

Published in final edited form as:

J Speech Lang Hear Res. 2013 August ; 56(4): 1249–1259. doi:10.1044/1092-4388(2012/12-0185).

Shared Etiology of Phonological Memory and Vocabulary Deficits in School-Age Children

Robin L. Peterson¹, Bruce F. Pennington¹, Stefan Samuelsson², Brian Byrne^{2,3}, and Richard K. Olson^{2,4}

¹University of Denver

²Linköping University

³University of New England, Australia

⁴University of Colorado at Boulder

Abstract

Purpose—The goal of this study was to investigate the etiologic basis for the association between deficits in phonological memory (PM) and vocabulary in school-age children.

Method—Children with deficits in PM or vocabulary were identified within the International Longitudinal Twin Study (ILTS). The ILTS includes 1,045 twin pairs from the United States, Australia, and Scandinavia aged 5 to 8 years. We applied the DeFries-Fulker regression method to determine whether problems in PM and vocabulary tend to co-occur because of overlapping genes, overlapping environmental risk factors, or both.

Results—Among children with isolated PM deficits, we found significant bivariate heritability of PM and vocabulary weaknesses both within and across time. However, when probands were selected for a vocabulary deficit, there was no evidence for bivariate heritability. In this case, the PM-vocabulary relationship appeared to owe to common shared environmental experiences.

Conclusions—The findings are consistent with previous research on the heritability of specific language impairment and suggest that there are etiologic subgroups of children with poor vocabulary for different reasons, one more influenced by genes and another more influenced by environment.

Introduction

Extensive research demonstrates a link between phonological memory (PM), the component of working memory responsible for keeping phonological information active for brief periods, and vocabulary skill (Baddeley, Gathercole, & Papagno, 1998). In typically developing children, the correlation between PM and vocabulary averages approximately 0.3-0.5 (Baddeley, et al., 1998). Convergent evidence for the role of PM in broader language acquisition comes from children with several neurodevelopmental disorders, including specific language impairment (SLI), Down syndrome, and boys with Fragile X. In each case, research indicates that problems learning vocabulary and grammar relate to underlying PM weaknesses (Bishop, North, & Donlan, 1996; Gathercole & Baddeley, 1990a; Gray, 2004; Jarrold, Baddeley, & Phillips, 1999; Jarrold, Thorn, & Stephens, 2009; Lee, Pennington, & Keenan, 2010; Montgomery, 1995; Pierpont, Richmond, Abbeduto, Kover, & Brown, 2011). However, the etiologic basis for the link between PM and vocabulary deficits is not well

understood. In this study, we applied the DeFries-Fulker regression method in the International Longitudinal Twin Study (ILTS) to determine whether impairments in PM and vocabulary co-occur because of overlapping genes, overlapping environmental risk factors, or both. The ILTS includes twins aged 5 to 8 from the United States, Australia, and Scandinavia.

At the neurocognitive level, the relationship between PM and vocabulary has been understood within the context of two contrasting theoretical frameworks. The phonological storage framework of Baddeley, Gathercole, and colleagues (Baddeley, et al., 1998; Gathercole, 2006) offers a bottom-up explanation, and emphasizes the causal role of PM in vocabulary acquisition. According to this viewpoint, the phonological loop evolved to allow learning of novel phonological forms. Supporting evidence comes from neuropsychological patients and studies of second language learners, as well as from longitudinal and experimental studies in children with both typical and atypical development (Baddeley, Papagno, & Vallar, 1988; Cheung, 1996; Gathercole & Baddeley, 1990b; Gathercole, Willis, Emslie, & Baddeley, 1992; Masoura & Gathercole, 1999; Michas & Henry, 1994). A competing theoretical viewpoint holds that vocabulary exerts top-down effects on the quality of phonological representations, which in turn influence performance on PM tasks, particularly nonword repetition (Metsala, 1999; Snowling, Chiat, & Hulme, 1991). On balance, the empirical evidence indicates that the PM-vocabulary relationship is reciprocal (Brown & Hulme, 1996; Gathercole, et al., 1992; Gupta & Tisdale, 2009), though the primary direction of effect varies with development. Bottom-up influences appear to predominate earlier in language acquisition (Cheung, 1996; Gathercole, et al., 1992; Jarrold, Baddeley, Hewes, Leeke, & Phillips, 2004). In a phenotypic study in the ILTS utilizing a path modeling approach, we found evidence for the bottom-up effect in population-based samples from each country, as well as in children from the U.S. with phonologically-based language disorders (dyslexia or speech sound disorder). The top-down effect was only evident in samples with intact language and relatively more advanced literacy development. We argued that over the early school years, the influence of PM on vocabulary learning waned with language development, while the top-down effect emerged as a consequence of learning to read (Peterson, Pennington, Samuelsson, Byrne, & Olson, in submission).

Etiologic Overlap of Language Sub-skills

Understanding the etiology of various language sub-skills, as well as their etiologic relationships, has recently been a focus of increased research attention (Byrne et al., 2009; Dale, Harlaar, Hayiou-Thomas, & Plomin, 2010; Hayiou-Thomas, Dale, & Plomin, 2012; Hayiou-Thomas, Harlaar, Dale, & Plomin, 2010; Hayiou-Thomas et al., 2006; Olson et al., 2011). Studies from the ILTS and Twins Early Development Sample (TEDS) have converged in documenting a number of important findings. At preschool age, most language skills (including vocabulary and grammar) are predominantly influenced by shared environment (environmental experiences shared by members of a twin pair), and to a lesser extent, genes. However, within several years, the pattern reverses, and the effects of genes predominate. An exception to this pattern has been reported for some phonological measures, including PM. In this case, genetic effects outweigh environmental influences from the start, and remain the dominant influence as children mature.

These findings thus raise some challenges for both the bottom-up and top-down accounts of the PM-vocabulary relationship. Since both frameworks offer causal explanations, both require a shared etiology (genetic, environmental, or both) for individual differences in PM and vocabulary, while the differing pattern of twin results in preschool children is suggestive of distinct etiologies. However, only a multivariate approach can definitively address this issue. We are aware of no previous study that has specifically examined shared etiology between PM and vocabulary deficits. However, a small number of studies have addressed

closely related questions, with mixed results. Bishop, Adams, and Norbury (2006) found no evidence for shared heritability of PM and grammar deficits in 6 year olds using the DeFries-Fulker method. Hayiou-Thomas and colleagues (Hayiou-Thomas, et al., 2006) reported analyses across a wider range of individual differences in 4-year-olds from the TEDS sample. The genetic correlation between nonword repetition and an expressive vocabulary measure was nonsignificant, while the genetic correlation between nonword repetition and several other language measures (including a measure of grammar) were large and significant. Factor analysis indicated that nonword repetition loaded on a “speech” factor with articulation, while all other language measures loaded together on a second, “general language” factor. The genetic correlation between speech and general language latent traits was 0.64, and the shared environment correlation between the two latent traits was perfect (1.0). To summarize, there is evidence that the etiologic influences operating on a wide range of language constructs within a time point are largely shared. However, specific support for a common etiology for deficits in PM and broad language is scant.

