

A behaviour-genetic analysis of orthographic learning, spelling and decoding

Brian Byrne and William L. Coventry

University of New England

Richard K. Olson, Jacqueline Hulslander, Sally Wadsworth,
John C. DeFries, Robin Corley and Erik G. Willcutt

University of Colorado

Stefan Samuelsson

Linköping University and Stavanger University

As part of a longitudinal twin study of literacy and language, we conducted a behaviour-genetic analysis of orthographic learning, spelling and decoding in Grade 2 children (225 identical and 214 fraternal twin pairs) in the United States and Australia. Each variable showed significant genetic and unique environment influences. Multivariate analyses revealed very high genetic correlations among the variables, indicating that the same genes are involved in their aetiology. These genes are partly independent of those contributing to intelligence. A further analysis indicated that the covariation between decoding and orthographic learning is mediated by shared genes rather than by a direct causal path. The authors argue that a learning parameter, most directly assessed by orthographic learning in this study, underlies all three literacy variables. The results are also discussed in relation to Share's self-teaching hypothesis, which may require modification.

In this article we report on a behaviour-genetic analysis of orthographic learning in Grade 2 children. We examine the genetic and environmental influences on it and on its relations with spelling achievement and decoding. The study is part of an ongoing international longitudinal twin study (ILTS) of early literacy and language development (Byrne et al., 2002, 2005, 2006, 2007; Samuelsson et al., 2005, 2007; Willcutt et al., 2007).

Orthographic learning is the process whereby words' spelling patterns are committed to memory such that they can be retrieved automatically rather than decoded from grapheme–phoneme correspondences (Ehri, 2005). The mature reading system is characterised by a high degree of rapid, automatised recognition of printed words. For languages like English, with substantial numbers of 'irregular' words like *yacht* and *women* whose pronunciations are only partly determined by spelling patterns, the ability

to recognise such words is often employed as a sign that the reader has achieved a mature system.

English affords the opportunity to examine orthographic learning ‘on line’ because it offers more than one way to spell most words. In this study we take advantage of that opportunity to study possible genetic and environmental determinants of variation in orthographic learning, something not yet undertaken as far as we are aware. Typically, studies that can detect genetic and environmental influences, such as those using twins as we do here, do not examine learning in progress, focusing instead on ‘static’ measures such as existing spelling ability. In our project, young twin children read short passages containing novel words, such as *laif*, that can be spelled in two or more ways (e.g., *lafe*), and their success in subsequently spelling those words was used as a measure of orthographic learning. These data were then subject to analyses based on the relative within-pair similarities of identical and fraternal twins. Identical, or monozygotic (MZ), twins share all their genes; fraternal, or dizygotic (DZ), twins share half their segregating genes (genes that make people different). Genetic influence is evident to the degree to which within-pair similarity of MZ twins is higher than that of DZ twins. The analysis can also quantify and differentiate between environmental influences that originate in factors that twins share, such as family and school, and those that they do not, such as random accidents, different teachers and so on.

Learning and achievement measures of the ‘same’ process, such as orthographic learning and spelling achievement, do not necessarily fully overlap. For example, in the domain of literacy Byrne, Fielding-Barnsley and Ashley (2000), Byrne, Shankweiler and Hine (in press) and Hindson et al. (2005) have shown that the rate at which preschool children grasp the idea of phoneme identity (that two words can start [or end] with the same phoneme) accounts for variance in subsequent reading achievement on top of that accounted for by the final level of phoneme identity achieved during lessons designed to teach that idea; learning rate and achievement level capture processes different enough to independently predict future reading levels. Thus it is an open question whether orthographic learning, measured by success in acquiring new spelling patterns, and orthographic achievement, measured by a conventional spelling test using previously encountered words, are one and the same. This question extends beyond the phenotypic level, a question that could in principle be answered by evaluating the size of the correlation between orthographic learning and spelling. A behaviour-genetic study can in addition reveal the degree to which the same genetic and environmental factors affect the variables.

In the ILTS, we have already observed considerable overlap in the genetic determination of literacy variables. One metric that captures the overlap is the genetic correlation between variables. Genetic correlations represent the degree to which the genetic influences on one trait are the same as the genetic influences on another trait. The correlation can be low even when the two traits, independently, are highly heritable, or high even when both traits show only modest heritability. In Samuelsson et al. (2005), we report a genetic correlation of .68 between phonological awareness and print knowledge among preschool children. In Byrne et al. (2007), we report genetic correlations in Grade 1 between a combined word and nonword identification measure on the one hand and spelling, rapid naming and reading comprehension on the other of .90, .69 and .97, respectively. (Each of these four Grade 1 variables was substantially heritable, with estimates of genetic variance of .82, .71, .56 and .76, respectively.) However, there is also evidence of genetic independence among our heritable variables. For example, as the US

sample of twins progresses from Kindergarten to Grade 1, there is both genetic continuity (same genes at both school grades) and genetic change (new genes coming into play) in the determination of word identification skill (Byrne et al., 2007). Outside of our study, Bishop, Adams and Norbury (2006) have shown that there are distinct genetic influences on two aspects of developmental language impairment, limitations in phonological short-term memory and verb tense marking. Thus the question of whether orthographic learning is genetically correlated with spelling achievement is an open one – there is evidence for both genetic dependence and independence among variables in the domains of literacy and language.

