Unraveling genetic and environmental components of early literacy: a twin study

A. G. Bus · D. Out

Published online: 26 January 2008
© The Author(s) 2008

Abstract Even though the acquisition of early literacy skills obviously depends on stimuli and incentives in children's environment we may expect that genes define the constraints for acquiring some or all early literacy skills. Therefore behavior genetic analyses were carried out on twin data including 27 identical and 39 same sex dizygotic twins, 4 years old. We focused on a series of early literacy skills selected because they are developing in the focal age range: writing the proper name and mama, rhyming, and alphabetic knowledge. The data for each skill were subjected to structural equation modeling. First, it turned out that even in this young group genetic differences are the primary force in shaping early developing writing skills. Second, bivariate analyses showed that similar environmental and genetic influences are involved across various skills. Third, from secondary analyses on the available twin studies on early literacy it was apparent that the findings of this small-scaled study correspond to outcomes of other large-scale twin studies.

Keywords Emergent literacy · Twin study · Genetics · Environmental component · Structural equation modeling

Although young children acquire early literacy skills like name writing, rhyming and alphabetic knowledge through an environment enriched with written language and literacy-related activities, genetics may be important as well. Of course an environment involving adults who model and elicit literate activities is a minimum condition for literacy development (e.g., Aram & Levin, 2004; Bus, 2001; Bus, van IJzendoorn, & Pellegrini, 1995; Sénéchal, 2006; Sénéchal & LeFevre, 2002). However, not all children may profit to the same extent from environmental

A. G. Bus (✉) · D. Out
Department of Education and Child Studies, Leiden University, Wassenaarseweg 52,
Leiden 2333 AK, The Netherlands
e-mail: bus@fsw.leidenuniv.nl

influences. Whether children benefit from an environment enriched with literacy may also depend on biologically endowed traits for developing components of reading ability or on genetically influenced engagement in print related activities. How sensitive young children are to pick up implicit clues about names in everyday life may depend on a biologically endowed trait for memorizing orthographic knowledge. Likewise children who struggle with executive functions in general or in print related tasks in particular are more likely to withdraw from practice and fail to develop higher levels of skills that come with practice (Crain-Thoreson & Dale, 1992; Olson, 2004). For instance, we can imagine that the necessary practice for name writing may largely depend on children's spontaneous engagement that in its turn provokes more adult scaffolding. This type of genotype-environment correlations may contribute substantially to the total genetic influence which may result in a strong genetic component for at least some aspects of early literacy (McGue, Bouchard, Iacono, & Lykken, 1993). A strong genetic influence would have major implications for prevention programs. Where experience reflects genetically influenced tendencies, children may be vulnerable for failure despite a stimulating environment. Put differently, if genes define the constraints for acquiring some or all early literacy skills some children need more intense support than is required for other children with no inherited difficulties in becoming literate.

Twin studies are an excellent way to estimate to what extent a genetic component is important for the development of individual differences in early literacy (Plomin, DeFries, McClearn, & McGuffin, 2001). Evidently, there is clear support for genetic influences or genotype-environment influences on a particular trait when identical (Monozygotic, MZ) twins resemble each other more than fraternal (Dizygotic, DZ) twins do. We applied behavioral-genetic analyses to our twin data to assess the extent of genetic influences, shared environmental influences and non-shared environmental influences on literacy skills that are developing when children are about four years old. To estimate the influence of the shared environment component (c^2), the genetic component (h^2), and the unique environment component (e^2) in this study the data for each skill were subjected to structural equation modeling by use of the *Mx* statistical modeling package (Neale, Boker, Xie, & Maes, 1999).

