Available online at www.sciencedirect.com

SCIENCE DIRECT®

ACADEMIC PRESS

J. Experimental Child Psychology 84 (2003) 97–123

Journal of Experimental Child Psychology

www.elsevier.com/locate/jecp

Genetic and environmental influences on individual differences in printed word recognition

Javier Gayán[a,*] and Richard K. Olson[b]

[a] *Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK*
[b] *Department of Psychology, University of Colorado, Boulder, CO, USA*

Received 21 March 2002; revised 30 October 2002

Abstract

The genetic and environmental etiologies of individual differences in printed word recognition and related skills were explored in 440 identical and fraternal twin pairs between 8 and 18 years of age. A theoretically driven measurement model identified five latent variables: IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding. Cholesky decomposition models on these five latent constructs revealed the existence of both common and independent genetic effects, as well as non-shared environmental influences. There was evidence for moderate genetic influences common between IQ, phoneme awareness, and word-reading skills, and for stronger IQ-independent genetic influences that were common between phoneme awareness and word-reading skills, particularly phonological decoding. Phonological and orthographic coding skills in word recognition had both significant common and significant independent genetic influences, with implications for "dual-route" and "connectionist" reading models, subtypes of reading disabilities, and the remediation of reading disabilities.

Keywords: Word reading; Phonological and orthographic skills; Twin modeling; Behavior genetics

---

There are substantial individual differences in children's accuracy and fluency in printed word recognition. These differences are strongly correlated with the ultimate

---

[*] Corresponding author. Fax: 1-419-828-7278.
E-mail address: gayan@well.ox.ac.uk (J. Gayán).

doi:10.1016/S0022-0965(02)00181-9

goal of reading, comprehension of text (Perfetti, 1985; Shankweiler et al., 1999). Children with disabilities in reading comprehension are usually in the low tails of normal distributions for word-reading accuracy and fluency, while the best reading comprehenders are typically in the upper tails. The importance of accurate and fluent recognition of printed words has driven extensive research on the development of its component skills of phonological decoding and orthographic coding (Barker, Wagner, & Torgesen, 1992; Olson, Kliegl, Davidson, & Foltz, 1985), on "connectionist" modeling of the development of those skills (Harm & Seidenberg, 1999; Seidenberg & McClleland, 1989), and on correlated language and memory skills (Catts & Kamhi, 1999; Olson, 1994; Wagner & Torgesen, 1987).

Large environmental differences in language background, quality of reading instruction, and continued support for reading are obvious causes of word-reading ability differences in diverse societies such as in the United States (Baker, Scher, & Mackler, 1997; Whitehurst & Lonigan, 1998). Yet substantial individual differences remain within sub-populations with relatively normal and homogeneous levels of reading support. The present research used data from identical and fraternal twins to explore the genetic and environmental bases of these individual differences in word reading, component skills in orthographic coding and phonological decoding, general intelligence (IQ), and an important analytic language skill called phoneme awareness. While estimates of genetic and environmental influence were obtained for each skill independently, more theoretically interesting analyses explored the balance of shared and independent genetic and environmental influences between the skills. We will describe and discuss the importance of these different skills for models of reading development and individual differences after briefly reviewing the rationale for our behavioral genetic analyses.

Data from identical and fraternal twins were used in the present research because identical twins share all their genes, while fraternal twins only share half of their segregating genes on average. Thus, any extra similarities of identical twins over fraternal twins can be of genetic origin (Plomin, DeFries, McClearn, & McGuffin, 2001). Behavioral-genetic models can be applied to twin data to give quantitative estimates of the balance of genetic and environmental influences on individual differences in a trait like word-reading ability, and on its relations with other variables (Neale & Cardon, 1992). In addition, it is possible to separate the influences of environmental factors that are shared by the twins from those that are not shared. For example, since identical twins have the same genes and grow up in the same home and school environment, any differences between identical twins must be due to non-shared environmental influences and/or test error.

Latent-trait modeling with multiple measures of each trait was used in the present analyses to estimate non-shared environmental influences without contamination from test error. Multiple measures were used to estimate latent traits for two important component skills in word recognition called phonological decoding and orthographic coding (Olson et al., 1985; Olson, Forsberg, & Wise, 1994a). Phonological decoding is typically measured by having participants read pronounceable nonwords aloud (e.g., tegwop, framble). This provides an index of the readers' knowledge of common grapheme/phoneme correspondences, and of their ability to use this knowledge to decode phonologically both nonwords and words they might encounter for the first time in print. Accurate and fluent phonological decoding may be an important "self-teaching" mechanism for learning new words (Share, 1995). Children with specific reading disability tend to be uniquely deficient in phonological decoding (Rack, Snowling, & Olson, 1992).

However, the frequent irregularities of English grapheme/phoneme correspondences across different words indicate that phonological decoding is not the only mechanism for reading English. Children must also learn the specific orthographic patterns for words, compared to other orthographic patterns that would sound the same but are nonwords (i.e., rain versus rane), and they must be able to distinguish the meaning of homophones based on their specific spelling patterns instead of how they sound (i.e., bear versus bare). This skill in rapidly remembering and interpreting the specific orthographic patterns for words is called orthographic coding. It is likely that the development of orthographic coding is partly dependent on phonological decoding (Ehri, 1989), but it is also possible that orthographic coding has partly independent genetic and environmental etiologies and different relations to other cognitive skills.

The reading-related cognitive skills explored in the present analyses include Wechsler full-scale IQ (Wechsler, 1974, 1981), and the language skill phoneme awareness. IQ is a very broad and general assessment of cognitive ability. The components of IQ most closely related to word reading are the verbal subscales (Olson et al., 1994a), and verbal IQ can be substituted for full-scale IQ in the present analyses with nearly identical results. (Full-scale IQ is used here because its reliability is well established.) There have been many studies of the relation between IQ and reading, with frequent findings of moderate correlations with word recognition, and stronger correlations with reading comprehension (Conners & Olson, 1990). However, there has been little research on the genetic and environmental etiologies of these correlations, or of the correlations between IQ, phoneme awareness, phonological decoding, and orthographic coding. Previous evidence from a subset of the present twin sample suggested that there is a significant genetic correlation between IQ and a composite measure of word recognition, reading comprehension, and spelling (Alarcón & DeFries, 1997). In the present study we explored more specific relations between IQ, phoneme awareness, and component word-reading skills.

Phoneme awareness is a very specific reading-related cognitive skill. It is defined here as the ability to isolate and manipulate the segments of speech at the level of the single phoneme. For example, we measured how well children were able to play a "Pig Latin" game wherein they moved the first consonant or consonant cluster of a spoken word to the end and add the /ay/ sound (e.g., "pig" becomes "igpay"). In another task called phoneme deletion, they were asked to delete a spoken phoneme (e.g., /r/) from a spoken nonword (e.g., /prot/), and say the resulting word ("pot"). There is extensive evidence that phoneme awareness is strongly correlated with word-reading skills, particularly in younger children (Wagner, Torgesen, & Rashotte, 1994), and phoneme awareness is uniquely deficient in most children with specific reading disability (Rack et al., 1992).

Awareness of phonemes may not be the only important phonological language skill in very early reading development. Some research has suggested the additional importance of young children's awareness for the larger ends of spoken syllables that define their rhyme relations with other spoken syllables (bone, groan, throne) (Bryant, 2002; Goswami, 2002). However, other research has suggested that phoneme awareness is a better predictor of early reading skill and no additional variance is accounted for by measures of rhyme sensitivity (Hulme et al., 2002). Unfortunately, we do not have sufficient measures of rhyme sensitivity to address this question in our older sample. Our exclusive focus on the language skill of phoneme awareness in the present analyses was dictated by its lack of ceiling effects in our older participants and the availability of three phoneme-level measures for latent trait modeling.

We addressed two main questions about the genetic and environmental relations between phoneme awareness and other variables. The first question concerned the amount of independent genetic and environmental variance in phoneme awareness after removing genetic and environmental variance related to IQ. There is much evidence that phoneme awareness does have independent phenotypic variance related to reading after controlling for IQ, but the balance of genetic and environmental influences on this independent variance has not been explored. The second question concerned the IQ-independent relations between phoneme awareness, word recognition, phonological decoding, and orthographic coding. Only one behavioral-genetic study has approached this question with a relatively small sample of 6-year-old twins (Hohnen & Stevenson, 1999). They found that after controlling for other language skills including verbal subscales from an IQ test, the relation between phoneme awareness and word recognition was entirely due to shared environment influences. Here, we explore whether the same pattern holds in the present older and much larger sample.

