

Cognitive modelling and the behaviour genetics of reading

Anne Castles

University of Melbourne

Timothy Bates and Max Coltheart

Macquarie University

Michelle Luciano and Nicholas G. Martin

Queensland Institute of Medical Research

While it is well known that reading is highly heritable, less has been understood about the bases of these genetic influences. In this paper, we review the research that we have been conducting in recent years to examine genetic and environmental influences on the particular reading processes specified in the *dual-route* cognitive model of reading. We argue that a detailed understanding of the role of genetic factors in reading acquisition requires the delineation and measurement of precise phenotypes, derived from well-articulated models of the reading process. We report evidence for independent genetic influences on the lexical and nonlexical reading processes represented in the dual-route model, based on studies of children with particular subtypes of dyslexia, and on univariate and multivariate genetic modelling of reading performance in the normally reading population.

It is well established that the skill of reading has a strong genetic basis. Heritability data from behaviour genetic studies indicate that at least half of the variance in reading ability is genetic (DeFries, Fulker & LaBuda, 1987; Olson, Wise, Conners & Rack, 1989). These data suggest that an enhanced understanding of the nature of these genetic influences on reading has the potential to play an important role both in uncovering the factors that influence the normal acquisition of reading ability and in illuminating the pathways by which genes and environment act to cause developmental reading disorder.

However, the task of uncovering such factors is by no means an easy one. As Fisher and Defries (2002) note, the genetic analysis of reading and dyslexia is complex for many reasons. For example, there is no straightforward correspondence between genotype and phenotype, and phenotypic variations can depend on the developmental stage of the subject. Similarly, there is a lack of consensus on the definition of dyslexia, and on whether it is a single trait or a cluster of traits with distinct aetiologies. Thus, as is the

case for most cognitive and behavioural traits, the isolation of genetic risk factors presents a formidable challenge.

In our view, a key requirement for meeting this challenge is the utilisation of precisely defined and reliably measured reading behaviours, which articulate the cognitive components of reading and writing. This in turn requires a well-specified cognitive model of reading that makes clear how these components explain the variance of reading, and that predicts constraints on the observed test correlations, which can in turn be tested within quantitative genetic models.

With some notable exceptions (Gayán & Olson, 2001, 2003; Marlow et al., 2003; Olson et al., 1989; Olson, Forsberg & Wise, 1994), much of the behaviour genetic research on reading has treated this complex skill as a single phenotype, measured by omnibus standardised tests. Measurement of reading in this way is inconsistent with the cognitive evidence, which indicates that, regardless of which specific theoretical model is adhered to, there are a number of different processes involved in reading, and that these may develop at different rates, and may be differentially impaired, in different individuals (Castles & Coltheart, 1993; Castles, Datta, Gayán & Olson, 1999; Harm & Seidenberg, 1999; Manis, Seidenberg, Doi, McBride-Chang & Peterson, 1996).

In this paper, we review the research that we have recently been carrying out exploring the genetic bases of the particular reading subprocesses that are embodied in the *dual-route* model of reading (Coltheart, Curtis, Atkins & Haller, 1993; Coltheart, Rastle, Perry, Langdon & Ziegler, 2001). We have adopted this cognitive model as our preferred model, as it has been shown to be able to account for a large number of basic psycholinguistic phenomena in normal reading (Coltheart et al., 2001), has successfully predicted the existence of particular patterns of reading deficit, both in the acquired and the developmental dyslexic population (Castles & Coltheart, 1993) and is sufficiently explicit to have been implemented computationally (Coltheart et al., 2001).

Many of the results that we report are also potentially consistent with other computational cognitive models of reading, such as the PDP framework of Seidenberg, Plaut and colleagues (Plaut, McClelland, Seidenberg & Patterson, 1996; Seidenberg & McClelland, 1989). However, rather than attempting to discriminate between these models, which converge on many points, our aim here is more general: to show that a deeper understanding of the genetic bases of reading and its acquisition can be obtained if genetic variance is mapped onto theoretically driven and precisely defined reading phenotypes. This is because, as these phenotypes represent relatively pure and psychologically meaningful constructs, their use will increase the power of genetic linkage analyses over that possible from measures that confound the function of multiple processes (see e.g. Marlow et al., 2003), will decrease the likelihood of replication failures across different studies and will be more likely to map onto a manageable set of gene functions.