Insofar environmental components affect the early literacy skills that we assessed at the age of four, we wonder whether the same aspects of children's environment explain individual differences in skills that have a significant environmental component. The strong relation between the parents' educational level and children's literacy level when schooling begins suggests that higher-educated parents, contrary to the low-educated ones, may be more inclined to focus their children's attention on literacy and initiate relevant activities and games (Brooks-Gunn, Han, & Waldfogel, 2002). As a result, their children have a better start when reading instruction at school begins. Hence we expect considerable overlap in the shared environmental factors that relate to early literacy skills. In the same vein, we wonder whether there is also overlap in genetic components. Not all young children profit equally from modeling behavior or instruction by adults because some may fail genetically mediated child characteristics that play a role in becoming literate. For instance, children may lack in the capacity to memorize 'chunks' of letters or

not have sufficient control of the executive capacity to sustain attention and concentrate on execution of literacy-related tasks like name writing. It is plausible that self-selection of literacy-related experiences is likely to reflect underlying genetically influenced tendencies. A bivariate model was applied to test whether environmental or genetic components explain the observed phenotypic covariances between literacy skills and whether the same environmental or genetic component are relevant for both early literacy skills (Boomsma, Busjahn, & Peltonen, 2002).

Twin studies involving children in the age range of four to five (Lemelin et al., 2007; Samuelsson et al., 2005) so far mainly focus on broad school readiness measures (color and shapes, spatial recognition, numbers, and letters) or on components that are also assessed in older age groups such as phonological awareness and language skills. Researchers are less inclined to study emerging reading and writing skills that seem especially eligible for assessing differences in reading and writing skills developing in this early age range and that reflect the kind of learning about literacy that takes place in young children (Paris, 2005). Four-year-olds practice *writing* (the proper name and other names like mama) and they develop basic skills for *reading* (e.g., knowing that writing is composed of letters and that letters relate to sounds in spoken words). To check on the validity and replicability of our small twin study, we test whether the results of this study converge with the results of previous studies in a similar age range. Data allowed us to make this comparison for only one of the skills, namely phonological awareness. From a series of studies in the United States, Australia and Scandinavia focusing on young children it appears that phonological awareness has a strong environmental component, in sharp contrast with studies in older age groups. The latter studies convincingly demonstrate that genetics are important not only to reading outcomes, but also to components of reading ability, such as phonological awareness, decoding and spelling (Harlaar, Dale, & Plomin, 2007; Petrill, Deater-Deckard, Thompson, DeThorne, & Schatschneider, 2006).

Thus, this twin study with 4 years old children was designed to test the following hypotheses:

1. Are environmental influences main components in explaining early literacy including early reading and writing skills? It is conceivable that environmental influences hardly differentiate between children since nowadays literacy is a basic aspect of children's environment thereby explaining hardly any variance.
2. Are genetic influences essential for developing early reading and writing skills? Biologically endowed traits may interfere with benefiting from activities promoted by a literate environment. Taking advantage from environmental influences may be constrained by genetically determined behaviors such as executive functions resulting in a strong genetic component.
3. In so far as environmental influences contribute to the development of various aspects of early literacy we wonder whether there is overlap in environmental components involved in various aspects of early literacy. Even though we can imagine that separate literacy skills require specific instruction it is also imaginable that literate families promote a whole range of relevant activities thereby causing the family's literacy level to be a main component.

4. If there are genetic influences on the various aspects of literacy we ask whether they overlap. Genes may define the constraints for acquiring early reading and writing skills as a result of which there may not be much overlap between both types of knowledge. When for instance genetically influenced executive functions are essential, we may expect that genes define the constraints for acquiring some or all early literacy skills.
5. Are the outcomes of the present small study comparable with findings in larger studies in the same age range? To test whether the outcomes in this study are similar to those in three other studies, we carried out a secondary analysis on the extant data.

Method

Participants

All children were sampled through the Netherlands Twin Register (Boomsma, Orlebeke, & Van Baal, 1992). The families were predominantly middle-class. On a scale for maternal educational level, ranging from 1 to 7, the mean level was 4.7 ($SD = 1.6$), implying 6 years of secondary education. The mean age of the mothers was 32.1 years ($SD = 3.4$). The 27 identical (MZ) and 39 same sex fraternal twins (DZ), 54.9 months ($SD = 2.4$) and 55.9 (2.0) respectively, had entered kindergarten which starts in the Netherlands at the day children become four but did not yet receive reading instruction, as is common in the Netherlands.

Procedure

Twins were assessed in their homes by research assistants when they were on average about four and half years old. Both twins were tested individually during visits at the children's home. While one of the children was tested by the experimenter the other child was not present in the room. The literacy tasks were part of a broader set of tests and tasks not discussed here. Testing always started with writing the proper name, mama and some other words (to assess alphabetic knowledge) and ended with rhyming.