Genetic overlap seen among language variables might extend more broadly. Plomin and Kovas (2005) surveyed evidence for what they called ‘generalist genes’, ones that affect cognitive abilities well beyond literacy and language to domains such as mathematics (see also Plomin, Kovas & Howarth, 2007). It is clear from their evidence that the well-attested phenotypic correlation among cognitive skills is due, in part, to shared genetic influence, although not all genes are generalists, and environments tend to be ‘specialist’. Thus, a natural question that would face us should there be genetic overlap between our measures of orthographic learning and spelling (and our measure of decoding, as described next) would be how broad the influence of the common genes might be. Would there be substantial genetic correlation, for example, with cognitive skills as distant as those tapped by tests of performance IQ, such as Block Design? If so, we may be picking up the effects of IQ rather than processes specific to reading and spelling. We have data from the preschool phase of our assessments that allow us to address this question.

The method we have elected to use, learning the spelling of novel words, was developed by Share (1999) to test his self-teaching hypothesis (Share, 1995; see also Firth, 1972). This hypothesis states that decoding a novel word supplies an opportunity to learn that word’s spelling pattern. An implication is that variation in decoding should go hand-in-hand with variation in spelling achievement. Research using Share’s experimental paradigm has confirmed that (1) decoding novel words does result in learning their spelling patterns, even after a single exposure, (2) there is variation in children’s rate of orthographic learning and (3) this variation correlates with variation in decoding skill (e.g., Bowey & Muller, 2005; Cunningham, Perry, Stanovich & Share, 2002; Kyte & Johnson, 2006; Nation, Angell & Castles, 2007; Share, 1995, 1999, 2004). Thus, our study also allows us to examine the self-teaching hypothesis in a genetically sensitive design, posing the question, for example, of whether genetic determinants of decoding, which we document below, are the same as those of orthographic learning, assuming that we discover any genetic influence on orthographic learning at all. The answer to this question is relevant to the causal relation between decoding and orthographic learning. The self-teaching hypothesis holds that decoding is the sine qua non of orthographic learning (Share, 1995), but should they both be subject to influence by the same genes, the nature of the causal relation requires closer examination, as we discuss.

The ILTS is a project following literacy and language growth each year from preschool to Grade 2 in samples of twins from Australia, Norway, Sweden and the United States of America. Results are presented in the papers cited above, and are too extensive to fully summarise here. In brief, however, we have shown that before the start of formal education, there are detectable genetic influences on known precursors to literacy growth, including phonological awareness, rapid naming and print awareness. There are also detectable effects of shared (family, preschool) environment, particularly on print

awareness and various measures of spoken language such as vocabulary and morphology. School progress in literacy is substantially affected by genes, with evidence of both genetic continuity and genetic change across development. There are also signs of gene–environment interaction in that genetic effects in the first school year appear to be more pronounced in an educational system that offers more intensive and extensive literacy instruction in that year (Australia vs Scandinavia; Samuelsson et al., 2006).

At Grade 2, we administered a standardised test of spelling and also assessed how well the children could decode novel words and learn their orthographic patterns, following Share's (1999) techniques. Grade 2 data have not so far been reported, so here we detail these parts of the second-grade battery next.

Method

Participants

The Grade 2 sample included in this article only includes children from Australia and the United States. In Scandinavia, literacy is not taught until Grade 1 and hence the samples are not directly comparable in terms of school experience (Samuelsson et al., 2006). We have Grade 2 data on 225 MZ pairs and 214 same-sex DZ pairs, 80 and 54, respectively, in Australia, and 145 and 160 in the United States. Mean ages were 98.8 months ($SD = 5.0$) for MZ pairs and 99.7 (4.8) months for DZ pairs. The mean ages of the samples when they were assessed in the preschool phase were 58.3 months ($SD = 3.0$) for MZ pairs and 58.4 (2.9) months for DZ pairs.

Materials and procedure

The children in the ILTS are tested individually, each member of a twin pair at the same time by a different tester, either in their school premises or at home (see Byrne et al., 2007; Samuelsson et al., 2007). The Grade 2 battery, which we do not describe here in detail, includes measures of word and nonword identification efficiency, reading comprehension, vocabulary and orthographic choice, as well as measures specific to this article.