We begin by briefly reviewing the dual-route model of reading, and outlining the logic of the twin study designs that we have used. We then describe a series of studies exploring the heritability of particular reading phenotypes derived from the dual-route model. Finally, we discuss how future research may pinpoint more specifically the developmental trajectories that lead from genetic imprints at conception through to the particular reading profiles displayed by children once they have reached school age and beyond.

The dual-route model of reading

The dual-route model proposes that reading aloud is achieved via a coordinated system of processing modules that support two simultaneous routes, a lexical route, and a nonlexical route, as shown in Figure 1 (Coltheart, 1978; Coltheart et al., 1993, 2001; Morton & Patterson, 1987). Both routes begin with a system for the analysis of visual input to construct letter representations. In the lexical route, the output of this process is used to access a lexicon or memory store of local representations of previously seen written words. In the case of the nonlexical route, no previous experience of the letter string is required; instead, a rule-based process is used to convert the graphemes into phonemes. Both routes generate a phonological output that is directed to common systems for producing a spoken response.

The two routes are distinguished by their performance on two particular types of reading material. Irregular words such as *yacht* can only be read aloud correctly via the lexical route, as the nonlexical route ‘regularises’ these words because of its application of letter–sound rules. Nonwords, such as *gop*, can only be read aloud correctly by the rule-based nonlexical route, as there is no entry in the orthographic lexicon that matches this particular string of letters. Importantly, although each route depends on multiple modules, several of these are shared by both routes: it is the *orthographic lexicon* and the *letter–sound application* modules of Figure 1 that are unique to the lexical and nonlexical routes, respectively, and that therefore need to be isolated if pure phenotypes are to be

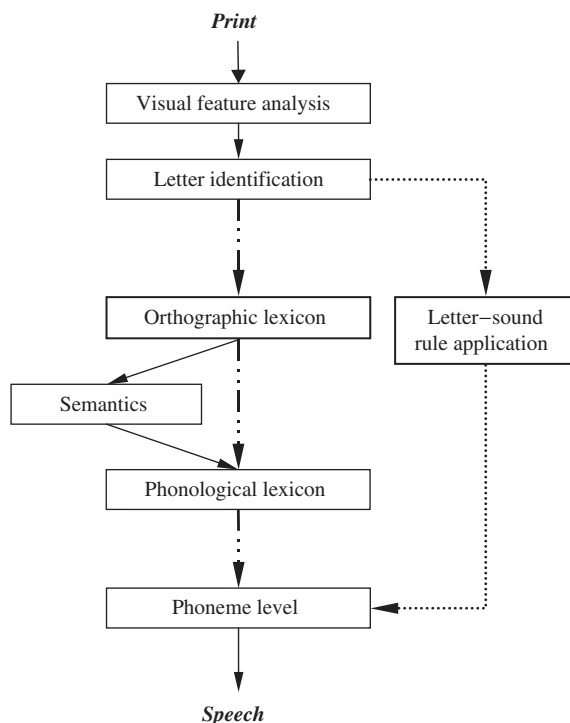


Figure 1. The dual-route model of reading. The series of processes constituting the lexical route are connected with dashed lines (—...—), while the nonlexical route components are connected with dotted lines (...). Inputs to and outputs from the system are shown in solid lines.

obtained. For this reason, in the behaviour genetic analyses reported below, these two components of reading were the particular focus of interest.

Behaviour genetic modelling of cognitive processes

In order to conduct behaviour genetic analyses of the basic cognitive processes represented in the dual-route model described above, the phenotypes involved must be assessed in a genetically informative sample. That is, individuals who differ in known ways in their genetic and environmental similarity to each other must be selected. In the studies described here, a twin design was used to this end. The power of twin designs lies in the fact that twins may be either *monozygotic* (MZ: developing from a single fertilised egg and therefore sharing all of their genes) or *dizygotic* (DZ: developing from two fertilised eggs and therefore sharing on average 50% of their segregating genes, the same level of genetic similarity as is found in nontwin siblings). These known differences in genetic similarity, together with a testable assumption of equal environments for MZ and DZ twins (Bouchard & Propping, 1993), allows exploration of the effects of genetic and environmental variance on a particular phenotype (Neale & Cardon, 1992).

