Published in final edited form as:

Sci Stud Read. 2007 February 1; 20(1-2): 51–75. doi:10.1007/s11145-006-9018-x.

Genetic and Environmental Influences on Vocabulary and Reading Development

Richard K. Olson,

University of Colorado at Boulder, Linköping University, Sweden

Janice M. Keenan,

University of Denver

Brian Byrne,

University of New England, Australia, Linköping University, Sweden

Stefan Samuelsson,

Linköping University, Sweden

William L. Coventry,

University of New England, Australia

Robin Corley,

University of Colorado at Boulder, Institute for Behavioral Genetics

Sally J. Wadsworth,

University of Colorado at Boulder, Institute for Behavioral Genetics

Erik G. Willcutt,

University of Colorado at Boulder, Institute for Behavioral Genetics

John C. DeFries,

University of Colorado at Boulder, Institute for Behavioral Genetics

Bruce F. Pennington, and

University of Denver

Jacqueline Hulslander

University of Colorado at Boulder, Institute for Behavioral Genetics

Abstract

Genetic and environmental relations between vocabulary and reading skills were explored longitudinally from preschool through grades 2 and 4. At preschool there were strong shared-environment and weak genetic influences on both vocabulary and print knowledge, but substantial differences in their source. Separation of etiology for vocabulary and reading continued for word recognition and decoding through grade 4, but genetic and environmental correlations between vocabulary and reading comprehension approached unity by grade 4, when vocabulary and word recognition accounted for all of the genetic and shared environment influences on reading comprehension.

Keywords

vocabulary; word recognition; reading comprehension; twins; genetic; genes; environment; longitudinal; development; simple view

Much research has shown that individual differences in the development of vocabulary are significantly correlated with individual differences in reading, more strongly so with reading comprehension than with word recognition (Braze, Tabor, Shankweiler, & Mencil, 2007; Nation & Cocksey, 2009; Storch & Whitehurst, 2002). In addition, there is evidence for partial independence in the phenotypic contributions of oral language (including vocabulary) and decoding to reading comprehension as early as the second grade (Kendeou, van den Broek, White, & Lynch, 2009). A fundamental question about these individual differences is the relative influence from genetic and environmental factors on vocabulary, on different reading skills, and on their correlation. Are the same genetic and environmental influences at play for these skills, or are there significant differences in etiology between skills? The present study addresses this question longitudinally through the early stages of reading development by comparing the similarities of identical and fraternal twins beginning in preschool, and subsequently in a follow-up assessment near the end of the grade 2 when children are “learning to read,” and then again at the end of grade 4 when children are “reading to learn” (Chall, 1983).

The present analyses are based on data from the International Longitudinal Twin Study (ILTS) of early reading development. The ILTS includes samples of identical and same-sex fraternal twins from Australia, Colorado, and Scandinavia. The combined sample currently exceeds 2000 twins who have been tested in preschool. The twins were initially assessed on a broad range of pre-reading skills including letter name knowledge and vocabulary in the preschool year before admission to kindergarten, and before learning to read could interact with those skills. Subsequently the twins were tested on reading and related skills at the end of kindergarten, first, and second grade in all countries, and at the end of fourth grade in the U.S. sample. Vocabulary was assessed in preschool and again at the end of grades 2 and 4.

In the present study we focused on the ILTS measures of vocabulary and pre-reading knowledge about print (e.g., letter names) in pre-kindergarten, and on vocabulary, word recognition, decoding (nonword reading), and reading comprehension at the end of grades 2 and 4. We explored the genetic and environmental influences on each of the skills separately at each of these test occasions. We also explored the developmental genetic and environmental correlations for vocabulary from pre-kindergarten through grades 2 and 4, and we compared the genetic and environmental correlations between vocabulary and the reading measures at each test occasion. Thus we were able to test hypotheses about developmental changes in the etiology of individual differences in vocabulary and of the relations between vocabulary, word recognition, decoding, and reading comprehension.

Research with older children has previously explored the genetic and environmental relations between individual differences in word recognition, reading comprehension, and listening comprehension (Betjemann, Keenan, Olson, & DeFries, in press; Betjemann et al., 2008; Harlaar & Petrill, 2009; Keenan, Betjemann, Wadsworth, DeFries, & Olson, 2006). These studies have revealed correlated but also independent genetic influences on word recognition, reading comprehension, and listening comprehension. Moreover, the genetic influences on reading comprehension independent of those on word reading were completely mediated by genetic influences on listening comprehension, a result that was consistent with Gough and Tunmer’s (1986) “simple view” of reading comprehension. In the present study we also explore the independent genetic and environmental influences on

reading comprehension after controlling for those on word recognition and decoding at the end of grades 2 and 4, but here we ask if they can be accounted for by individual differences in genetic and environmental influences on vocabulary, and if these influences change across the early stages of reading development.

A recent behavior genetic analysis of vocabulary data in the ILTS sample at preschool yielded estimates with 95% confidence intervals for genetic ($a^2 = .19, .09-.31$), shared environment ($c^2 = .53, .42-.62$), and unique environment ($e^2 = .26, .23-.31$) (Byrne et al., 2009). Thus the genetic influence on vocabulary was quite low but still statistically significant, and the majority of individual differences were due to variation in family environment at preschool. In contrast, Byrne et al. also reported that vocabulary at the end of second grade yielded a much higher estimate of genetic influence and lower estimates of shared and non-shared environment ($a^2 = .44 (.31-.59)$, $c^2 = .36 (.22-.49)$, $e^2 = .19 (.16-.23)$). In addition, the longitudinal genetic and environmental relations between vocabulary at preschool and at grade 2 were assessed in a developmental Cholesky model. In spite of the substantial increase in genetic and decrease in family environment influence from preschool to grade 2, the genetic and environmental influences were largely correlated (ie., due to the same genes and environmental factors), and there was no significant evidence of new sources for genetic or shared environment influences on vocabulary in the second grade beyond those influencing vocabulary in preschool.

The present study expands on the analyses in Byrne et al. (2009) in several ways. It includes a larger sample, and it extends the longitudinal assessment of vocabulary and reading skills to the end of grade 4, approximately transitioning Chall's (1983) stages of "learning to read" through grade 2 and "reading to learn" by the end of grade 4. At preschool we analyzed the relations between vocabulary and knowledge about print (mostly letter name and sound knowledge, since few children could read). In grades 2 and 4 we analyzed our measures of vocabulary, word recognition, decoding (nonword reading), and reading comprehension. Where multiple measures of a skill were available, we modeled the shared variance between those measures as latent traits to reduce the influence of measurement error and thus increase power to detect genetic and environmental influences (Loehlin, 2004).