A possible resolution for this confusing pattern of results is offered by research examining the heritability of specific language impairment (SLI). SLI is defined by poor overall language development, including grammar and vocabulary. Children with the disorder show robust phonological deficits, and PM has been identified as both an endophenotype and candidate causal deficit (Bishop, 2006). Research initially suggested high heritability of SLI (Bishop, 2006), but in a study from the TEDS sample, Hayiou-Thomas and colleagues (Hayiou-Thomas, Oliver, & Plomin, 2005) found heritability was negligible. Bishop & Hayiou-Thomas (2008) demonstrated that the difference owed to varying diagnostic criteria. Specifically, heritability of SLI is low when the disorder is defined only on the basis of poor language scores, but much higher when diagnosis requires children to have had a history of speech-language therapy. Children with SLI are most likely to attract clinical concern when they have comorbid problems with speech production, which in turn are associated with phonological processing deficits (Bishop & Hayiou-Thomas, 2008; Peterson, Pennington, Shriberg, & Boada, 2009; Zhang & Tomblin, 2000). This work thus suggests there are etiologic subgroups of SLI: one with low language of environmental origin, and a second with genes that cause poor speech, language, and phonological processing. Possibly, we will find that there is a subgroup of children with PM and vocabulary deficits that share a genetic basis, and a second subgroup where these deficits are linked by common environmental risks.

Longitudinal twin studies can investigate etiologic relationships of different measures across time in order to address causal direction. Dionne and colleagues (Dionne, Dale, Boivin, & Plomin, 2003) used this approach to investigate the association between vocabulary and grammar in young children. Etiologic results indicated shared genetic and shared common environment operating on both early vocabulary and later grammar as well as on early grammar and later vocabulary, with magnitude of the effects being similar in both directions. The authors concluded that results supported bidirectional “bootstrapping” of these constructs in toddlers. We will use similar logic to investigate the degree to which shared etiologies contribute to deficits in PM and vocabulary across time. These longitudinal results will have implications for discriminating between the theoretical models of PM-vocabulary relations. For example, stronger evidence for a shared etiology for early PM deficits and later vocabulary deficits than for early vocabulary deficits and later PM deficits would be consistent with the phonological storage framework. Conversely, finding the opposite pattern would be consistent with a top-down account.

Background Specific to the Current Investigation: Previous ILTS Studies

Several previous studies have investigated the univariate etiology of verbal memory and vocabulary across the full range of individual differences in the ILTS. These studies utilized

the same measures of age 5 vocabulary, age 7 PM, and age 8 vocabulary as the current study, while the age 5 verbal memory composite used in earlier studies was slightly broader than in the current study. Previous studies have utilized both latent traits and composites, with similar results.

Samuelsson and colleagues reported on the behavior genetics of the age 5 constructs (Samuelsson et al., 2005). Verbal memory showed large effects of genes ($h^2 = .57$; 95% confidence interval (CI) = .35-.79) and modest effects of shared environment ($c^2 = .29$; 95% CI = .08-.48). At this age, vocabulary showed the opposite pattern, with moderate effects of genes ($h^2 = .32$; 95% CI = .06-.56) and a large effect of shared environment ($c^2 = .60$; 95% CI = .38-.81). In a later study (Byrne et al., 2007), age 7 Sentence Memory had moderate genetic effects ($h^2 = .35$, 95% CI = .05-.67), and the effect of shared environment remained modest and was not statistically significant ($c^2 = .24$; 95% CI = .00-.49). Finally, for age 8 vocabulary (Byrne et al., 2009), the effects of genes and shared environment were similar in magnitude and both in the moderate range ($a^2 = .44$; 95% CI = .31-.59; $c^2 = .36$; 95% CI = .22-.49). Thus, as described above, these results agree with findings from the TEDS in documenting that 1) genes have a smaller influence on preschool vocabulary than on measures of verbal short-term memory and 2) genetic influence on vocabulary increases over the school years.

Research Questions

The primary goals of the current study are to understand the etiologic relationship between PM and vocabulary deficits, both within and across time. We addressed the following specific questions:

1. What are the univariate etiologies of PM deficits and vocabulary deficits in school-age children? Do a similar balance of genetic and environmental influences cause poor PM and vocabulary as contribute to individual differences across the full range?
2. What is the etiologic relationship between PM and vocabulary deficits at age 5? Do these problems co-occur primarily because of shared genes or shared environmental risk factors? Is there evidence for etiologic subgroups?
3. What is the etiologic relationship between PM and vocabulary deficits across the school years? Is there stronger evidence for a common etiology for the bottom-up effect (early PM and later vocabulary deficits) or for the top-down effect (early vocabulary and later PM deficits)? Alternatively, are both effects similar in magnitude, thus suggesting reciprocal PM-vocabulary relations during this period?

Method

Participants

This study included participants from the four countries in the ILTS (United States, Australia, Sweden, and Norway), with Sweden and Norway grouped together as Scandinavia for analyses. The twin pairs recruited from the Colorado Twin Registry in the U.S., the National Health and Medical Research Council's Australian Twin Registry, and from the Medical Birth Registries in Norway and Sweden. No payment was given for participation in Australia and Scandinavia, but the parents in the U.S. sample received a payment of \$100 for participation. Parents of the Colorado twins were approached by mail or phone and 86% agreed to participate in the study. The twins' parents in Australia and Scandinavia were all approached by mail with a participation rate of 62% in Australia and 60% in Scandinavia. Identical and same-sex fraternal twin pairs were initially tested prior to entry into kindergarten (mean age = 58.83 months), with follow-up tests at the end of

kindergarten, 1st and 2nd grade (mean ages of 75.53 months, 88.50 months, and 100.49 months, respectively). All children spoke the native language of their home country as their first language (i.e., English, Swedish, or Norwegian). This study was approved by the Institutional Review Boards at the participating institutions, and parents gave informed consent for their children's participation.

