

Genetic and Environmental Sources of Covariance Between Reading Tests Used in Neuropsychological Assessment and IQ Subtests

Mark Wainwright,^{1,2,3} Margaret J. Wright,¹ Gina M. Geffen,² Laurie B. Geffen,² Michelle Luciano,¹ and Nicholas G. Martin¹

Received 3 July 2003—Final 12 Nov. 2003

In this study, we examined genetic and environmental influences on covariation among two reading tests used in neuropsychological assessment (Cambridge Contextual Reading Test [CCRT], [Beardsall, L., and Huppert, F. A. (1994). *J. Clin. Exp. Neuropsychol.* **16**:232–242], Schonell Graded Word Reading Test [SGWRT], [Schonell, F. J., and Schonell, P. E. (1960). *Diagnostic and attainment testing*. Edinburgh: Oliver and Boyd.] and among a selection of IQ subtests from the Multidimensional Aptitude Battery (MAB), [Jackson, D. N. (1984). *Multidimensional aptitude battery*, Ontario: Research Psychologists Press.] and the Wechsler Adult Intelligence Scale-Revised (WAIS-R) [Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. San Antonio: The Psychological Corporation]. Participants were 225 monozygotic and 275 dizygotic twin pairs aged from 15 years to 18 years (mean, 16 years). For Verbal IQ subtests, phenotypic correlations with the reading tests ranged from 0.44 to 0.65. For Performance IQ subtests, phenotypic correlations with the reading tests ranged from 0.23 to 0.34. Results of Structural Equation Modeling (SEM) supported a model with one genetic General factor and three genetic group factors (Verbal, Performance, Reading). Reading performance was influenced by the genetic General factor (accounting for 13% and 20% of the variance for the CCRT and SGWRT, respectively), the genetic Verbal factor (explaining 17% and 19% of variance for the CCRT and SGWRT), and the genetic Reading factor (explaining 21% of the variance for both the CCRT and SGWRT). A common environment factor accounted for 25% and 14% of the CCRT and SGWRT variance, respectively. Genetic influences accounted for more than half of the phenotypic covariance between the reading tests and each of the IQ subtests. The heritabilities of the CCRT and SGWRT were 0.54 and 0.65, respectively. Observable covariance between reading assessments used by neuropsychologists to estimate IQ and IQ subtests appears to be largely due to genetic effects.

KEY WORDS: Heritability; intelligence; neuropsychology; reading; twins.

INTRODUCTION

Investigations using the classical twin design to estimate genetic and environmental contributions to covariance among cognitive abilities are increasingly

common (Plomin and Craig, 2001). These studies have an important role in elucidating the genetic and environmental architectures that underlie mental abilities. From a neuropsychological perspective, clarification of the genetic and environmental structures that underpin the observable correlation between reading measures used in neuropsychological assessment of premorbid intelligence and IQ subtests contributes to an understanding of the basis of this important relationship.

Previous research from this laboratory using a somewhat smaller sample has examined the genetic

¹ Queensland Institute of Medical Research, Brisbane, Australia.

² Cognitive Neuropsychology Laboratory, School of Psychology and Medicine, University of Queensland, Brisbane, Australia.

³ To whom correspondence should be addressed at Queensland Institute of Medical Research, Post Office Royal Brisbane Hospital, Herston, Brisbane, QLD, 4029, Australia. Tel: +617 3362 0272. Fax: +617 3362 0101. e-mail: markWa@qimr.edu.au

basis of the observable correlations among IQ subtests (Luciano *et al.*, 2003). The current paper has a different focus in investigating the genetic and environmental sources of covariation among two reading tests used in neuropsychological assessment and a selection of IQ subtests. The two reading tests were the Cambridge Contextual Reading Test (CCRT) (Beardsall and Huppert, 1994), which is a contextualized version of the widely used National Adult Reading Test (NART) (Nelson, 1982), and the Schonell Graded Word Reading Test (SGWRT) (Schonell and Schonell, 1960). IQ subtests were drawn from the Multidimensional Aptitude Battery (MAB) (Jackson, 1984) and the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler, 1981). This study differs from previous behavioral genetic research investigating the relationship between reading measures and IQ in that it uses a neuropsychological test of reading that was explicitly designed to correlate with IQ measures and acts as a proxy measure of word knowledge.

Contrary to a modular view of intellectual abilities that is popular within cognitive neuroscience, it has been recognized that common genetic factors influence multiple specific cognitive skills (Petrill, 1997; Plomin and Craig, 2001). Consistent with the dominant theoretical perspective of the phenotypic structure of intelligence, several studies have suggested that a genetic General factor as well as genetic group and specific influences (termed Bifactor models, i.e., models with two levels) (Jensen, 1998) underlie observable relationships among cognitive measures (Rijsdijk, Vernon, and Boomsma, 1998; Rijsdijk, Vernon, and Boomsma, 2002). Possible alternatives to a Bifactor structure are a Spearman model (General factor plus specifics) or a Thurstone model (group factors plus specifics) (see Jensen and Weng, 1994).

The CCRT is a neuropsychological instrument used for estimating premorbid IQ for patients with suspected brain dysfunction resulting from neurological insult. The test, like the NART from which it is derived, requires reading aloud of irregular words. However, whereas the NART requires reading single words in isolation, the CCRT embeds the test items within meaningful sentences. The utility of the CCRT in estimating premorbid intelligence rests on three primary assumptions. First, that the ability to accurately read words aloud is related to IQ in neurologically unimpaired populations. Second, that following brain insult, word reading ability is less diminished than other cognitive skills. Third, that as all words are irregular, the application of grapheme–phoneme conversion (gpc) rules yields incorrect pronunciations, and thus correctly

read words are part of the reader's internal lexicon (Franzen, Burgess, and Smith-Seemiller, 1997). The contextualization of words is assumed to maximize the correlation between IQ and word reading by providing the best opportunity for word recognition (Beardsall and Huppert, 1994).

The SGWRT is used as an adjunct to the NART when assessing patients suspected of being poor readers. It was the first reading test used to assess premorbid IQ. However, it was realized that as a test designed specifically for children, it yielded a low ceiling when assessing adults (Franzen *et al.*, 1997). The test is composed of both regular and irregular words and is not contextualized.