Specifically, because twins raised in their biological families share the family environment of their co-twin, and have only a random similarity with the environment of other twin pairs, and because the genetic similarity of MZ and DZ twins differs in a known quantitative fashion (100% in MZ, 50% in DZ), the variance in any trait measured in a population of MZ and DZ twins can be analysed in terms of: (a) genetic effects, (b) effects of the environment that are shared by twins and (c) residual unshared environmental effects, including measurement error. These latter environmental effects include events experienced by only one twin, such as illnesses, individual educational experiences or particular peer-group affiliations.

Using twins as our genetically informative sample, we have been pursuing a range of methods for exploring the heritability of different sub-processes in the reading system, particularly the lexical and nonlexical processes. We outline the approaches we have adopted, and some basic results, below.

Genetic and environmental influences on lexical and nonlexical reading processes

The heritability of reading deficits in 'subtypes' of dyslexia

One fruitful approach that we, and others, have adopted to explore the genetic basis of different sub-processes of the normal reading system has been to explore the heritability of reading impairments in children with particular patterns or subtypes of dyslexia. It has been established for some time that dyslexia is not a homogeneous disorder (Boder, 1973; Castles & Coltheart, 1993; Manis et al., 1996; Mitterer, 1982; Stanovich, Siegel & Gottardo, 1997). Although most poor readers tend to be impaired on both word and nonword reading tasks, a proportion of children identified with impairments in reading display particular difficulties with nonword reading, while their irregular word reading is relatively spared. This pattern of deficits, interpreted in dual-route terms as a specific impairment in acquiring the nonlexical processing route, is commonly referred to as *developmental phonological dyslexia* (Campbell & Butterworth, 1985; Castles &

Coltheart, 1993). Other reading-impaired children display particularly poor performance on reading irregular words, while their nonword reading is spared, suggesting a specific lexical acquisition impairment in dual-route terms. Reduced reading function in this form is typically referred to as *developmental surface dyslexia* (Castles & Coltheart, 1996; Hanley, Hastie & Kay, 1992). If genetic influences on the reading impairment in these two types of poor readers were found to differ, some support would be provided for the dissociability of lexical and nonlexical reading processes, and for variation in the genetic and environmental factors involved in their process of acquisition.

A particular kind of behaviour genetic analysis, designed to examine the heritability of deviant group membership when twins are selected from the extreme low end of the normal distribution on the variable of interest (in this instance, reading ability), is required for these kinds of studies (Defries & Fulker, 1985). Here, the twin of a pair who is deviant enough on the reading dimension to be called reading disabled is nominated as the *proband* and the other member of the pair is called the *co-twin*. The heritability of the probands' group membership is then assessed by comparing the amount of regression towards the mean for the MZ and DZ co-twins. If the probands' group reading deficit was entirely because of genetic influence (and there was no test error), MZ co-twins would show no regression to the normal population mean, because they share 100% of their genes. DZ co-twins would regress halfway to the population mean on average, as they share half of their segregating genes. On the other hand, if the probands' reading deficit was entirely because of shared environment influence, both MZ twins and DZ twins would show no regression towards the population mean. Therefore, by comparing the relative MZ and DZ co-twin regression with the mean, estimates of the proportional genetic and shared environment influence on the probands' group membership in the low tail of the reading dimension can be obtained.

Castles et al. (1999) used this kind of analysis to examine the heritability of reading deficits in a group of phonological dyslexics and a group of surface dyslexics. The phonological dyslexic subgroup consisted of 322 children who were at least one standard deviation below the age average on a single-word reading test and whose scores on a nonword reading aloud task were significantly lower than their scores on an irregular word reading aloud task. The surface dyslexic group also consisted of 322 children who were at least one standard deviation below average on the single-word reading test, but who in this case had scores on the irregular word task that were significantly lower than their scores on the nonword task.

The results that we obtained are summarised in Table 1. As can be seen, for the children with phonological dyslexia, the estimated heritability component, based on the relative regression to the mean for MZ compared with DZ twins, was highly significant, while the shared environmental influence was smaller (but still significant). Thus, there

Table 1. Genetic and environmental influences on word recognition impairments in surface and phonological dyslexia, based on data from Castles et al. (1999).

	A genetic influence	C shared environment influence	E unshared environment and error
Phonological dyslexia	.67**	.27*	.06
Surface dyslexia	.31*	.63**	.06

*Significant at $p < .05$.

**Significant at $p < .001$.

was evidence for both genetic and shared environmental components to word recognition deficits in the phonological dyslexic group, but the genetic component made the greater contribution of the two. The proportion of genetic and shared environmental influence on word recognition deficits in the surface dyslexic subgroup differed markedly from that of the phonological dyslexics: in this case, the shared environmental component made the greater contribution of the two, although, again, both genetic and shared environmental influences were significant.