The spelling achievement measure was the spelling subtest from the Wide Range Assessment Test-Revised (WRAT-R; Jastak & Wilkinson, 1984). Children are asked to spell up to 45 words, ranging from easy items such as *go* and *cat* to difficult ones like *belligerent* and *occurrence*. Testing ceases after 10 consecutive errors, and scoring is simply correct or incorrect for each word as a whole (range 0–45).

Orthographic learning was modelled on a procedure devised by Share (1999, 2004). Children read short texts aloud, each of which contains a novel word that could be spelled at least two ways. They are told to notice the spelling of the new word because they would be asked to spell it later. Here is an example:

The new word is vade.

There is a hairy monster called a vade. The vade is very big. If you see a vade you should run away.

We created 15 such items (see Appendix). The children read them in blocks of three, after which they were asked to write out the spelling of the three novel words. If they

faltered or erred on the target word (*vade*, for example) in reading a text, the correct pronunciation was supplied by the tester, thus ensuring that each child either generated the phonology for the novel words for themselves or heard it from the tester. All children received the same version of the test, in the interests of uniformity and therefore comparability. There were five blocks of three texts, with spelling scores ranging from 0 to 15. Decoding score was the mean number of mispronunciations of the novel word (maximum four – see Nation et al. [2007], who also employed this measure).

Details of the preschool assessment phase are presented by Byrne et al. (2002) and Samuelsson et al. (2005). In brief, testing occupied approximately 45 minutes on each of 5 days spread across 1 or 2 weeks, and covered phonological awareness, familiarity with letters and other aspects of text, oral language, verbal and visuospatial memory and learning, and rapid naming. The tests relevant to this article are Block Design and Vocabulary from the Wechsler Preschool and Primary Scale of Intelligence-Revised, WPPSI-R (Wechsler, 1989). These tests were administered on the first and fifth days, respectively.

Results

In the first part of the results, we focus on the Grade 2 data (orthographic learning, spelling and decoding). Then, because of the high genetic correlations that we report, we turn to the data relevant to the generality issue by incorporating into the analyses the aspects of IQ that we assessed in the preschool phase. Finally, we focus specifically on the genetic relations between decoding and orthographic learning because doing so addresses the status of the self-teaching hypothesis.

Phenotypic analyses

In Table 1 we present descriptive statistics for spelling, orthographic learning and decoding. The mean differences between countries failed to reach significance even with these large sample sizes. Spelling and orthographic learning correlated .75. Decoding errors correlated $-.69$ with spelling and $-.61$ with orthographic learning.

Our specially created measure of orthographic learning was relatively brief at just 15 items (to fit our assessment schedule) and therefore potentially of low reliability. However, the index of internal consistency, Cronbach's α , was satisfactory at .78. Cronbach's α for our decoding measure was .87.

Behaviour-genetic analyses

We pooled the data across countries to increase power for our behaviour-genetic analyses. We first standardised the spelling and reading error scores within country, and,

Table 1. Means (and standard deviations) for orthographic learning, spelling and decoding.

	United States	Australia	Total
Orthographic learning (max. = 15)	8.91 (3.62)	8.42 (3.66)	8.76 (3.63)
Spelling (max. = 45)	17.49 (5.70)	18.26 (5.86)	17.72 (5.76)
Decoding errors (max. = 4)	0.54 (0.47)	0.51 (0.57)	0.53 (0.50)

Max., maximum.

following standard practice in behaviour genetics, also adjusted the scores for age and gender and truncated outliers to $\pm 3 SD$. The within-pair MZ and DZ correlations are shown in Table 2. In the case of each variable, the MZ correlation was considerably higher than its DZ counterpart, evidence for significant heritability. This is confirmed in the values for the proportions of total variance attributable to additive genetic factors (a^2), which are also presented in Table 2 along with shared (c^2) and unique (e^2) environment components in an ACE model computed using the Mx statistical package (Neale, Boker, Xie & Maes, 2002). Shared environment was not a significant component for any of the variables. Unique environment, which includes measurement error, was significant in all cases.

Having established that orthographic learning, spelling and decoding are subject to genetic influence, the focus of the study shifts to the genetic relationship among these measures. This relationship can be examined through a Cholesky decomposition of the variables, from which the genetic, shared environment and unique environment correlations can be ascertained. Cholesky decomposition is similar in principle to hierarchical regression in that the effects of an 'independent' variable on a 'dependent' variable are determined after the effects of a correlated independent variable have been removed. When applied to twin data (Neale et al., 2002), the model can partition covariation into genetic, shared environment and unique environment components.

Table 3 shows the model for the three variables. To illustrate how to interpret the table, consider the genetic factor coefficients. The loadings of factor A_1 on the three measured variables of .83, .64 and .73 show that a single genetic factor affects all measures, indicating a common source of genetic variation. The proportion of the variance accounted for by this genetic source for each variable is estimated from the squares of the

Table 2. Intra-class correlations and Mx modelling results (95% confidence intervals) for orthographic learning, spelling and decoding.