The phenotypic correlation of word reading tests with IQ in neurologically intact populations is well established (see Franzen *et al.*, 1997). An estimate of the proportion of variance in WAIS-R Verbal IQ accounted for by the NART was 72% in a large cross-validation sample of participants aged from 16 to 88 years (Crawford *et al.*, 1989). From the same study, estimates of 66% and 33% were found for WAIS-R Full Scale IQ and Performance IQ, respectively. Similar estimates have been provided from other studies (e.g., Griffin *et al.*, 2001; Lastine-Sobecks, Jackson, and Paolo, 1998). Beardsall (1998) reported the proportion of variance in WAIS-R Verbal IQ accounted for by the CCRT as 61%. NART performance has also been found to load strongly (0.85) on a general Intelligence factor, with only Information and Vocabulary evidencing higher loadings (Crawford *et al.*, 1989).

Investigations of the genetic basis of the relationship between word reading ability and IQ have most often assessed American samples using the Peabody Individual Achievement Test (PIAT). The PIAT is similar to the SGWRT in assessing reading of regular and irregular words and is not contextualized. Estimates from these studies have suggested that between 21% to 80% of the phenotypic correlation between IQ and word reading ability (regular and irregular words considered together) is due to genetic factors. The majority of these estimates have tended from the middle to upper end of the range (Alarcón and DeFries, 1997; Brooks, Fulker, and DeFries, 1990; Cardon *et al.*, 1990; Wadsworth *et al.*, 1994). There has been no previous research that has looked at the genetic basis of the relationship between IQ and irregular word reading specifically, either contextualized or in isolation. This is of interest because irregular words may only be read accurately if the word is known to the respondent, whereas regular words may be accurately pronounced if the word is familiar or by the application of gpc rules. Thus, the

accurate reading of regular words may invoke two distinct processes (whole word recognition when a regular word is known and gpc when a regular word is unknown), whereas accurate reading of irregular words relies solely on whole word recognition.

Genetic studies investigating the relationship between reading and Full-Scale IQ have indicated that word reading performance is influenced by both a genetic Reading factor and a genetic Intelligence factor (Alarcón and DeFries, 1990; Brooks, Fulker, and DeFries, 1990). In a study that delineated Verbal and Performance IQ, Cardon *et al.* (1990) found a stronger genetic correlation between word reading (PIAT) and Verbal IQ (0.96) than between word reading and Performance IQ (0.45). They further reported that 78% and 67% of the phenotypic correlations between word reading and Verbal and Performance IQ, respectively, were due to hereditary influences. Consistent with theories of intelligence that distinguish between verbal and nonverbal skills, evidence that reading is more strongly related to verbal than nonverbal measures, and that reading is influenced by genetic Intelligence and Reading factors, the Bifactor model to be tested here will stipulate that the reading measures will load onto a genetic Reading factor, a genetic Verbal factor, and a genetic General factor.

How much of the variance in NART or CCRT performance can be attributed to genetic influences has not been previously examined. However, there have been a small number of studies looking at the heritability of irregular word reading using other tests. In an early study, Stevenson (1991) did not find a significant genetic influence on irregular word reading. In contrast, Castles *et al.* (1999) reported a significant heritability of 0.31 for isolated irregular word reading in a group defined as having surface dyslexia (specific difficulty with reading irregular words), with common environment accounting for 63% of the variance. Other research has used tasks assessing orthographic coding skills such as distinguishing true words from pseudo-homophones (e.g., rain, rane), which are presumed to have cognitive processes in common with those used for irregular word reading. However, it should be noted that performance on the CCRT requires not only satisfactory orthographic skills but knowledge of increasingly uncommon words, which is the presumed basis of its relationship with IQ. Though an initial study (Olson *et al.*, 1989) found no significant genetic influence on orthographic coding, subsequent research with larger samples suggested a heritability of approximately 0.6 (Gayan *et al.*, 1997; Olson, Forsberg, and Wise, 1994).

The purpose of this paper is to examine the suitability of models in which the genetic factors that influence Verbal IQ subtest performance also influence performance on reading measures used explicitly by neuropsychologists to estimate IQ. Comparisons of fit for Bifactor, Thurstone, and Spearman models will be conducted. It is hypothesized that a Bifactor model (i.e., with a General and group factors) will provide the best fit of the data based on current theorizing and previous empirical studies. As the current sample is composed of adolescents, it is also expected that in accord with previous research investigating the structure of intellectual abilities in children and adolescents, a single common environmental factor will fit the common environment structure (e.g., Luciano *et al.*, 2003). Evidence that the same genetic and environmental factors contribute to variance in neuropsychological reading and IQ subtest performance would add theoretical weight to the use of reading measures as proxies for IQ and, in particular, Verbal IQ.

METHODS

Participants

Data collection is proceeding as part of a continuing study of cognition, the Brisbane Memory, Attention and Problem Solving (MAPS) twin study. Current results are reported for 500 twin pairs (225MZ, 275DZ). The majority of twins were drawn from a study of naevi susceptibility conducted through the Queensland Institute of Medical Research (QIMR), with additional twins ascertained through mail-outs to schools within southeast Queensland. For same-sex twin pairs, zygosity was determined by blood groups (typed by ABO, Rh [CcDEe] and MNSs by the Red Cross Blood Transfusion Service) and DNA typing of nine independent DNA microsatellite polymorphisms ($PIC > 0.7$) by polymerase chain reaction (PCR), multiplexed both by PCR and electrophoresis (ABI Profiler System). Probability of error in zygosity assignment is less than 10^{-3} . Each participant was tested as close as possible to their 16th birthday with most in their penultimate year of secondary education. Ages ranged from 15.42 years to 18.16 years with a mean of 16.23 years and a standard deviation of 0.33 years. Participants were excluded if there was a parental report of either twin having a history of significant head injury, neurological or psychiatric illness, substance abuse or dependence, or current use of medication with known effects on the central nervous system (not including short-term treatment). Participants had normal or

corrected-to-normal vision ($>6/12$ Snellen equivalent). Informed parental (twins <18 years old) or participant (twins >17 years old) consent was obtained. Payment was made to participants based on task performance for a component of the overall study (Delayed Response Task) (see Luciano *et al.*, 2001).