The subgroup selections in the Castles et al. (1999) study were based on irregular word and nonword reading, but similar heritabilities have been reported when alternative measures of component processes in reading have been used in assessing the heritability of group deficits in reading-impaired samples, such as the orthographic and phonological choice tasks of Olson and colleagues (Gayán & Olson, 2001; Olson et al., 1989, 1994). Together, these results strongly suggest that both lexical and nonlexical reading processes have an inherited component, and that there may be some differences in the nature and extent of the genetic factors influencing the acquisition of each route. We have since extended these findings by exploring genetic and environmental influences on specific reading processes in unselected reader samples.

Heritability of reading subprocesses in unselected samples

The results described above provide a picture of the degree to which genetic factors are involved when particular processes in reading are severely impaired in particular individuals. However, it is entirely plausible that both the magnitude and the mechanism of genetic influence on *deviant* group membership could differ from those for individual differences across the normal range (Gayán & Olson, 2003). Given that the dual-route model is a description of the normal reading process, a valuable extension of the above findings would be to show that these same genetic influences on reading subprocesses are evident in an unselected sample of normal readers.

Recently, because of the greater accessibility of software for carrying out advanced behaviour genetic modelling, data of this kind have begun to be reported (Bates, Castles, Coltheart, Gillespie, Wright & Martin, 2004; Gayán & Olson, 2003). This work can be broadly divided into *univariate analyses*, where the genetic and environmental influences on each of the reading subprocesses are modelled separately, and *multivariate analyses*, where more than one subprocess is examined within a single model.

Univariate analyses. We began our exploration of genetic and environmental influences on reading in an unselected sample by obtaining basic, univariate estimates of the heritabilities of regular word, irregular word and nonword reading in a large sample of normal adolescent and adult twins (Bates et al., 2004). The participants were 470 twin pairs (125 MZ; 345 DZ) participating in a large DNA study (mainly unrelated to reading) being conducted by the Queensland Institute of Medical Research. The participants were 18 years old on average ($sd = 2.7$). Their regular word, irregular word and nonword reading accuracy were assessed via telephone testing using a 120-word extended version of the tests Castles & Coltheart (1993), with additional items added to increase the difficulty of this test for an older sample. More details on the sample and methodology can be found in Bates et al. (2004).

Univariate genetic models were then fitted to each of the three sets of reading data separately, by the method of maximum likelihood in Mx (Neale, Boker, Xie & Maes,

2002). This process decomposes the total variance in the observed measures into that attributable to *A*, *C* and *E*, with *A* referring to additive genetic effects, *C* referring to common or shared environmental effects (effects common to siblings within a family, and varying between families) and *E* representing the residual variance not otherwise explained, and comprising unique environmental effects and measurement error.

The results indicated that the heritabilities of these three reading measures in an unselected sample are as high as, or even higher than, those found in impaired-reader populations. Additive genetic effects (*A*) were found to account for 73% of the variance in lexical reading, as measured by irregular word reading and 71% of the variance in nonlexical reading, as measured by nonword reading. Regular word reading, which is proposed to reflect both lexical and nonlexical processing, was found to have a 61% additive genetic component. These extremely high heritabilities are consistent with other recent estimates in unselected reader samples, with Gayán and Olson (2003) reporting heritabilities of .87 and .80 for 'orthographic' and 'phonological' reading, respectively, in an unselected group of readers from 8 to 18 years old.

For all three item types, the remaining variance was best modelled as a unique environment influence (including measurement error: *E*) rather than as differences between families (*C*). Thus, we found no evidence of an influence of shared environment on reading in this sample. This somewhat surprising finding is again consistent with the findings of Gayán and Olson (2003), who reported nonsignificant shared environment effects of 1–5% for orthographic and phonological reading in their sample. It may be that the ages of the participants in these two studies account for this result: behaviour genetic studies of young readers suggest that while shared environment exerts a significant influence on early reading (and may also do so on the reading of older dyslexics), by about age nine these effects are replaced with increases in genetic effects and unshared environment (Byrne et al., 2002; Tiu, Wadsworth, Olson & DeFries, 2004).