Methods

Participants

The sample at preschool consisted of 497 monozygotic (MZ) twin pairs, (225 US, 155 Australian, and 117 Scandinavian [Norway and Sweden combined]), and 500 same-sex dizygotic (DZ) pairs (264, 117, and 119 respectively), total 997 pairs. The sample at Grade 2 consisted of 406 MZ twin pairs, (222 US, 120 Australian, and 64 Scandinavian) and 424 DZ pairs (260, 92, and 72 respectively), total 830 pairs. The sample at grade 4 consisted of 176 MZ and 213 DZ pairs, all from the US. The twins were recruited from birth records in Colorado and Scandinavia, and from the (volunteer) Australian Twin Registry in Australia. Zygosity was determined from DNA extracted from cheek swabs, or in a minority of cases from selected items from the Nichols and Bilbro (1966) questionnaire. All twins were learning to read in their first language (English for Australia and US, Norwegian or Swedish in Scandinavia).

Mean ages (*SD* in months) at preschool were 4 years 10 months (2.3), 4 years 9 months (3.4), and 5 years (1.7) for the US, Australian, and Scandinavian samples, respectively. Respective ages at grade 2 were 8 years 5 months (3.8), 7 years 11 months (4.5), and 8 years 9 months (3.7). The mean age for the grade 4 sample in the US was 10 years 5 months (3.9). The age differences between countries at preschool and grade 2 are all significant ($p < .001$). Scandinavian children start school at a later age than those in the other countries, and the US

children were tested in the summer after the school year while Australian children were tested during the final three or four months of the school year (owing to the shorter summer vacation in Australia of six weeks versus twelve in the US).

Measures

The measures included in the present analyses and described here are from larger test batteries that were administered in the ILTS at preschool and at the end of grades 2 and 4.

Preschool

Vocabulary: A confrontation naming test of expressive vocabulary, the Hundred Pictures Naming Test (HPNT, Fisher & Glennister, 1992), required subjects to name each picture shown individually within 5 seconds. An error was scored if they failed to respond within the 5 second limit. Test-retest reliability is reported in the test manual as .97, and Cronbach alpha = .89. We also administered the vocabulary subtest from the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI, Wechsler, 1989). This test requires children to give definitions of 25 words spoken by the tester. Test-retest reliability in 4.5 year olds is reported in the test manual to be .83. For both vocabulary tests, analyses were based on the number of items correct.

Print Knowledge: The print knowledge assessment included measures of letter recognition in a four-choice format from names or sounds pronounced by the experimenter, Concepts about Print (Clay, 1975), and familiarity with common signs such as *stop*. Test-retest reliability is not available, but a low-bound estimate of test-retest reliability for a composite of these measures is provided by the MZ twin correlation in the present study ($r = .82$).

Grade 2

Vocabulary: The Boston Naming Test (Kaplan, Goodglass, & Weintraub, 2001) is a confrontation naming measure in which the child is required to name pictures of 60 concrete objects, ranging from common ones like bed to rarer ones like abacus. We administered all 60 items. The score was the number of items correctly named. Test-retest reliability is not available. A low-bound estimate of test-retest reliability is provided by the MZ twin correlation in the present study ($r = .81$).

Word Recognition and Decoding: We used the *Test of Word Reading Efficiency (TOWRE)*, Torgesen, Wagner, & Rashotte, 1999). Children read a list of words and a list of nonwords as quickly as possible, with the score being the number correctly read in 45 seconds. There are two equivalent forms of the test, Forms A and B, and we administered both to optimize the reliability of the scores and for latent trait modeling in the present study. Test-retest reliability for children aged 6 – 9 years is reported in the test manual as .97 for word and .90 for nonword standard scores. Standard scores were used in the analyses.

Reading Comprehension: The Woodcock Passage Comprehension test from the Woodcock Reading Mastery Test-Revised uses a “cloze” procedure that requires the child to orally provide a single missing word, denoted by an underlined space in a sentence, to assess their ability to understand short passages consisting of one or two short sentences that are read silently (Woodcock, 1989). Internal reliability is reported in the test manual to be .88 between 5 and 18 years of age. Standard scores were used in the analyses.

Grade 4

Vocabulary: The Boston Naming Test of expressive vocabulary administered in grade 2 was also administered in grade 4. A second vocabulary test was introduced in grade 4, the

Peabody Picture Vocabulary Test – III (PPVT-III), to assess receptive vocabulary (Dunn & Dunn, 1997). It requires participants to choose among 4 pictures to represent the meaning of spoken target items. Test-retest reliability is reported in the test manual to be .92. Standard scores were used in the analyses. The phenotypic correlation between the two vocabulary measures was $r = .69$.

Word Recognition and Decoding: The TOWRE measures of word and nonword reading administered in grade 2 were also administered in grade 4, but only with form A. In grade 4 we added the Woodcock-Johnson III Tests of Achievement untimed measures of Letter-Word Identification and Word Attack (nonword reading) (Woodcock, McGrew, & Mather, 2001). One year test-retest reliabilities are reported in the test manual as .85 for Letter-Word Identification and .81 for word attack at age 8–10. Standard scores were used in the analyses.

Reading Comprehension: The Woodcock Passage Comprehension test administered in grade 2 was also administered in grade 4. The battery at grade 4 also included the Gates-MacGinitie test of reading comprehension (level 4, form S) (MacGinitie, MacGinitie, Maria, & Dreyer, 2000). This measure includes a series of passages that participants read silently and then answer four multiple choice questions to assess their comprehension while the passage is still visible. We used the raw-score squared in our analyses to correct for skew. Test-retest reliability is not available. A low-bound estimate of test-retest reliability for the Gates-MacGinitie is .77 based on the MZ twin correlation in the present study. The phenotypic correlation between the two comprehension measures was $r = .74$.