Measures

Multidimensional Aptitude Battery

The MAB is a multiple-choice test of general intelligence and is well-suited for projects using large numbers of participants (Vernon, 2000). It was based on the WAIS-R and yields Full Scale IQ, Verbal IQ, and Performance IQ scores. The scales correlate well with their WAIS-R counterparts, with Jackson (1984) reporting correlations 0.94 for Verbal, 0.79 for Performance, and 0.91 for Full Scale IQ between the MAB and WAIS-R. The scales also have acceptable test-retest reliabilities being equal to or above 0.95 for the Verbal, Performance, and Full Scales as reported by Jackson (1984). Slightly lower, albeit still satisfactory test-retest reliability coefficients (Verbal IQ, 0.88; Performance IQ, 0.84; Full Scale IQ, 0.89) were reported by Luciano (2001) using test-retest data for 50 twin pairs drawn from the same overall sample used here.

The Battery is composed of 10 subtests with the Verbal and Performance scales composed of five tests each. For this study, as with previous research from this laboratory (Luciano *et al.*, 2003), five tests from the MAB were administered. The three verbal tests were Vocabulary, Information, and Arithmetic. The two Performance tests were Spatial and Object Assembly. The Vocabulary and Information tests are similar to WAIS-R tests, assessing word knowledge and fund of information across a broad range of domains. The Arithmetic test gives mathematical problems in words that may be solved using pencil and paper. This differs from WAIS-R arithmetic, which requires retention in memory of the problem as well as ongoing working and thus loads on a working memory factor. The Spatial test assesses mental rotation, whereas Object Assembly requires imagined reconstitution of disconnected parts of common objects. For each subtest there was a 7-minute time limit. For the three verbal tests, all participants were tested using a computerized version of the MAB. For the two MAB Performance tests, 326 twin pairs completed a pencil and paper version, with the remaining pairs using a computerised version. The Digit Symbol-Substitution (DS-S) test from the WAIS-R was also administered.

Cambridge Contextual Reading Test

The CCRT is composed of 36 short English passages containing one or two sentences in which are embedded the 58 target words. The number of target words embedded in a passage ranges from one to six. Participants were asked to read aloud, at normal tempo, the sentences that were presented on printed sheets. Participants were unaware of which words were being assessed for the CCRT. There was no termination rule based on errors so that each participant read all words on the list. Two trained research assistants conducted the assessments, marking which CCRT words were pronounced incorrectly on a separate scoring sheet as well as making note of any non-CCRT reading errors or omissions (data not used in this study). The CCRT score was the number of words correctly pronounced. All reading responses were audio recorded with participant's consent. Periodic assessment of the audio taped responses and scoring found the marking of the CCRT by the research assistants to be highly accurate.

Schonell Graded Worded Reading Test

The SGWRT (Schonell and Schonell, 1960) is composed of 100 words arranged in 20 rows of five words each in ascending order of difficulty. Participants began reading at the 11th row. If all words on that row were read correctly, participants were given credit for all preceding words and read forward through the rows until there were 10 consecutive errors or the test was completed. If an error was made on the 11th row, the preceding row was read and so on until no errors were made on a row. Once an error-free row had been read, participants read forward through the rows until either there were 10 consecutive errors or the test was completed. Similar marking procedures, audio recording provisions, and accuracy of marking checks as used for the CCRT were conducted with the SGWRT.

Procedure

Assessment of word reading and general cognitive ability formed part of a larger testing schedule for the twins that also involved behavioral and electrophysiological assessment of working memory and information processing speed using computerized tasks. Testing involved two sessions designated A and B with one twin undertaking session A while the co-twin undertook session B. Order of testing session was counter-balanced according to birth order. Upon completion of the first session, there was a brief break followed by each twin sitting the complementary session. The

CCRT, SGWRT, MAB, and DS-S were administered in session A. Total testing time for sessions A and B combined was between $3\frac{1}{2}$ and 4 hours. All assessment was done in the morning. After testing, there was gathering of information regarding physical characteristics (e.g., height, eye color) and collection of blood.

Statistical Analyses

Data Screening and Transformation

Reading tests and DS-S entered the testing protocol at a later stage than the MAB IQ subtests. As such, there were fewer participants for these measures. In addition, there were a small number of missing data due to computer or experimenter error. Data were analyzed using the MX package. All data were screened for normality and univariate and multivariate outliers. Ten univariate outliers were excluded from analyses according to a Z score criterion of 3.29 (Tabachnick and Fidell, 1996). No multivariate outliers were detected. A negative skew was apparent in the SGWRT scores. These were reflected and a square root transformation taken that yielded satisfactory normality and kurtosis. This resulted in lower scores indicating better test performance, and as such, negative correlations were obtained with the transformed SGWRT. This negative sign is ignored in the reporting to avoid confusion.

Testing Equality of Means, Variances, and Covariances According to Zygosity, Sex, and Education

To establish regularity in sampling and measurement, the equality of means and variances according to birth order and zygosity was tested. A series of nested models with increasingly restrictive equality constraints upon means and variances was assessed, with each model being compared to a preceding less restrictive model using a χ^2 statistic. A significant change in fit was interpreted as a given constraint hypothesis being unlikely to be true.

Equality of means according to sex and education were also assessed. A male mean was specified for each variable as a deviation from the female mean. A χ^2 statistic was then used to determine whether the deviation parameter could be dropped without a significant change in fit. There were some differences in the duration of formal education experienced between twin pairs (students may leave school at 15 years old—most co-twins within a pair had received the same duration of schooling). As the amount of formal education undertaken was expected to affect test (especially verbal)

performance, a weighted regression parameter that tested effects for time spent in formal education was also specified.

Equality of covariance between monozygotic female (MZF) and monozygotic male (MZM) and between dizygotic female (DZF) and dizygotic male (DZM) was assessed, as was equality of covariance between same-sex DZ twin pairs and opposite-sex DZ twin pairs (DZO) to assess differences in the magnitude of genetic effects and whether different genes were being expressed according to sex.