The magnitude of the heritabilities in these univariate analyses supports the view that much of the normal individual variance in both lexical and nonlexical reading is heritable. These findings are potentially very important for understanding the basis both of dyslexia and of the normal reading process, for they tell us that the genetic and environmental influence that we find to be associated with reading impairment is not just representative of a particular 'clinical' subgroup of the population, but is in fact implicated in normal, skilled reading, with the clinical group simply representing the extreme end of a normal distribution.

What these data do not tell us, however, is whether the genetic and environmental influences that we find on particular reading subprocesses are *independent* of each other. Are the sets of genes that are responsible for the variance in lexical reading skills observed in our sample the same as those responsible for the variance in nonlexical reading skills? Similarly, is the environmental aetiology distinct, or common to both reading processes? To address these issues, multivariate modelling is required, and we will now turn to a discussion of these analyses.

Multivariate analyses. Lexical and nonlexical reading processes, while distinct, have much in common with each other. Inspection of Figure 1 makes this fact apparent: the two routes share the same processes at several points, including at the level of visual feature analysis, letter identification and phonemic output. Therefore, it is highly likely that the two routes will share some of the same genetic and environmental influences. However, it is also possible that there will be a subset of genetic and environmental

factors that are specific to each process. At a behavioural level, there is evidence to suggest that the cognitive, perceptual and linguistic determinants of the acquisition of the two routes differ to some degree, with phonological language skills being particularly implicated in the development of the nonlexical route and environmental factors such as print exposure playing a more significant role in lexical acquisition (see e.g. Cunningham, Perry & Stanovich, 2001; Rack, Snowling & Olson, 1992). We will discuss how particular hypotheses about the acquisition of the two routes can be explored in behaviour genetic analyses in a later section, but a first aim was to determine whether evidence could indeed be found for both shared and unique genetic and environmental variance being associated with each of the two reading routes.

In some recent work, we have examined genetic and environmental influences on the irregular and nonword reading data from our large unselected reader sample within a single, multivariate model (Bates, Castles, Luciano, Wright, Coltheart & Martin, *in press*). Specifically, we used a Cholesky decomposition procedure to test for shared and independent genetic and environmental effects on the lexical and nonlexical phenotypes (Neale & Cardon, 1992). The Cholesky procedure functions in a manner very similar to the hierarchical regression analysis widely used in behavioural research. The independent contribution of a predictor variable is assessed after accounting for its shared variance with all of the other predictor variables. However, in contrast to nongenetic studies, the data from MZ and DZ twins allow for a further decomposition of the shared and independent variance into different genetic and environmental components. Models are therefore fitted in which the total variances (and covariances between variables) are parameterised in terms of the *A*, *C* and *E* factors described previously.

Our results, described in detail in Bates et al. (*in press*), showed evidence for independent genetic and environmental effects on both the lexical and nonlexical reading phenotypes. Approximately 59% of the variance in irregular and nonword reading was accounted for by a general additive genetic factor (*A*). However, each phenotype was also separately influenced by significant independent additive genetic components, accounting for 11% of the variance in nonword reading and 14% of the variance in irregular word reading. The remaining sources of variance were attributable to unique environmental effects, including measurement error (*E*): there was a small general *E* effect that accounted for 6% of the variance in irregular and nonword reading, and larger independent *E* effects, accounting for 24% of the variance in nonword reading and 19% of the variance in irregular word reading. Consistent with the univariate results, the model could be simplified by dropping the common environment (*C*) parameter without any significant loss of fit, suggesting that shared environmental factors were not influencing reading performance in this sample.

Gayán and Olson (2003) report broadly similar results in their younger-reader sample, using latent trait constructs instead of individual measures of lexical and nonlexical skills. Both phonological decoding and orthographic coding skills in word recognition were found to have significant common as well as significant independent genetic influences. Although the genetic correlation between phonological decoding and orthographic decoding was high (.80, after controlling for IQ), there remained a substantial and significant 36% of independent genetic variance between the two traits. Thus, although linkage analyses have not identified specific genes associated with these traits to date (Fisher & Defries, 2002), it would appear from these results that such genes should be identifiable.

In summary, the multivariate behavioural-genetic modelling reported here points to the need to postulate the existence of distinct sets of genes exerting control over the acquisition

of different components of the reading system. In cognitive dual-route terms, there appear to be some genes that influence only the ability to acquire nonlexical reading skills and others that influence only the ability to acquire lexical reading skills. The question then becomes: exactly what are these distinct genetic influences, and how do they differentially affect the process of acquisition of the lexical and nonlexical routes? In the next section, we will consider some possible ways of tackling this important question.