None of the children were enrolled in kindergarten at the first assessment, but most were attending a preschool program (94% in Australia, 85% in the U.S., and 90% in Scandinavia), with the average attendance being 17.3 hours per week in Australia, 16.8 hours in the U.S., and 25.7 hours in Scandinavia. Most of the preschool programs in both Australia and the U.S. included at least some literacy related activities but this was quite variable and difficult to quantify. Literacy activities were explicitly excluded from the national preschool curriculum in both Norway and Sweden. There were no significant differences in parents' mean years of education in the U.S. (14.1), Australia (13.4) and Scandinavia (13.9). It should be noted that the Australian Twin Registry is a voluntary registry and currently registers only about 15% of twins. However, on national SES indicators with a mean of 100 and standard deviation of 15, the Twin Registry scores were Education = 101.4 (11.9), Income = 100.5 (11.6), Marital Status = 102 (10.3), Occupation 101.5 (7.5), and Single-parenthood = 100.6 (6.1). Although these numbers indicate a slight bias toward higher SES and less variance for families in the Australian Twin Registry, the difference is small.

Data from initial testing were available for a total of 1,045 twin pairs (264 from Australia, 489 from Colorado, and 292 from Scandinavia). Not all twin pairs had completed the study, however. Age 7 data were available for 956 twin pairs and age 8 data for 870 twin pairs. Table 1 characterizes the sample by gender, country, and zygosity.

Procedure

At study entry, participants completed a battery of cognitive, language, and preliteracy tests over a period of 5 days. Members of a twin pair were tested separately, either in quiet rooms at their preschool or in their homes. Follow-up testing at the ends of Kindergarten, 1st, and 2nd grades included a battery of cognitive, language, and literacy tests. Testing for each follow-up visit took place in a single session lasting approximately one hour at participants' homes or schools.

Measures

The most comprehensive language and cognitive evaluation took place during the initial testing, with two measures of each PM and vocabulary at this time. Due to time constraints, fewer measures were administered during follow-up testing, and each language construct was reassessed at just one later time point (PM at age 7 and vocabulary at age 8). Measures changed across time to avoid ceiling effects and to assess newly developing skills.

Phonological Memory—PM was assessed at age 5 with The Children's Nonword Repetition Test (Gathercole, Willis, Baddeley, & Emslie, 1994) as well as WPPSI-R Sentence Memory (Wechsler, 1989). Raw scores were corrected for age and standardized within gender and country, and an age 5 PM composite was created by averaging the standardized scores, then re-standardizing so that the composite had a mean of 0 and standard deviation of 1. PM was reassessed at age 7 with WPPSI Sentence Memory. Again, raw scores were corrected for age and standardized within gender and country. Previous research has demonstrated that sentence repetition is constrained by PM (Willis & Gathercole, 2001)¹.

Vocabulary—Vocabulary knowledge was assessed at age 5 with WPPSI Vocabulary (Wechsler, 1989) and a confrontation naming task (Fisher & Glennister, 1992). Raw scores were corrected for age and standardized within gender and country, and an age 5 vocabulary composite was created by averaging the standardized scores, then re-standardizing so that the composite had a mean of 0 and standard deviation of 1. Vocabulary knowledge was reassessed at age 8 with the Boston Naming test (Kaplan, Goodglass, & Weintraub, 2001), and raw scores were corrected for age and standardized within gender and country.

History of Speech Sound Disorder (SSD)—Children with histories of SSD were identified within the U.S. portion of the ILTS (total $n = 978$) via a parent questionnaire. This questionnaire was mailed to all families in the Colorado LTS, and the response rate was 65.5%. A child was considered to have a positive history of SSD if the parent endorsed that the child had received speech/language therapy and that the child had had difficulties with articulation. A child was considered to have a negative history of SSD if the parent answered “no” to both the speech/language therapy question and the articulation difficulties question. If the questionnaire was not returned, or if the parent answered “yes” to one of the two SSD criteria and “no” to another, SSD history was considered ambiguous and the child was removed from the relevant analyses. We also included a number of exclusionary criteria to ensure that parents were describing idiopathic SSD, to remain consistent with previous etiologic research. Exclusionary criteria included parent-endorsed poor hearing, cleft palate, risky birth, or more serious developmental problems (i.e., autism or intellectual disability). Overall, this procedure identified 80/554 children as having a positive history of SSD. Thus, lifetime prevalence of SSD in this sample is estimated at 14.4%, which is similar to rates reported by previous epidemiological studies (Beitchman, Nair, Clegg, & Patel, 1986; Peckham, 1973). Of the 80 children with positive SSD histories, 38 came from twin pairs in which both members of the pair met criteria while 42 were children whose co-twin did not meet SSD criteria. The SSD pairwise concordance rate for monozygotic (MZ) twins was 0.93, while the pairwise concordance rate for dizygotic (DZ) twins was 0.83. Inclusion of these data will allow us to examine whether there is stronger evidence for a genetic basis for poor language among children with a history of SSD than among children without such a history.

Analyses

Three sets of analyses were conducted in order to investigate the etiology of PM and vocabulary deficits both within a time point (at age 5) and across time. For all analyses, we first identified the subset of participants with a deficit in at least one of the constructs. A deficit was defined as a score that fell at least 1.25 standard deviations below the population mean, and probands were selected for having a deficit in either PM or vocabulary (i.e., we selected the bottom 10 percent of the population on each variable.)

The DeFries-Fulker (DF) method (DeFries & Fulker, 1985; DeFries & Fulker, 1988) is a regression-based approach suitable for extremes analysis. Probands who fall below a cut-off value in a particular variable are selected. The scores of MZ and DZ co-twins are then predicted from proband scores in a regression model. The logic of the method is as follows: to the extent that having a poor score is due to genes, the scores of DZ co-twins should regress further back toward the population mean than the scores of MZ co-twins. Effects of shared environment can also be estimated with this method, and are based on the extent to which co-twins resemble probands, regardless of zygosity. Effects of non-shared

¹Further, in a related phenotypic study in an independent sample, we assessed PM with a sentence repetition task as well as the more “traditional” PM measures of nonword repetition and digit span. Results from path analyses describing the relationship between PM and vocabulary from age 5 to 8 were essentially identical whether PM skill was estimated with sentence repetition only or with a composite of nonword repetition and digit span (Peterson et al., in submission).

environments include measurement error and are based on the extent to which MZ probands and their co-twins differ.

