL = phoneme deletion; LAC = Lindamood auditory conceptualization; TWR = timed word recognition; PWR = PIAT word recognition; OCHL = Orthographic choice latency; HCHL = homonym choice latency; OPDA = oral phonological decoding accuracy; OPDL = oral phonological decoding latency; SPDL = silent phonological decoding latency.



Indeed, genetic correlations between these tasks are mostly substantial (Table 4). Genetic effects on deficits in WR also seem to affect to a large degree reading component skills such as OC (.81–.94), and even somewhat more PD (.97–.99), and to a lesser degree language skills such as PA (.53–.70). These results suggest a large genetic overlap between deficits in WR and those in the reading component tasks (perhaps a little lower for OC), and a substantial overlap between WR and PA.

Genetic influences on OC deficits correlate about .28–.39 with genetic influences on deficits in PA tasks, about .67–.73 with PD, and about .81–.94 with WR. These results suggest that genetic effects on deficits in orthographic skills also create deficits in the other reading and language skills, more so in WR tasks, somewhat less in PD, with the smallest effect on PA skills. These genetic correlations vary in size depending on the task, and suggest some independent genetic influences on OC deficits. Genetic factors affecting deficits in PD tasks correlate .64–.67 with those in PA, revealing that the genetic effects influencing PA affect both reading component skills, but PD seems more genetically related to PA than OC. Finally, two PA skills, PTP and PDL, as expected, exhibit a large genetic correlation of .85.

In addition, we computed genetic correlations between the accuracy and latency versions of several tasks. The genetic correlation between TWR and PWR was estimated as .98, an expected result given that they are very similar tasks, the main difference consisting in a 2-sec initiation requirement in TWR. The genetic correlations between the accuracy and latency components range from a low .23 for HCH to a substantial .71 for OPD. These estimates suggest that common genetic effects may contribute to both reading accuracy and fluency, although large contributions of independent genetic effects must also be present.

Similarly to the univariate analyses, we also tested for the possibility that the bivariate genetic etiology of group deficits changed across the age range. However, the interaction between age and the bivariate genetic effect was small and nonsignificant for all pairs of measures.

## DISCUSSION

Data from identical and same-sex fraternal twins were analyzed to assess the balance of genetic and environmental influences on group deficits in several reading and related language skills. The reading skills were isolated WR, PD (both oral and silent nonword reading), and OC (participants' sensitivity to words' specific orthographic patterns). Each of the reading skills was tested with multiple measures that included assessments of both accuracy and speed of processing. The language skill was PA, tested with three measures of accuracy in the deletion or manipulation of phonemes in speech. Two sets of behavioral genetic analyses addressed questions about genetic and environmental influences on group deficits in these reading and related language skills. The first set included univariate analyses that separately as-

sessed the genetic and environmental etiology for group deficits in each measure. The second set of analyses used a bivariate approach to assess the degree of shared genetic influence between selected pairs of measures. In the first part of the discussion we consider the results from the univariate analyses for accuracy and speed in the different measures. The second part of the discussion focuses on results from the bivariate analyses. The third concluding section discusses and summarizes the broad implications of the results for theory, etiology, and remediation in RD.

### Univariate Analysis

When probands with a school history for reading problems were selected to be at least 1.5 *SDs* below the mean of the control population on a measure, the pattern of significantly greater DZ than MZ cotwin regression toward the control mean revealed substantial genetic influences on the group deficits in all measures except SPDL. This one exception was qualified by the fact that the sample size was quite small and the standard error of the heritability estimate was relatively large. In general, genetic influences accounted for slightly more than half of the group deficits. Shared environment influences on group deficits were generally much smaller than genetic influences across the measures in this sample, although they were generally significant.

*Word recognition.* It is apparent from the standard errors of the individual genetic and environmental estimates in Table 3 that differences between the measures would not be statistically significant. Nevertheless, there are some interesting trends for WR compared to the other measures. The PWR and TWR measures showed nearly identical patterns of genetic and environmental influence. They had the lowest nonshared environmental influence and the highest shared environmental influence in comparison to other measures, and their heritability levels tended to be slightly lower than for accuracy deficits in the other measures. This suggests that shared environmental factors such as print exposure might have slightly greater influence on WR deficits than on deficits in more basic orthographic and phonological component processes in reading and language.