Tests

We selected a series of reading and writing tests, partly overlapping with other twin studies carried out in the same age range (rhyming), and partly unique for this study (writing tasks):

- (1) Writing the proper name
- (2) Writing mama

- (3) A dictation of words that were selected because they were likely to be unknown: KAAS [cheese] and ZON [sun]
- (4) Rhyme production as indicator of phonological awareness

Testing literacy skills when children are this young is tricky to the extent that tests are susceptible to ceiling and floor effects (Paris, 2005). We carefully selected tests that were eligible for differences in this age range and at the same time reflected the kind of activities that interest children in this young age group. Based on previous studies we estimated that each of these skills would be normally distributed in a group of 4-year-olds. Writing the proper name starts around two-and-half and is often completed around five (Levin, Both-de Vries, Aram, & Bus, 2005). Alphabetic writing, i.e., children use conventional letters that sometimes represent sounds of the dictated words, develops between three and six years of age (Levin & Bus, 2003). In earlier stages children mainly produce written-like forms such as pseudo-cursive scribble or strings of pseudo-letters. Writing other names such as mama is more similar to the proper name than to other words (Both-de Vries, 2006). For name writing children need to memorize 'chunks' of letters. Based on various studies we assumed that rhyming would be a better indicator for phonological awareness in this age range than more complex skills such as phoneme identification or deletion. Hindson et al. (2005), for instance, report that 4-year-olds scored at chance level on phoneme identification whereas scores on rhyming tasks by the same group remained above chance level. The minimum intercoder reliability (Spearman correlations) for the three writing tasks was .96. The correlation between KAAS and ZON was substantial ($r = .83$) which indicates that the error variance is low. Cronbach alpha for the rhyming task was .78.

The test scores confirmed that the present sample was midway to mastery of the selected tests and scales. On average children scored about two out of four words correctly on the rhyming test ($M = 56.2\%$, $SD = 39.8$). The scores were bimodal with 25% unable to produce any rhyming word and 28% scoring all four items. Writing the proper name, writing mama and writing the other dictated words was coded on a 6-point scale: (1) pseudo-cursive, (2) conventional letters not related to sounds in the spoken word, (3) one phonetic letter, (4) two phonetic letters, (5) invented spelling (readable but not yet correct), (6) conventional spelling (Levin & Bus, 2003). In accordance with previous studies (e.g., Levin et al., 2006) scores for name writing ($M = 3.15$, $SD = 1.55$) excelled those for mama ($M = 2.07$; $SD = 1.27$) and other words ($M = 1.99$; $SD = 1.05$). Name writing was often conventional (33%) whereas writing mama and other words was rarely conventional, 10% and 3% respectively. The distributions of the scores on the 6-point scale were normal. These variables were therefore treated as continuous variables in the subsequent analyses. Inspection of the box plots did not reveal any outliers.

Data-analysis

Univariate models were fitted on raw data in order to quantify and test the significance of additive genetic influences (A), shared (C) and unshared environmental influences (E) for each of the four literacy skills (Plomin et al., 2001). The variance/covariance

Table 1 The variance/covariance between identical (MZ) twins

	Twin 1	Twin 2
Twin 1	$A + C + E$	$A + C$
Twin 2	$A + C$	$A + C + E$

Note: A, genetic influences; C, shared environmental influences; E, unshared environmental influences

between identical (MZ) twins is parameterized in Table 1. The covariance between Twin 1 and Twin 2 is $A + C$, because identical (MZ) twins share 100% of the same genes and, by definition, 100% of the shared environment. For fraternal (DZ) twins, the variance is the same as for MZ twins ($A + C + E$), but the covariance between DZ twins is parameterized as $0.5A + C$, because fraternal twins share on average 50% of their genes and 100% of the shared environment.