In the first set of analyses, we ran univariate DF regressions to examine the relative contributions of genes, shared environment, and nonshared environment (h^2_g , c^2_g , and e^2_g , respectively) to a deficit in PM or vocabulary. We used the basic DF equation: $C = B_1P + B_2R + K$, where C stands for co-twin's score on the relevant construct, P stands for the proband's score, and R represents the coefficient of relationship (1.0 for MZ twins and 0.5 for DZ twins). Data were scaled so that B_1 gave an unbiased estimate of heritability (h^2_g), and estimates of c^2_g and e^2_g were then derived with an adaptation of Falconer's method. In these and subsequent DF analyses, twin pairs in which both members met the extreme selection criteria were double entered, and standard errors of the regression coefficients were conservatively corrected for the number of double-entered pairs (Rodgers & Kohler, 2005; Stevenson, Pennington, Gilger, Defries, & Gillis, 1993).

In the second set of analyses, we reran the univariate regressions with the inclusion of country as a covariate in order to test if heritability varied as a function of culture. We used the extended regression equation: $C = B_1P + B_2R + B_3Culture + B_4P*Culture + B_5Culture*R + K$. The B_5 coefficient tested whether the heritability of the construct varied as a function of culture. We treated culture as a dichotomous variable, and for each construct, we ran three sets of regressions (Australia vs. U.S., U.S. vs. Scandinavia, and Australia vs. Scandinavia).

Finally, we tested bivariate heritability and shared environmentality of PM and vocabulary both within a time point (at age 5) and across development. The goal of these analyses was to further understand the basis for the phenotypic correlation between PM and vocabulary—does it owe to shared genes, shared environmental experiences or both? The bivariate model is similar to the basic univariate model, except that proband selection is based on one construct, and this value is used to predict the co-twin's score on a second construct. For example, to determine whether genes that contribute to a deficit in PM at age 5 also influence vocabulary at the same age, we ran the equation: $C_{Voc5} = B_1P_{PM5} + B_2R + K$. We scaled data so that B_1 gave an unbiased estimate of $h^2_{g,xy}$. This value indicates the proportion of a poor score on the co-twin measure (age 5 vocabulary) that is due to genes that also influence the selection measure (age 5 PM). In addition, we computed bivariate shared environmentality ($c^2_{g,xy}$) based on Falconer's method.

In all analyses, we checked that the assumptions of multiple regression were met including homoscedasticity, normality of residuals, lack of multicollinearity, and lack of influential cases. These assumptions were generally met, with a few exceptions described in more detail below.

Missing data

There were a small number of missing data at the two earlier time points (age 5 and age 7). Generally, data were missing because of time constraints. Data for both PM variables at age 5 were missing for 12 children (of 2,090 total), and data for both vocabulary variables at age 5 were missing for 4 children (of 2,090 total). Five children (of 1,912 total) were missing data for sentence repetition at age 7. No data were missing for vocabulary at age 8.

Results

Before beginning etiologic analyses, we examined demographic information for the sample at age 5, characterized according to whether participants had no PM or vocabulary deficit, a PM deficit only, a vocabulary deficit only, or both deficits. Table 2 provides this information

for the 2,078 children for whom both PM and vocabulary data were available at age 5. We were interested in determining whether we could identify factors that might be associated with etiologic subgroups for language-related difficulties. For these purposes, the most important comparison is between the PM deficit only and vocabulary deficit only groups. Ages did not differ across the four groups. There were also no gender differences, which is unsurprising since scores were normed within gender. There were significant differences in parent education ($F(3, 1877) = 18.65, p < .001$); the parents of children with neither deficit had completed more education than parents of the other three groups, who did not differ from one another. Within the U.S. sample, we further examined the variables of race and history of SSD. There is little racial diversity in the Australian or Scandinavian samples. In the U.S., the proportion of Caucasian children was not equivalent across proband groups ($X^2(3)=48.41, p<.001$), and this difference remained significant when comparing just the PM deficit only and vocabulary deficit only groups ($X^2(1)=14.55, p=.001$). The PM deficit only group was 91.4% Caucasian ($n=64$), 1.4% African American ($n=1$) and 7.1% Non-White Hispanic ($n=5$). The Vocabulary deficit only group was 56.0% Caucasian ($n=37$), 5.2% Asian American ($n=3$), 19.0% African American ($n=11$), 10.0% Non-White Hispanic ($n=6$), and 1.7% Pacific Islander ($n=1$). The group with both PM and vocabulary deficits was 70.0% Caucasian ($n=30$), 6.8% Asian American ($n=3$), 11.4% African American ($n=5$), and 13.6% Non-White Hispanic ($n=6$). Rates of speech therapy for articulation problems also varied ($X^2(3)=18.60, p<.001$). Children in the PM deficit only group had nearly twice the rate of SSD compared to children in the vocabulary deficit only group. Although this direct comparison did not reach statistical significance, it is notable that the PM deficit only group had a higher rate of SSD than controls, while the vocabulary deficit only group did not.

Univariate DF results

Univariate results are summarized in Table 3. In general, our findings concerning the etiology of PM and vocabulary deficits agree with previous results from the ILTS across the full range of individual differences. We found that deficits in both constructs were significantly heritable at all time points, indicating that a portion of a poor PM or vocabulary score is due to genetic influence. At ages 5 and 7, heritability of a PM deficit was large while the effect of shared environment was small and nonsignificant. The pattern differed for age 5 vocabulary. In that case, heritability was moderate, and there was a similar, statistically significant effect of shared environment. By age 8, heritability of a vocabulary deficit was large; the effect of shared environment remained moderate and statistically significant. Estimates of nonshared environment, which includes measurement error, were generally modest.

Univariate DF results with covariate of country

Next, we tested whether heritability of the four constructs varied across countries by including country as a covariate in the univariate DF equations. These results should be interpreted with some caution, since splitting the total sample into countries resulted in relatively small proband groups (n twin pairs for Australia, United States, and Scandinavia as follows: 5 PM: 59, 114, and 68; age 5 vocabulary: 59, 106, and 75; age 7 PM: 47, 108, and 69; age 8 vocabulary: 43, 104, and 50). Furthermore (and not surprisingly), there was evidence for multicollinearity of the key interaction term with other variables in the model. The most likely result of this multicollinearity is an inflated standard error, and hence reduced power to detect significant cross-cultural effects. There was no evidence for cross-country differences in heritability of age 5 vocabulary, age 7 PM, or age 8 vocabulary. For age 5 PM, there was a significant difference between Australia and Scandinavia ($\beta = 0.61, p = 0.046$), while the difference between Australia and the U.S. reflected a trend ($\beta = 0.43, p =$

0.098). Heritability of a PM deficit was highest in Australia ($h^2_g = 1.00, .51, \text{ and } .41$ in Australia, the U.S., and Scandinavia, respectively).