*Orthographic and phonological skills.* A preliminary study of a small twin sample found that the group deficit in OCH was not significantly heritable (Olson et al., 1989). However, it was highly heritable in the present analyses with a much larger twin sample, as were related orthographic measures of HCH and PSP. Children must have seen the target words in their print environment to make an informed choice in the orthographic tasks, and this might be more associated with en-

vironmental differences in print exposure (Olson, Forsberg, & Wise, 1994). However, there was no indication of any greater environmental influence on the group deficit in OC, compared to that found for other measures including PD and PA. This does not mean that exactly the same genes and environmental influences are acting on all measures: The bivariate analyses discussed later showed that significant overlap was found in genetic etiology, but some differences may also occur across some of the measures.

*Latency and accuracy.* The results for latency and accuracy in the reading measures were generally similar, although there was a pattern of higher nonshared environment influence and slightly lower genetic influence for group deficits in latency. The speed with which participants performed the tasks may have been more influenced than accuracy by idiosyncratic interpretations of task demands to “answer as quickly as you can.” There was almost no difference in the pattern of genetic and environmental influences on PWR and TWR, so the additional demand for speed as well as accuracy in the TWR task did not seem to add to the total genetic influence on group deficits. However, the bivariate analyses discussed in the next section suggested some partial independence for the source of genetic influences on accuracy and latency within the orthographic and phonological tasks.

## Bivariate Analysis

From the foregoing univariate analyses, we observed similarly high genetic influences on group deficits across the measures, including measures of orthographic and phonological skills, which theory suggests may depend on partly independent underlying processes (Coltheart, Curtis, Atkins, & Haller, 1993; Olson et al., 1985, 1999). Equally heritable group deficits in two reading-related skills could possibly involve at least partly independent genetic pathways for those deficits. The present bivariate analyses assessed the genetic correlations between measures by selecting the proband on one measure and assessing differences in MZ and DZ cotwin regression on a second measure. The genetic correlations estimated from these analyses can illustrate the balance between common and independent genetic influences within different pairs of measures.

*Genetic correlations between deficits in orthographic and phonological skills.* In general, a substantial proportion of the genetic effects causing reading and language deficits is shared among the reading and language tasks analyzed, especially between WR and PD, and to a lesser extent between WR and OC, and between WR and PA.

Evidence for partial genetic independence of orthographic- and phonological-skill deficits emerged from the less-than-perfect genetic correlations between OCHA and accuracy in the PD tasks (.67–.73). Additional evidence for the partially independent genetic etiology for deficits in orthographic and phonological reading skills came from their different genetic correlations with deficits in PA. Genetic correlations between OCHA and PTP (.39) or PDL (.28) were substantially lower than the genetic correlations between OPD and PTP (.64) or PDL (.67). In summary, the present pattern of results suggests that although there were significant genetic correlations between measures of OC and PD, the less-than-perfect correlations indicated also some independent genetic etiology. Moreover, the difference between orthographic and phonological correlations with PA suggested that at least part of their independence in genetic pathways was through genetic influences on PA: Deficits in PA were more closely linked both phenotypically and in genetic etiology with deficits in PD. We would like to find some other language or perceptual skill that is more phenotypically and genetically related to deficits in OC. Wolf (1999) and others have suggested that slow rapid automatized naming (RAN) of letters and numbers is uniquely linked to deficits in OC. However, previous phenotypic factor analyses including poor readers from the present twin sample have not supported this hypothesis (Olson, Forsberg, & Wise, 1994), and recent bivariate genetic analyses between group deficits in RAN and PD or OC yielded very similar estimates in magnitude (Compton, Davis, DeFries, Gayán, & Olson, 2001).