A third major question concerned the genetic and environmental etiologies of the correlated and independent variance for phonological decoding and orthographic coding. Phonological decoding and orthographic coding both involve reading processes, so it is reasonable to hypothesize that they would share some genetic and environmental influences. On the other hand, there may also be some genetic and environmental influences that are specific to each skill. It is clear that a reader must have seen a word like rain in their print environment to make a correct choice between that word and its phonological foil (rane). Indeed, there is evidence that this skill is uniquely related to various measures of print exposure, which suggests the importance of environmental influences (Cunningham, Perry, & Stanovich, 2001; Olson et al., 1994a; Stanovich & West, 1989). However, it is also possible that there are unique and specific genetic influences on individual differences in orthographic coding that are not shared with phonological decoding. If so, this would have implications for general models of word reading and for individual differences in reading process profiles, including subtypes of reading disability, that will be considered in the Discussion.

Answers to all of the above questions may vary depending on participant characteristics such as gender, age, and range of reading skill. Therefore, we tested the significance of measurement and genetic model differences for males versus females, older children versus younger children, and twins with school history for reading difficulty in at least one member of each pair, versus pairs with no school history for reading problems.

In summary, the present series of analyses assessed the proportional impact of genetic and environmental factors on individual differences in word reading, specifically focusing on discovering effects common or specific to word-reading and related skills. Latent constructs of IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding were estimated in a measurement model from observed indicator variables. These latent variables were error-free representations of the constructs that could be included in a more complex structural model. Subsequent structural equation analyses separated the common and independent genetic/environmental factors influencing the three reading constructs (word recognition, phonological decoding, and orthographic coding). We also included IQ and phoneme awareness in the analyses to determine the degree of genetic covariation and independence between general cognitive ability, phonological language skills, and the component reading skills. The specific nature of genetic influence was explored by comparing the fit of models assuming additive versus non-additive genetic transmission. Finally, we tested for differences in the pattern of genetic and environmental effects by age, gender, and reading level (relatively lower versus higher reading performance).

Methods

Participants

Participants analyzed in the present study were 440 pairs of twins, 257 identical (monozygotic, MZ) and 183 same-sex fraternal (dizygotic, DZ), from the Colorado Learning Disabilities Research Center (DeFries et al., 1997). Twins were identified from school records in 27 Colorado school districts.

One group of twin pairs was initially selected to have at least one member with a school or parent-identified history of reading problems (the low-average-performance group). School history for reading problems was very loosely defined by the presence of any low to low-average reading test scores. While school test data were the main criteria for school history, these data were of widely varying quality and were not always available, particularly among the younger twins. Therefore, we also considered parental references to reading problems and participation in remedial reading classes. The selection for school history of reading problems was loose enough that individual differences in reading and related skills among the twins with school history, including their co-twins with or without school history, approximated a normal distribution. In addition, a second group of twins with no specific school history of reading problems was identified and subsequently tested in the laboratory (the high-average-performance group). This group did include some low readers based on laboratory testing, so their distributions on the experimental variables were approximately normal, though with means that were approximately one standard deviation above those of the low-average-performance group. Note that the high-average and low-average groups are not constrained to include only poor readers or only good readers, since the resulting truncated distributions would not be suitable for the present genetic models of individual differences that require normally distributed data (Neale & Cardon, 1992).

In order to be included in this study, twins must have achieved a verbal and performance IQ score of at least 65, show no evidence of neurological, emotional, or behavioral problems, and have no uncorrected sensory deficits. In addition, English was the first language for all twins. Twin pairs were selected so that both twins in each pair had data on all variables analyzed.

The current sample included 283 pairs of low-average-performance twins (157 MZ and 126 DZ) and 157 pairs of high-average-performance twins (100 MZ and 57 DZ). The number of twin pairs in each group, broken down by age, gender, and zygosity, are shown in Table 1. The mean age for the twins at the time of testing was 10.56 years (range = 7.78–18.58), and the mean was used to divide the sample into young (7.78–10.56) and old (10.57–18.58) groups.

Measures

Participants from the twin families were administered a large battery of tests over two 2.5 h sessions. Tests included in this battery were, among others, the Wechsler Intelligence Scale for Children-Revised (Wechsler, 1974) or the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), the Peabody Individual Achievement Test

Table 1
Number of male (M), female (F), and total (T) monozygotic (MZ) and dizygotic (DZ) twin pairs in the different age (young and old) and reading (high and low average performance) groups

	Young M + F = T	Old M + F = T	Total M + F = Total
Low-average group			
MZ	39 + 37 = 76	44 + 37 = 81	83 + 74 = 157
DZ	35 + 27 = 62	36 + 28 = 64	71 + 55 = 126
Total	74 + 64 = 138	80 + 65 = 145	154 + 129 = 283
High-average group			
MZ	21 + 29 = 50	25 + 25 = 50	46 + 54 = 100
DZ	15 + 17 = 32	13 + 12 = 25	28 + 29 = 57
Total	36 + 46 = 82	38 + 37 = 75	74 + 83 = 157
Total			
	M + F = T	M + F = T	M + F = Total
MZ	60 + 66 = 126	69 + 62 = 131	129 + 128 = 257
DZ	50 + 44 = 94	49 + 40 = 89	99 + 84 = 183
Total	110 + 110 = 220	118 + 102 = 220	228 + 212 = 440

(PIAT) (Dunn & Markwardt, 1970), and other experimental tasks developed to assess performance on certain reading and language skills such as time-limited word recognition, phonological decoding, orthographic coding, and phoneme awareness (Olson et al., 1994a). For the present study, we analyzed the following measures (see Olson, Forsberg, Wise, & Rack, 1994b, for a more complete description of the tasks):

IQ was evaluated by the the Wechsler Intelligence Scale for Children-Revised (Wechsler, 1974) or the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), according to the age of the participant. Full-scale IQ (FIQ) scores were analyzed, though Verbal IQ scores yielded very similar results in the present analyses.

Three measures of *phoneme awareness* (*PA*) were included in the test battery. The 45 trial phoneme transposition (PTP) task is a "pig-latin" game, in which participants were required to take the first sound off the front of a word, put it at the end, and add the sound /ay/. For example, "rope" would become "ope-ray" (Olson, Wise, Conners, Rack, & Fulker, 1989). The 68-trial phoneme deletion (PDL) task presented participants with a spoken nonword 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 word (e.g., "say prot," "now say prot without the /r/ sound"—"pot") (Olson et al., 1994a). Some of these trials started with a word, which after removal of a phoneme, became a nonword. Finally, 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. Participants' scores in all three tasks were based on percent correct responses.

Word recognition (*WR*) was measured with two different tasks. The Peabody Individual Achievement Test (PIAT) for word recognition (PWR) (Dunn & Markwardt, 1970) had participants read across rows of increasingly difficult, unrelated words until they reach an error criterion. There was no time constraint in this standardized measure. An experimental time-limited word recognition test (TWR) (Olson et al., 1989, 1994a) assessed word-recognition accuracy when single words were presented on a computer screen and the participants' correct response was initiated within 2 seconds.

Phonological decoding (*PD*) was measured by the oral reading of 45 1-syllable nonwords (NW1) and 40 2-syllable nonwords (NW2) (e.g., ter, strale, lobsel), presented one at a time on a computer screen, and by a 65 trial silent phonological nonword reading task (PHO) in which participants quickly choose, from three pronounceable letter strings (e.g., coam–baim–goam), the one that would sound like a common word if read aloud (e.g., coam—comb) (Olson et al., 1989; Olson et al., 1994a). *Z*-scores for accuracy and median correct reaction time were combined to produce a composite score for each of the three tasks.