Bivariate DF results

Finally, we tested bivariate heritability and shared environmentality of PM and vocabulary both within a time point (at age 5) and across time. The cross-time analyses investigated three meaningful variable pairs (age 5 PM/age 8 vocabulary, age 7 PM/age 8 vocabulary, and age 5 vocabulary/age 7 PM). Results are reported in Table 4. These results can be understood within the context of the phenotypic association for each of the variable pairs, also reported in Table 4 as the Phenotypic Group Correlation (Oliver, Dale, & Plomin, 2004). In general, phenotypic associations clustered in the moderate to large range.

There was evidence for a shared etiology for PM and vocabulary deficits, though there were also negative findings. Further, the pattern of results varied according to whether probands were selected for a PM deficit or vocabulary deficit. There was significant bivariate heritability only for age 7 PM and age 8 vocabulary. In that case, $h^2_{g,xy}$ was .27, indicating that 27% of a poor vocabulary score at age 8 is due to genes that also influenced a PM deficit at age 7. In the other two analyses in which selection was on PM, bivariate heritability was not statistically significant. However, the magnitude of $h^2_{g,xy}$ was similar to the age 7 PM/age 8 vocabulary analysis, and reflected a trend for age 5 PM/age 5 vocabulary. In contrast, when probands were selected for a vocabulary deficit, results suggested that common shared environmental experiences underlie the phenotypic overlap of PM and vocabulary. Bivariate shared environmentality was moderate and statistically significant for age 5 vocabulary and age 7 PM, indicating that 34% of a poor PM score at age 7 is due to environmental experiences that also influenced a vocabulary deficit at age 5. Further, $c^2_{g,xy}$ was very similar in magnitude and reflected a trend for the within-time analysis at age 5.

Finally, to better understand the univariate results for vocabulary, which suggested increasing heritability over time, we computed bivariate heritability of vocabulary deficits from age 5 to age 8. Results suggested that stability in poor vocabulary over time likely owes to both genetic and environmental influences. Specifically, there was a trend for bivariate heritability ($h^2_{g,xy} = 0.25, p = .07, CI = 0, 0.51$); a similar effect size for bivariate shared environmentality just crossed the threshold for statistical significance ($c^2_{g,xy} w = 0.29, p = .04, CI = .01, .57$).

The asymmetry in our findings suggests that there may be etiologic subgroups of children in whom PM and vocabulary weaknesses co-occur for different reasons. This explanation might also help account for some of our non-significant findings, to the extent that children from different subgroups are grouped together in a given set of probands. To test this possibility, we ran the bivariate analyses again in children who had only a PM deficit or only a vocabulary deficit at age 5. Results are reported in Table 5.

In the PM deficit only subgroup, bivariate heritability of PM and vocabulary weaknesses were moderate statistically significant both within time and across time, while bivariate shared environmentality was null. Although these children did not, by definition, meet the cut-off for a vocabulary deficit, their mean vocabulary scores were still below average. These results indicate that approximately 40% of their poor vocabulary scores were due to genes that also influenced their PM deficit. In contrast, in the vocabulary deficit only subgroup, there was not compelling evidence for a shared etiology for PM and vocabulary weaknesses. Bivariate heritability was essentially null both within and across time, and bivariate shared environmentality was modest and not statistically significant in either case (though it reflected a trend for age 5 vocabulary/age 5 PM).

Discussion

This study addressed three questions: 1) Do deficits in PM and vocabulary arise from the same types of etiologic influences as PM and vocabulary scores in the full population? 2) Does the phenotypic correlation between PM and vocabulary arise primarily from shared genes, shared environments, or both? 3) How does the etiologic relation of earlier PM to later vocabulary compare to that of earlier vocabulary to later PM?

On the first question, it appears that a similar balance of genes and environment cause weaknesses in PM and vocabulary as cause the full range of individual differences. Our findings echo earlier conclusions from the ILTS and TEDS that both constructs are heritable and that shared environment contributes more to individual differences in vocabulary than to PM (Hayiou-Thomas et al., 2006; Samuelsson et al., 2005). Also, consistent with findings from both samples, there was some suggestion that genetic influence on vocabulary grew stronger with age, though our confidence intervals for estimates at the two time points were large and overlapping. We found few cross-cultural etiologic differences, in agreement with previous findings in the ILTS across the full range of individual differences. We did find differentially high heritability of age 5 PM deficits in Australia. Previous studies have documented that performance of the Australian cohort is superior to that of the other countries (Samuelsson et al., 2005; Byrne et al., 2009), likely reflecting a selection bias at least in part, since only the Australian sample is drawn from a volunteer registry. The reduced environmental range associated with this self-selection could relate to higher heritability in Australia, particularly at younger ages (i.e., before the start of formal schooling, which acts to reduce environmental variance in all countries).

Regarding the second question, the etiologic basis for the PM-vocabulary association appeared to vary for different groups of children. There was clear evidence for a genetic basis to the PM-vocabulary relationship among one subgroup. Genes that influenced membership in the PM deficit only proband group also impacted co-twins' vocabulary scores. However, genes that influenced membership in the vocabulary proband group did not carry over to co-twins' PM scores. Instead, shared environmental experiences explained the phenotypic correlation in this case. Notably, however, the effect for bivariate environmentality was weaker (and nonsignificant) in the vocabulary deficit only subgroup. Within the U.S. sample, children in the PM deficit only group had higher rates of SSD histories compared to controls, consistent with previous research demonstrating that a history of speech therapy is one marker for language problems of genetic origin (Bishop & Hayiou-Thomas, 2008). Our findings thus agree with this earlier study documenting etiologic subgroups of SLI and suggest that there may be at least two routes to a poor vocabulary score: one through genes that also influence PM, and one through a relatively impoverished environment.