The parameters of the genetic influences, shared and unshared environmental influences were estimated using maximum likelihood estimation procedures. We tested in succession an AE-model, CE-model, and E-model, each time omitting one or two of the three components. The fit of these simplified models was compared to the fit of the full ACE-model using likelihood-ratio tests. Each model was given an overall $-2 \log$ likelihood goodness of fit; the difference in fit between two models was represented by χ^2 . A significant χ^2 indicates that the fit of the simpler model is significantly worse compared to the ACE model and that preference should be given to the full model. When the change of fit is non-significant, the simpler and more parsimonious model is preferable to the full ACE model. Our estimates for the heritability, shared and unshared environmental influences of the phenotype were based on this final model.

Bivariate genetic analysis focuses on the covariance (correlation) between two traits and estimates the genetic and environmental contributions to the observed covariance (Boomsma et al., 2002). The essence of bivariate genetic analyses is the comparison of the cross-trait/cross-twin correlation between MZ twins and DZ twins. For these analyses, the phenotypic covariation between two early literacy skills is decomposed into genetic and environmental components (a Cholesky decomposition). When the covariance between one literacy skill in one twin and the other literacy skill in the co-twin is higher for MZ than for DZ twins, genetic factors mainly explain the observed phenotypic covariation between the two literacy skills. In a bivariate design it is possible to investigate the exact extent to which the same genetic or environmental influences are involved in both literacy skills. This is indicated by the correlation between latent components of the skills (genetic correlation r_g , shared environment correlation r_c , unshared environment correlation r_e).

Results and discussion

Univariate analyses

For each of the four literacy skills, several univariate models were evaluated. For instance, correlations for rhyming were: $r_{MZ} = .816$ ($p < .01$) and $r_{DZ} = .652$

Table 2 Fit comparison of nested models for four literacy skills

Model	Fit statistic (-2ll)	df	Change in fit (χ^2)	Change in df	p-value
<i>Rhyming</i>					
ACE	1248.442	123			
CE	1250.104	124	1.662	1	0.197
AE	1253.671	124	5.229	1	0.022
E	1294.856	125	46.414	2	<.001
<i>Writing the proper name</i>					
ACE	606.999	122			
CE	610.150	123	3.151	1	.0076
AE	606.999	123	0	1	–
E	623.039	124	16.041	2	<.001
<i>Writing mama</i>					
ACE	577.942	112			
CE	584.129	113	6.187	1	0.013
AE	577.942	113	0	1	–
E	592.732	114	14.790	2	0.001
<i>Writing dictated words</i>					
ACE	515.536	109			
CE	515.803	110	0.268	1	0.605
AE	518.344	110	2.808	1	0.094
E	534.516	111	18.908	2	<0.001

($p < .01$); see Table 3. The analyses summarized in Table 2 show that the fit of the CE model is not significantly worse compared to the ACE model whereas both AE and E models show a significantly poorer fit to the data. Therefore, the CE model was preferred to the AE and E models. Based on this final model, 73% of the variance in rhyming scores was due to shared environmental influences whereas the remaining variance was explained by unique environmental influences and measurement error. The genetic component was not significant for rhyming. Likewise, the CE model was preferred to the AE and E models for writing words. For the proper name and mama the AE model was best fitting. Both the CE and AE models did not yield a significantly worse fit but since the de chi-square did not increase for the AE model in comparison with the ACE model the AE model seems the best choice.

Table 3 presents estimates and 95% confidence intervals of the A, C, and E component for each of the assessed skills. For individual differences in writing the proper name and writing mama the genetic component was most important, but for rhyming and alphabetic knowledge the shared environmental component explained most of the observed variance. Note that as a consequence of the small sample the 95% confidence intervals were substantial. At first sight it is rather surprising that the AE-model was the most parsimonious model for writing the proper name and writing mama whereas the CE-model was accepted as the final model for alphabetic knowledge manifested in the dictated words. Apparently variation in writing the

Table 3 Outcomes of univariate modeling of four literacy skills

	<i>r</i> MZ	<i>r</i> DZ	A (%)	C (%)	E (%)
Rhyming	.816	.652	–	73 (58–82)	27 (18–42)
Alphabetic knowledge	.599	.541	–	61 (39–76)	39 (24–61)
Name writing	.633	.275	63 (37–79)	–	37 (21–63)
Writing mama	.719	.214	73 (44–86)	–	27 (14–56)

Note: 95% confidence intervals in parentheses

proper name and mama mainly depends on genetic predispositions whereas individual differences in writing other words are mainly explained by shared environmental influences. This contrast makes sense when we realize that in this early stage alphabetic writing appears to be greatly influenced by adults modeling and instructing writing (Sénéchal, 2006; Sénéchal & LeFevre, 2002), whereas the ability to write names mainly results from practice initiated by children. Genetic predispositions for memorizing orthographic knowledge may result in differences in amount of practice, thereby causing a genotype-environment correlation to be effective.