Orthographic coding (*OC*) is defined as the ability to recognize words' specific orthographic patterns. Three specific measures were administered to the sample in order to assess orthographic coding skills. One measure, the 80 trial forced-choice task (orthographic choice, OCH), required the rapid recognition of a target word versus a phonologically identical background foil that was not a word (e.g., rain rane;

sammon salmon) (Olson et al., 1985, 1989). A second measure, the 65 trial homonym choice (HCH) task, required that participants hear a sentence such as "Which is a fruit?" and subsequently choose between a pair of homophones on the computer screen (pair pear) (Olson et al., 1994a). A third measure was the 84 trial PIAT spelling subcomponent (PSP) (Dunn & Markwardt, 1970) which required the untimed forced choice of a target word (e.g., cloudy) which was presented orally, among four orthographically and often phonologically similar alternatives printed on a card (e.g., clowdy cloady cloudey cloudy). This task came to an end if the participant answered incorrectly in five out of seven consecutive trials. Both orthographic choice and homonym choice were scored as the participants' percentage of correct answers, and PIAT spelling as the number of correct responses, that is, the number of trials answered minus incorrect answers.

Analysis

Linear structural equations were used to analyze the covariance structure of the data. Statistical inference from structural equation modeling generally assumes multivariate normality and asymptotic theory (SAS, 1989). Asymptotic theory refers to the law of large numbers, whereby large samples improve the consistency of parameter estimates. The total of 880 participants analyzed in the present study validates the assumption of asymptotic theory. Univariate and multivariate tests of normality were performed using several statistical procedures from two major statistical packages, SAS (SAS, 2000) and SPSS (SPSS, 2000).

Instead of analyzing the observed variables directly, a latent variable model was designed from theory and tested for fit to the data. The advantages of path analysis with latent variables are several, since it provides evidence for the construct validity of the manifest variables and, most importantly, it provides reliable (measured without error) variables, that can be subsequently analyzed in a structural or causal model (Hatcher, 1994).

The measurement model was designed from prior theoretical and empirical evidence, such as factor analysis (Olson et al., 1994a). The measurement model identified 5 correlated latent constructs from 12 observed indicator variables (see Fig. 1). The first latent construct was intended to be a general cognitive factor (IQ) and it was measured uniquely by full-scale IQ (FIQ), entering the published reliability of this task as the factor loading. (Nearly identical results were obtained with verbal IQ because the verbal subscales were most strongly correlated with all reading measures and phoneme awareness.) A second construct was phoneme awareness (PA), a linguistic ability hypothesized to have a crucial role in the development of reading skills. Phoneme awareness was measured by the phoneme transposition (PTP) and phoneme deletion (PDL) tasks, as well as by the Lindamood auditory conceptualization (LAC). Word recognition (WR) was the third latent variable in the measurement model, with timed word recognition (TWR) and PIAT word recognition (PWR) as measurement variables. Phonological decoding (PD) was observed by 1-syllable (NW1) and 2-syllable nonword reading (NW2), as well as the silent nonword task (PHO). Finally, orthographic coding (OC) was the fifth construct, employing orthographic choice (OCH), homonym choice (HCH), and PIAT spelling (PSP) as indicator variables. The measurement model was fitted to the data under two different statistical packages, SAS (SAS, 2000) and Mx (Neale, 1999).

Once the measurement portion of the model was consolidated, Cholesky decomposition models were applied to test for common and independent genetic and environmental effects on the latent constructs (Neale, 1999; Neale & Cardon, 1992). In principle, the Cholesky procedure is similar to hierarchical regression analyses in non-genetic studies, where the independent contribution of a predictor variable is assessed after accounting for its shared variance with other predictor variables. In the present analyses, data from identical and fraternal twins reared together allow for the further decomposition of shared and independent variance among the latent variables into different genetic and environmental components.

Up to three sources of individual differences can be estimated simultaneously from the twin data analyzed here (Neale & Cardon, 1992). We can estimate the relative influence on individual differences from additive genetic effects, shared environmental effects, and non-shared environmental effects. Additive genetic effects (A) refer to the effects of genes that add up additively (i.e., not due to interactions). Shared environmental effects (C) are those which make a pair of twins resemble each other, regardless of their genetic similarity, such as common family and school environments. Non-shared environmental effects (E) include those effects that result in

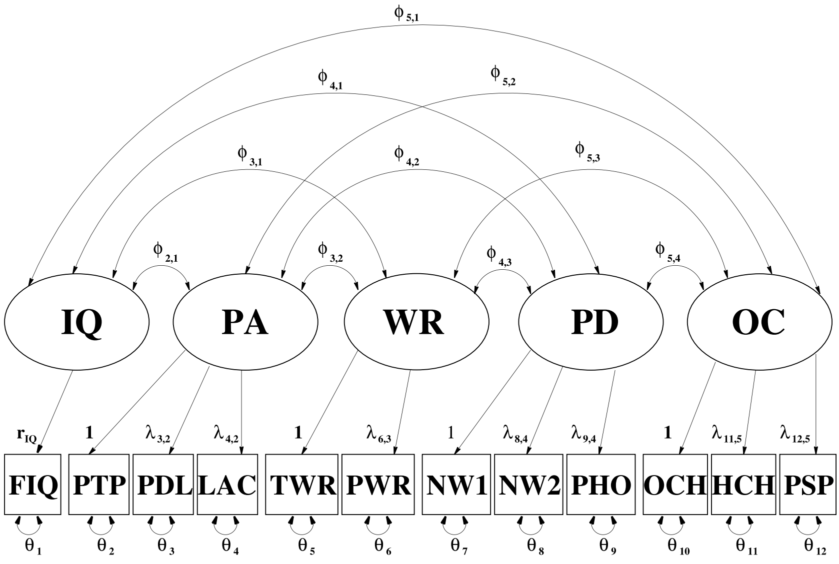

Fig. 1. Measurement model of correlated latent variables and indicator variables (PA, phoneme awareness; WR, word recognition; PD, phonological decoding; OC, orthographic coding; FIQ, full-scale IQ; PTP, phoneme transposition; PDL, phoneme deletion; LAC, Lindamood auditory conceptualization; TWR, time-limited word recognition; PWR, PIAT word recognition; NW1, 1-syllable nonword oral reading; NW2, 2-syllable nonword oral reading; PHO, silent phonological nonword reading; OCH, orthographic choice; HCH, homonym choice; PSP, PIAT spelling).

differences within twin pairs, regardless of their genetic similarity. These effects can be directly estimated from any differences within identical twin pairs, since they share both their genes and their family environment. Non-shared environmental influences typically include test error variance, but our use of latent constructs allows for the separation of this error variance from other more interesting non-shared environmental influences. For illustration, a general bivariate ACE Cholesky model is shown in Fig. 2. This model has an additive genetic factor (A_1) common to variables 1 and 2, with respective factor loadings a_{11} and a_{21}. This model also features an independent additive genetic factor (A_2) which affects variable 2 specifically (factor loading a_{22}). The shared environment (C) and non-shared environment (E) influences are represented similarly. In this manner, the observed variance and covariance among measures is decomposed into common and independent effects: the pattern of covariation is represented by factors common to all variables, common to all except the first variable, common to all except the first two variables, and so on until a specific factor loads onto the last variable (see Fig. 5 for another example).

Another source of individual differences comes from non-additive genetic effects. These effects arise from genes with dominant or recessive genetic transmission, or from epistatic interactions, and thus are also called dominance effects (D). So, we can specify the Cholesky model in a second way, to estimate the separate influences of additive (A) and non-additive (D) genetic effects as well as the effects of the non-shared environment (E). In the present analyses, we fit both ACE and ADE Cholesky models, to see if there is significant evidence for non-additive genetic influence. Unfortunately, due to model limitations (Neale & Cardon, 1992), it is not possible to simultaneously estimate all four possible factors (additive and non-additive genetic, shared and non-shared environment) from data of twins reared together.

Due to the absence of a meaningful order of the members of each twin pair in the study, the covariance matrices employed in these analyses were created using double entry of members of all twin pairs. That is, each member of a twin pair was entered twice, once as Twin 1 and again as Twin 2. This method provides a good control for twin order since the resulting covariance matrix should approximate the average of

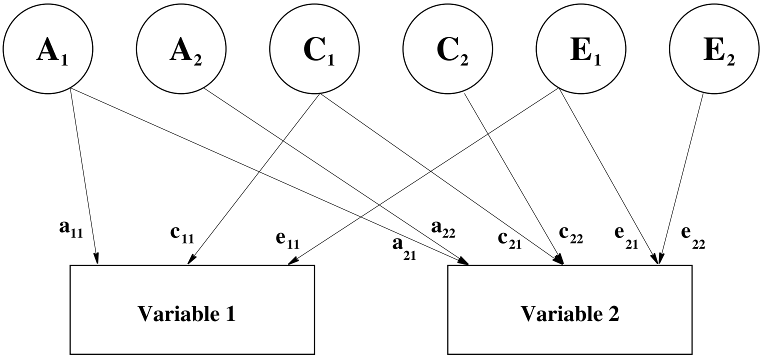

Fig. 2. A general ACE Cholesky model of additive genetic (A), shared environment (C), and non-shared environment (E).

infinite randomizations of twin order. Covariance matrices computed using double entry result in fewer observed statistics, so the degrees of freedom for model fitting need to be adjusted accordingly. Mx (Neale, 1999) easily accommodates this degrees of freedom correction for double entered data.