The amount of language input provided to children is a well-documented environmental influence on vocabulary development (e.g., Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991). We did not have direct measures of language input, but did find that parents of children with vocabulary deficits had lower levels of education than parents of children with no language deficits. However, the meaning of this finding is ambiguous. Although parent education levels are one proxy for the quality of the linguistic environment parents provide to their children, educational attainment is itself genetically influenced (Baker, Treloar, Reynolds, Heath, & Martin, 1996; Miller, Mulvey, & Martin, 2001; Szanton, Johnson, Thorpe, & Whitfield, 2009). Further, though poor PM was more influenced by genes and poor vocabulary more by shared environment, we found no differences in parent education between the PM deficit only subgroup and the vocabulary deficit only subgroup. Within the U.S. sample, a disproportionate number of racial minority children, and African American

children in particular, were identified as having a vocabulary deficit. This finding could reflect measurement error, cultural bias in the vocabulary measures in the current study, a genuine etiologic risk factor that differentially impacts African American children, or some combination of these factors. The current study design does not allow us to disentangle these possibilities.

Our third research question concerned the relative magnitude of the estimates for shared etiology of the bottom-up effect (PM influencing vocabulary acquisition) versus the top-down effect (vocabulary influencing PM) over time. This question bears on the competing cognitive models accounting for the PM-vocabulary association. Within the full proband groups, these effects were similar in magnitude. Within the proband subgroups selected for “pure” PM or vocabulary deficits, however, the bottom-up effect appeared stronger. Overall, these results agree with previous phenotypic work, including our own earlier study in the ILTS, and suggest that the relationship between PM and vocabulary is reciprocal during the early school years, but the bottom-up effect is likely dominant for children with phonological problems. The bottom-up effect appears to be primarily driven by genes, while the top-down effect appears to be primarily driven by shared environment. Future studies of both normal and disordered language development are needed to confirm this interesting result.

The current study included a number of limitations. First, we made the simplifying assumptions common to twin studies, including equal environments for MZ and DZ twins, generalizability to non-twin samples, and no significant assortative mating. The DF equations used in this study presume additive genetic effects only. We did not test for gene-gene or gene-environment interactions, which would be fruitful questions for future research. Second, our sample size was relatively small, and we combined samples across countries with different methods for compiling their twin registries. Although current results provided only limited evidence for cross-cultural etiologic differences, with greater power, we might have detected more such effects. Third, our measures changed across time, and later assessment of PM and vocabulary occurred at two different points. Although our measures were selected to be optimal for each age, the different tests do emphasize slightly different skills. Thus, our conceptualization of PM and vocabulary are by necessity relatively coarse. Research using finer-grained measures has found that the nature of the PM-vocabulary relationship depends partly on how these constructs are assessed (Jarrold et al., 2009). Learning to distinguish a novel phonological form from close distractors depended heavily on verbal short-term memory, while establishing a mapping between a novel phonological form and a referent did not. It is conceivable that the former process is more subject to genetic influence while the latter relates predominantly to environment. A future study could use more specific measures within a genetically-sensitive design to investigate this possibility.

Despite these limitations, we believe this study makes an important contribution to research investigating the etiologic overlap of language sub-skills. We provide the first specific evidence that weaknesses in PM and vocabulary can owe to a common underlying cause. We also provide confirmatory evidence that the etiology of SLI is heterogeneous, and that genetic risk factors appear to account for poor language development in only a subgroup of children. These results have important theoretical and clinical implications. Theoretically, we support a critical prediction of models (including the phonological storage framework) that seek to explain the PM-vocabulary association. Clinically, results underscore the importance of considering both familial-genetic history and current environment when estimating children’s level of risk for poor language development. Children with either type of risk factor are more likely to have SLI, and thus should benefit from more in-depth language assessment and treatment, if warranted.

Acknowledgments

Funding was provided by the Australian Research Council (DP0663498 and DP0770805), the National Institute for Child Health and Human Development (HD27802, HD38526, and HD49027), and grants from the Research Councils of Norway and Sweden. We thank the Colorado and Australian Twin Registries, our testers, and the children and parents involved.

References

- Baddeley A, Gathercole S, Papagno C. The phonological loop as a language learning device. *Psychological Review*. 1998; 105:158–173. [PubMed: 9450375]
- Baddeley AD, Papagno C, Vallar G. When long-term learning depends on short-term storage. *Journal of Memory and Language*. 1988; 27:586–595.
- Baker LA, Treloar SA, Reynolds CA, Heath AC, Martin NG. Genetics of educational attainment in Australian twins: Sex differences and secular changes. *Behavior Genetics*. 1996; 26:89–102. doi: 10.1007/bf02359887. [PubMed: 8639155]
- Beitchman JH, Nair R, Clegg M, Patel PG. Prevalence of speech and language disorders in 5-year-old kindergarten children in the Ottawa-Carleton region. *Journal of Speech & Hearing Disorders*. 1986; 51:98–110. [PubMed: 3702369]
- Bishop DVM. What causes specific language impairment in children? *Current Directions in Psychological Science*. 2006; 15:217–221. [PubMed: 19009045]
- Bishop DVM, Adams CV, Norbury CF. Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain & Behavior*. 2006; 5:158–169.
- Bishop DVM, Hayiou-Thomas ME. Heritability of specific language impairment depends on diagnostic criteria. *Genes Brain and Behavior*. 2008; 7:365–372. doi: 10.1111/j.1601-183X.2007.00360.x.
- Bishop DVM, North T, Donlan C. Nonword repetition as a behavioural marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry*. 1996; 37:391–403. [PubMed: 8735439]
- Brown, GDA.; Hulme, C. Nonword repetition, STM, and word age-of-acquisition: A computational model. In: Gathercole, SE., editor. *Models of Short-Term Memory*. Psychology Press; East Sussex, UK: 1996. p. 129-148.
- 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. doi: 10.1016/j.jneuroling.2008.09.003. [PubMed: 20161176]
- 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*. 2007; 20:77–102.
- Cheung H. Nonword span as a unique predictor of second-language vocabulary language. *Developmental Psychology*. 1996; 32:867–873.
- Dale PS, Harlaar N, Hayiou-Thomas ME, Plomin R. The Etiology of Diverse Receptive Language Skills at 12 Years. *Journal of Speech Language and Hearing Research*. 2010; 53:982–992. doi: 10.1044/1092-4388(2009/09-0108).
- DeFries JC, Fulker DW. Multiple regression analysis of twin data. *Behavior Genetics*. 1985; 15:467–473. [PubMed: 4074272]
- DeFries JC, Fulker DW. Multiple regression analysis of twin data: Etiology of deviant scores versus individual differences. *Acta Geneticae Medicae et Gemellologiae: Twin Research*. 1988; 37:205–216.
- Dionne G, Dale PS, Boivin M, Plomin R. Genetic evidence for bidirectional effects of early lexical and grammatical development. *Child Development*. 2003; 74:394–412. [PubMed: 12705562]
- Fisher, JP.; Glennister, JM. *The Hundred Pictures Naming Test*. Australian Council for Educational Research; Hawthorn, Australia: 1992.