Bivariate analyses

Are similar or different aspects of the shared-environment involved in learning to rhyme and in alphabetic writing? Are the same genes involved when attempts are made to write the proper name or mama? To explore these questions we analyzed rhyming and alphabetic writing simultaneously in a bivariate analysis, in which a CE model was selected for both early literacy skills. In contrast, for the bivariate analysis of writing the proper name and mama, an AE model was specified. Main results are summarized in Table 4. Under cross-trait/cross-twin correlations we report the correlations between the first literacy skill in the first twin and the second literacy skill in the second twin and, vice versa, the first literacy skill in the second twin and the second literacy skill in the first twin. Probably as a consequence of the small sample the correlations are somewhat different but they all go in the same

Table 4 Results of bivariate analyses to estimate genetic and environmental contributions to covariance between early literacy skills

	Cross trait cross twin correlation		Percentage of covariance explained by		Correlations	
	MZ	DZ	C/A (%)	E (%)	C/A	E
Rhyme-symbolic	.389/–.083	.610 ($p < .01$)/.357 ($p < .01$)	C = 99	1	$r_C = .62$	–.01
Name-mama	.734 ($p < .01$)/.565 ($p < .01$)	.284/–.003	A = 86	14	$r_A = .84$	.32

direction. The second column indicates which part of the covariance can be attributed to shared environment (C), genetic factors (A), or to unique environment including measurement error (E). In the third column we report the strength of the correlation between latent factors representing genetics, shared and unique environment. These correlations indicate to what extent the effects can be ascribed to the same environmental factors or genes.

As can be derived from Table 4, there was considerable overlap in the shared environmental factors related to individual differences in alphabetic knowledge and in rhyming ($r_c = .62$). Both alphabetic knowledge and rhyming are stimulated by similar environmental factors. Parents who teach rhymes and elicit rhyming also model how to write words with letters. There was considerable overlap in the genetic factors that influence writing the proper name and mama ($r_g = .84$). This commonality is in line with the hypothesis that there are genetically influenced constraints to practice writing (Samuelsson et al., 2005). Some children may be more inclined to withdraw from practicing than their peers.

Secondary analyses

We also tested to what extent the present results based on a small sample of MZ and DZ twins reveal findings similar to other comparable twin studies. We selected studies where similar measures were applied to young subjects in the same age range. From four available studies reporting appropriate information about children in the preschool age, three studies, one from Scandinavia, one from Australia and one from the USA, were included in these secondary analyses. A fourth study (Kovas et al., 2005) was excluded because the focus in this study is on language impaired children. To test whether the outcomes in the three remaining studies and our study are similar, we carried out a secondary analysis (cf. Bartels, van den Berg, Sluyter, Boomsma, & de Geus, 2003). After testing whether the studies could be combined, the effects of the genetic component (Table 5, model 6) and shared environment component (model 5) were evaluated. Phonological awareness was assessed in all four studies and therefore most appropriate for a secondary analysis. Other measures in the Samuelsson et al. (2005) study such as print awareness did not overlap with measures in our study and were therefore unsuitable for secondary analysis.