	MZ_r	DZ_r	a^2	c^2	e^2
Orthographic learning	.55	.36	.43 (.26, .62)	.15 (.00, .31)	.42 (.35, .49)
Spelling	.77	.37	.74 (.59, .81)	.03 (.00, .17)	.23 (.18, .28)
Decoding errors	.68	.31	.68 (.50, .74)	.00 (.00, .16)	.32 (.26, .39)

MZ, monozygotic; DZ, dizygotic; a^2 , additive genetic factors; c^2 , shared environment components; e^2 , unique environment components.

Table 3. Cholesky model of additive genetic (A), shared environment (C) and non-shared environment (E) factor loadings on decoding, orthographic learning and spelling.

Factor	Variable		
	Decoding errors	Orthographic learning	Spelling
A_1	.83 (.71, .86)	.64 (.48, .77)	.73 (.63, .82)
A_2		.15 (.00, .40)	.45 (.00, .54)
A_3			.00 (.00, .41)
C_1	.01 (.00, .44)	.39 (.00, .56)	.18 (.00, .41)
C_2		.01 (.00, .46)	.00 (.00, .35)
C_3			.00 (.00, .23)
E_1	.56 (.51, .62)	.14 (.07, .22)	.16 (.10, .22)
E_2		.63 (.58, .68)	.20 (.15, .25)
E_3			.40 (.37, .45)

coefficients, .69, .41 and .53, respectively. Factors A_2 and A_3 are not significant because the confidence intervals for all values contain zero (although it is possible that with a larger sample size the loading of factor A_2 on spelling might become significant). Thus, for example, orthographic learning is not affected by genes other than ones it shares with decoding – there is no significant independent genetic influence as the confidence interval of .00–.40 indicates. Not one of the C loadings is significant, consistent with the lack of significant shared environment effects noted earlier. The E matrix shows that unique environment influences the variables independently, possibly reflecting independence in measurement error. There is a small amount of overlap in unique environment influence.

Estimates of the genetic, shared environmental and non-shared environmental correlations between the measures (i.e., the extent to which the same genes or environmental factors are influencing decoding, orthographic learning and spelling, r_A , r_C and r_E , respectively) are obtained from the loadings. The genetic correlation, for example, is estimated from the ratio of the phenotypically standardised genetic correlation to the product of the square roots of the two heritabilities (i.e., r_A for orthographic learning and spelling = $(.64 \times .73) + (.15 \times .45) / (\sqrt{.43} \times \sqrt{.74}) = .95$). The genetic and unique environment correlations are presented in Table 4. There are very high genetic correlations among all three variables. Thus a single genetic factor accounts for all of the statistically significant genetic variance in the three measures. The unique environment correlations are significant though small to modest.

In Table 5 we report the Cholesky model when our available measure of nonverbal IQ, Block Design from the preschool testing phase, is added first. Block Design, with a univariate heritability of .49 ($.70^2$), shares genetic variance with the three Grade 2

Table 4. Genetic (above diagonal) and unique environment (below diagonal) correlations among decoding, orthographic learning and spelling.

Measure	1	2	3
1. Decoding errors	–	.97	.85
2. Orthographic learning	.22	–	.95
3. Spelling	.33	.47	–

Table 5. Cholesky model of additive genetic (A), shared environment (C) and non-shared environment (E) factor loadings on Block Design, decoding, orthographic learning and spelling.

Factor	Variable			
	Block Design	Decoding errors	Orthographic learning	Spelling
A_1	.70 (.56, .83)	.27 (.08, .48)	.27 (.12, .41)	.26 (.10, .42)
A_2		.78 (.63, .82)	.57 (.39, .71)	.67 (.55, .78)
A_3			.15 (.00, .41)	.45 (.00, .54)
A_4				.00 (.00, .41)
C_1	.39 (.00, .55)	.06 (.00, .32)	.00 (.00, .23)	.06 (.00, .28)
C_2		.05 (.00, .40)	.41 (.00, .56)	.22 (.00, .43)
C_3			.00 (.00, .46)	.00 (.00, .35)
C_4				.00 (.00, .22)
E_1	.60 (.55, .65)	.04 (.00, .10)	.08 (.00, .15)	.08 (.02, .14)
E_2		.56 (.51, .62)	.14 (.06, .22)	.15 (.10, .21)
E_3			.62 (.57, .67)	.19 (.13, .25)
E_4				.40 (.36, .44)

measures, with loadings of .26 or .27 on factor A_1 . However, there is substantial genetic influence on those measures that is independent of Block Design, with loadings ranging from .57 to .78 on factor A_2 . The genetic correlations between Block Design and decoding, orthographic learning and spelling were .33, .42 and .30, respectively. These are more modest than the genetic correlations among the three Grade 2 measures reported in Table 4, ranging from .85 to .97.