Procedure

Preschool testing was conducted in the twins' homes or preschools during 5 approximately hour-long sessions over a 1 – 2 week period. Testing at the end of grade 2 was conducted in a single session lasting about 1 hour in the twins' homes or schools. Testing at the end of grade 4 was conducted in a 1.5 to 2 hour session in the twins' homes. Two testers separately assessed each twin in a pair at the same time, except in Scandinavia where just one tester per pair was available.

Analyses

For reasons discussed by Byrne et al. (2009) and Samuelsson et al. (2008), Australian children tended to outperform US children and US children outperformed Scandinavian children from preschool through grade 2. Therefore we standardized scores within country before combining the samples for more powerful statistical analyses, in addition to adjusting scores for sex, age, and age squared, as is customary in behavior genetic analyses (McGue & Bouchard, 1984). Age- and gender- adjusted scores above or below 3 standard deviations from the mean in each country were truncated to 3 standard deviations.

All behavior genetic analyses were based on the multivariate Cholesky decomposition model (Neale, Boker, Xi, & Maes, 2006). The model estimates the relative influences from genetic, shared family environment (influences that make twins in a pair similar regardless of genetic similarity), and unique environment (influences that make twins in a pair differ regardless of genetic similarity and shared environment). It does this for each variable in the model, and it also estimates the degree to which genetic and environmental influences on a variable are correlated with those influences on other variables in the model. All estimates are based on the fact that identical twins share all their genes, while fraternal twins share half of their segregating genes on average. In addition, it is assumed that shared family environment influences are equally similar for identical and fraternal twins, and the genetic

influences from different genes are additive (Plomin, DeFries, McClearn, & McGuffin, 2008).

We first conducted a developmental Cholesky analysis on vocabulary, using latent traits based on the multiple vocabulary measures we had at preschool and grade 4, and the single vocabulary measure at grade 2. Then we turned to the relations between vocabulary and reading, starting with a two-factor Cholesky model with print knowledge and vocabulary at preschool, followed by four-factor models of vocabulary, word recognition, decoding, and reading comprehension in grades 2 and 4. Finally, we analyzed the genetic and environmental relations between preschool vocabulary and the reading measures in grade 4. Contrasts between parameter estimates within the Cholesky models were tested by the significance ($p < .05$) of loss of model fit when the different estimates were constrained to be equal (Neale et al., 2006). Confidence intervals for the genetic and environmental parameters and correlations were all 95%.

Results

Developmental Cholesky results for vocabulary at preschool and the end of grades 2 and 4

The developmental Cholesky results for vocabulary in pre-kindergarten, grade 2, and grade 4 are presented in Table 1. We describe these results in some detail here as an example for how to interpret the Cholesky results in subsequent tables. The A factors are for the genetic coefficients, C factors for the shared environment coefficients, and the E factors are for the non-shared environment coefficients. The estimates of a^2 , c^2 , and e^2 are the sums of the respective A, C, and E squared path coefficients, or the single squared path coefficient for the first factor. For example, consider the A factor structure for genetic influence on pre-kindergarten vocabulary. The pre-kindergarten latent trait for vocabulary is entered first in the Cholesky model under A1 with a significant genetic path coefficient of .54, and its genetic influence is the square of that path coefficient ($a^2 = .29$). Grade 2 vocabulary (Boston Naming) is entered second, and its path coefficient under A1 (.56) represents the genetic variance it shares with pre-kindergarten vocabulary, which is significantly greater than 0 at $p < .05$. The path coefficient for grade 2 vocabulary under A2 (.35) is not significantly greater than 0. It represents the genetic variance remaining after removing the genetic variance it shares with the pre-kindergarten latent trait. Genetic influence on grade 2 vocabulary ($a^2 = .44$) is the sum of the squared A1 and A2 path coefficients. Although the genetic influence appears to be stronger for grade 2 compared to preschool vocabulary, the difference was not statistically significant ($p > .05$). Finally, the genetic influence on the grade 4 latent trait for vocabulary is significantly shared with pre-kindergarten vocabulary (.65), but not significantly with grade 2 vocabulary after controlling for pre-kindergarten vocabulary (.39), and it has virtually no remaining genetic influence after controlling for genetic influence on pre-kindergarten and grade 2 vocabulary (.00 under A3). Thus, even though there is a significant quantitative increase in the a^2 estimates from preschool to grade 4 ($\Delta \chi^2 = 7.02$, $\Delta df = 1$, $p < .05$), it is largely the same qualitative genetic influences that are involved at each grade. These results suggest that the modest genetic influences on individual differences in pre-kindergarten vocabulary are amplified at later ages.

Similarly, pre-kindergarten shared environment under C1 accounts for all of the significant shared environment influence across the early grades: no significant qualitatively new shared environment influences emerge for vocabulary in grades 2 and 4 after controlling for pre-kindergarten vocabulary. Also, there was a significant decrease in shared environment from pre-kindergarten to grade 4 ($\Delta \chi^2 = 5.91$, $\Delta df = 1$, $p < .05$).

For non-shared environment, the pattern is a little different. The modest non-shared environment influences at pre-kindergarten are significantly shared with vocabulary in

grades 2 and 4, but Boston Naming has additional significant non-shared environment influence (.42*) that is independent from the latent traits in pre-kindergarten and grade 4. This is most likely due to measurement error for the single Boston Naming measure.

Another way of viewing the genetic and environmental relations between vocabulary across the grades is through their correlations, which are computed by post-multiplying the A, C, and E Cholesky matrices by their transpose and standardizing the resulting matrices (Gayán & Olson, 2003). The resulting correlations are independent from the magnitude of the a^2 , c^2 , and e^2 estimates. For example, the genetic correlation (r_a) could be high even though one or both of the variables have very low genetic influence. As suggested by the lack of significance after the A1 and C1 factors in Table 1, both the genetic and shared environment correlations presented in Table 2 are quite high from preschool through the end of grade 4. This indicates a high consistency in the specific qualitative genetic and environmental influences that are responsible for the varying quantitative estimates of genetic influence on vocabulary from pre-kindergarten through grade 4.