The order in which variables are entered in a Cholesky decomposition does not affect the estimated variation (i.e., heritability estimate) and covariation among the variables (i.e., genetic correlation). Thus, a first goal of the present analyses was to estimate the relative influence of genetic (additive and non-additive) and environmental (shared and non-shared) effects on the latent constructs. A second goal was to provide estimates of the genetic and environmental correlations among these constructs, which index the degree to which the same genetic (or environmental) factors are influencing two traits.

The order of the variables in a Cholesky model, nonetheless, affects the size of the Cholesky factor loadings. When variables are entered in a useful order in a Cholesky model, specific factor loadings can tell us about independent effects once common effects with other variables have been controlled for. Therefore, in an attempt to examine the specificity and covariation of genetic and environmental influences on reading skills after the effects shared with IQ and phoneme awareness have been partialed out, the possible meaning of the estimated factors is discussed, with the variables entered in the following order: IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding. Secondary analyses focused on variance in IQ shared with word reading after controlling for phoneme awareness, and on the shared and independent variance in phonological decoding and orthographic coding after controlling for IQ.

The phenotypic covariance and correlation matrices of identical and fraternal twin pairs were initially computed separately for low- and high-average reading, young and old, and males and females. Since the analyses by age, gender, and reading level yielded only minor differences, we will begin with a brief review of these results and then will focus primarily on the detailed results from the combined sample.

Results

The results are reported in five main sections. The first section describes high- and low-average reading group differences and their overlapping distributions on all measures. The second section reports results from tests for reading group, age group, and gender group differences in measurement and structural genetic models of relations among the measures. The resulting minor group differences in measurement models and the lack of significant group differences in the structural genetic models justified our focus in the third section on the measurement model and correlations among measures in the combined sample. In the fourth and largest section, we explore the genetic and environmental implications of the twin data, first through simple comparisons of identical and fraternal twin correlations for individual measures, and second through more informative ACE Cholesky models based on the twin covariance matrices for latent traits. Finally, in the fifth section, we report results from tests of ADE Cholesky models that raise the possibility of non-additive genetic effects for some of the traits.

High- and low-average group mean differences and distributions

Although the school history selection criterion for reading problems was rather loose and often pertained to only one twin of a pair in the low-average group, the low-average group means were well below the high-average group means, averaging .98 *SD* across all measures (see Table 2). However, it is also important to recognize the groups' substantial overlap for their normal distributions. To illustrate these overlapping group distributions, plotting normally distributed scores with a group difference of 1 *SD* would result in the overlapping distributions for simulated data shown in Fig. 3. Interestingly, variability in some of the reading and language measures in the low-average group tended to be larger than in the high-average group.

One notable result from Table 2 is that the high- versus low-average reading group difference in IQ is similar to differences for the three reading measures and phoneme awareness. Does this mean that reading differences are simply due to differences in general cognitive ability, or are there IQ-independent influences on reading and related cognitive skills? Before turning to this and related questions in the fourth section, we need to see if it is reasonable to combine the twin sample across reading group, age, and gender to provide more statistically powerful tests of genetic and environmental influences.

Table 2
Mean score, standard deviation (*SD*), and number of individuals (*N*) for variables in this study in the low- and high-average reading performance groups, and effect size of difference between the two groups

	Low-average group			High-average group			Effect
Task	Mean	*SD*	*N*	Mean	*SD*	*N*	Size
Full-scale IQ	100.30	11.68	566	112.81	11.21	314	1.09
Phoneme transposition	−1.10	1.60	566	0.00	1.00	314	0.79
Phoneme deletion	−1.22	1.50	566	0.00	1.00	314	0.92
Lindamood auditory conceptualization	−1.00	1.18	566	0.00	1.00	314	0.90
Timed word recognition	−0.95	1.05	566	0.00	1.00	314	0.92
PIAT word recognition	94.20	11.69	566	107.40	8.32	314	1.24
1-Syllable nonword reading	−1.30	1.42	566	0.00	1.00	314	1.02
2-Syllable nonword reading	−1.22	1.17	566	0.00	1.00	314	1.10
Silent phonological nonword reading	−1.65	1.52	566	0.00	1.00	314	1.24
Orthographic choice	−0.83	1.20	566	0.00	1.00	314	0.74
Homonym choice	−0.78	1.10	566	0.00	1.00	314	0.73
PIAT spelling	92.72	11.48	566	104.77	11.20	314	1.06

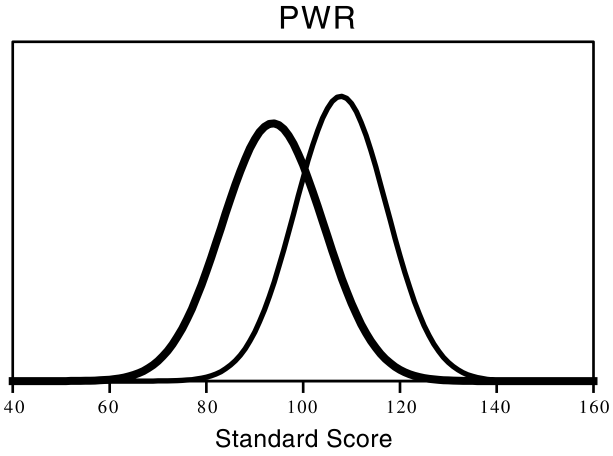

Fig. 3. Simulated representation of low- and high-average group distributions for PIAT word recognition (PWR) standard scores.

Tests for group differences in measurement and structural genetic models

Prior to model fitting, all variables were age-regressed and rank normalized within each of the four age and reading ability groups. To test for group differences in either the measurement (loadings of latent constructs on indicator variables) or the structural (genetic and environmental parameter estimates) part of the model, 16 covariance matrices were created, along four dimensions: zygosity (MZ and DZ), gender (male and female), age (young and old), and average reading performance (low and high). The ACE Cholesky on latent variable model was then fitted to these 16 covariance matrices, providing parameter estimates for each of these gender, age, and reading groups. This complex model resulted in an adequate fit to the data, as indexed by the Non-Normed Fit Index (NNFI, .87), or the Tucker Lewis Index (TLI, .84). Nonetheless, we only used this full model to test for group differences by equating parameter estimates across groups and observing the change in fit of the model.

Using this rationale, very weak evidence for gender differences was found ($p = .05$). In a similar manner, age differences were found significant ($p = .03$), and average reading performance groups also differed significantly ($p < .0001$). Careful analysis of the age and reading performance dimensions revealed that the older/lower performance group had similar characteristics to the younger/higher performance group. These two groups had similar levels of performance in time-limited word recognition mimicking a reading-level comparison. In addition, the older/higher performance group could be equated to the two groups above, suggesting that only the younger/lower performance group was significantly different than the rest of the groups. In depth analysis of these groups, testing parameter estimate differences between groups, reveal that there were no significant differences at the structural level. The difference was at the measurement level, where the only parameter estimate significantly different between the groups was the loading on orthographic choice ($p = .029$), which was larger in the younger/lower performance group (.76) than in the rest of the groups (.68). The reason for this difference could be that the test used to measure orthographic choice skills is more sensitive in the range of the younger/lower participants than for the other participants due to a ceiling effect. The loading on Lindamood auditory conceptualization had the opposite behavior, being smaller in younger poor readers (.66) versus all other groups (.77), although this difference did not achieve statistical significance ($p = .056$).

In summary, at the measurement level, there were small but statistically significant (due to very large sample size) group differences related to gender, age, and reading group. However, there were no significant group differences at the structural level for the pattern of genetic and environmental influences. This result justified combining the sample across the gender, age, and reading groups to increase statistical power for subsequent tests of the measurement model and for behavior genetic analyses.