Testing Equality of Correlations Between Reading and IQ Subtests for Able and Less Able Readers

Because the MAB is a written test requiring satisfactory reading skills to both understand the test instructions and complete test items, it was possible that less able readers would be disadvantaged on the MAB. This could result in a confounding of the phenotypic correlations between the reading tests and the intelligence subtests as well as influencing the estimate of genetic correlations for these measures. However, if the reading tests were equally predictive of IQ subtest scores for both able and less able readers, then it suggests that reading skill per se did not influence IQ subtest performance.

Due to the sample being composed of twins, the analyses were conducted in MX in the following way, which takes into account the relatedness of the participants. Designations as either able (> -1 SD) or less able (< -1 SD) readers on both the CCRT and SGWRT were made, and the equality of correlations between able and less able readers for both reading tests and the IQ subtests were examined according to zygosity and birth order. Twin pairs were divided into six concordance groups, being MZ concordant able readers (both able readers), MZ concordant less able readers, MZ discordant (one able reader, one less able reader), and likewise for DZ twins. For both reading tests, there were six bivariate regressions assessed (the six IQ subtests). Within each bivariate regression model, there were 12 correlations based on the crossing of birth order (twin 1, twin 2) with the six concordance groups. In the initial modeling, the 12 correlations within each regression model were free to vary with the subsequent constraint of all 12 correlations being equal. A significant change in χ^2 between the unrestricted and restricted models was interpreted as indicating that reading score was not equally predictive of IQ subtest score for able and less able readers. As six regression models for each reading measure were assessed, a Bonferroni correction was applied.

General Statistical Approach and Model Fitting

Nonadditive effects were not considered as they are conflated with common environment, and initial inspection of MZ and DZ correlations suggested common environment rather than dominance influences were involved (DZ correlation greater than half MZ correlation) (Neale and Cardon, 1992). A model with a Cholesky decomposition into A, C, and E was used as a baseline against which competing models were assessed. Previous research (Eaves, Heath, and Martin, 1984; Luciano *et al.*, 2003) suggested that a single C factor should be present, and for this reason, only models of the form ACE were assessed with A and C structures fitted simultaneously.

There are two principal variants of Bifactor models. These are hierarchical models in which the influence of a General factor is via the group factors, which in turn directly influence the variables, and nested models in which the General factor acts independently and directly on the variables. Though it is clear that these two forms of models are not identical and posit different structures by which a General factor exerts its influence, it has been noted that the general interpretation of each of the models is the same (Gustafsson and Balke, 1993) and that both are equally satisfactory for estimating the influence of a General factor (Jensen and Weng, 1994). Empirical assessment of the identification of a hierarchical bifactor model with a somewhat complex factor structure (reading measures loading onto General, Verbal, and Reading factors) indicated that additional constraints beyond those necessary for identification of a nested Bifactor model would be required (constraints among the loadings of the group factors on the General factor) making interpretation of a hierarchical model restrictive and somewhat uninformative. For this reason and due to the same general interpretation of nested and hierarchical models (Gustafsson and Balke, 1993), a nested Bifactor model was selected to be assessed in comparison with Spearman and Thurstone models.

In the nested Bifactor model, the two reading measures were included in both the genetic Verbal factor and the genetic Reading factor. Constraints were required to permit model identification. Within the genetic Reading factor, the pathways to the two reading measures were constrained to be equal. Additionally, DS-S was not incorporated in the genetic Performance factor. Though the invoking of constraints is not considered ideal, there was theoretical justification for choosing this constraint above other possibilities given that DS-S assesses processing speed (Kaufman and Lichtenberger, 1999), whereas Object and Spatial

assess visuospatial manipulation of plane forms. Furthermore, DS-S was drawn from a different test battery. Moreover, it was not considered a significant limitation in interpreting the model as it was the relationship of the reading measures with the subtests and in particular the Verbal subtests that was of primary interest. The pathways from the genetic Performance factor to Object Assembly and Spatial were also constrained to be equal. For the Thurstone model, the reading tests also loaded on both a Verbal and Reading factor.

RESULTS

Testing Equality of Means, Variances, and Covariances According to Zygosity, Sex, and Education

For all tests of equality of means, variances, and covariances, Bonferroni corrections were applied to an initial significance level of 0.05. No significant differences were found in means and variances according to zygosity for any of the variables. Table I shows the means and standard deviations for males and females. No significant sex differences in means were found for either of the reading measures or Vocabulary. Of the remaining tests, males had significantly higher mean scores for Information, Arithmetic, Spatial, and Object Assembly. Females obtained a higher mean for DS-S. As such, a deviation for males was incorporated in means modeling for these variables. For duration of formal education, only the parameter for Object Assembly could be dropped without a significant loss of fit. For all other variables, greater time spent in formal education was associated with significantly higher test scores. For these variables, regressions for duration of formal education were included in the means modeling. For all variables, correlations for MZM and MZF did not differ significantly. Likewise, there were no significant differences between DZM, DZF, and opposite sex pairs. Thus, sex limitation was not incorporated into the modeling. The stronger correlations for MZ than DZ indicated that genetic effects were relevant to variance in the measures.

Testing Equality of Covariances for Able and Less Able Readers

Of the 12 regressions assessing the equality of prediction between neuropsychological reading measures and IQ subtests, only for the SGWRT-Vocabulary and SGWRT-Object Assembly regressions was there a significant loss of fit. Thus, it did not appear that

Table I. Means and Standard Deviations for Males and Females and Correlations According to Zygosity

Tests	Males		Females		Correlations				
					MZF	MZM	DZF	DZM	DZO
					Twin pairs (range of <i>n</i>)				
	<i>n</i>	Mean (SD)	<i>n</i>	Mean (SD)	112–117	90–108	64–69	59–61	129–137
Vocabulary	505	17.32 (4.87)	466	17.06 (4.95)	0.64	0.53	0.43	0.28	0.37
Information	507	19.95 (5.09)	469	21.04 (6.05)	0.78	0.72	0.45	0.48	0.51
Arithmetic	508	11.88 (2.69)	470	12.39 (2.94)	0.61	0.65	0.29	0.26	0.39
CCRT	481	30.98 (7.65)	470	29.80 (8.11)	0.78	0.81	0.53	0.54	0.50
SGWRT	481	85.00 (10.83)	430	83.00 (13.17)	0.78	0.82	0.60	0.40	0.48
Spatial	508	28.75 (9.49)	470	31.63 (8.66)	0.61	0.62	0.14	0.44	0.28
Object	509	12.50 (3.76)	475	13.15 (3.69)	0.63	0.64	0.49	0.34	0.34
Digit Symbol	488	63.81 (9.89)	442	54.85 (10.30)	0.65	0.67	0.26	0.43	0.36