Genetic and environmental influences on pre-kindergarten vocabulary and print knowledge: overlap and independence

The genetic and environmental relations between the latent traits for pre-kindergarten vocabulary and print knowledge (mainly knowledge of letter names and sounds) are presented in Table 3. It is interesting that although the sizes of the genetic and shared environmental estimates are very similar for vocabulary and print knowledge, there are significantly different genetic and shared environmental influences at play for the two latent traits. This is indicated by the significant A2 and C2 factors for print knowledge after controlling for the genetic and shared environment influences on vocabulary. The genetic ($r_a = .62, .34-.91$) and shared environmental ($r_c = .59, .49-.69$) correlations for vocabulary and print knowledge are consistent with the pattern of results in Table 3: the genetic and shared environmental correlations are significantly greater than 0, but there is also evidence for substantial independent genetic and shared-environment influences since the genetic and shared environmental correlations are significantly less than 1.

Cholesky results for vocabulary and reading skills at the end of grade 2

Cholesky results are presented in Table 4 for the genetic and environmental relations between vocabulary, as measured by Boston Naming, latent traits for word recognition and decoding (nonword reading), and reading comprehension as measured by the Woodcock Passage Comprehension test at the end of grade 2. Vocabulary (A1 factor) shares significant genetic influence with word recognition, decoding, and comprehension. Word recognition under the A2 factor shares substantial genetic influence with decoding and comprehension after controlling for genetic influences on vocabulary. Significant genetic influence remains for decoding under the A3 factor, but this is not shared with comprehension after controlling for vocabulary and word recognition. Finally, there is a small but significant genetic influence on comprehension under factor 4, after controlling for genetic influences on the first three factors.

The shared environment influences on vocabulary (C1) are significantly shared with word recognition and comprehension but not decoding, and there is no further significant shared environment influence on word recognition, decoding, or comprehension after controlling for shared environment influences on vocabulary. In contrast, most of the non-shared environment influence shown on the diagonal is specific to each factor, a common result in multivariate genetic models, particularly when the models include individual measures, Boston Naming and Woodcock Passage Comprehension in this case, that may include substantial unrelated measurement error.

The grade 2 genetic and environmental correlations are presented in Table 5. They illustrate the significantly stronger relation between vocabulary and comprehension compared to vocabulary and word recognition and decoding ($\Delta \chi^2 = 15.04$, $\Delta df = 1$, $p < .05$), the very high genetic correlation between word recognition and decoding, and their substantial genetic correlations with comprehension. From this pattern of results at the end of second grade, it seems that the genetic and shared environmental etiologies of word recognition and decoding are more tightly linked than vocabulary to comprehension ($\Delta \chi^2 = 7.03$, $\Delta df = 1$, $p < .05$; $\Delta \chi^2 = 5.68$, $\Delta df = 1$, $p < .05$). Next we ask if this pattern holds for vocabulary and reading comprehension at the end of grade 4, or do individual differences in vocabulary share more genetic and environmental influences with reading comprehension in these older children.

Multivariate Cholesky for grade 4 vocabulary, word recognition, decoding, and comprehension latent traits

The Cholesky model results in Table 6 based on latent traits for all variables show significant shared genetic influence for vocabulary (A1) with word recognition, decoding, and comprehension. The corresponding genetic correlations with vocabulary in Table 7 were .70 with word recognition, .53 with decoding, and .97 with comprehension. Judging from the non-overlapping confidence intervals, the genetic correlation with vocabulary was significantly higher for comprehension than for word recognition or decoding. A similar pattern was found for the shared-environment correlations, though the very low amount of shared environment influence on the reading measures yielded very large and overlapping confidence intervals. The very modest non-shared environment influences only loaded significantly on the E1 factor for vocabulary in Table 6, and the non-shared environment correlations in Table 7 were all quite high, due to the removal of measurement error through latent-trait modeling.

One particular contrast of interest was the genetic correlations between vocabulary and reading comprehension versus word recognition and reading comprehension. Vocabulary had a significantly stronger genetic correlation with reading comprehension in the latent trait model ($\Delta \chi^2 = 4.01$, $\Delta df = 1$, $p < .05$).

After controlling for genetic influences shared with vocabulary, there were substantial and significant independent genetic influences on word recognition under the A2 factor that were shared with decoding, and there were very modest but statistically significant independent genetic influences that were shared between word recognition and comprehension. To confirm that vocabulary was contributing independent genetic variance to reading comprehension after controlling for word recognition, we also ran the four-factor Cholesky model with word recognition entered as the first factor and vocabulary as the second factor. The independent genetic influence on vocabulary was highly significant and it was largely shared with comprehension. As in the Cholesky model with vocabulary entered first, there was no significant genetic or environmental influence on reading comprehension after controlling for word recognition and vocabulary.

In our final Cholesky analysis, we substituted the pre-kindergarten vocabulary latent trait for the grade 4 vocabulary latent trait and modeled its genetic and environmental relations with the grade 4 reading latent traits. We predicted that the model results would be similar to those in Tables 6 and 7 because the developmental Cholesky results for vocabulary in Tables 1 and 2 showed that there were no qualitatively new significant genetic or shared environment influences on grade 4 vocabulary after controlling for pre-kindergarten and grade 2 vocabulary. In fact, the Cholesky results with pre-kindergarten vocabulary were very similar to those shown in Table 6 and 7 for grade 4 vocabulary, even though the genetic influence on pre-kindergarten vocabulary in the sample restricted to children from the US

with grade 4 data ($a^2 = .15$) was lower than for grade 4 vocabulary ($a^2 = .44$). Also, the phenotypic correlations between pre-kindergarten vocabulary and grade 4 word recognition ($r = .45$), decoding ($r = .29$), and comprehension ($r = .65$) were lower than those with grade 4 vocabulary ($r = .62$, $.44$, and $.84$ respectively). Nevertheless, the genetic correlations between pre-kindergarten vocabulary and grade 4 reading were still strongest with comprehension ($r_a = .89$) compared to word recognition ($r_a = .85$) and decoding ($r_a = .72$). A similar pattern was present for the shared environment correlations ($r_c = .99$, $.80$, and $.64$ respectively). Taken together, pre-kindergarten vocabulary and grade 4 word recognition accounted for all of the genetic and shared environment influences on grade 4 reading comprehension, a longitudinal-vocabulary version of the “simple view.” (The full Cholesky modeling results for this last analysis are available from the first author.)