Behaviour-genetic modelling and theories of reading acquisition

Reading is a learned skill and, without being taught, children will not learn to read. Therefore, the genetic and environmental influences on lexical and nonlexical reading that we have identified are likely to reflect perceptual, cognitive or linguistic factors affecting the process of acquisition of these reading skills over a period of time. In some cases, these will be general factors, influencing the acquisition of both processes. In other cases, our evidence suggests that they must be specific factors, differentially affecting the success of acquisition of one process or the other. The cognitive and psycholinguistic research literature is replete with theories about what these factors might be. In this section, we consider how behaviour-genetic modelling might be applied to the task of testing and discriminating between such theories.

We would argue that the data that we and others have obtained on the independence of genetic influences on lexical and nonlexical reading already allow for some observations to be made on existing theories. Consider, for example, the hypothesis of Stanovich, Siegel and Gottardo (1997) in relation to the bases of surface and phonological dyslexia. They propose that the differences in these two types of reading impairment can be accounted for in terms of different degrees of 'core phonological' impairment, combined with a lack, or otherwise, of exposure to print. Specifically, developmental phonological dyslexics are proposed to suffer from a basic and severe phonological language deficit, which impairs their ability to learn letter-sound correspondences. Developmental surface dyslexics are proposed to suffer from the same core phonological deficit, but in a much milder form, and in this case overlaid by a lack of exposure to print.

Our data, and those of Gayán and Olson (2003), would seem to indicate that, at the very least, this cannot be the full story in relation to surface and phonological dyslexia. Only one potentially biologically based factor is being proposed here: the core phonological deficit. The other factor, print exposure, is a shared or unique environmental influence. Yet, our data show that there are two *independent* genetic influences on lexical and nonlexical reading. If the core phonological deficit were the only genetic factor influencing both kinds of reading deficit, to a greater or a lesser degree, the modelling would have indicated a single general additive factor, with different path loadings for irregular words and nonwords. So, while both phonological skills and print exposure are likely to influence the acquisition of lexical and nonlexical reading routes, and to be implicated when they fail to develop in different types of dyslexia, the behaviour-genetic data tell us that a further genetic factor needs to be identified before the acquisition of lexical skills, and thus the basis of surface dyslexia, can be fully explained.

Advances in testing theories of reading acquisition have also already been made by Gayán and Olson (2003), through their extension of the multivariate modelling techniques described above to include key language and cognitive measures. Given that it is unlikely that there will be a genetic specialisation for reading itself, it is these

preexisting cognitive skills that may prove most useful in uncovering causal pathways to the successful acquisition of lexical and nonlexical skills. In addition to measuring latent traits for reading sub-processes, Gayán and Olson also measured IQ, an important covariate that may allow general versus specific genetic influences on reading processes to be identified (see Cardon, Dilalla, Plomin, DeFries & Fulker, 1990; Knopik et al., 2002; Stevenson, Graham, Fredman & McLaughlan, 1987), and phoneme awareness in their sample. This allowed them to test a number of basic hypotheses about the genetic relationship between reading processes and these variables, and, in some instances, to examine causal theories.

For example, the dominant theory of the acquisition of nonlexical or phonological decoding skills is that it is underpinned by the ability to perceive and manipulate speech sounds in spoken language; that is, the core phonological skills described above (e.g. Rack et al., 1992; Wagner, Torgesen & Rashotte, 1994; for a discussion of the causal hypothesis, see Castles & Coltheart, 2004). These language skills are thought likely to have a biological basis. Therefore, finding of a high degree of genetic covariation between phonological language skills and nonlexical reading skills would provide some support for this causal hypothesis. Gayan and Olson report moderate genetic associations between their phoneme awareness measure and all of their reading measures, but the magnitude of the genetic correlation was found to be significantly larger for phonological decoding (.79) than for orthographic decoding (.55). These results thus support the hypothesis that there may be a genetically based causal influence from young children's development in phonemic awareness to their later phonological decoding skills, and that this causal pathway may be less relevant (although still significant) to the acquisition of lexical or orthographic coding skills.

In future studies, we plan to extend this line of research to examine a range of other hypotheses about the cognitive processes that may underpin the acquisition of lexical and nonlexical reading skills. For example, we expect to be able to delineate the role of the environmental variable of print exposure more precisely. If print exposure does indeed have a specific effect on the acquisition of lexical skills, we would expect performance on this measure to load on the same unique environmental factor that has been found to explain significant variance in irregular word reading. It should not, however, load on the unique environmental factor that we have identified for nonword reading.