We repeated the Cholesky analysis employing a broader measure of cognitive skills by adding the WPPSI Vocabulary measure to Block Design to create an abbreviated IQ variable. The combination of these two subtests of the WPPSI is one of the short forms recommended by Sattler (1992) as a suitable measure of IQ for research purposes when time and/or resources prevent the administration of the full battery. The pattern of results was quite similar to that in Table 5, except that the univariate heritability of the combined measure, Vocabulary plus Block Design, was just .25 (.50²). This resulted from the substantial shared environment influence on early vocabulary and its consequently lower heritability noted in earlier analyses (Byrne et al., 2002; Samuelsson et al., 2005). The important point, however, is that the loadings on factor A_2 , which indicate genetic influence on decoding, orthographic learning and spelling that is independent of the genetic influence on our broader IQ measure, are similar in magnitude to the analogous figures in Table 5, at .72, .55 and .59 (compare with .78, .57 and .67 from Table 5). The genetic correlations of this IQ measure with decoding, orthographic learning and spelling were .49, .49 and .50, higher than for Block Design but still considerably less than the genetic correlations among the literacy variables. The increase is likely due to the inclusion of a language component in the IQ composite.

In the final analyses, we contrasted a Cholesky decomposition model of the relation between decoding and orthographic learning with what is known as a direction-of-causation (DOC) model (Duffy & Martin, 1994; Heath et al., 1993). The Cholesky model predicts that common genetic and environmental sources account for the covariation between decoding and orthographic learning. The DOC model predicts that levels of decoding and orthographic learning are each determined by independent genetic and environmental sources and that the covariation between them is accounted for by a direct causal influence of decoding on orthographic learning. The models are shown in Figure 1.

Our attempt to fit a DOC model resulted in a poorer fit relative to the Cholesky; $\Delta - 2LL = 5.89$, $\Delta df = 1$, $p = .015$. Thus, the common-gene account of the covariation between decoding and orthographic learning is supported.

Discussion

Summary of results

The primary goals of this component of the ILTS were: (1) to ascertain the genetic and environmental influences on orthographic learning, as operationalised by techniques developed by Share (1999); and (2) to examine the genetic and environmental relations between orthographic learning, spelling and decoding. Orthographic learning was moderately affected by genes, with significant non-shared environment effects. As expected from previous results in the ILTS (Byrne et al., 2005, 2006, 2007), spelling achievement was also subject to genetic influence, as was decoding. The genetic correlations close to 1 show that learning the spelling patterns of novel words depends on

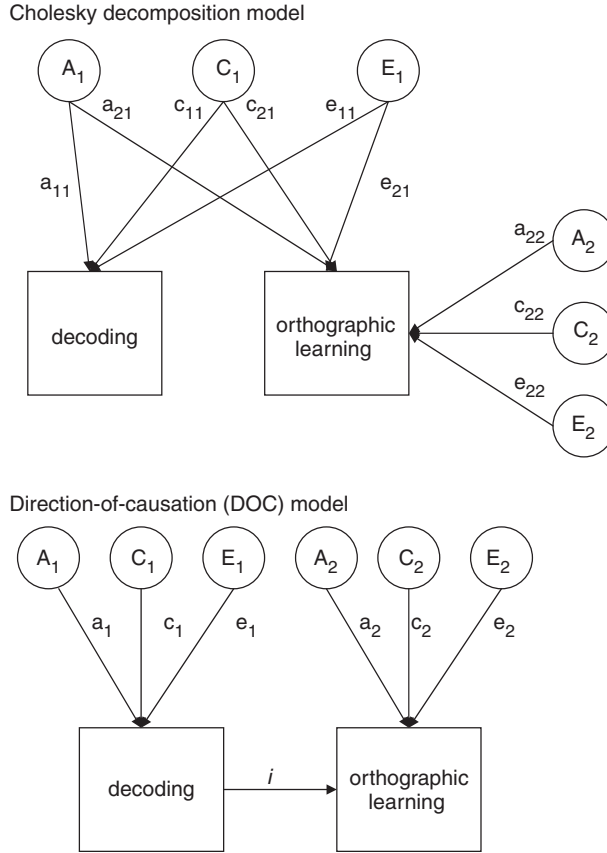


Figure 1. Cholesky decomposition model and direction of causation (DOC) model, shown for one member of a twin pair. Latent variables are represented by circles, measured variables by rectangles. In the Cholesky model, ACE factors load on both decoding and orthographic learning as well as only orthographic learning. In the DOC model, ACE factors load independently on decoding and orthographic learning, with a path, i , that represents the causal effect of decoding on orthographic learning.

genes that also influence existing spelling achievement and decoding skill. Those genes accounted for virtually all of the genetic variance in the three measures. A model in which the covariation between decoding and orthographic learning was explained by common genes provided a better fit to the relation between those two variables than one that assumed a direct causal path from decoding to orthographic learning.