Discussion

The present study explored the genetic and environmental etiologies of early vocabulary and reading development in identical and fraternal twins beginning in pre-kindergarten, before children’s reading could influence vocabulary, then through grade 2 when children are “learning to read,” and finally through grade 4 when children are “reading to learn” (Chall, 1983). We first examined the developmental stability of the levels of genetic and environmental influence on vocabulary and the degree to which those influences were correlated across the early grades. We found that individual differences in vocabulary were much more influenced by shared family environment than by genes when the child was in pre-kindergarten. However, this pattern changed significantly in grades 2 and 4, when genetic and shared-environment influences were more equally influential.

Our finding that genetic influence on vocabulary increases and shared environment decreases as children get older is echoed by Hart et al. (2009) who reported a more modest and non-significant trend toward increasing genetic influence and decreasing shared environmental influence on vocabulary (Boston Naming) from approximately the end of kindergarten through grade 3. Our greater age range and our inclusion of an initial vocabulary assessment before children went to school may be responsible for the stronger and significant developmental change in genetic and environmental etiology in the present study. It is likely that there is a greater range of shared environment for vocabulary in the twins’ homes at pre-kindergarten than in the schools, where there is often a common curriculum across schools (cf., Hart & Risley, 1995).

The main focus of our study was to understand the development of genetic and environmental relations between vocabulary and reading. We found a clear separation in the specific genetic and environmental origins for individual differences in pre-kindergarten vocabulary and print knowledge, even though the estimates of the relative strengths of genetic and environmental influences on the two latent traits were very similar. Their genetic and environmental correlations were significantly greater than 0, but they were also significantly less than 1, reflecting substantial genetic and environmental independence. In our subsequent analyses of vocabulary and reading in grades 2 and 4, we explored the degree of genetic and environmental independence as a function of grade and type of reading skill.

When children are in the beginning stages of learning to read in first grade, genetic influences on word reading and decoding skills are not significantly different from those on reading comprehension, as indicated by genetic correlations close to 1 (Byrne et al., 2007). One reason may be that the wide variation in early beginning readers’ word decoding accuracy is the main influence on comprehension of printed text. In the present study we found the beginning of a separation between the genetic etiologies of word recognition and

reading comprehension at the end of grade 2, when their genetic correlation was significantly less than 1. However, compared to vocabulary, both word recognition and decoding had higher genetic correlations with reading comprehension, and there was significant additional genetic influence on reading comprehension after controlling for genetic influences from vocabulary, word recognition, and decoding.

At the end of grade 4, the availability of multiple measures of vocabulary and reading comprehension allowed modeling of the measures' shared variance as latent traits. This provided estimates of genetic and environmental influences on vocabulary and reading comprehension that are relatively free of measurement error and item-specific variance. Two results from these latent-trait analyses are particularly notable. First, reading comprehension had a significantly stronger genetic correlation with vocabulary than with word recognition or decoding at the end of grade 4, reversing the pattern observed at the end of grade 2. Second, there was substantial genetic influence on reading comprehension that was independent from word recognition, and all of this independent genetic influence was mediated by genetic influence on vocabulary.

Previous behavior-genetic studies of reading comprehension have reported independent genetic contributions from word recognition and listening comprehension that accounted for all of the genetic influence on reading comprehension, thus providing a genetic basis for Gough and Tunmer's (1986) "simple view" of reading comprehension (Keenan et al., 2006). The results of the present study suggest a vocabulary basis for the "simple view". However, the composites of reading comprehension and listening comprehension measures used by Keenan et al. included much longer passages than the one or two sentences in the Woodcock Passage Comprehension cloze test, or the four to six generally short sentences in the Gates-MacGinitie test. Comprehension of the more lengthy passages used in Keenan et al. may have placed greater demands on attention and working-memory processes. Indeed, a recent report by Keenan et al. (2010) has shown that genetic influences on comprehension processes can be partly independent from those on vocabulary and word recognition, depending on the test used to assess reading comprehension. For longer passages that require constructing a mental model, vocabulary does not account for all of the word-recognition-independent genetic influence on comprehension; the addition of listening comprehension after vocabulary was needed to account for all the genetic influence on reading comprehension for more extended texts (Keenan et al., 2010).

The final question we addressed in our analyses was whether the genetic and environmental relations between vocabulary and reading at grade 4 are maintained longitudinally from pre-kindergarten vocabulary to grade 4 reading. They were, and in addition, the joint genetic and shared-environmental influences from word recognition and pre-kindergarten vocabulary accounted for all of the genetic and shared-environmental influences on reading comprehension at the end of grade 4.

Our findings extend those from previous longitudinal studies using data from the Twins Early Development Study (TEDS) in England and Wales to analyze the genetic and environmental relations between pre-kindergarten oral language and later reading (Harlaar, Hayiou-Thomas, Dale, & Plomin, 2008; Hayiou-Thomas, Harlaar, Dale, & Plomin, 2006). The TEDS study by Harlaar et al. modeled a latent trait for parent ratings of twins' combined vocabulary and syntax at ages 2, 3, and 4 years, and a latent trait for global teacher ratings of reading achievement on a 5 point scale at 7, 9, and 10 years of age that did not explicitly separate word recognition and reading comprehension. In spite of the differences in measurement from our study, the phenotypic correlation for their vocabulary/syntax and global-reading latent traits ($r = .40$) was very similar to ours ($r = .45$) between pre-kindergarten vocabulary and grade 4 word recognition, and the genetic and shared-

environmental correlations were similar as well. However, their correlations were substantially lower than what we found for pre-kindergarten vocabulary and grade 4 reading comprehension, possibly because their teacher ratings of reading comprehension were based more on word decoding accuracy.

The findings of our study and the TEDS study clearly demonstrate the long reach of genetic and environmental influences on pre-kindergarten vocabulary to later reading development. In addition, our study highlighted the uniquely strong longitudinal genetic and environmental correlations between pre-kindergarten vocabulary and grade 4 reading comprehension. Thus, it is clear that individual differences in vocabulary need to be carefully considered when we contemplate interventions for raising reading skills, particularly reading comprehension.