Measurement model for individual differences in the combined sample

The measurement model was designed from theory and fitted to the overall covariance matrix. Fit indices such as the Goodness of Fit Index (GFI, .94), Root Mean Squared Error Approximation (RMSEA, .07), Bentler's Comparative Fit Index (CFI, .97), Normed Fit Index (NFI, .96), and Non-Normed Fit Index (NNFI, .97) suggested a good fit of the model to the data. All significance tests for factor loadings, residual variances (except for the Full-scale IQ residual variance, which was constrained by the reliability of the measure), and latent factor correlations were significant. Ninety-five percent confidence intervals of parameter estimates are shown in Fig. 4. The reliability of indicator variables and latent factors and the discriminant validity of the latent constructs were tested and found to be at least adequate (Hatcher, 1994).

Phenotypic correlations among the 12 indicator variables were all positive and significant, in the full sample of 880 individuals. Full-scale IQ was correlated with the phoneme awareness and reading variables between .16 and .41; the phoneme awareness measures (phoneme transposition, phoneme deletion, and Lindamood auditory conceptualization) correlated .62–.74 among themselves, and .24–.62 with the reading measures, and the reading variables exhibited correlations in the .40–.82 range. From the fitted measurement model, latent variable correlations were all positive and significant, ranging from .33 (IQ and phonological decoding) to .95 (word recognition and phonological decoding). Ninety-five percent confidence intervals of these correlations are presented on the top section of Fig. 4.

Genetic and environmental influences estimated from additive models

Preliminary evidence from twin correlations for individual measures. Before reviewing results from Cholesky models with latent traits, simple comparisons of identical (MZ) and fraternal (DZ) twin correlations can provide rough indications of genetic and environmental influences for individual measures. Univariate MZ

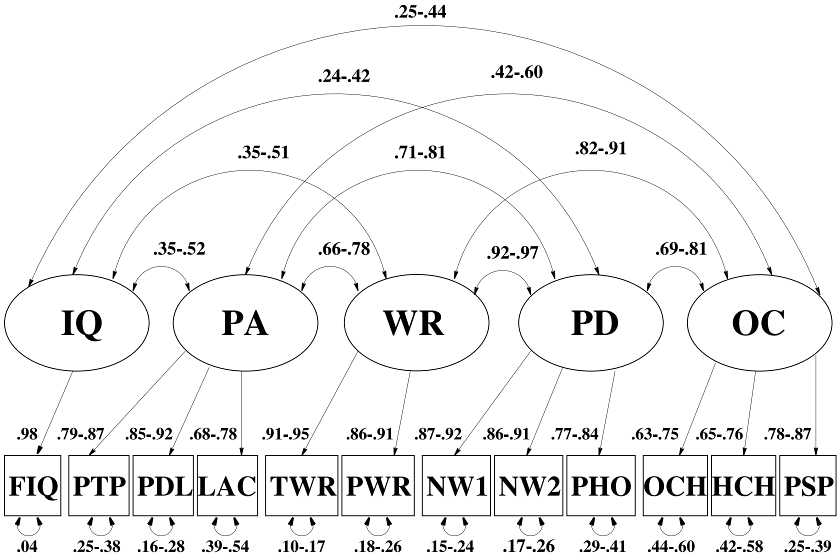

Fig. 4. Ninety-five percent confidence interval of parameter estimates of measurement model (PA, phoneme awareness; WR, word recognition; PD, phonological decoding; OC, orthographic coding; FIQ, full-scale IQ; PTP, phoneme transposition; PDL, phoneme deletion; LAC, Lindamood auditory conceptualization; TWR, time-limited word recognition; PWR, PIAT word recognition; NW1, 1-syllable nonword oral reading; NW2, 2-syllable nonword oral reading; PHO, silent phonological nonword reading; OCH, orthographic choice; HCH, homonym choice; PSP, PIAT spelling).

twin orrelations (.50–.79) were approximately .3–.4 units larger than DZ twin correlations (.14–.45), suggesting the presence of genetic influences in all of the indicator variables. Correlation-based estimates of heritability for most indicator measures were between 54 and 76%, obtained simply by doubling the difference between the MZ and DZ twin correlations. However, the heritability of Lindamood auditory conceptualization was estimated at 34% only, and that of timed word recognition at 90%. For many variables (Full-scale IQ, phoneme transposition, phoneme deletion, Lindamood auditory conceptualization, and 1-syllable nonword reading), DZ twin correlations were larger than half the MZ twin correlations, suggesting that shared environmental effects may have an impact on individual differences in these traits. The difference between the MZ correlation and the heritability provides an estimate of shared environment effects of 0–18%. For the word recognition measures (timed word recognition and PIAT word recognition), for some of the phonological decoding measures (2-syllable nonword reading and silent phonological nonword reading), and especially for the orthographic coding measures (orthographic choice, homonym choice, and PIAT spelling), DZ correlations were smaller than half the MZ correlations, suggesting the presence of non-additive genetic effects (2–48%). Finally, non-shared environmental influences, along with error of measurement, were estimated at 21–50% of the variance.

Fit of the initial ACE Cholesky model with latent traits. The results from MZ and DZ twin correlations provided rough estimates of genetic influence on the individual measures. In the following subsections, more powerful and statistically appropriate ACE Cholesky models were fitted to the MZ and DZ covariance matrices to estimate genetic and environmental influences on latent traits and on the correlations between the traits. The initial model entered IQ first, followed by phoneme awareness, word recognition, phonological decoding, and orthographic coding as shown in Fig. 5. Subsequent models changed the order of variables to address further questions about the independence of genetic and environmental influences for selected variables (Fig. 5 shows only one covariance matrix, while the full model included three covariance matrices factorized for genetic, shared environment, and non-shared environment.) For the initial model, fit indices such as RMSEA (.02), NFI (.93), NNFI (.96), and TLI (.95) suggested a good fit of the model to the data.

Standardized factor loadings for the ACE Cholesky model are shown in Table 3. These loadings present a general overview of the genetic and environmental covariation among the latent traits. Most of these loadings were positive, and a few were negative, suggesting effects of opposing direction. For example, factor A_4 had a positive loading on phonological decoding and a negative loading on orthographic coding. Nonetheless, none of the negative loadings were significantly different from zero.

Genetic and environmental influences on the latent traits. Estimates of the proportion of phenotypic variance explained by genetic (h^2), shared environmental (c^2), and non-shared environmental (e^2) factors were computed from the factor loadings by post-multiplying each standardized factor-loading matrix by its transpose. For example, the heritability of individual differences on IQ can be computed as $.83^2 = .69$. For phoneme awareness, this calculation is: $h^2 = .51^2 + .76^2 = .83$. These parameter estimates and their 95% confidence intervals are presented in Table 4. Genetic factors have significant effects and account for about 69–87% of individual differences on IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding, the remaining variance (13–31%) being explained by environmental factors. Shared environment has a small impact, accounting for 1–13% of the variance, which is not significantly different from zero. Non-shared environmental factors have a similar, though significant, effect, explaining about 9–19% of

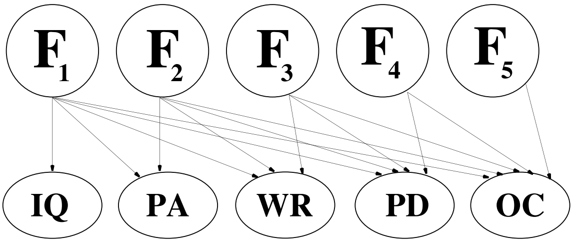

Fig. 5. Latent variable Cholesky model (PA, phoneme awareness; WR, word recognition; PD, phonological decoding; OC, orthographic coding).