As described above, structural equation modeling was employed to fit several nested models to the four MZ and four DZ correlations of phonological awareness by maximum likelihood estimation procedures. The fit of the model with fixed A- and C-components (models 2, 3, and 4) was not significantly worse compared to the fit of the full model in which the A and C components were all different (model 1), indicating that the estimates of the A and C component in these four studies are very similar and that the results can be therefore combined. In subsequent analyses, an AE-model, CE-model and E-model were fitted in order to test the significance of the A and C components. Furthermore, it was tested whether both the A and C component significantly contributed to individual differences. It was found that leaving out the A or C components resulted in a significantly worse fit to the data

Table 5 Analyses of phonological awareness including three studies reported in Samuelsson et al. (2005) and our data

Model	χ^2	df	$\Delta \chi^2$	Δ df	p	Compared to model	A	C	E
1. ACE, all different	0.000	16					34–60%	20–48%	18–26?
2. ACE, A equal	0.958	19	0.958	3	0.812	1	48%	26–35%	17–26%
3. ACE, C equal	1.343	19	1.343	3	0.719	1	42–52%	31%	17–26%
4. ACE, A equal, C equal	5.408	22	5.408	6	0.493	1	47%	32%	21%
5. AE, A equal, no C	21.864	23	16.456	1	<.001	4	80%	–	20%
6. CE, C equal, no A	53.408	23	47.999	1	<.001	4	–	67%	33%
7. E, no C and no A	458.092	24	452.684	2	<.001	4	–	–	100%

compared to an ACE model which suggests that both genetic factors and the shared environment contribute significantly to individual differences in rhyming. The results confirm our finding that the shared environment is an important factor but genetic influences are significant as well. As this estimate was based on a substantially larger group, the estimates in Table 5 may be more valid indices of genetic and shared environmental influences of phonological awareness.

Conclusions

Shared environmental influences are significant and substantial for two out of the four skills that develop early: rhyming and alphabetic knowledge. This result suggests that special activities are required to elicit rhyming and to make children aware that words are composed of letters and how letters relate to sounds. The present outcomes show *that* environmental components explain these early literacy skills but not *which* facets of the environment are essential to promote learning. We hypothesize that children acquire alphabetic knowledge when adults are likely to explain how to write words (Sénéchal, 2006; Sénéchal & LeFevre, 2002). Learning to rhyme, on the other hand, may depend on the presence of songs and on games initiated by caregivers or on other incentives that focus children's attention on sounds instead of meaning. For instance, Maclean, Bryant, and Bradley (1987) revealed evidence for the idea that knowledge of songs and lullabies predicts the development of phonological skills. From bivariate analyses appears that a similar environment promotes rhyming and alphabetic knowledge at the age of four. The most plausible interpretation of this finding may not be that both skills result from similar activities. We may rather assume that families that are more similar in educational level are more inclined to promote relevant activities (Parker, Boak, Griffin, Ripple, & Peay, 1999).

Genetic differences and commonalities are important in shaping familial resemblance in reading skills of older children (e.g., Harlaar et al., 2007; Petrill

et al., 2006). The present study gives evidence for the hypothesis that genetic differences are also important in the earliest stages of becoming literate. We found that early developing skills such as writing names mainly depend on genetic factors. Heritability appears to account for 63–73% respectively of the individual differences in writing the proper name and mama. It is not surprising that the bivariate analysis reveals considerable overlap in genes related to both writing skills ($r_c = .84$). This finding may indicate that the same genetically endowed traits are involved in memorizing ‘chunks’ of letters. This finding may also make sense in light of the theory that genetics and some experiential effects are passively correlated from the earliest stages of development.

Apparently, an environment where writing is omnipresent and adults model writing does not meet children’s needs to promote their ability to write names. Though environmental influences are indispensable the finding that genetic factors are most important for both writing the proper name and mama indicates that not all children benefit to the same extent from an environment enriched with writing. It seems therefore a plausible assumption that genetically influenced predispositions affect young children’s engagement in name writing activities. In a similar vein of argumentation, Samuelsson et al. (2005) argued that not all early literacy-related experiences equally depend on environmental influences. They report that the frequency of reading mainly depends on shared environment whereas the frequency of looking in books or asking to be read to also depends on a genetic component. Analogously we hypothesize that not all children benefit to the same extent from an environment where writing is present in plenty due to genetically influenced predispositions to be more or less engaged in writing names.