Another result with potential educational implications for improving reading comprehension is the stark contrast in the balance of genetic and shared environment influences on vocabulary versus those influences on the word recognition and decoding measures in grades 2 and 4. The very high genetic and low shared-environment influences on individual differences in reading across the early grades that were found in the present study and by others (Byrne et al., 2006, 2009; Harlaar, Spinath, Dale, & Plomin, 2005; Petrill, Deater-Deckard, Thompson, & DeThorne, 2006) may be due in part to a limited environmental range across teachers, classrooms, and schools in the twin samples that have been studied. Therefore, high genetic influence does not necessarily deny the potential for the success of intensive interventions focused directly on word recognition accuracy and fluency, be they school-wide or on reading disabilities (Byrne et al., 2010; Olson, Byrne, & Samuelson, 2009). On the other hand, the high shared-environment influence on vocabulary does not necessarily imply that it will be easy to broadly raise vocabulary in school-wide or targeted intervention programs. Vocabulary reflects the broad linguistic knowledge base that children have accumulated thus far, so short-term interventions that focus on learning specific words may have little effect on broad vocabulary against this background, though they may have a strong influence on reading comprehension within specific targeted domains. Still, the relatively strong influence of shared environment on vocabulary and its high shared-environment correlation with reading comprehension suggest that vocabulary should be an important focus of environmental interventions for deficits in reading comprehension.

The partly independent genetic and environmental etiology of vocabulary from word recognition and decoding may be an important reason why interventions focused on word recognition and decoding have had relatively little impact on reading comprehension once children have learned basic decoding skills (Edmonds et al., 2009; Olson & Wise, 1992; Torgesen et al., 2001). It is becoming increasingly clear that individual differences in background knowledge largely reflected in vocabulary are an important source of variance in reading comprehension in the later grades. Imparting this knowledge should be seen as the primary goal for improving literacy, beginning with children's pre-kindergarten language environments, to "learning to read," and then on to "reading to learn."

Finally, the results of the present study clearly imply that there are genetically based differences in the neural mechanisms associated with children's vocabulary development and their development in decoding and word recognition. These differences are implied by genetic correlations with vocabulary that were significantly less than 1.0 at preschool with print knowledge and at the end of the second and fourth grades with decoding and word recognition. Previous molecular-genetic research has identified genes associated with deficits in measures of reading that have been dominated by decoding, spelling, and word recognition (Fisher & DeFries, 2002). Future molecular-genetic research on reading should be expanded to include vocabulary, listening comprehension, and reading comprehension of

extended text to identify genes that are uniquely associated with individual differences and deficits in these highly related skills. Future neurological research should follow suit to understand how these genes influence different brain processes related to the development of vocabulary, listening comprehension, and reading comprehension.

Acknowledgments

Funding was provided by the Australian Research Council (DP0663498 and DP0770805), the National Institute for Child Health and Human Development (HD27802 and HD38526), and grants from the Research Councils of Norway and Sweden. We thank the Colorado and Australian Twin Registries, our testers, and the twins and parents involved.

References

- Betjemann RS, Willcutt EG, Olson RK, Keenan JM, DeFries JC, Wadsworth SJ. Word reading and reading comprehension: Stability, overlap, and independence. *Reading and Writing: An Interdisciplinary Journal* 2008;21:539–558.
- Betjemann R, Keenan J, Olson RK, DeFries JC. How Choice of Reading Comprehension Test Influences the Outcomes of Genetic Analyses. *Scientific Studies of Reading*. (in press).
- Braze D, Tabor W, Shankweiler DP, Mencil WE. Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities* 2007;40:226–243. [PubMed: 17518215]
- Byrne B, Coventry WL, Olson RK, Samuelsson S, Corley R, Willcutt EG, ...DeFries JC. Genetic and environmental influences on aspects of literacy and language in early childhood: Continuity and change from preschool to Grade 2. *Journal of Neurolinguistics* 2009;22:219–236. [PubMed: 20161176]
- Byrne B, Coventry WL, Olson RK, Wadsworth SJ, Samuelsson S, Petrill SA, Willcutt EG, Corley R. “Teacher effects” in early literacy development: Evidence from a study of twins. *Journal of Educational Psychology* 2010;102:32–42. [PubMed: 20204169]
- Byrne B, Olson RK, Samuelsson S, Wadsworth S, Corley R, DeFries JC, Willcutt E. Genetic and environmental influences on early literacy. *Journal of Research in Reading* 2006;29:33–49.
- Byrne B, Samuelsson S, Wadsworth S, Hulslander J, Corley R, DeFries JC, ...Olson RK. Longitudinal twin study of early literacy development: Preschool through Grade 1. *Reading and Writing: An Interdisciplinary Journal* 2007;20:77–102.
- Chall, JS. Stages of reading development. New York: McGraw-Hill; 1983.
- Clay, M. The early detection of reading difficulties: A diagnostic survey. Auckland, New Zealand: Heinemann; 1975.
- Dunn, LM.; Dunn, M. Peabody Picture Vocabulary Test-III. Circle Pines, MN: American Guidance Service; 1997.
- Edmonds MS, Vaughn S, Wexler J, Reutebuch C, Cable A, Tackett KK, Schnakenberg JW. A synthesis of reading interventions and effects on reading comprehension outcomes for older struggling readers. *Review of Educational Research* 2009;79:262–300. [PubMed: 20072704]
- Fisher SE, DeFries JC. Developmental dyslexia: genetic dissection of a complex cognitive trait. *Nature Reviews Neuroscience* 2002;3:767–780.
- Fisher, JP.; Glennister, JM. The hundred pictures naming test. Hawthorn, Australia: Australian Council for Educational Research; 1992.
- Gayán J, Olson RK. Genetic and environmental influences on individual differences in printed word recognition. *Journal of Experimental Child Psychology* 2003;84:97–123. [PubMed: 12609495]
- Gough PB, Tunmer WE. Decoding, reading, and reading disability. *Remedial and Special Education* 1986;7:6–10.
- Harlaar N, Hayiou-Thomas M, Dale PS, Plomin R. Why do preschool language abilities correlate with later reading? A twin study. *Journal of Speech, Language, and Hearing Research* 2008;51:688–705.