Note: MZF, monozygotic females; MZM, monozygotic males; DZF, dizygotic females; DZM, dizygotic males; DZO, dizygotic opposite-sex; CCRT, Cambridge Contextual Reading Test; SGWRT, Schonell Graded Word Reading Test.

differences in ability to read task instructions or items of itself markedly affected IQ subtest performance. However, it should be noted that the relatively small sample size of less able readers may have resulted in limited power to detect a significant difference.

Phenotypic Analyses

Phenotypic correlations are shown in Table II. The strongest correlations within the matrix were observed among the verbal measures. The reading measures correlated most strongly with each other and with Information and Vocabulary. The lowest phenotypic correlations were between DS-S and the two reading measures and Vocabulary.

Genetic Analyses and Model Fitting

Table III shows the various models that were fit to the data. A Cholesky ACE model was specified as a base against which theoretically driven models were

assessed. A nested Bifactor model with a genetic General factor, three genetic group factors (Verbal, Performance, Reading) and specific genetic factors, in conjunction with a general common environment factor was fitted first. Table III shows that the nested bifactor model provided a satisfactory fit in comparison with the Cholesky decomposition. Neither the Thurstone model nor the Spearman model provided a satisfactory fit in comparison with the nested Bifactor model. The influence of unique environmental effects was left as a Cholesky decomposition in the preferred nested Bifactor model.

The values of path coefficients from the nested Bifactor model are shown in Figure 1. For the sake of clarity, only parameter estimates above 0.1 are shown for the E structure. The genetic General, genetic Verbal, and genetic Reading factors account for 13%, 17%, and 21% of the variance for the CCRT and 20%, 19%, and 21% of the variance for the SGWRT, respectively. Relatively small genetic specific influences (less than 11%) were present for the majority of subtests.

Table II. Phenotypic Correlations Among the Reading Tests and IQ Subtests

Tests	Vocabulary	Information	Arithmetic	CCRT	SGWRT	Spatial	Object
Information	0.69						
Arithmetic	0.44	0.54					
CCRT	0.60	0.65	0.44				
SGWRT	0.62	0.65	0.48	0.85			
Spatial	0.28	0.36	0.42	0.29	0.32		
Object	0.33	0.42	0.44	0.32	0.34	0.56	
Digit Symbol	0.26	0.33	0.36	0.23	0.27	0.34	0.34

Note: CCRT, Cambridge Contextual Reading Test; SGWRT, Schonell Graded Word Reading Test.

Table III. Competing Models for Factor Structure of Genetic, Common Environment, and Unique Environment Influences

Model	−2LL	df	vs	Δ−2LL	Δ−df	p value
1. ACE Cholesky Decomposition	40923.66	7510				
2. Nested Bifactor model	40957.07	7551	1	33.41	41	0.79
3. Thurstone model	41005.52	7557	2	48.45	6	<0.001
4. Spearman model	41026.79	7558	2	69.72	7	<0.001

Note: The best fitting model is highlighted in bold. A value below 0.05 represents a significant change in χ^2 and thus a significant loss of fit.

However, 32% of the variance for DS-S is explained by a genetic specific influence. This large value was due to DS-S not being included within the Performance factor. The common environment factor had a statistically significant influence on all measures except DS-S.

Somewhat stronger loadings on this factor for Verbal IQ subtests and reading measures are evident ranging from explaining 9% of the variance for Arithmetic to 25% of the variance for CCRT. For the Performance measures, the common environment factor accounted