It is a remarkable outcome that memorizing names is a biologically endowed trait whereas shared environmental influences are significant for the precursors of decoding, namely sensitivity to sounds (rhyming) and alphabetic knowledge. This pattern fits the finding in older age groups that memorizing orthographic knowledge often constrains the development of reading and writing skills. Many older children make spelling mistakes that are ‘phonetic’, i.e. the spelling is wrong, but it does convey the sound of the words the children are trying to write (Frith, 1980). They make mistakes because they concentrate on the letter-sound relationship and do not remember or use their memory for whole words or for ‘chunks’ of letters. When they read words these children attend mostly to salient letters that have a systematic relationship with the word’s pronunciation. In some degree this strategy may suffice for reading, which can proceed using only partial cues, but is insufficient to ensure satisfactory spelling performance (Snowling, 2000). In other words, the present findings indicate that from the very start of literacy a biologically endowed trait for memorizing orthographic knowledge may interfere with reading and writing development.

Pooling four studies carried out in the United States, Australia, Scandinavia, and the Netherlands, all exploring the influence of the shared-environment component, the genetic component, and the unique-environment component, the same model applies to all four studies when we focus on phonological awareness. Despite the small number of twins involved in the Dutch study reported here the results match

with the three other studies. Although we could not execute secondary analyses for all four variables that were assessed in this study, the conclusion that the results for phonological awareness were similar to the outcomes of other twin studies in the same age range adds to the validity and replicability of the present findings.

Limitation and future directions

A relatively homogeneous sample like the one in this study (mainly composed of higher educated families) may insufficiently highlight environmental influences. The environmental component may become stronger when we would include families that are more variable in the presence of written language and in modeling and instruction by parents and other adults. On the other hand, in so far the present study revealed an environmental influence in the present age range we do not expect that this influence remains as children grow older. Effects of genes may ‘amplify’ as children are in a stage that they have more control over activities and can take the first step towards practicing. Because self-selection of activities is likely to reflect underlying genetically influenced tendencies, one would expect the heritability of literacy-related skills to increase with age as may appear from longitudinal studies starting at the age of four and continuing until formal reading instruction begins. Further research is also needed to test the hypothesis that genetically influenced predispositions like executive capacities or their memory for ‘chunks’ of letters determine to what extent children are engaged in writing names thereby explaining the finding that some early literacy skills are mainly genetically influenced.