- Harlaar, N.; Petrill, SA. A genetic basis for the Simple View of Reading?. Paper presented at the Pacific Coast Research Conference; San Diego, CA. 2009 Feb.
- Harlaar N, Spinath FM, Dale PS, Plomin R. Genetic influences on early word recognition abilities and disabilities: A study of 7-year-old twins. *Journal of Child Psychology and Psychiatry* 2005;46:373–384. [PubMed: 15819646]
- Hart, B.; Risley, TR. Meaningful differences in the everyday experience of young American children. Baltimore: Brookes; 1995.
- Hart SA, Petrill SA, DeThorne LS, Deater-Deckard K, Thompson LA, Schatschneider C, Cutting LE. Environmental influences on the longitudinal covariance of expressive vocabulary: measuring the home literacy environment in a genetically sensitive design. *Journal of Child Psychology and Psychiatry* 2009;50:911–919. [PubMed: 19298476]
- Hayiou-Thomas ME, Harlaar N, Dale PS, Plomin R. Genetic and environmental prediction from different aspects of preschool language and non-verbal ability to 7-year reading. *Journal of Research in Reading* 2006;29:50–74.
- Kaplan, E.; Goodglass, H.; Weintraub, S. Boston Naming Test. 2. Baltimore, MD: Lippincott Williams & Wilkins; 2001.
- Keenan JM, Betjemann R, Wadsworth SJ, DeFries JC, Olson RK. Genetic and environmental influences on reading and listening comprehension. *Journal of Research in Reading* 2006;29:75–91.
- Keenan, JM.; Priebe, SJ.; Miller, AC.; Meenan, C.; Hua, A.; Olson, RK. Speaking up for listening comprehension. Paper presented at the meeting of the Society for the Scientific Study of Reading; Berlin, Germany. 2010 Jul.
- Kendeou P, van den Broek P, White MJ, Lynch JS. Predicting reading comprehension in early elementary school: the independent contributions of oral language. *Journal of Educational Psychology* 2009;101:765–778.
- Loehlin, JC. Latent variable models: An introduction to factor, path, and structural equation analysis. 4. Mahwah, NJ: Lawrence Erlbaum Associates; 2004.
- MacGinitie, WH.; MacGinitie, RK.; Maria, K.; Dreyer, LG. Gates-MacGinitie Reading Tests. Itasca, IL: Riverside Publishing; 2000.
- McGue M, Bouchard TJ Jr. Adjustment of twin data for the effects of age and sex. *Behavior Genetics* 1984;14:325–343. [PubMed: 6542356]
- Nation K, Cocksey J. The relationship between knowing a word and reading it aloud in children's word reading development. *Journal of Experimental Child Psychology* 2009;103:296–308. [PubMed: 19410262]
- Neale, MC.; Boker, SM.; Xie, G.; Maes, HH. Mx: Statistical Modeling. 7. VCU Box 900126, Richmond, VA 23298: Department of Psychiatry; 2006.
- Nichols RC, Bilbro WC. The diagnosis of twin zygosity. *Acta Genetica* 1966;16:265–275.
- Olson, RK.; Byrne, B.; Samuelsson, S. Reconciling strong genetic and strong environmental influences on individual differences and deficits in reading ability. In: Pugh, K.; McCardle, P., editors. *How Children Learn To Read: Current Issues and New Directions in the Integration of Cognition, Neurobiology and Genetics of Reading and Dyslexia Research and Practice*. Hillsdale NJ: Lawrence Earlbaum/Taylor-Francis; 2009. p. 215-233.
- Olson RK, Wise BW. Reading on the computer with orthographic and speech feedback: An overview of the Colorado Remedial Reading Project. *Reading and Writing: An Interdisciplinary Journal* 1992;4:107–144.
- Petrill SA, Deater-Deckard K, Thompson LA, DeThorne LS. Reading skills in early readers: Genetic and shared environmental influences. *Journal of Learning Disabilities* 2006;39:48–55. [PubMed: 16512082]
- Plomin, R.; DeFries, JC.; McClearn, GE.; McGuffin, P. Behavioral genetics. 5. New York, NY: Worth Publishing; 2008.
- Samuelsson S, Byrne B, Olson RK, Hulslander J, Wadsworth S, Corley R, ...DeFries JC. Response to early literacy instruction in the United States, Australia, and Scandinavia: A behavioral-genetic analysis. *Learning and Individual Differences* 2008;18:289–295. [PubMed: 19122888]

- Storch SA, Whitehurst GJ. Oral Language and Code-Related Precursors to Reading: Evidence From a Longitudinal Structural Model. *Developmental Psychology* 2002;38:934–947. [PubMed: 12428705]
- Torgesen J, Alexander A, Wagner R, Rashotte C, Voeller K, Conway T. Intensive remedial instruction for children with severe reading disabilities: Immediate and long-term outcomes from two instructional approaches. *Journal of Learning Disabilities* 2001;34:33–58. [PubMed: 15497271]
- Torgesen, JK.; Wagner, RK.; Rashotte, CA. *Test of word reading efficiency*. Austin, TX: Pro-ed; 1999.
- Wechsler, D. *Manual for the Wechsler Preschool and Primary Scale of Intelligence-Revised*. New York: Psychological Corporation; 1989.
- Woodcock, RW. *Woodcock Reading Mastery Tests*. Circle Pines, MN: American Guidance Service; 1989.
- Woodcock, RW.; McGrew, KS.; Mather, N. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishers; 2001.

Table 1

Genetic and environmental influences on pre-kindergarten, grade 2, and grade 4 vocabulary.

	A1	A2	A3	a ²
Pre-Kindergarten Latent Trait	.54*			.29 (.17–.43)
Grade 2 Boston Naming	.56*	.35		.44 (.33–.57)
Grade 4 Latent Trait	.65*	.39	.00	.57 (.40–.76)
	C1	C2	C3	c ²
Pre-Kindergarten Latent Trait	.81*			.65 (.52–.77)
Grade 2 Boston Naming	.53*	.30		.37 (.25–.47)
Grade 4 Latent Trait	.58*	.19	.20	.41 (.23–.58)
	E1	E2	E3	e ²
Pre-Kindergarten Latent Trait	.23*			.05 (.01–.11)
Grade 2 Boston Naming	.12*	.42*		.19 (.16–.22)
Grade 4 Latent Trait	.13*	.04	.00	.02 (.002–.06)

Note. 95% confidence intervals in parentheses.

* Significantly greater than zero.

N (pairs) = 997 for pre-kindergarten, 830 for grade 2, and 389 for grade 4.

Table 2

Genetic and environmental correlations for pre-kindergarten, grade 2, and grade 4 vocabulary.