- Gathercole SE. Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*. 2006; 27:513–543.
- Gathercole SE, Baddeley AD. Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language*. 1990a; 29:336–360.
- Gathercole SE, Baddeley AD. The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*. 1990b; 81:439–454.
- Gathercole SE, Willis CS, Baddeley AD, Emslie H. The Children's Test of Nonword Repetition: A test of phonological working memory. *Memory & Cognition*. 1994; 2:103–127.
- Gathercole SE, Willis CS, Emslie H, Baddeley AD. Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental Psychology*. 1992; 28:887–898.
- Gray S. Word learning by preschoolers with specific language impairment: predictors and poor learners. *Journal of Speech, Language, and Hearing Research*. 2004; 47:1117–1132.
- Gupta P, Tisdale J. Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of Memory and Language*. 2009; 61:481–502. doi: 10.1016/j.jml.2009.08.001.
- Hayiou-Thomas ME, Dale PS, Plomin R. The etiology of variation in language skills changes with development: a longitudinal twin study of language from 2 to 12 years. *Developmental Science*. 2012; 15:233–249. doi: 10.1111/j.1467-7687.2011.01119.x. [PubMed: 22356179]
- Hayiou-Thomas ME, Harlaar N, Dale PS, Plomin R. Preschool speech, language skills, and reading at 7, 9, and 10 years: Etiology of the relationship. *Journal of Speech Language and Hearing Research*. 2010; 53:311–332. doi: 10.1044/1092-4388(2009/07-0145).
- Hayiou-Thomas ME, Kovas Y, Harlaar N, Plomin R, Bishop DVM, Dale PS. Common aetiology for diverse language skills in 4 1/2-year-old twins. *Journal of Child Language*. 2006; 33:339–368. [PubMed: 16826830]
- Hayiou-Thomas ME, Oliver B, Plomin R. Genetic influences on specific versus nonspecific language impairment in 4-year-old twins. *Journal of Learning Disabilities*. 2005; 38:222–232. doi: 10.1177/00222194050380030401. [PubMed: 15940960]
- Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Early vocabulary growth – relation to language input and gender. *Developmental Psychology*. 1991; 27:236–248. doi: 10.1037/0012-1649.27.2.236.
- Jarrold C, Baddeley AD, Hewes AK, Leeke TC, Phillips CE. What links verbal short-term memory performance and vocabulary level? Evidence of changing relationships among individuals with learning disability. *Journal of Memory and Language*. 2004; 50:134–148.
- Jarrold C, Baddeley AD, Phillips C. Down syndrome and the phonological loop: The evidence for, and importance of, a specific verbal short-term memory deficit. *Down Syndrome: Research & Practice*. 1999; 6:61–75.
- Jarrold C, Thorn ASC, Stephens E. The relationships among verbal short-term memory, phonological awareness, and new word learning: Evidence from typical development and Down syndrome. *Journal of Experimental Child Psychology*. 2009; 102:196–218. [PubMed: 18707692]
- Kaplan, E.; Goodglass, H.; Weintraub, S. *Boston Naming Test*. 2nd ed.. Lippincott, Williams, & Wilkins; Philadelphia: 2001.
- Lee NR, Pennington BF, Keenan JM. Verbal short-term memory deficits in Down syndrome: phonological, semantic, or both? *Journal of Neurodevelopmental Disorders*. 2010; 2:9–25. doi: 10.1007/s11689-009-9029-4. [PubMed: 22127838]
- Masoura EV, Gathercole SE. Phonological short-term memory and foreign language learning. *International Journal of Psychology*. 1999; 34:383–388.
- Metsala JL. Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*. 1999; 91:3–19.
- Michas IC, Henry LA. The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology*. 1994; 12:147–163.
- Miller P, Mulvey C, Martin N. Genetic and environmental contributions to educational attainment in Australia. *Economics of Education Review*. 2001; 20:211–224. doi: 10.1016/S0272-7757(00)00018-2.

- Montgomery JW. Examination of phonological working memory in specifically language-impaired children. *Applied Psycholinguistics*. 1995; 16:355–378.
- Oliver B, Dale PS, Plomin R. Verbal and nonverbal predictors of early language problems: an analysis of twins in early childhood back to infancy. *Journal of Child Language*. 2004; 31:609–631. doi: 10.1017/s030s000904006221. [PubMed: 15612392]
- Olson RK, Keenan JM, Byrne B, Samuelsson S, Coventry WL, Corley R, Hulslander J. Genetic and Environmental Influences on Vocabulary and Reading Development. *Scientific Studies of Reading*. 2011; 15:26–46. doi: 10.1080/10888438.2011.536128. [PubMed: 21132077]
- Peckham CS. Speech defects in a national sample of children aged seven years. *British Journal of Disorders of Communication*. 1973; 8:2–8. [PubMed: 4759956]
- Peterson RL, Pennington BF, Samuelsson S, Byrne B, Olson RK. Phonological memory and vocabulary growth in school-age children with and without language-based disorders. in submission.
- Peterson RL, Pennington BF, Shriberg LD, Boada R. What influences literacy outcome in children with Speech Sound Disorder? *Journal of Speech Language and Hearing Research*. 2009; 52:1175–1188. doi: 10.1044/1092-4388(2009/08-0024).
- Pierpont EI, Richmond EK, Abbeduto L, Kover ST, Brown WT. Contributions of phonological and verbal working memory to language development in adolescents with fragile X syndrome. *Journal of Neurodevelopmental Disorders*. 2011; 3:335–347. doi: 10.1007/s11689-011-9095-2. [PubMed: 21993552]
- Rodgers JL, Kohler HP. Reformulating and simplifying the DF analysis model. *Behavior Genetics*. 2005; 35:211–217. doi: 10.1007/s10519-004-1020-y. [PubMed: 15685433]
- Samuelsson S, Byrne B, Quain P, Wadsworth S, Corley R, DeFries JC, Olson R. Environmental and genetic influences on prereading skills in Australia, Scandinavia, and the United States. *Journal of Educational Psychology*. 2005; 97:705–722.
- Snowling MJ, Chiat S, Hulme C. Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics*. 1991; 12:369–373.
- Stevenson J, Pennington BF, Gilger JW, DeFries JC, Gillis JJ. Hyperactivity and spelling disability – testing for shared genetic etiology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 1993; 34(7):1137–1152. doi: 10.1111/j.1469-7610.1993.tb01779.x.
- Szanton SL, Johnson B, Thorpe RJ, Whitfield K. Education in Time: Cohort Differences in Educational Attainment in African-American Twins. *PLOS ONE*. 2009;4. doi: e7664 10.1371/journal.pone.0007664.
- Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence--Revised. Psychological Corporation; New York: 1989.
- Willis CS, Gathercole SE. Phonological short-term memory contributions to sentence processing in young children. *Memory*. 2001; 9:349–363.
- Zhang XY, Tomblin JB. The association of intervention receipt with speech-language profiles and social-demographic variables. *American Journal of Speech-Language Pathology*. 2000; 9:345–357.