Practical implications

The present results are at odds with the prevailing idea that children acquire early literacy skills mainly through environmental influences. Particularly where name writing skills are involved the genetic component matters more than environmental influences. Besides a stimulating environment, genetic differences are important in defining the constraints for acquiring early literacy skills at the age of four. For instance, despite of supportive environmental influences, children may discontinue practicing because they fail the executive capacities to start, plan and complete activities like name writing or because their attempts to memorize the orthography of words are unsuccessful. The present finding demonstrating the relevance of genetics to early literacy skills like name writing has far-reaching consequences for interventions at the age of four. Where genetic differences are the main constraint, we may need to provide more instructional support than is required for children with no heritable constraints. To ensure greater and more intense practice of writing skills children may need a support system of prompts, hints, and feedback that normally is not provided with a high intensity in educational settings of young children. Computer-assisted early reading experiences may therefore be a potentially powerful solution for those children who need frequent and intensive scaffolding of their learning (Meltzer, 2007).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Aram, D., & Levin, I. (2004). The role of maternal mediation of writing to kindergartners in promoting literacy achievements in school: A longitudinal perspective. *Reading and Writing: An Interdisciplinary Journal*, 17, 387–409.
- Bartels, M., van den Berg, M., Sluyter, F., Boomsma, D. I., & de Geus, E. J. C. (2003). Heritability of cortisol levels: Review and simultaneous analysis of twin studies. *Psychoneuroendocrinology*, 28, 121–137.
- Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature Reviews Genetics*, 3, 872–882.
- Boomsma, D. I., Orlebeke, J. F., & Van Baal, G. C. M. (1992). The Dutch twin register: growth data on weight and height. *Behavior Genetics*, 22, 247–251.
- Both-de Vries, A. (2006). It's all in the name. Early writing: From imitating print to phonetic writing. Unpublished Dissertation, Leiden University.
- Brooks-Gunn, J., Han, W.-J., & Waldfogel, J. (2002). Maternal employment and child cognitive outcomes in the first three years of life: The NICHD study of early child care. *Child Development*, 73, 1052–1072.
- Bus, A. G. (2001). Early book reading in the family: A route to literacy. In S. Neuman & D. Dickinson (Eds.), *Handbook on research in early literacy* (pp. 179–191). New York: Guilford Publications.
- Bus, A. G., Van IJzendoorn, M. H., & Pellegrini, A. D. (1995). Storybook reading makes for success in learning to read. A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65, 1–21.
- Crain-Thoreson, C., & Dale, P. S. (1992). Do early talkers become early readers? Linguistic precocity, preschool language, and emergent literacy. *Developmental Psychology*, 28, 421–429.
- Frith, U. (1980). *Cognitive processes in spelling*. London: Academic Press.
- Harlaar, N., Dale, P. S., & Plomin, R. (2007). From learning to read to reading to learn: Substantial and stable genetic influence. *Child Development*, 78, 116–131.
- Hindson, B., Byrne, B., Shankweiler, D., Fielding-Barnsley, R., Newman, C., & Hine, D. (2005). Assessment and early instruction of preschool children at risk for reading disability. *Journal of Educational Psychology*, 97, 687–704.
- Kovas, Y., Hayiou-Thomas, M. E., Oliver, B., Dale, P. S., Bishop, D. V. M., & Plomin, R. (2005). Genetic and environmental origins of diverse language skills in 41/2-year-old twin boys and girls. *Child Development*, 76, 632–651.
- Lemelin, J. P., Boivin, M., Forget-Dubois, N., Dionne, G., Séguin, J. R., Brendgen, M., Vitaro, F., Tremblay, R. E., & Pêrusse, D. (2007). The genetic-environment etiology of cognitive school readiness and later academic achievement in early childhood. *Child Development*, 78, 1855–1869.
- Levin, I., & Bus, A. G. (2003). How is emergent writing based on drawing? Analyses of children's products and their sorting by children and mothers. *Developmental Psychology*, 39, 891–905.
- Levin, I., Both-de Vries, A. C., Aram, D., & Bus, A. G. (2005). Writing starts with own name writing: From scribbling to conventional spelling of Israeli and Dutch children. *Applied Psycholinguistics*, 26, 463–477.
- Maclean, M., Bryant, P., & Bradley, L. (1987). Rhymes, nursery rhymes, and reading in early childhood. *Merrill-Palmer Quarterly*, 33, 255–281.
- McGue, M., Bouchard, T. J., Iacono, W. G., & Lykken, D. T. (1993). Behavioral genetics of cognitive ability: A life-span perspective. In R. Plomin & G. E. McClearn (Eds.), *Nature nurture psychology* (pp. 59–76). Washington, DC: American Psychological Association.
- Meltzer, L. (2007). *Executive function in education: From theory to practice*. New York: The Guilford Press.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (1999). *Mx: Statistical modeling* (5th ed.). Department of Psychiatry.
- Olson, R. K. (2004). SSSR, environment, and genes. *Scientific Studies of Reading*, 8, 111–124.
- Paris, S. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, 40, 184–204.

- Parker, F. L., Boak, A. Y., Griffin, K. W., Ripple, C., & Peay, L. (1999). Parent-child relationship, home learning environment, and school readiness. *School Psychology Review*, 28, 423–425.
- Petrill, S. A., Deater-Deckard, K., Thompson, L. A., Schatschneider, C., & DeThorne, L. S. (2006). Longitudinal genetic analysis of early reading: The Western Reserve Reading Project. *Reading and Writing*, 20, 127–146.
- Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavior genetics* (4th ed.). New York: Worth.
- Samuelsson, S., Byrne, B., Quain, P., Wadsworth, S., Corley, R., DeFries, J. C., Willcutt, E., & Olson, R. (2005). Environmental and genetic influences on prereading skills in Australia, Scandinavia, and the United States. *Journal of Educational Psychology*, 97, 705–722.
- Sénéchal, M. (2006). Testing the home literacy model: parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Scientific Studies of Reading*, 10, 59–87.
- Sénéchal, M., & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skills: A five-year longitudinal study. *Child Development*, 73, 445–460.
- Snowling, M. J. (2000). *Dyslexia* (2nd ed.). Oxford: Blackwell.