Another key issue that we wish to explore is the role of semantics. One possibility that needs to be considered is that the independent genetic variance that we have found to be associated with irregular word reading in fact reflects the genetic effect of semantics. That is, as irregular words activate semantics, while nonwords do not, this semantic factor accounts for the necessity to propose independent genetic influences on these two reading skills (a hypothesis consistent with connectionist models such as those of Plaut et al., 1996). This possibility can be explicitly examined within a twin sample by separately assessing semantic performance in addition to irregular word and nonword reading (for instance by verbally asking subjects to define words). If irregular word reading cannot be achieved without semantics, this measured semantic performance should load on the same genetic factor that explains irregular word reading. If, on the other hand, semantics is a cognitive module independent of lexical processing, then a good fit to these three variables should require a new, additional genetic factor.

Another way of tackling this question would be to examine the genetic and environmental influences on *regular*, as compared with irregular, word reading. Both word types would be expected to be influenced by semantics. However, as regular words

can also be read via the same route as nonwords, the reading of regular words should be influenced by a second genetic factor – that found to be associated with nonlexical reading. Therefore, regular word reading, as compared with irregular word reading, should show significant loadings on at least two independent genetic factors and these should be the same genetic factors found to influence irregular and nonword reading. Using the same logic, we can also explore whether the independent genetic influences that we identify may be associated with oral vocabulary skill, which would obviously be involved in regular and irregular word reading, but not nonword reading.

Analyses of this kind, combined with careful attention to the connection between cognitive theories on the one hand and genetically informative data on the other, should prove fruitful in uncovering the complex causal pathways to success and failure in children learning to read.

References

- Bates, T.C., Castles, A., Coltheart, M., Gillespie, N., Wright, M. & Martin, N.G. (2004). Behaviour genetic analyses of reading and spelling: A component processes approach. *Australian Journal of Psychology*, 56, 115–126.
- Bates, T.C., Castles, A., Luciano, M., Wright, M., Coltheart, M. & Martin, N.G. (in press). Genetic and environmental bases of reading and spelling: A genetic dual route model. *Reading and Writing*.
- Boder, E. (1973). Developmental Dyslexia: A diagnostic approach based on three typical reading–spelling patterns. *Developmental Medicine and Child Neurology*, 15, 663–687.
- Bouchard, T.J.J. & Propping, P. (Eds.) (1993). *Twins as a tool for behavioral genetics*. *Life Sciences Research Report*, Vol. 53. Chichester: John Wiley & Sons.
- 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, 49–74.
- Campbell, R. & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject: A developmental case with associated deficits of phonemic processing and awareness. *Quarterly Journal of Experimental Psychology A*, 37A, 435–475.
- Cardon, L.R., Dilalla, L.F., Plomin, R., DeFries, J.C. & Fulker, D.W. (1990). Genetic correlations between reading performance and IQ in the Colorado adoption project. *Intelligence*, 14, 245–257.
- Castles, A. & Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition*, 47, 149–180.
- Castles, A. & Coltheart, M. (1996). Cognitive correlates of developmental surface dyslexia: A single case study. *Cognitive Neuropsychology*, 13, 25–50.
- Castles, A. & Coltheart, M. (2004). Is there a link from phonological awareness to success in learning to read? *Cognition*, 91, 77–111.
- Castles, A., Datta, H., Gayán, J. & Olson, R.K. (1999). Varieties of developmental reading disorder: Genetic and environmental influences. *Journal of Experimental Child Psychology*, 72, 73–94.
- Coltheart, M. (1978). Lexical access in simple reading tasks. In G. Underwood (Ed.), *Strategies of information processing*. London: Academic Press.
- Coltheart, M., Curtis, B., Atkins, P. & Haller, M. (1993). Models of reading aloud: Dual-route and parallel-distributed-processing approaches. *Psychological Review*, 100, 589–608.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108, 204–256.
- Cunningham, A.E., Perry, K.E. & Stanovich, K.E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing: An Interdisciplinary Journal*, 14, 549–568.
- DeFries, J.C. & Fulker, D.W. (1985). Multiple regression analysis of twin data. *Behavior Genetics*, 15, 467–473.
- DeFries, J.C., Fulker, D.W. & LaBuda, M.C. (1987). Evidence for a genetic aetiology in reading disability of twins. *Nature*, 329, 537–539.
- Fisher, S.E. & DeFries, J.C. (2002). Developmental dyslexia: Genetic dissection of a complex cognitive trait. *Nature Reviews, Neuroscience*, 3, 767–780.