*Genetic correlations between deficits in accuracy and latency.* Genetic correlations between accuracy and latency measures were first explored for TWR and PWR. In the univariate analyses, these two measures had very similar levels of heritability for their group deficits. In the bivariate analyses, the two measures also had a nearly perfect (.98) genetic correlation. Both WR measures required accurate performance, and the additional constraint of response initiation within 2 sec may not have been stringent enough to reveal any independent genetic etiology for WR speed on top of that for WR accuracy. The orthographic and phonological choice tasks included direct measures of response latency. The latency variables were based only on correct responses to reduce contamination from rapid guessing. We limit our discussion to bivariate accuracy and latency results from OCH ( $r_g = .66$ ) and OPD ( $r_g = .71$ ) because these measures had significant bivariate heritabilities for deficits in latency and accuracy (see bottom section of Table 4). The genetic correlations showed that there were significant shared genetic influences on accuracy and latency deficits in these measures. On the other hand, there also appeared to be some independent genetic influences because the correlations were well below unity. We can only speculate about possible reasons for partly separate genetic influences on accuracy and latency within the tasks. For example, there might be ge-

netically based temperament differences in participants' response to the task demands for speed (some might be more cautious or some might be more competitive in their desire to respond quickly), and these temperament differences may have relatively little influence on accuracy. There also might be constitutional differences in participants' basic processing speed that are partly independent from accuracy (Wolf, 1999).

### Implications of the Results for Theory, Etiology, and Remediation in Reading Disabilities

Theories in acquired and developmental RD have generally embraced the concept that there is more than one pathway to poor reading. This view is reflected in contemporary explorations of deficits in accuracy versus speed in component reading skills (Compton et al., 2001; Wolf, 1999) and in the longer history of attempts to characterize individual poor readers by their unique patterns of orthographic and phonological deficits (Baron, 1979; Boder, 1976; Castles & Coltheart, 1993; Manis et al., 1999; Olson et al., 1985). The results of the study herein suggest that differences in speed and accuracy, and orthographic and phonological performance profiles may have a partly genetic basis: The genetic effects that influenced group deficits in accuracy and speed in the reading tasks were partly common and partly independent. The genetic effects that influenced accuracy in OC and PD were also partly common and partly independent. The partial independence in genetic etiology for orthographic and phonological reading deficits was further supported by higher genetic correlations between measures of PD and the language skill of PA. This independence may ultimately be supported at a molecular genetic level if specific genes are found to have stronger effects on deficits in specific component reading and language skills. Recent linkage studies suggest that such process-specific genes may exist, although given the large genetic correlations among reading and language skills obtained in this study, genes with large effects are most likely common to most of these reading skills (Fisher et al., 1999; Gayán et al., 1999; Grigorenko et al., 1997; Grigorenko, Wood, Meyer, & Pauls, 2000).

The foregoing discussion shows that there is not likely to be any single genetic pathway for reading disability. We must also emphasize that there is no single answer to the overall proportions of genetic and environmental influences on group deficits in reading. The univariate analyses generally showed stronger genetic rather than environmental influences on group deficits in our measures of reading and related language skills, but these estimates are specific to the range of genetic and environmental variation within the sample. Twins in volunteer samples may not always reflect the full range of shared environment in the population. If the environmental variation in the general population is larger than that in the present sample, genetic influences on group deficits could be lower and environmental influences higher than in the present sample. In fact, one common environmental impediment to read-

ing development, English as a second language, was deliberately excluded from the present sample. It seems likely that if these children had been included in the sample, the estimates of influence from shared environment would have been higher and influences from genetic factors would have been reciprocally lower.

A second point should be understood regarding the present estimates of relatively small environmental influences on group deficits in reading development: The present behavioral–genetic analyses assume the same degree of shared environment influence for MZ and DZ twins. However, the MZ twins' greater genetic similarity may lead them to select more similar microenvironments for reading within their shared family environment. For example, if a pair of MZ twins shared a genetically based difficulty in learning to read, they might both choose to read less than normally developing children would in that environment. In contrast, if one DZ twin had a genetically based reading problem whereas the other did not, their genetic differences could lead to very different levels of print exposure and reading practice within the same home or school. It is important to recognize the possibility that at least part of the genetic influence on reading deficits may be through differences in environmental selection, and this may have implications for remediation.