Table 3
Genetic (A), shared environment (C), and non-shared environment (E) factor loadings on IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding, computed from full ACE Cholesky model

	A factors					*C* factors					*E* factors				
Task	A_1	A_2	A_3	A_4	A_5	C_1	C_2	C_3	C_4	C_5	E_1	E_2	E_3	E_4	E_5
IQ	.83	–	–	–	–	.36	–	–	–	–	.43	–	–	–	–
Phoneme awareness	.51	.76	–	–	–	−.05	.27	–	–	–	.12	.28	–	–	–
Word recognition	.49	.51	.60	–	–	−.01	.04	.18	–	–	.11	.25	−.19	–	–
Phonological decoding	.44	.55	.51	.20	–	−.17	.13	.05	.00	–	.11	.27	−.24	−.08	–
Orthographic coding	.41	.34	.70	−.23	.22	−.04	−.01	−.04	.00	.00	.10	.18	−.11	.26	.00

Table 4
Genetic (h^2), shared environmental (c^2), and non-shared environmental (e^2) variance estimates (and 95% confidence intervals) for IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding, computed from ACE Cholesky model

Task	h^2	c^2	e^2
IQ	.69(.49–.84)	.13(.00–.31)	.19(.15–.24)
Phoneme awareness	.83(.62–.94)	.08(.00–.27)	.09(.05–.14)
Word recognition	.85(.69–.92)	.04(.00–.19)	.11(.08–.15)
Phonological decoding	.80(.62–.88)	.05(.00–.21)	.15(.11–.20)
Orthographic coding	.87(.75–.94)	.01(.00–.11)	.12(.06–.20)

individual differences. These non-shared environment estimates should be free of error variance, since they were obtained from analysis with latent constructs.

Genetic and non-shared environment correlations for the latent traits. Table 5 presents estimates and 95% confidence intervals for genetic correlations and non-shared environmental correlations. Since shared environment influences were not significant, none of the shared environment correlations were significantly different from either zero or one. A genetic correlation (r_g) can be computed directly from the model by post-multiplying the genetic Cholesky matrix by its transpose, and standardizing the resulting matrix, and indexes the degree to which the same genetic factors are influencing two

Table 5
Genetic (above diagonal) and non-shared environmental (below diagonal) correlations (and 95% confidence intervals) for IQ, phoneme awareness (PA), word recognition (WR), phonological decoding (PD), and orthographic coding (OC), computed from ACE Cholesky model

	IQ	PA	WR	PD	OC
IQ	–	.56(.38–.72)	.53(.37–.68)	.49(.30–.67)	.44(.29–.61)
PA	.39(.18–.59)	–	.75(.65–.86)	.79(.70–.89)	.55(.42–.70)
WR	.33(.15–.49)	.82(.60–.99)	–	.97(.93–1.00)	.92(.85–.99)
PD	.29(.13–.45)	.75(.54–.96)	.98(.89–1.00)	–	.82(.74–.94)
OC	.27(.04–.50)	.58(.25–.92)	.66(.42–.88)	.49(.23–.74)	–

traits. Likewise, a non-shared environmental correlation (r_e) estimates the common variance among two traits which can be explained by the same non-shared environments.

Genetic correlations (Table 5, above diagonal) between IQ and the reading and language measures were moderate, ranging from .44 to .56, with large confidence intervals. Genetic correlations between phoneme awareness and the reading measures were moderate to large in size, and significant, ranging from the smaller .55 for orthographic coding, to the larger .79 for phonological decoding. These two genetic correlations were significantly different from each other, suggesting a larger genetic relationship between phoneme awareness and phonological decoding. Finally, genetic correlations among the reading measures were large (.82–.97), significantly different from zero, but sometimes not significantly different from one.

Non-shared environmental correlations (Table 5, below diagonal) were significant but small between IQ and the other phenotypes, and only moderate between orthographic coding and the other reading and language skills. The most striking result was the large and significant relationships among phoneme awareness, word recognition, and phonological decoding, estimated between .75 and .98, and sometimes not significantly different from one.

Important structural relations among the traits. The order and the number of latent variables in the model were varied in different analyses to test two theoretically interesting questions about the structural relations among the traits. First, to test for the variance in IQ shared with word reading after controlling for phoneme awareness, the latent traits were entered in the following order: phoneme awareness, IQ, word recognition, phonological decoding, and orthographic coding. Interestingly, after controlling for phoneme awareness, there was no significant genetic effect shared by IQ and word recognition. In fact, after controlling for phoneme awareness, there was no significant shared genetic effect for IQ and phonological decoding or orthographic coding either.

Second, to test the hypothesis of genetic independence between phonological decoding and orthographic coding, we also fitted a model with only the latent traits of IQ, phonological decoding and orthographic coding, in this order, arguing that perhaps the genetic covariation with phoneme awareness and word recognition was masking a specific genetic effect for orthographic coding in the full model. Indeed, there was evidence for significant independent genetic effects for phonological decoding and orthographic coding when only controlling for IQ. This model estimated the genetic correlation between phonological decoding and orthographic coding as .80, leaving a substantial and significant 36% of independent genetic variance between these two traits.

Results from an ADE Cholesky model of non-additive genetic effects

The presence of non-additive genetic effects can not be ruled out. Due to methodological constraints (Neale & Cardon, 1992), behavioral genetic models of twins reared together cannot test for additive and non-additive genetic effects, and shared and non-shared environmental effects, all simultaneously. The pattern of MZ and DZ covariation of the twin data partially suggests the possibility of non-additive genetic effects. Because shared environmental effects are not significant and estimated at low values, we can substitute non-additive genetic effects instead of shared environment in the model. The fit of such an ADE Cholesky on latent variable model was nearly identical to that of the ACE Cholesky model.

Estimates of genetic additive (a^2), genetic non-additive (d^2), and non-shared environmental (e^2) variance (and 95% confidence intervals) for IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding, computed from the ADE Cholesky model, are presented in Table 6. Non-additive genetic effects are small and non-significant for IQ, phoneme awareness, and phonological decoding. Nonetheless, they are moderate and significant for word recognition. Finally, orthographic coding skills show a large and significant non-additive genetic influence. Therefore, it is possible that non-additive genetic factors create individual differences in these reading and cognitive skills, especially for orthographic coding skills. When non-additive genetic influences are in play, it is likely that ACE models with their additive assumption will overestimate genetic influence and underestimate shared environment influence.

Discussion

The main goal of the present study was to understand the genetic and environmental etiology of relations between different cognitive and word-reading skills. The skills were modeled as latent traits with multiple indicator variables, and models were compared for differences related to gender, age, and reading ability.

While pursuing the main goal of understanding the etiology of relations between the latent traits, we also obtained estimates of the proportional influences from genes, shared environment, and non-shared environment for each trait. The discussion will begin with a brief review of these results, along with previous estimates from different types of analyses for the heritability of group deficits. Results from both types of analyses will be qualified by questions related to genetic model assumptions and sample selection. Then we will turn to more central questions about the genetic and environmental etiology of relations between the different word-reading and related cognitive skills.

Table 6
Genetic additive (a^2), non-additive (d^2), and non-shared environmental (e^2) variance estimates (and 95% confidence intervals) for IQ, phoneme awareness, word recognition, phonological decoding, and orthographic coding, computed from ADE Cholesky model

Task	a^2	d^2	e^2
IQ	.74(.38–.85)	.08(.00–.43)	.18(.15–.23)
Phoneme awareness	.77(.35–.93)	.14(.00–.56)	.09(.05–.14)
Word recognition	.51(.04–.87)	.38(.02–.85)	.11(.08–.15)
Phonological decoding	.56(.07–.86)	.28(.00–.78)	.15(.11–.20)
Orthographic coding	.18(.00–.76)	.72(.12–.94)	.11(.05–.18)

Previous behavior-genetic analyses of heritability (h_g^2) for group deficits

The Colorado twin study of normal readers and children with reading disabilities has employed two different statistical approaches to estimating genetic and environmental influences on reading and related skills. Previous behavioral-genetic analyses of phoneme awareness, word recognition, phonological decoding, and orthographic coding have focused on the heritability of group deficits in these skills by employing a subset of participants included in the present Cholesky analyses of individual differences (Gayán & Olson, 2001; Olson et al., 1989, 1994a). Probands (affected twins) were selected to be at least 1.5 *SD* units below the local population mean for specific measures. The heritability of the group deficit in each measure was estimated from a model developed by DeFries and Fulker (1985) that compared identical (MZ) and fraternal (DZ) co-twin regression to the population mean. Greater DZ co-twin regression provided evidence for a genetic etiology. Using this approach, Gayán and Olson (2001) recently estimated group-deficit heritabilities (h_g^2) of .46–.72 for phoneme awareness, .57–.59 for word recognition, .60–.71 for phonological decoding, and .55–.67 for orthographic coding, depending on the specific measure used for each skill. The genetic origins of these deficits was further specified by DNA analyses. Significant evidence was found for linkage of deficits in phoneme awareness, phonological decoding, and orthographic coding to markers on the short arm of chromosome 6 (Gayán et al., 1999). Two independent laboratories have replicated these results for similar measures (Fisher et al., 1999; Grigorenko, Wood, Meyer, & Pauls, 2000).