As a result of the finding of high genetic overlap among the literacy variables, a secondary goal was to determine the generality of this genetic factor. There was a degree of genetic overlap between IQ (g) and the literacy measures, but the genetic influence on the three literacy variables was to a substantial degree independent of IQ, even when vocabulary contributed to the IQ measure.

Interpretation

We have a measure of learning (orthographic learning) that is subject to genetic influence and two measures of current knowledge (spelling achievement and decoding skill) that

are subject to the same genetic influence, as shown by the high genetic correlations. We favour a parsimonious account of this pattern which posits as basic a genetically determined learning-rate factor, most directly tapped by orthographic learning in our study. According to our hypothesis, this factor has, historically, influenced both spelling knowledge and decoding skill in the children in our sample. Thus, spelling is the 'crystallised' product of this factor. So too is decoding skill, reflecting prior responsiveness to learning opportunities, abundant in the school contexts in which children in the sample operate. In brief, this is a common genetic factor model with one of our variables, orthographic learning, as the most direct measure of that factor and the other variables, spelling achievement and decoding skill, as products of its prior operation.

As well as accommodating our data, an account such as this can explain the predictive power of responsiveness measures across a broad time span, because such measures index learning rate relatively directly. For the same reason, this account can accommodate the additional variance that responsiveness explains on top of more static measures of achievement (Byrne et al., 2005; Byrne et al., in press; Hindson et al., 2005). It also accommodates the abundant evidence for a correlation between reading and other abilities that patently depend on learning, such as vocabulary and many measures of working and long-term memory (for a summary see Snowling, 2000). It is also consistent with the work of Snowling, Gallagher and Frith (2003), who favour an associative learning-based account of deficiencies in children at risk for reading disability, exemplified by greater difficulties in memorising verbal material such as nursery rhymes.

We believe that the relative genetic independence of the literacy measures from other cognitive skills, roughly captured under the term *g* and comprising classic measures of performance and verbal IQ, enhances our learning account to the extent that variability is not 'just' IQ. But we are unable to be more specific about the breadth of the hypothesised learning rate factor. It may be limited to something as narrow as the mapping of speech onto print, extend to literacy and language more generally or cover a broader range of content areas. But whatever its generality, this account, if supported in future research, could have the effect of directing the search for genes contributing to literacy development towards those that appear important for learning processes, including those identified in animal models (e.g., Tsien, 2000). There have been recent successes in identifying human genes linked to severe reading disability (e.g., Cope et al., 2005; Hannula-Jouppi et al., 2005; Meng et al., 2005; Paracchini et al., 2006), and animal models have been invoked in advancing our understanding of the precise functions of some of these genes (Meng et al., 2005; Paracchini et al., 2006). The possibility that some are important for detecting and/or storing environmental contingencies, i.e. for learning, is worth consideration.

Although we have demonstrated a high genetic overlap among the three literacy variables and that this is independent of IQ to a substantial extent, we cannot rule out explanations other than our favoured learning rate parameter one on the basis of our data. There could be genetically determined variation in motivation to read, for example. However, the extensive research pointing to covariation between reading and a wide variety of memory-based tasks, mentioned above, tends to be consistent with our hypothesis.

The self-teaching hypothesis

The self-teaching hypothesis (Share, 1995) may need qualification in light of these results. Because our analyses indicated that decoding and orthographic learning are

affected by common genes, there is no need to postulate a direct causal path between them, as the model does in asserting that decoding is the sine qua non for orthographic learning. The relationship would be analogous to that between smoking and depression, described as non-causal by Kendler et al. (1993) because both are subject to common genetic influence. However, we cannot claim that our results generalise to all children in all school grades. Recently, Hulslander, Olson, Willcutt and Wadsworth (2007) have suggested that the self-teaching function may operate in early grades but decline in importance as decoding levels rise above some threshold required for adequate (if still (relatively) impaired) reading. So the hypothesis may hold for children younger than those in our sample. It may also hold for children in Grade 2 or higher with decoding skills below some yet-to-be-established threshold. Simply on grounds of intuition, it seems plausible that children whose very low levels of decoding skill inhibit independent reading are unlikely to have the opportunities to commit word-specific patterns to memory; they would simply lack exposure to reasonable numbers of new words. These possibilities invite more research.

In summary, we have shown that there exists a high phenotypic and genetic overlap among decoding, spelling and orthographic learning. The genes that influence this trio of measures are partly independent of genes that affect intelligence more broadly. We suggest that a common factor, best described as a learning parameter, links these aspects of literacy. We also suggest that in the continuing search for the genetic bases of reading ability and disability the biology of learning processes receive appropriate attention.