Evidence for genetic influences on deficits in reading and related skills should not discourage our best efforts toward environmental intervention and remediation. Many examples of genetically influenced problems exist, such as myopia and diabetes, for which extraordinary environmental intervention can be quite effective. We alluded earlier to the possible influence of genetically based constraints on reading practice. Much evidence shows that increased reading practice in a supportive environment, as with computer-assisted reading (Olson, Wise, Ring, & Johnson, 1997; Wise, Ring, & Olson, 2000), can be helpful for the remediation of RD. Evidence also exists that intervention directed toward highly heritable deficits in phonological reading and language skills can have significant benefits for those skills and for word reading accuracy in younger children with reading disability (Torgesen et al., 1999; Wise et al., 2000). The present evidence for partly independent genetic etiologies in different reading-related skills (accuracy vs. latency, orthographic vs. phonological processing) supports the idea that remediation should be individually designed to help with children's specific deficits. Ultimately we hope to use information from disabled readers' specific behavioral profiles and reading-related genes to select the most effective methods for early intervention and remediation.

#### ACKNOWLEDGMENTS

This work was supported in part by program project and center grants from the NICHD (HD-11681 and HD-27802), and RO1 HD-22223.

The contributions of staff members of the many Colorado school districts that participate in our research, and of the twins and their families, are gratefully acknowledged.

## REFERENCES

- Bakwin, H. (1973). Reading disability in twins. *Developmental Medicine and Child Neurology*, *15*, 184–187.
- Barker, T. A., Wagner, R. K., & Torgesen, J. K. (1992). The role of orthographic processing skills on five different reading tasks. *Reading Research Quarterly*, *27*, 334–345.
- Baron, J. (1979). Orthographic and word specific mechanisms in children's reading of words. *Child Development*, *50*, 60–72.
- Boder, E. (1976). School failure—Evaluation and treatment. *Pediatrics*, *58*, 394–403.
- Bruce, D. J. (1964). The analysis of word sounds by young children. *British Journal of Psychology*, *34*, 158–170.
- Castles, A. E., & Coltheart, M. C. (1993). Varieties of developmental dyslexia. *Cognition*, *47*, 149–180.
- Chall, J. S. (1967). Learning to read: The great debate. New York: McGraw-Hill.
- Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud—Dual-route and parallel-distributed-processing approaches. *Psychological Review*, *100*, 589–608.
- Compton, D. L., Davis, C. J., DeFries, J. C., Gayán, J., & Olson, R. K. (2001). Genetic and environmental influences on reading and RAN: An overview of results from the Colorado twin study. In M. Wolf (Ed.), *Conference proceedings of dyslexia research foundation conference in extraordinary brain series: Time, fluency, and developmental dyslexia* (pp. 277–303). Baltimore: York Press.
- DeFries, J. C., Fulker, D., & LaBuda, M. (1987). Evidence for a genetic aetiology in reading disability of twins. *Nature*, *329*, 537–539.
- DeFries, J. C. (1985). Colorado reading project. In D. B. Gray & J. F. Kavanagh (Eds.), *Biobehavioral measures of dyslexia* (pp. 107–122). Parkton, MD: York Press.
- DeFries, J. C., Filipek, P. A., Fulker, D. W., Olson, R. K., Pennington, B. F., Smith, S. D., & Wise, B. W. (1997). Colorado Learning Disabilities Research Center. *Learning Disabilities: A Multidisciplinary Journal*, *8*, 7–19.
- DeFries, J. C., & Fulker, D. W. (1985). Multiple regression analysis of twin data. *Behavior Genetics*, *15*, 467–473.
- DeFries, J. C., & Fulker, D. W. (1988). Multiple regression analysis of twin data: Etiology of deviant scores versus individual differences. *Acta Geneticae Medicae et Gemellologiae*, *37*, 205–216.
- DeFries, J. C., & Gillis, J. J. (1991). Etiology of reading deficits in learning disabilities: Quantitative genetic analysis. In *Neuropsychological foundations of learning disabilities* (pp. 29–47). New York: Academic.
- Dunn, L. M., & Markwardt, F. C. (1970). *Examiner's manual: Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Service.
- Ehri, L., & Wilce, L. (1980). The influence of orthography on readers' conceptualization of the phonemic structure of words. *Applied Psycholinguistics*, *1*, 371–385.
- Fisher, J. H. (1905). Case of congenital word blindness (Inability to learn to read). *Ophthalmological Review*, *24*, 315–318.
- Fisher, S. E., Marlow, A. J., Lamb, J., Maestrini, E., Williams, D. F., Richardson, A. J., Weeks, D. E., Stein, J. F., & Monaco, A. P. (1999). A quantitative trait locus on chromosome 6p influences different aspects of developmental dyslexia. *American Journal of Human Genetics*, *64*, 146–156.
- Frith, U., Wimmer, H., & Landerl, K. (1998). Differences in phonological recoding in German- and English-speaking children. *Scientific Studies of Reading*, *2*, 31–54.
- Gayán, J., & Olson, R. K. (1999). Reading disability: Evidence for a genetic etiology. *European Child and Adolescent Psychiatry*, *8*(Suppl. 3), S52–S55.
- Gayán, J., Smith, S. D., Cherny, S. S., Cardon, L. R., Fulker, D. W., Brower, A. M., Olson, R. K., Pennington, B. F., & DeFries, J. C. (1999). Quantitative trait locus for specific language and reading deficits on chromosome 6p. *American Journal of Human Genetics*, *64*, 157–164.