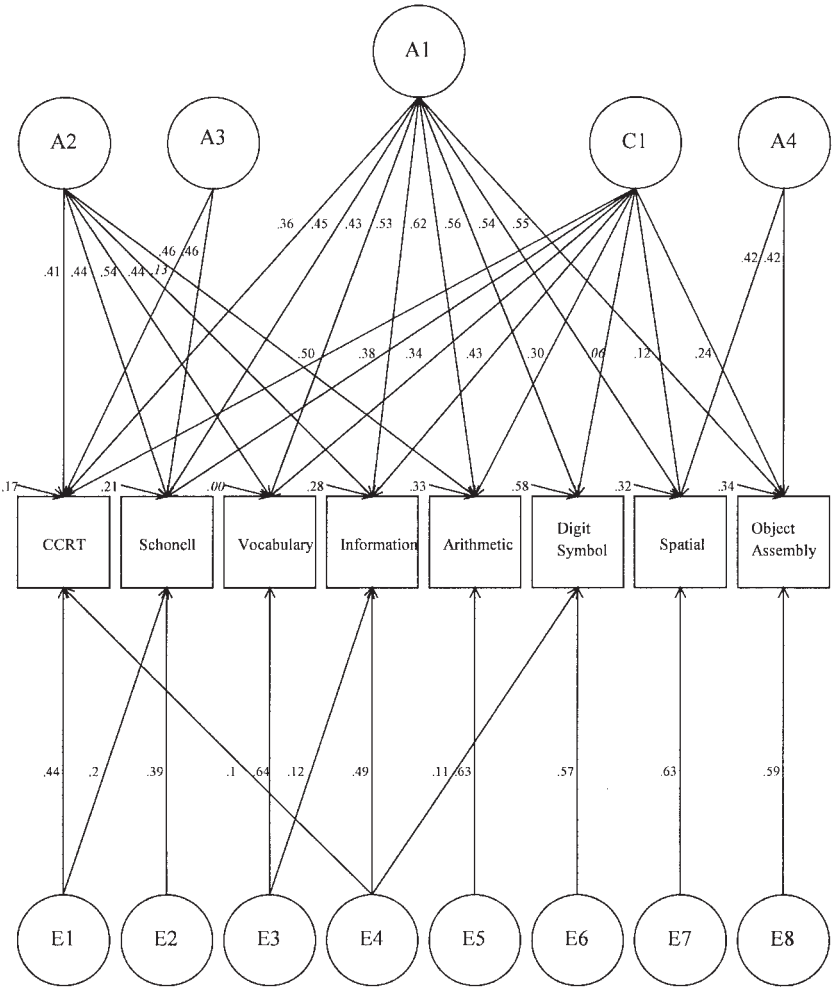


Fig. 1. Path diagram depicting the additive genetic, common environmental, and unique environmental factor loadings for the CCRT, SGWRT, and the IQ subtests. A1, genetic General factor; A2, genetic Verbal factor; A3, genetic Reading factor; A4, genetic Performance factor; C1, Common environment factor; E1 to E8, unique environment factors. Arrows above the variables = specific genetic effects. Estimates in italics may be dropped from the model.

Table IV. Genetic Correlations Among the Reading Tests and IQ Subtests

	Vocabulary	Information	Arithmetic	CCRT	SGWRT	Spatial	Object
Information	0.91						
Arithmetic	0.68	0.72					
CCRT	0.75	0.68	0.53				
SGWRT	0.77	0.72	0.58	0.94			
Spatial	0.45	0.51	0.62	0.36	0.40		
Object	0.44	0.51	0.62	0.35	0.40	0.81	
Digit Symbol	0.44	0.50	0.61	0.35	0.39	0.50	0.50

Note: CCRT, Cambridge Contextual Reading Test; SGWRT, Schonell Graded Word Reading Test.

for only between 1% to 6% of the variance. Unique environment effects are principally specific to each of the measures and range from 14% for the SGWRT to 41% for Vocabulary.

Genetic correlations are shown in Table IV. Reference to Figure 1 makes clear why in this model, strong genetic correlations are evident between the CCRT and the Vocabulary and Information and likewise between the SGWRT and Vocabulary and Information. It is because there is relatively little genetic variance in these measures that does not overlap due to common factors. For instance, the only factorial source of genetic variance in the CCRT that is not shared with Vocabulary is through the Reading factor, as both the CCRT and Vocabulary load onto the genetic General and genetic Verbal factors, and there is limited specific influence for either. The correlations indicate that there is approximately 46% to 56% overlap in genetic effects influencing neuropsychological reading test performance and Vocabulary and Information assessed by the MAB. A very strong genetic correlation was also found between the two reading measures, with almost complete overlap in genetic effects with only small specific genetic effects influencing the reading measures.

Table V reports the proportions of observable covariation attributable to genetic effects and the heritabilities (diagonal) for the individual test performances. Approximately between half and two-thirds of population variance for these cognitive tests are attributable to genetic influences. Proportions of covariances between reading scores and IQ subtests attributable to genetic influence ranged from 58% (CCRT-Information) to 94% (SGWRT-DS-S). Thus, for all relationships among neuropsychological reading tests and IQ subtests, no less than half of the covariation was due to genetic influences.

DISCUSSION

The results support a nested Bifactor model with one General, three group, and specific factors of genetic influence in preference to either a Thurstone or Spearman model. The model implies that word pronunciation on reading tests used in neuropsychological assessment is affected by a set of genes that influence general cognitive ability and that the phenotypic relationship between these reading measures and IQ

Table V. Proportions of Phenotypic Correlations Among the Reading Tests and IQ Subtests Accounted for by Genetic Factors

	Vocabulary	Information	Arithmetic	CCRT	SGWRT	Spatial	Object	Digit Symbol
Vocabulary	0.47							
Information	0.68	0.56						
Arithmetic	0.77	0.72	0.51					
CCRT	0.63	0.58	0.63	0.54				
SGWRT	0.70	0.67	0.69	0.65	0.65			
Spatial	0.85	0.80	0.80	0.67	0.76	0.58		
Object	0.70	0.68	0.77	0.61	0.72	0.85	0.59	
Digit Symbol	0.93	0.90	0.98	0.90	0.94	0.89	0.92	0.65

Note: The diagonal represents the heritabilities for the individual subtests. CCRT, Cambridge Contextual Reading Test; SGWRT, Schonell Graded Word Reading Test.

subtests is partly due to the effect of these genes. This is consistent with the phenotypic factor analysis conducted by Crawford *et al.* (1989) that showed NART performance loading onto a general Intelligence factor.

There is also evidence that an independent set of genes (Verbal factor) influences performance on the reading tests as well as Vocabulary and Information subtests. This influence may be due to performance on these tests having common reliance on language facility including verbal comprehension and encoding, retention, and access to long-term declarative knowledge (Kaufman and Lichtenberger, 1999). The absence of a significant loading for Arithmetic on this genetic factor is presumably due to its substantial dependence on mathematical reasoning rather than language skills despite the problems being stated in words. Evidence of a separate factor for word reading is in accord with previous findings (Alarcón and DeFries, 1997; Brooks, Fulker, and DeFries, 1990) and indicates that word pronunciation on reading tests used by neuropsychologists is influenced by a genetic factor that is particular to reading skill independent of genetic influences on IQ.

In general, previous research has reported phenotypic correlations between the NART or CCRT and Full Scale IQ, Verbal IQ, and Performance IQ scores rather than with subtest performance. The phenotypic correlations reported here between the reading measures and IQ subtests are somewhat lower than those typically reported between IQ scores and reading measures (Crawford *et al.*, 1989; Griffin *et al.*, 2001; Lastine-Sobecks, Jackson, and Paolo, 1998). This may be partly due to the lower reliability of IQ subtests compared to summary IQ scores (Wechsler, 1981). Additionally, studies previously reporting NART/CCRT and IQ correlations have typically used samples composed of older participants than reported here (Beardsall, 1998; Crawford *et al.*, 1989; Griffin *et al.*, 2001; Lastine-Sobecks, Jackson, and Paolo, 1998).