It is theoretically possible that the amount and specific mechanisms of genetic influence on deviant group membership (h_g^2) could differ from those for individual differences across the normal range (h^2). However, the evidence from the present analyses of individual differences suggests rather similar results for h_g^2 and h^2. Heritability estimates for group deficits (h_g^2) in the specific measures from Gayán and Olson (2001) are not quite as high as those for individual differences (h^2) in the present analyses, but this may be at least partly due to the use of more reliable latent-trait estimates for h^2. Moreover, as described below, the pattern of bivariate relations between group deficits in phoneme awareness, phonological decoding, and orthographic coding found by Gayán and Olson mirror those found for genetic correlations for individual differences in these variables. The results from Gayán and Olson and the present study are at least consistent with a similar genetic etiology for group deficits and individual differences across the normal range.

Heritability (h^2) for individual differences in latent traits

In contrast to previous behavioral and DNA analyses of genetic influences on group deficits, the present analyses focused on the genetic etiology of individual differences across a broad range of reading ability, based on MZ and DZ twin correlations for latent traits. The estimates discussed below were derived from the entire sample of

MZ and DZ twins. This approach was supported by the fact that there were only minor differences in measurement and structural models between sub-samples divided on gender, age, or reading-ability group. For the four age-by-reading-ability groups, only the younger low-average ability group was significantly different from the others, and this was largely due to a stronger measurement loading for the orthographic choice task in that group. There were no significant model differences at the structural level, so analyses of the larger combined sample were justified.

For the combined sample of 257 identical and 183 fraternal twin pairs, the estimates of heritability for individual differences (h^2) were quite high for reliability-adjusted IQ (.69), and for the latent traits of phoneme awareness (.83), word recognition (.85), phonological decoding (.80), and orthographic coding (.87). However, there was evidence for some non-additive genetic effects (e.g., dominance or epistasis) for word recognition, phonological decoding, and particularly orthographic coding that may have inflated estimates of genetic influence in the additive (ACE) model. Non-additive genetic influences are indicated when DZ twin correlations are less than half the size of MZ twin correlations, as they were for some measures of these three latent traits. Thus, a genetic model that included non-additive genetic transmission provided a good fit to the data with large non-additive genetic estimates, particularly for orthographic coding. Unfortunately, it is not possible to simultaneously model additive genetic, non-additive genetic, and shared and non-shared environment effects with data from twins reared together (Neale & Cardon, 1992). While the present results do suggest that genetic influences on the latent traits in this sample are strong, they may not be as strong as the values estimated from an additive genetic model.

A second qualification to the very high genetic and low shared-environment estimates is the restricted environmental range of the twin sample. There were few twins from the very poor areas of Denver where average reading performance and school quality is well below national norms, and no twins had English as their second language. The latter constraint was deliberate, but the former partly reflects a very low volunteer rate from Denver's poor districts and schools. The proportional influence of genes on individual differences in a population will always depend partly on the range of relevant environmental influences. The restriction of environmental range in the present sample probably allowed genes to have proportionally greater influence and shared environment less influence on individual differences in reading and related cognitive skills than would be found in the general population (Turkheimer et al., 2001).

Genetic and environmental influences on relations between latent traits

Regardless of the exact levels of genetic influence on each of the latent traits, there are deeper and theoretically more interesting questions about genetic and environmental influences on relations between the traits. These relations were explored first by estimates of genetic correlations between the traits, and second by estimating the independent genetic variance in a trait after controlling for genetic variance shared with other traits. Three main questions were explored: (a) How does the broad range of cognitive skills represented by IQ compare with the more specific language skill of phoneme awareness in their relations to word-reading skills? (b) Does phoneme awareness relate differentially to the skills of word recognition, phonological decoding, and orthographic coding? (c) What is the balance of shared and independent genetic influences on phonological decoding and orthographic coding?

IQ and phoneme awareness relations with word-reading skills. In answer to the first question, there were moderate genetic correlations for phoneme awareness and word-reading skills with IQ (.44–.56), and generally stronger genetic correlations between phoneme awareness and word-reading skills (.55–.79). These results suggest that the stronger phenotypic correlations that have often been reported between phoneme awareness and word-reading skills in previous non-genetic studies are largely due to common genetic influences. We also found that after controlling for genetic influences shared with IQ, there still were highly significant genetic covariance estimates between phoneme awareness and the word-reading skills. Thus, the present results support the partly biological basis of a phonological "module" that is critically important for the development of printed word recognition. The present results differ from those reported by Hohnen and Stevenson (1999) for twins who were beginning readers in the first grade, where the relations between phoneme awareness and reading were due only to shared environment influences, after controlling for genetic influences shared with verbal IQ and other language skills. It is not clear if the different results are due to our older age range or to the additional language measures that were controlled in the Hohnen and Stevenson study.

After controlling for phoneme awareness, IQ had no significant genetic correlation with word-reading skills. The results might be different for reading comprehension. There is much evidence that when readers have moved beyond the beginning stages of development in word-reading skills, their comprehension of extended texts is more strongly related to broad IQ tests that include assessments of verbal knowledge and verbal working memory (Conners & Olson, 1990; Nation & Snowling, 1999; Perfetti, Marron, & Foltz, 1996). Unfortunately, we did not have a reliable or broad assessment of reading comprehension in the present study that would have supported a direct comparison with the present models for word reading. This omission has been rectified in our current research, but it will be several years before the new twin sample will be large enough to compare behavioral-genetic models for individual differences in word-reading and comprehension skills.

Differential phoneme awareness relations with phonological decoding and orthographic coding. In answer to the second question, the genetic correlations between IQ and phonological decoding (.49) and IQ and orthographic coding (.44) were very similar, but the genetic correlations between phoneme awareness and phonological decoding (.79) and phoneme awareness and orthographic coding (.55) were significantly different. This pattern of results mirrors that found by Gayán and Olson (2001) when they compared the genetic correlation for group deficits in phoneme awareness and phonological decoding (.64–.67) with the genetic correlation for group deficits in phoneme awareness and orthographic coding (.28–.39). Thus, for both individual differences across the normal range and for group deficits, phoneme awareness had a stronger genetic overlap with phonological decoding than with orthographic coding.

The present behavioral genetic results, along with the significant predictive relation from pre-readers' phoneme awareness to later phonological decoding and word recognition found by Wagner et al. (1994), suggest that there may be a genetically based causal influence from young children's development in phoneme awareness to their later phonological decoding skills and word recognition. This question is currently being explored more directly in an international longitudinal study of pre-school twins in the US, Australia, and Norway (Byrne et al., 2002).

Shared and independent genetic influences on phonological decoding and orthographic coding. The third question concerns the possibility of independent genetic influences on phonological decoding and orthographic coding. We have seen that these two traits have a significant genetic correlation, but do they also have significant independent genetic etiologies after controlling for genetic influences shared with the other skill? The answer is yes, as might be predicted from their different genetic correlations with phoneme awareness discussed above. For example, after controlling for genetic influences shared with IQ and phonological decoding, we found a significant independent genetic influence on orthographic coding, accounting for about 35% of the orthographic coding genetic variance. This independent genetic influence was not significant in the full Cholesky model after also controlling for phoneme awareness and word recognition, so it seems that orthographic coding is a partly independent and genetically influenced component skill in word recognition. (Otherwise, the independent genetic variance in orthographic coding might have been due to some extraneous choice-task or spelling variance that was independent from word recognition.) Similarly, significant independent genetic variance in phonological decoding was demonstrated after controlling for genetic influences on IQ and orthographic coding, but not after also controlling for phoneme awareness and word recognition.

How do these results converge with extant models of reading and reading disability? "Dual-Route" models of word processing postulate partly separate "phonological" and direct "orthographic" or "visual word form" routes to the lexicon (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001). Connectionist modelers eschew the Dual Route notions of separate routes to a lexicon, but they do include highly interconnected processing units that are associated with phonological, orthographic, or semantic information for words (Harm & Seidenberg, 1999). The present results do not settle the Dual-Route versus Connectionist arguments, but they do support a partly biological/genetic basis for the distinction in both types of models between phonological and orthographic representations.