Genetic		
	1.	2.
1. Pre-Kindergarten Latent Trait	--	
2. Grade 2 Boston Naming	.85 (.64–1.0)	--
3. Grade 4 Latent Trait	.86 (.63–1.0)	1.0 (.89–1.0)
Shared Environment		
	1.	2.
1. Pre-Kindergarten Latent Trait	--	
2. Grade 2 Boston Naming	.87 (.75–1.0)	--
3. Grade 4 Latent Trait	.91 (.76–1.0)	.93 (.87–1.0)
Unique Environment		
	1.	2.
1. Pre-Kindergarten Latent Trait	--	
2. Grade 2 Boston Naming	.27 (.06–.75)	--
3. Grade 4 Latent Trait	.96 (.26–1.0)	.52 (.10–1.0)

Note. 95% Confidence intervals in parentheses.

Table 3

Genetic and environmental influences on preschool vocabulary and print knowledge.

	A1	A2	a ²
Vocabulary Latent Trait	.50*		.25 (.12–.38)
Print Latent Trait	.31*	.40*	.26 (.16–.36)
	C1	C2	c ²
Vocabulary Latent Trait	.85*		.72 (.58–.84)
Print Latent Trait	.49*	.67*	.70 (.60–.78)
	E1	E2	e ²
Vocabulary Latent Trait	.20*		.04 (.01–.08)
Print Latent Trait	.22*	.00	.05 (.03–.08)

Note. 95% confidence intervals in parentheses.

* Significantly greater than zero. N (pairs) = 997.

Table 4
Genetic and environmental influences on grade 2 vocabulary and reading skills.

	A1	A2	A3	A4	a ²
Boston Naming	.67*				.45 (.34–.58)
Word Recognition Latent Trait	.30*	.85*			.81 (.69–.87)
Decoding Latent Trait	.32*	.77*	.30*		.78 (.65–.86)
Woodcock Passage Comprehension	.47*	.52*	.03	.35*	.61 (.47–.73)
C1	C2	C3	C4	c ²	
Boston Naming	.60*				.36 (.24–.47)
Word Recognition Latent Trait	.18*	.13			.05 (.001–.17)
Decoding Latent Trait	.14	.22	.00		.07 (.00–.19)
Woodcock Passage Comprehension	.24*	.24	.00	.00	.11 (.01–.24)
E1	E2	E3	E4	e ²	
Boston Naming	.43*				.19 (.16–.22)
Word Recognition Latent Trait	.06*	.37*			.14 (.12–.17)
Decoding Latent Trait	.03	.28*	.27*		.15 (.13–.18)
Woodcock Passage Comprehension	.07*	.17*	.04	.49*	.27 (.24–.32)

Note. 95% confidence intervals in parentheses.

* Significantly greater than zero. N (pairs) = 830.

Table 5

Genetic and environmental correlations for grade 2 vocabulary and reading skills.

	Genetic		
	1.	2.	3.
1. Boston Naming	--		
2. Word Recognition Latent Trait	.34 (.19–.47)	--	
3. Decoding Latent Trait	.36 (.21–.50)	.94 (.92–.96)	--
4. W. Passage Comprehension	.60 (.45–.76)	.82 (.77–.84)	.80 (.75–.86)
	Shared Environment		
	1.	2.	3.
1. Boston Naming	--		
2. Word Recognition Latent Trait	.80 (.14–1.0)	--	
3. Decoding Latent Trait	.54 (.004–1.0)	.94 (.01–1.0)	--
4. W. Passage Comprehension	.70 (.29–1.0)	.99 (.24–1.0)	.98 (.02–1.0)
	Unique Environment		
	1.	2.	3.
1. Boston Naming	--		
2. Word Recognition Latent Trait	.17 (.06–.27)	--	
3. Decoding Latent Trait	.07 (.003–.18)	.71 (.64–.77)	--
4. W. Passage Comprehension	.14 (.05–.23)	.34 (.25–.43)	.29 (.19–.38)

Note. W. = Woodcock. 95% Confidence intervals in parentheses.

Table 6
Genetic and environmental influences on grade 4 vocabulary and reading skills.

	A1	A2	A3	A4	η^2
Vocabulary Latent Trait	.66*				.44 (.30–.63)
Word Recognition Latent Trait	.62*	.63*			.77 (.57–.92)
Decoding Latent Trait	.47*	.75*	.00		.78 (.58–.90)
Comprehension Latent Trait	.90*	.22*	.00	.00	.86 (.63–.95)
C1	C2	C3	C4	ϵ^2	
Vocabulary Latent Trait	.74*				.54 (.36–.68)
Word Recognition Latent Trait	.23*	.29			.14 (.01–.34)
Decoding Latent Trait	.12	.27	.00		.09 (.00–.27)
Comprehension Latent Trait	.30*	.08	.00	.00	.09 (.01–.31)
E1	E2	E3	E4	ϵ^2	
Vocabulary Latent Trait	.14*				.02 (.003–.04)
Word Recognition Latent Trait	.28*	.10			.09 (.06–.13)
Decoding Latent Trait	.34*	.15	.00		.14 (.09–.19)
Comprehension Latent Trait	.21*	.00	.00	.00	.04 (.01–.10)

Note. 95% confidence intervals in parentheses.

* Significantly greater than zero. N (pairs) = 389.

Table 7

Genetic and environmental correlations for grade 4 vocabulary and reading skills.

	Genetic		
	1.	2.	3.
1. Vocabulary Latent Trait	--		
2. Word Recognition Latent Trait	.70 (.51–.83)	--	
3. Decoding Latent Trait	.53 (.30–.70)	.98 (.96–.99)	--
4. Comprehension Latent Trait	.97 (.87–1.0)	.85 (.77–.92)	.72 (.61–.81)
	Shared Environment		
	1.	2.	3.
1. Vocabulary Latent Trait	--		
2. Word Recognition Latent Trait	.63 (.21–1.0)	--	
3. Decoding Latent Trait	.42 (.003–1.0)	.97 (.01–1.0)	--
4. Comprehension Latent Trait	.97 (.56–1.0)	.80 (.19–1.0)	.63 (.01–1.0)
	Unique Environment		
	1.	2.	3.
1. Vocabulary Latent Trait	--		
2. Word Recognition Latent Trait	.94 (.51–1.0)	--	
3. Decoding Latent Trait	.91 (.41–1.0)	1.0 (.95–1.0)	--
4. Comprehension Latent Trait	1.0 (.22–1.0)	.94 (.57–1.0)	.91 (.53–1.0)

Note. 95% Confidence intervals in parentheses.