The genetic correlations between the reading measures and Verbal IQ subtests are weaker than the genetic correlation of 0.96 between reading recognition and Verbal IQ reported by Cardon *et al.* (1990). The disparity is due to the inclusion of a genetic Reading factor in the preferred model in this study which contributes significantly to the heritability of word recognition independent of shared genetic effects between reading and Verbal IQ subtests. Nevertheless, the genetic correlations reported here are substantial. The genetic correlations between the reading and Performance IQ subtests found here are comparable with the estimate of Cardon *et al.* (1990) of a genetic correlation of 0.45 between reading recognition and Performance IQ.

There was evidence that genetic influences contribute a larger proportion of the covariance between Performance IQ tests and reading measures than for Verbal IQ tests and reading measures. This is due to the common environment factor having a somewhat stronger influence on the correlations among Verbal IQ tests and reading measures, possibly through formal education or the intellectual environment within the home, while having less effect on the correlations among Performance tests.

This study is the first to report heritability for the CCRT. Though early studies of irregular word reading did not report significant genetic influence, the estimate here of 0.54 is consistent with estimates from later studies of orthographic coding that suggested a heritability of approximately 0.6 (Gayan *et al.*, 1997; Olson, Forsberg, and Wise, 1994). However, these estimates may not be directly comparable as the CCRT requires knowledge of uncommon words in addition to basic orthographic skills. The estimate obtained here is larger than that obtained by Castles *et al.* (1999) in their irregular word reading task. However, in that research, the reported heritability was for readers defined as having surface dyslexia, whereas a much broader range of reading ability was sampled here. Additionally, because the CCRT was specifically designed to correlate strongly with IQ, the larger heritability estimate may reflect a greater influence of general cognitive ability, which is known to be highly heritable (Jensen, 1998).

Though it has been argued that the classic twin method may overestimate heritabilities (Grayson, 1988; Hopper, 2000), it has also been suggested that the effects of assortative mating may inflate the estimates of common environmental influences at the expense of genetic effects. This is of relevance to this study as there is previous evidence that covariation among a variety of cognitive abilities may be influenced by a common environmental factor (Eaves, Heath, and Martin, 1984; Petrill *et al.*, 1996).

Eaves, Heath, and Martin (1984) originally proposed that a genetic General factor expressible as a linear function of a single common environmental factor subsuming all task performances suggests assortative mating. The nonsignificant loading for DS-S precluded this hypothesis being tested in this study. However, it also seems reasonable to consider the possibility that the common environment factor actually reflects the influence of the common environment rather than being an artifact of assortative mating, with the higher loadings for the verbal tests being consistent with the view that verbal skills are more shaped by environmental

influences than nonverbal skills, which are considered more fluid (Jensen, 1998).

There is evidence of unique environmental factors influencing multiple test performances, which may reflect correlated test error as well as unique experiences influencing multiple abilities. However, it is important to note that there is very great power to detect E effects, and thus unique environmental influences on covariation among the measures, while statistically significant, are of little substantive interest.

There was little difference between the two reading assessments in terms of the magnitude of genetic and environmental effects on the measures themselves or their relationships with other measures. This was probably largely due to the more difficult items on the SGWRT being irregular words, and it was with these more difficult items that variation in performance among the participants occurred. It should also be noted that despite normalization, the variance for the SGWRT was constricted, and as such the observed correlations for this measure may be depressed. In conjunction with researchers from Macquarie University, future research will be directed at examining genetic and environmental influences on variance and covariance among purer measures of basic reading skills such as single (regular and irregular) word and nonword reading and IQ and how these relationships differ in their genetic and environmental architectures from those found between neuropsychological reading tests and IQ subtests.

In summary, the current paper has suggested that in adolescents ranging from 15 to 18 years of age, the correlation between reading measures used in neuropsychological assessment and IQ subtests is substantially attributable to genetic factors through both a genetic General factor and a genetic Verbal factor. There is also a significant contribution to covariance particularly with Verbal IQ measures and neuropsychological reading measures from a common environment factor. The findings suggest that abilities measured by neuropsychological reading assessments and, in particular, Verbal IQ subtests are influenced to a substantial extent by the same biological and experiential underpinnings and adds theoretical weight to the use of reading measures as proxies for measuring intellectual ability.

ACKNOWLEDGMENTS

We greatly appreciate the support from the twins and their parents and their willingness to participate in this study. We are grateful to Marlene Grace, Ann Eldridge, and Kathleen Moore for recruitment of twin pairs and

data collection. This project is supported by grants from the Australian Research Council (Grants A 79600334, A 79906588, and A 79801419) and the Human Frontier Science Program (Grant RG0154/1998B).