The distinction between orthographic and phonological processing skills has been the basis for several "subtype" descriptions of individual differences among both normal (Baron, 1979; Baron & Strawson, 1976), and disabled readers (Boder, 1973; Castles & Coltheart, 1993; Manis, Seidenberg, Doi, McBride-Chang, & Petersen, 1996; Mitterer, 1982; Olson et al., 1985). Poor readers with relatively greater deficits in phonological skills have been called "phonological dyslexics," while those with relatively greater deficits in orthographic skills have been called "surface dyslexics." Baron and Strawson used the terms "Chinese" and "Phonecian" to describe similar profile differences among normal college students. Classroom training studies have shown that different early teaching methods (i.e., phonics versus whole word) produce short-term differences in the relative strengths of phonological and orthographic reading skills (Connelly, Johnston, & Thompson, 1999), and there is evidence that intense phonological training can shift a child's phonological-dyslexic profile to a surface-dyslexic profile (Wise, Ring, & Olson, 2000). Nevertheless, the present results suggest that some individual differences on the phonological-surface dimension may be due to independent genetic influences on phonological and orthographic skills in both normal and disabled readers.

Conclusion

The present evidence for strong genetic influences on individual differences in phoneme awareness and word-reading skills in this sample should not diminish the importance of good early reading instruction and continued engagement in reading. Reading is a culturally transmitted skill, and there are large differences in family, school, and broader cultural support for reading. Some of this important environmental variation was restricted in our twin sample because we were trying to understand the etiology of individual differences and group deficits that emerge in mostly normal educational environments. Our results show that when the environment for reading development is generally good and relatively homogeneous, genes have a strong influence on individual differences and group deficits in word-reading and related language skills.

Although some children in good general environments for reading development may fail because of a genetic liability, there is much evidence that their reading and phonological skills can be substantially improved with intensive intervention (Wise et al., 2000). Future behavioral and molecular genetic research related to reading development will support a better understanding of biological and environmental influences on reading deficits, early diagnosis of pre-readers at risk for reading failure, and more precisely targeted interventions that match the unique genetic and environmental etiologies of children's reading disabilities.

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

Alarcón, M., & DeFries, J. C. (1997). Reading performance and general cognitive ability in twins with reading difficulties and control pairs. *Personality and Individual Differences, 22*, 793–803.

Baker, L., Scher, D., & Mackler, K. (1997). Home and family influences on motivations for reading. *Educational Psychologist, 32*, 69–82.

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.

Baron, J., & Strawson, C. (1976). Use of orthographic and word-specific knowledge in reading words aloud. *Journal of Experimental Psychology: Human Perception and Performance, 2*, 386–393.

Boder, E. (1973). Developmental dyslexia: A diagnostic approach based on three atypical reading and spelling patterns. *Developmental Medicine and Child Neurology, 15*, 663–687.

Bryant, P. (2002). It doesn't matter whether onset and rime predicts reading better than phoneme awareness does or vice versa. *Journal of Experimental Child Psychology, 82*, 41–46.

Byrne, B., Delaland, C., Fielding-Barnsley, R., Quain, P., Samuelsson, S., Hoien, T., Corley, R., DeFries, J. C., Wadsworth, S., Willcutt, E., & Olson, R. K. (2002). Longitudinal twin study of early reading development in three countries: Preliminary results. *Annals of Dyslexia, 52*, 49–74.

Castles, A. E., & Coltheart, M. C. (1993). Varieties of developmental dyslexia. *Cognition, 47*, 149–180.

Catts, H., & Kamhi, A. G. (1999). *Language and reading disabilities*. Needham Hts: Allyn and Bacon.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204–256.

Connelly, V., Johnston, R. S., & Thompson, G. B. (1999). The influence of instructional approaches on reading procedures. In G. B. Thompson & T. Nicholson (Eds.), *Learning to read: Beyond phonics and whole language* (pp. 103–123). New York, NY: Teachers College Press.

Conners, F., & Olson, R. K. (1990). Reading comprehension in dyslexic and normal readers: A component-skills analysis. In D. A. Balota, G. B. Flores-d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 557–579). Hillsdale, NJ: Erlbaum.

Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing, 14*, 549–568.

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.

Dunn, L. M., & Markwardt, F. C. (1970). *Examiner's Manual: Peabody Individual Achievement Test*. Circle Pines, MN: American Guidance Service.

Ehri, L. C. (1989). The development of spelling knowledge and its role in reading acquisition and reading-disability. *Journal of Learning Disabilities, 22*, 356–365.

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.

Gayán, J., & Olson, R. K. (2001). Genetic and environmental influences on orthographic and phonological skills in children with reading disabilities. *Developmental Neuropsychology, 20*, 483–507.

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.

Goswami, U. (2002). In the beginning was the ryme? A reflection of Hulme, Hatcher, Nation, Brown, Adams, and Stuart. *Journal of Experimental Child Psychology, 82*, 47–57.

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.

Harm, M. W., & Seidenberg, M. S. (1999). Phonology, reading acquisition, and dyslexia: Insights from connectionist models. *Psychological Review, 106*, 491–528.

Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute Inc.

Hohnen, B., & Stevenson, J. (1999). The structure of genetic influences on general cognitive, language, phonological and reading abilities. *Developmental Psychology, 35*, 590–603.

Hulme, C., Hatcher, P. J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology, 82*, 2–28.

Lindamood, C. H., & Lindamood, P. C. (1979). *Lindamood auditory conceptualization test*. Hingham, MA: Teaching Resources Corporation.

Manis, F. R., Seidenberg, M. S., Doi, L. M., McBride-Chang, C., & Petersen, A. (1996). On the bases of two subtypes of development dyslexia. *Cognition, 58*, 157–195.

Mitterer, J. O. (1982). There are at least two kinds of poor readers: Whole-word poor readers and recoding poor readers. *Canadian Journal of Psychology, 36*, 445–461.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition, 70*, B1–13.

Neale, M. C. (1999). Mx: Statistical modeling (5th ed.). Box 126 MCV, Richmond, VA 23298: Department of Psychiatry.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publishers.

Olson, R. K. (1994). Language deficits in "specific" reading disability. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 895–916). New York: Academic Press.

Olson, R. K., Forsberg, H., & Wise, B. (1994a). Genes, environment, and the development of orthographic skills. In V. W. Berninger (Ed.), *The varieties of orthographic knowledge I: Theoretical and developmental issues* (pp. 27–71). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Olson, R. K., Forsberg, H., Wise, B., & Rack, J. (1994b). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243–277). Baltimore: Paul H. Brookes Publishing Co.

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 Press.

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.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Perfetti, C. A., Marron, M. A., & Foltz, P. W. (1996). Sources of comprehension failure: Theoretical perspectives and case studies. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties: Processes and intervention* (pp. 137–165). Mahwah, NJ: Erlbaum.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral genetics* (4th ed.). New York, NY: Worth Publishers.

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.

SAS (1989). *SAS/STAT User's Guide*, Version 6 (4th ed., Vol. 1). Cary, NC: SAS Institute Inc.

SAS (2000). *SAS Software*, Version 8 of the SAS System for Unix. Cary, NC: SAS Institute Inc.

Seidenberg, M. S., & McClleland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*, 523–568.

Shankweiler, D., Lundquist, E., Katz, L., Stuebing, K. K., Fletcher, J. M., Brady, S., Fowler, A., Dreyer, L. G., Marchione, K. E., Shaywitz, S. E., & Shaywitz, B. A. (1999). Comprehension and decoding: Patterns of association in children with reading difficulties. *Scientific Studies of Reading, 3*, 69–94.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151.

SPSS (2000). *SPSS for windows*, Release 10.0. Chicago, IL: SPSS Inc.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly, 24*, 402–433.

Turkheimer, E., Haley, A., D'Onofrio, B., Waldron, M., Emery, R. E., & Gottesman, I. I. (2001). Socioeconomic status modifies heritability of intelligence in impoverished children. *Behavior Genetics, 31*, 471 (Abstract).

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

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. (1974). *Examiner's Manual: Wechsler Intelligence Scale for Children-Revised*. New York: The Psychological Corporation.

Wechsler, D. (1981). *Examiner's Manual: Wechsler Adult Intelligence Scale-Revised*. New York: The Psychological Corporation.

Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Annals of Dyslexia, 69*, 848–872.

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.