- Gayán, J. & Olson, R.K. (2001). Genetic and environmental influences on orthographic and phonological skills in children with reading disabilities. *Developmental Neuropsychology*, 20, 487–511.
- Gayán, J. & Olson, R.K. (2003). Genetic and environmental influences on individual differences in printed word recognition. *Journal of Experimental Child Psychology*, 84, 97–123.
- Hanley, J.R., Hastie, K. & Kay, J. (1992). Developmental surface dyslexia and dysgraphia: An orthographic processing impairment. *Quarterly Journal of Experimental Psychology A*, 44A, 285–319.
- Harm, M.W. & Seidenberg, M.S. (1999). Phonology, reading acquisition, and dyslexia, insights from connectionist models. *Psychological Review*, 106, 491–528.
- Knopik, V.S., Smith, S.D., Cardon, L., Pennington, B., Gayan, J., Olson, R.K. et al. (2002). Differential genetic etiology of reading component processes as a function of IQ. *Behavior Genetics*, 32, 181–189.
- Manis, F.R., Seidenberg, M.S., Doi, L.M., McBride-Chang, C. & Peterson, A. (1996). On the bases of two subtypes of developmental dyslexia. *Cognition*, 58, 157–195.
- Marlow, A.J., Fisher, S.E., Francks, C., MacPhie, I.L., Cherny, S.S., Richardson, A.J. et al. (2003). Use of multivariate linkage analysis for dissection of a complex trait. *American Journal of Human Genetics*, 72, 561–570.
- Mitterer, J.O. (1982). There are at least two kinds of poor readers: Whole word poor readers and recoding poor readers. *Canadian Journal of Psychology*, 36, 445–461.
- Morton, J. & Patterson, K. (1987). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. Patterson & J.C. Marshall (Eds.), *Deep dyslexia*. (2nd edn, pp. 91–118). London: Routledge and Kegan Paul.
- Neale, M.C., Boker, S.M., Xie, G. & Maes, H.H. (2002). *Mx: Statistical modelling*. (6th edn). Richmond, VA: Department of Psychiatry.
- Neale, M.C. & Cardon, L.R. (1992). *Methodology for genetic studies of twins and families*, Vol. 67. Dordrecht: Kluwer Academic Publishers.
- Olson, R.K., Forsberg, H. & Wise, B. (1994). Genes, environment, and the development of orthographic skills. In V.W. Berninger (Ed.), *The varieties of orthographic knowledge, 1: Theoretical and developmental issues*, Vol. 8. (pp. 27–71). Dordrecht: Kluwer Academic Publishers.
- Olson, R.K., Wise, B., Conners, F. & Rack, J. (1989). Specific deficits in component reading and language skills: Genetic and environmental influences. *Journal of Learning Disabilities*, 22, 339–348.
- Plaut, D.C., McClelland, J.L., Seidenberg, M.S. & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103, 56–115.
- Rack, J.P., Snowling, M.J. & Olson, R.K. (1992). The nonword reading deficit in developmental dyslexia: A review. *Reading Research Quarterly*, 27, 28–53.
- Seidenberg, M.S. & McClelland, J.L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523–568.
- Stanovich, K.E., Siegel, L.S. & Gottardo, A. (1997). Converging evidence for phonological and surface subtypes of reading disability. *Journal of Educational Psychology*, 89, 114–127.
- Stevenson, J., Graham, P., Fredman, G. & McLoughlin, V. (1987). A twin study of genetic influences on reading and spelling ability and disability. *Journal of Child Psychology and Psychiatry*, 28, 229–247.
- Tiu, R.D.J., Wadsworth, S.J., Olson, R.K. & DeFries, J.C. (2004). Causal models of reading disability: A twin study. *Twin Research*, 7, 275–283.
- Wagner, R., Torgesen, J. & Rashotte, C. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, 30, 73–87.

Received 8 May 2005; revised version received 3 August 2005.

Address for correspondence: A/Prof Anne Castles, Department of Psychology, The University of Melbourne, Parkville, Vic. 3101, Australia. E-mail: acastles@unimelb.edu.au