- Grigorenko, E. L., Wood, F. B., Meyer, M. S., Hart, L. A., Speed, W. C., Shuster, A., & Pauls, D. L. (1997). Susceptibility loci for distinct components of developmental dyslexia on chromosomes 6 and 15. *American Journal of Human Genetics*, *60*, 27–39.
- Grigorenko, E. L., Wood, F. B., Meyer, M. S., & Pauls, D. L. (2000). Chromosome 6p influences on different dyslexia-related cognitive processes: Further confirmation. *American Journal of Human Genetics*, *66*, 715–723.
- Hallgren, B. (1950). Specific dyslexia (congenital word-blindness): A clinical and genetic study. *Acta Psychiatrica et Neurologica*, *65*(Suppl.), 1–287.
- Hermann, K. (1959). Reading disability: A medical study of word blindness and related handicaps. Springfield, IL: Charles C. Thomas.
- Hutzler, F., & Wimmer, H. (2001, June 1). *Computational modeling of learning to read German and English*. Paper presented at the 2001 meeting of the Society for the Scientific Study of Reading, Boulder, CO.
- Knopik, V. S. (2000). Differential genetic etiology of reading-related difficulties as a function of IQ, Department of Psychology, University of Colorado.
- Knopik, V. S., Alarcón, M., & DeFries, J. C. (1997). Comorbidity of mathematics and reading deficits: Evidence for a genetic etiology. *Behavior Genetics*, *27*, 447–453.
- LaBuda, M. C., DeFries, J. C., & Fulker, D. W. (1986). Multiple regression analysis of twin data obtained from selected samples. *Genetic Epidemiology*, *3*, 425–433.
- Lieberman, I. Y., Shankweiler, D., Fischer, F. W., & Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *Journal of Experimental Child Psychology*, *18*, 201–212.
- Light, J. G., & DeFries, J. C. (1995). Comorbidity of reading and mathematics disabilities: Genetic and environmental etiologies. *Journal of Learning Disabilities*, *28*, 96–106.
- Lindamood, C. H., & Lindamood, P. C. (1979). Lindamood auditory conceptualization test. Hingham, MA: Teaching Resources.
- Lundberg, I., Frost, J., & Peterson, O. (1988). Effects of an extensive program for stimulating phonological awareness. *Reading Research Quarterly*, *23*, 263–284.
- Manis, F. R., Seidenberg, M. S., Stallings, L., Jianisse, M., Bailey, C., Freedman, L., Curtin, S., & Keating, P. (1999). Development of dyslexia subgroups: A one year follow up. *Annals of Dyslexia*, *49*, 105–136.
- Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phonemes arise spontaneously? *Cognition*, *7*, 323–331.
- Nichols, R. C., & Bilbro, W. C. (1966). The diagnosis of twin zygosity. *Acta Genetica et Statistica Medica*, *16*, 265–275.
- Norrie, E. (1939). *Om ordblindhed*. Copenhagen.
- Olson, R. K., Datta, H., Gayán, J., & DeFries, J. C. (1999). A behavioral–genetic analysis of reading disabilities and component processes. In R. M. Klein & P. A. McMullen (Eds.), *Converging methods for understanding reading and dyslexia* (pp. 133–151). Cambridge, MA: MIT Press.
- Olson, R. K., Forsberg, H., & Wise, B. (1994). Genes, environment, and the development of orthographic skills. In V. W. Berninger (Eds.), *The varieties of orthographic knowledge: Vol. I: Theoretical and developmental issues* (pp. 27–71). Dordrecht, The Netherlands: Kluwer.
- Olson, R. K., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Eds.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243–277). Baltimore: Paul H. Brookes.
- Olson, R. K., Kliegl, R., Davidson, B. J., & Foltz, G. (1985). Individual and developmental differences in reading disability. In G. E. MacKinnon & T. G. Waller (Eds.), *Reading research: Advances in theory and practice* (pp. 1–64). New York: Academic.
- Olson, R. K., Wise, B., Conners, F., Rack, J., & Fulker, D. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities*, *22*, 339–348.