REFERENCES

- Alarcón, M., and DeFries, J. C. (1997). Reading performance and general cognitive ability in twins with reading difficulties and control pairs. *Personality and Individual Differences* **22**:793–803.
- Beardsall, L. (1998). Development of the Cambridge Contextual Reading Test for improving the estimation of premorbid verbal intelligence in older persons with dementia. *Br. J. Clin. Psychol.* **37**:229–240.
- Beardsall, L., and Huppert, F. A. (1994). Improvement in NART word reading in demented and normal older persons using the Cambridge Contextual Reading Test. *J. Clin. Exp. Neuropsychol.* **16**:232–242.
- Brooks, A., Fulker, D. W., and DeFries, J. C. (1990). Reading performance and general cognitive ability: a multivariate genetic analysis of twin data. *Personality and Individual Differences* **11**:141–146.
- Cardon, L. R., DiLalla, L. F., Plomin, R., DeFries, J. C., and Fulker, D. W. (1990). Genetic correlations between reading performance and IQ in the Colorado Adoption Project. *Intelligence* **14**:245–257.
- Castles, A., Datta, H., Gayan, J., and Olson, R. K. (1999). Varieties of developmental reading disorder: genetic and environmental influences. *J. Exp. Child Psychol.* **72**:73–94.
- Crawford, J. R., Parker, D. M., Stewart, L. E., Besson, J. A. O., and DeLacey, G. (1989). Prediction of WAIS IQ with the National Adult Reading Test: cross-validation and extension. *Br. J. Clin. Psychol.* **28**:267–273.
- Crawford, J. R., Stewart, L. E., Cochrane, R., Parker, D. M., and Besson, J. A. O. (1989). Construct validity of the National Adult Reading Test: a factor analytic study. *Personality and Individual Differences* **10**:585–587.
- Eaves, L. J., Heath, A. C., and Martin, N. G. (1984). A note on the generalized effects of assortative mating. *Behav. Genet.* **14**:371–376.
- Franzen, M. D., Burgess, E. J., and Smith-Seemiller, L. (1997). Methods of estimating premorbid functioning. *Arch. Clin. Psychol.* **12**:711–738.
- Gayan, J., Datta, H. E., Castles, A., and Olson, R. K. (1997). The etiology of group deficits in word decoding across levels of decoding and orthographic coding. Paper presented at the meeting of the Society for the Scientific Study of Reading, Chicago, 23 March 1997.
- Grayson, D. A. (1988). Twins reared together: minimizing shared environmental effects. *Behav. Genet.* **19**:593–604.
- Griffin, S. L., Rivera-Mindt, M., Rankin, E. J., Ritchie, A. J., and Scott, J. G. (2001). Estimating premorbid intelligence: comparison of traditional and contemporary methods across the intelligence continuum. *Arch. Clin. Neuropsychol.* **17**:497–507.
- Gustafsson, J.-E., and Balke, G. (1993). General and specific abilities as predictors of school achievement. *Multivariate Behavioural Research* **28**:407–434.
- Hopper, J. L. (2000). Why 'common environmental effects' are so uncommon in the literature. In T. D. Spector, H. Snieder, and A. J. MacGregor (eds.), *Advances in twin and sib-pair analysis*, London: Greenwich Medical Media, pp. 151–165.
- Jackson, D. N. (1984). *Multidimensional Aptitude Battery: Manual*. Ontario, Canada: Research Psychologists Press, Inc.
- Jensen, A. R. (1998). *The g factor: the science of mental ability*. Westport, CT: Praeger.
- Jensen, A. R., and Weng, L.-J. (1994). What is a good g. *Intelligence* **18**:231–258.

- Kaufman, A. S., and Lichtenberger, E. O. (1999). *Essentials of WAIS-III assessment*. New York: John Wiley & Sons, Inc.
- Lastine-Sobecks, J. L., Jackson, S. T., and Paolo, A. M. (1998). Identifying the pronunciation of irregularly spelled words: relation to verbal IQ. *Clin. Neuropsychol.* **12**:189–192.
- Luciano, M. (2002). *An information processing approach to the genetic study of psychometric intelligence: Links between processing speed, working memory and IQ*. Ph.D. Thesis, University of Queensland, Queensland, Australia.
- Luciano, M., Wright, M. J., Smith, G. A., Geffen, G. M., Geffen, L. B., and Martin, N. G. (2001). Genetic covariance among measures of information processing speed, working memory and IQ. *Behav. Genet.* **31**:581–592.
- Luciano, M., Wright, M. J., Geffen, G. M., Geffen, L. B., Smith, G. A., and Martin, N. G. (2003). A genetic two-factor model of the covariation among a subset of Multidimensional Aptitude Battery and WAIS-R subtests. *Intelligence* **31**:589–605.
- Neale, M. C., and Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Norwell, MA: Kluwer Academic Publishers.
- Nelson, H. E. (1982). *National Adult Reading Test*. Berkshire: NFER-Nelson Publishing Company.
- Olson, R. K., Forsberg, H., and Wise, B. (1994). Genes, environment and the development of orthographic skill. In V. W. Berninger (ed.), *The varieties of orthographic knowledge I: theoretical and developmental issues*. Dordrecht, The Netherlands: Kluwer Academic.
- Olson, R. K., Wise, B., Conners, F., Rack, J., and Fulker, D. (1989). Specific deficits in component reading and language skills: genetic and environmental influences. *J. Learn. Disabil.* **22**:339–348.
- Petrill, S. A. (1997). Molarity versus modularity of cognitive functioning? A behavioral genetic perspective. *Current Directions in Psychological Science* **6**:96–99.
- Petrill, S. A., Luo, D., Thompson, L. A., and Detterman, D. K. (1996). The independent prediction of general intelligence by elementary cognitive tasks: genetic and environmental influences. *Behav. Genet.* **26**:135–147.
- Plomin, R., and Craig, I. (2001). Genetics, environment and cognitive abilities: review and work in progress towards a genome scan for quantitative trait locus associations using DNA pooling. *Br. J. Psychiatry* **178**(Suppl. 40):s41–s48.
- Rijsdijk, F. V., Vernon, P. A., and Boomsma, D. (1998). The genetic basis of the relation between speed-of-information processing and IQ. *Behav. Brain Res.* **95**:77–84.
- Rijsdijk, F. V., Vernon, P. A., and Boomsma, D. (2002). Application of hierarchical genetic models to Raven and WAIS subtests: a Dutch twin study. *Behav. Genet.* **32**:199–210.
- Schonell, F. J., and Schonell, P. E. (1960). *Diagnostic and attainment testing*. Edinburgh: Oliver and Boyd.
- Stevenson, J. (1991). Which aspects of processing text mediate genetic effects. In B. F. Pennington (ed.), *Reading disabilities: genetic and neurological influences*. Dordrecht, The Netherlands: Kluwer Academic Publishers, pp. 31–43.
- Tabachnick, B. G., and Fidell, L. S. (1996). *Using multivariate statistics* (3rd Ed.). New York: Harper Collins.
- Vernon, P. A. (2000). Recent studies of intelligence and personality using Jackson's Multidimensional Aptitude Battery and Personality Research Form. In R. D. Goffin and E. Helmes (eds.), *Problems and solutions in human assessment: honoring Douglas N. Jackson at seventy*. Norwell, MA: Kluwer Academic Publishers, pp. 195–212.
- Wadsworth, S. J., DeFries, J. C., Fulker, D. W., and Plomin, R. (1994). Covariation among measures of cognitive ability and academic achievement in the Colorado adoption project: sibling analysis. *Personality and Individual Differences* **18**:63–73.
- Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale-Revised (WAIS-R)*. San Antonio, TX: The Psychological Corporation.

Edited by Dorret Boomsma