Table 1
Gender and Zygosity by Country within the ILTS

	Australia	United States	Scandinavia
Male twin pairs (n)			
MZ	83	97	71
DZ	56	146	71
Total	138	243	142
Female twin pairs (n)			
MZ	71	127	71
DZ	54	119	79
Total	126	246	150

Table 2
Demographic Information for Children Characterized According to Presence of PM and Vocabulary Deficits at Age 5

	Proband status at age 5			
	Control (no deficit)	PM deficit only	Vocabulary deficit only	PM and vocabulary deficits
n (overall)	1699	143	138	98
n (Australia)	429	38	38	21
n (United States)	796	70	58	44
n (Scandinavia)	474	35	42	33
Age in months	58.93 (2.78)	58.91 (3.00)	58.93 (2.85)	58.88 (2.60)
Parent education ¹	14.28 ^a (2.29)	13.48 ^b (2.16)	13.43 ^b (2.42)	12.88 ^b (1.92)
Caucasian ²	89.7% ^a (712/794)	91.4% ^a (64/70)	63.8% ^b (37/58)	68.2% ^b (30/44)
Speech therapy ³	12.3% ^a (58/470)	27.8% ^{b, c} (10/36)	14.3% ^b (4/28)	42.1% ^c (8/19)

¹ Average mother and father years of education.

² Within United States portion of sample only. These data are reported two ways: first as a percentage, and then as the number of Caucasian children divided by the total number of children with relevant data available.

³ History of speech therapy for articulation problems within United States portion of sample only. These data are reported two ways: first as a percentage, and then as the number of children with a history of speech therapy divided by the total number of children with relevant data available.

Table 3
Results of Univariate DeFries-Fulker Analyses of PM and Vocabulary Deficits

Selection measure	PM5	Voc5	PM7	Voc8
n pairs	236	240	223	197
MZ pairs				
Proband mean	-1.76	-1.94	-1.81	-1.74
Co-twin mean	-1.30	-1.49	-1.24	-1.48
DZ pairs				
Proband mean	-1.74	-1.80	-1.83	-1.73
Co-twin mean	-0.74	-1.05	-0.69	-1.05
h^2_g (95% CI)	.62 (.35-.90)	.37 (.08-.66)	.62 (.35-.89)	.50 (.20-.79)
c^2_g (95% CI)	.12 (0-.40)	.39 (.10-.69)	.06 (0-.35)	.36 (.05-.66)
e^2_g	.26	.23	.32	.15

Notes. Heritability and shared environment estimates whose confidence intervals do not include 0 (reported in bold) are statistically significant at the $p < .05$ level. PM5 = age 5 phonological memory composite; Voc5 = age 5 vocabulary composite; PM7 = age 7 sentence repetition; Voc8 = age 8 Boston Naming Test.

Table 4
Results of Bivariate DeFries-Fulker Analyses of PM and Vocabulary Deficits

Selection measure (X)	PM5	PM5	PM7	Voc5	Voc5
Co-twin measure (Y)	Voc5	Voc8	Voc8	PM5	PM7
n pairs	240	212	210	235	225
Phenotypic Group	0.61	0.43	0.42	0.52	0.42
Correlation					
MZ pairs					
Proband mean (X)	-1.76	-1.76	-1.81	-1.94	-1.94
Co-twin mean (Y)	-1.01	-0.72	-0.74	-0.81	-0.66
DZ pairs					
Proband mean (X)	-1.73	-1.73	-1.83	-1.80	-1.80
Co-twin mean (Y)	-0.75	-0.52	-0.50	-0.72	-0.68
$h^2_{E,XY}$ (95% CI)	.28 [^] (0-.58)	.21 (0-.58)	.27 (.01-.53)	.04 (0-.46)	.00 (0-.27)
$c^2_{E,XY}$ (95% CI)	.29 [^] (0-.60)	.20 (0-0.51)	.14 (0-.41)	.38 [^] (0-.81)	.34 (.06-.63)

Notes. Estimates whose confidence intervals do not include 0 (reported in bold) are statistically significant at the p<.05 level.

[^] p < .1. For abbreviations, see Table 2.

Table 5
Results of Bivariate DeFries–Fulker Analyses of PM and Vocabulary with Probands Selected for "Pure" PM or Vocabulary Deficits at Age 5

Selection measure (X)	PM5 (PM deficit only)	PM5 (PM deficit only)	Voc5 (vocabulary deficit only)	Voc5 (vocabulary deficit only)
Co-twin measure (Y)	Voc5	Voc8	PM5	PM7
n pairs	142	125	136	225
MZ pairs				
Proband mean (X)	-1.71	-1.71	-1.78	-1.94
Co-twin mean (Y)	-0.63	-0.52	-0.49	-0.66
DZ pairs				
Proband mean (X)	-1.63	-1.63	-1.68	-1.80
Co-twin mean (Y)	-0.29	-0.13	-0.44	-0.68
$h^2_{g,xy}$ (95% CI)	.38 (.06–.70)	.45 (.07–.83)	.02 (0–.32)	0 (0–.27)
$c^2_{g,xy}$ (95% CI)	0 (0–.33)	0 (0–.40)	.25 [^] (0–.56)	.23 (0–.59)

Notes. Estimates whose confidence intervals do not include 0 (reported in bold) are statistically significant at the $p < .05$ level.

[^]
 $p < .1$. For abbreviations, see Table 2.