Acknowledgements

The research is being conducted with the support of the Australian Research Council (A79906201), the National Institutes of Health (2 P50 HD27802 and 1 RO1 HD38526), the Research Council of Norway (154715/330), Stavanger University and the Swedish Research Council (345-2002-3701). We are grateful to the many twins, their families and the twins' teachers for their participation, and to the Australian Twin Registry of the National Health and Medical Research Council. Annette Stevenson greatly assisted with the design of the orthographic learning task, and we also thank our project coordinators and testers: in Sweden (Inger Fridolfsson), Norway (Bjarte Furnes), the United States (Kim Corley, Rachael Cole, Pat Davis, Barb Elliot, Kari Gilmore, Amy Rudolph, Ingrid Simecek and Angela Villella) and in Australia (Frances Attard, Fiona Black, Rosemary Brown, Marnie Church, Nicole Church, Maretta Coleman and Cara Newman).

References

- Bishop, D.V.M., Adams, C.V. & Norbury, C.F. (2006). Distinct genetic influences on grammar and phonological short-term memory deficits: Evidence from 6-year-old twins. *Genes, Brain and Behaviour*, 5, 158–169.
- Bowey, J.A. & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 92, 203–219.
- Byrne, B., Delaland, C., Fielding-Barnsley, R., Quain, P., Samuelsson, S., Høien, T. et al. (2002). Longitudinal twin study of early reading development in three countries: Preliminary results. *Annals of Dyslexia*, 52, 49–73.

- Byrne, B., Fielding-Barnsley, R. & Ashley, L. (2000). Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. *Journal of Educational Psychology*, 92, 659–667.
- Byrne, B., Olson, R.K., Samuelsson, S., Wadsworth, S., Corley, R., DeFries, J.C. et al. (2006). Genetic and environmental influences on early literacy. *Journal of Research in Reading*, 29, 33–49.
- Byrne, B., Samuelsson, S., Wadsworth, S., Hulslander, J., Corley, R., DeFries, J.C. et al. (2007). Longitudinal twin study of early literacy development: Preschool through Grade 1. *Reading and Writing: An Interdisciplinary Journal*, 20, 77–102.
- Byrne, B., Shankweiler, D. & Hine, D. (in press). Reading development in children at risk for dyslexia. In M. Mody & E. Silliman (Eds.), *Language impairment and reading disability: Interactions among brain, behavior, and experience*. New York: Guilford Press.
- Byrne, B., Wadsworth, S., Corley, R., Samuelsson, S., Quain, P., DeFries, J.C. et al. (2005). Longitudinal twin study of early literacy development: Preschool and Kindergarten phases. *Scientific Studies of Reading*, 9, 219–235.
- Cope, N., Harold, D., Hill, G., Moskvina, V., Stevenson, J., Holmans, P. et al. (2005). Strong evidence that KIAA0319 on chromosome 6p is a susceptibility gene for developmental dyslexia. *American Journal of Human Genetics*, 76, 581–591.
- Cunningham, A.E., Perry, K.E., Stanovich, K.E. & Share, D.L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology*, 82, 185–199.
- Duffy, D.L. & Martin, N.G. (1994). Inferring the direction of causation in cross-sectional twin data: Theoretical and empirical considerations. *Genetic Epidemiology*, 11, 483–502.
- Ehri, L.C. (2005). Development of sight word reading: Phases and findings. In M.J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook*. (pp. 135–154). Oxford: Blackwell.
- Firth, I. (1972). Components of reading disability. Unpublished doctoral dissertation, University of New South Wales, Sydney, Australia.
- Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kääriäinen, H. et al. (2005). The axon guidance receptor gene *ROBO1* is a candidate gene for developmental dyslexia. *Public Library of Science: Genetics*, 1(4) e50, 0467–0474.
- Heath, A.C., Kessler, R.C., Neale, M.C., Hewitt, J.K., Eaves, L.J. & Kendler, K.S. (1993). Testing hypotheses about direction of causation using cross-sectional family data. *Behavior Genetics*, 23, 29–50.
- Hindson, B.A., Byrne, B., Fielding-Barnsley, R., Newman, C., Hine, D. & Shankweiler, D. (2005). Assessment and early instruction of preschool children at risk for reading disability. *Journal of Educational Psychology*, 94, 687–704.
- Hulslander, J., Olson, R.K., Willcutt, E. & Wadsworth, S. (2007). Longitudinal stability of reading-related skills and their prediction of reading growth across the later grades. Manuscript submitted for publication.
- Jastak, S. & Wilkinson, G.S. (1984). *The wide range achievement test-revised: Administration manual*. Wilmington, DE: Jastak Associates Inc.
- Kendler, K.S., Neale, M.C., MacLean, C.J., Heath, A.C., Eaves, L.J. & Kessler, R.C. (1993). Smoking and major depression: A causal analysis. *Archives of General Psychiatry*, 50, 36–43.
- Kyte, C.S. & Johnson, C.J. (2006). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, 93, 166–185.
- Meng, H., Smith, S.D., Hager, K., Held, M., Liu, J., Olson, R.K. et al. (2005). DCDC2 is associated with reading disability and modulates neuronal development in the brain. *Proceedings of the National Academy of Sciences*, 102, 17053–17058.
- Nation, K., Angell, P. & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, 96, 71–84.
- Neale, M.C., Boker, S.M., Xie, G. & Maes, H.H. (2002). *Mx: Statistical modeling*. (6th edn). Richmond, VA: Department of Psychiatry, VCU.
- Paracchini, S., Thomas, A., Castro, S., Lai, C., Paramasivam, M., Wang, Y. et al. (2006). The chromosome 6p22 haplotype associated with dyslexia reduces the expression of KIAA0319, a novel gene involved in neuronal migration. *Human Molecular Genetics*, 15, 1659–1666.
- Plomin, R. & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin*, 131, 592–617.
- Samuelsson, S., Byrne, B., Quain, P., Wadsworth, S., Corley, R., DeFries, J.C. et al. (2005). Environmental and genetic influences on prereading skills in Australia, Scandinavia, and the United States. *Journal of Educational Psychology*, 97, 705–722.
- Samuelsson, S., Byrne, B., Wadsworth, S., Corley, R., DeFries, J.C., Willcutt, E. et al. (2007). Genetic and environmental influences on prereading skills and early reading and spelling development in the United States, Australia, and Scandinavia. *Reading and Writing: An Interdisciplinary Journal*, 20, 51–75.