- Olson, R. K., Wise, B. W., Ring, J., & Johnson, M. (1997). Computer-based remedial training in phoneme awareness and phonological decoding: Effects on the post-training development on word recognition. *Scientific Studies of Reading, 1*, 235–253.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115.
- Rack, J. P., Snowling, M. J., & Olson, R. K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly, 27*, 28–53.
- Rodgers, B. (1983). The identification and prevalence of specific reading retardation. *British Journal of Educational Psychology, 53*, 369–373.
- Rosner, J., & Simon, D. (1971). The auditory analysis test: An initial report. *Journal of Learning Disabilities, 4*, 384–392.
- Shaywitz, S. E., Escobar, M. D., Shaywitz, B. A., Fletcher, J. M., & Makuch, R. (1992). Evidence that dyslexia may represent the lower tail of a normal-distribution of reading-ability. *New England Journal of Medicine, 326*, 145–150.
- Stanovich, K. E., & Siegel, L. S. (1994). Phenotypic performance profile of children with reading disabilities: A regression-based test of the phonological-core variable-difference model. *Journal of Educational Psychology, 86*, 24–53.
- Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402–433.
- Thomas, C. J. (1905). Congenital word blindness and its treatment. *Ophthalmoscope, 3*, 380–385.
- Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., & Conway, T. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*, 579–593.
- Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33*, 223–238.
- Wadsworth, S. J., Olson, R. K., Pennington, B. F., & DeFries, J. C. (2000). Differential genetic etiology of reading disability as a function of IQ. *Journal of Learning Disabilities, 33*, 192–199.
- Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1994). The development of reading related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*, 73–87.
- Wechsler, D. (1981). *Examiner's manual: Wechsler Intelligence Scale for Children—Revised*. New York: Psychological Corporation. (Original work published 1974)
- Wise, B. W., Ring, J., & Olson, R. K. (2000). Individual differences in gains from computer-assisted remedial reading with more emphasis on phonological analysis or accurate reading in context. *Journal of Experimental Child Psychology, 77*, 197–235.
- Wolf, M. (1997). A provisional, integrative account of phonological and naming speed deficits in dyslexia: Implications for diagnosis and intervention. In B. Blachman (Ed.), *Foundations of reading acquisition and dyslexia* (pp. 67–92). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Wolf, M. (1999). What time may tell: Towards a new conceptualization of developmental dyslexia. *Annals of Dyslexia, 49*, 3–28.
- Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests—Revised*. Circle Pines, MN: American Guidance Service.
- Zerbin-Rüdin, E. (1967). Congenital word-blindness. *Bulletin of the Orton Society, 17*, 47–56.