- Samuelsson, S., Olson, R.K., Byrne, B., Willcutt, Wadsworth, S. & DeFries, J.C. (2006). *Genetic and environmental influences on early literacy development in Australia, Scandinavia, and the United States*. In R. K. Olson (Chair), Genetic and environmental influences on early reading, math, ADHD, and related skills. Symposium conducted at the 13th annual conference of the Society for Scientific Study of Reading, Vancouver, Canada, July.
- Sattler, J.M. (1992). *Assessment of children*. (3rd edn). San Diego, CA: Jerome M. Sattler Publishers Inc.
- Share, D.L. (1995). Phonological recoding and self-teaching: *Sine qua non* of reading acquisition. *Cognition*, 55, 151–218.
- Share, D.L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72, 95–129.
- Share, D.L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, 87, 267–298.
- Snowling, M.J. (2000). *Dyslexia*. Oxford: Blackwell.
- Snowling, M.J., Gallagher, A. & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development*, 74, 358–373.
- Tsien, J.Z. (2000). Linking Hebb's coincidence-detection to memory formation. *Current Opinion in Neurobiology*, 10, 266–273.
- Wechsler, D. (1989). *Manual for the Wechsler preschool and primary scale of intelligence – revised*. New York: Psychological Corporation.
- Willcutt, E., Betjemann, R., Wadsworth, S., Samuelsson, S., Corley, R., DeFries, J.C. et al. (2007). Preschool twin study of the relation between attention-deficit/hyperactivity disorder and prereading skills. *Reading and Writing: An Interdisciplinary Journal*, 20, 103–125.

Appendix

1. The new word is Laif.
The coldest town in the world is Laif. Laif is in Greenland. The people who live in Laif need very hot houses.
2. The new word is vade.
There is a hairy monster called a vade. The vade is very big. If you see a vade, you should run away.
3. The new word is lale.
Farmers grow a fruit called lale. Lale trees are red. Children like to eat lale when they are hungry.
4. The new word is slear.
There is a very big animal called a slear. The slear has big feet. If you have to carry a slear, you will need help.
5. The new word is berl.
Farmers grow a plant called berl. Berl likes to grow in wet places. Summer is the time to cut berl.
6. The new word is Feap.
There are lots of rocks on the planet Feap. Feap is very far away. To get to Feap, you need a rocket.
7. The new word is pewt.
There is a small hut called a pewt. A pewt has tiny windows. Some children found a pewt with three puppies inside.
8. The new word is staip.
There is a bird called a staip. The staip eats nuts and seeds. You can see a staip in the zoo.

9. The new word is noar.
Farmers make a warm drink called noar. Noar is made from milk. Noar is good in winter.
10. The new word is mese.
Some planets have sticky red stuff called mese. Mese drips out of the rocks. People do not like to stand on mese.
11. The new word is krent.
There is a very shy animal called a krent. The krent only comes out in the daytime. If you see a krent, it will run away
12. The new word is scep.
There is a kind of boat called a scep. The scep is made from a log. If you ride in a scep, you will be safe.
13. The new word is smitch.
There is a bug called a smitch. The smitch has lots of teeth. If a smitch bites you it will hurt.
14. The new word is fope.
There is a new ball called a fope. The fope is made of rubber. Some kids will get a fope for Christmas.
15. The new word is zupe.
There is a giant fish called the zupe. A zupe can jump very high. Some children like to swim with the zupe.

Received 6 August 2007.

Address for correspondence: Brian Byrne, School of Behavioural, Cognitive and Social Sciences, University of New England, Armidale, NSW 2351, Australia. E-mail: *bbyrne@une.edu.au*