

Unraveling the Relation Between Reading Comprehension and Print Exposure

Florina Erbeli 
Texas A&M University

Elsje van Bergen
Vrije Universiteit Amsterdam

Sara A. Hart
Florida State University

The purpose of this study was to test the directionality of influence between reading comprehension (RC) and print exposure (PE), thereby estimating genetic and environmental effects of this relation. The sample consisted of 910 twins in fourth through ninth grades ($M_{\text{age}} = 12.33$ years, $SD = 1.41$) from the Florida Twin Project on Reading, Behavior, and Environment. Using direction-of-causation model in a twin design, results supported a direction of influence running from RC to PE. This relation was underpinned by genetic and environmental factors of RC as well as PE. Implications for reading education are discussed.

Reading represents an important public health outcome. Skilled readers have, on average, better health (e.g., DeWalt, Berkman, Sheridan, Lohr, & Pignone, 2004) and higher level of income (e.g., McLaughlin, Speirs, & Shenassa, 2014) compared to less skilled readers. Proficient reading comprehension (RC) is acquired through extended reading practice (National Reading Panel, 2000; Stanovich, 1986). However, does extended print exposure (PE) in and of itself make children better reading comprehenders? The National Reading Panel (2000) argues that “it could be that if you read more, you will become a better reader; but it also seems possible that better readers simply choose to read more. So which is it?” (National Reading Panel, 2000, p. 3/21) The idea that RC and PE are closely interrelated has prompted researchers to study the relation between the two constructs extensively, mostly from a correlational perspective though (Mol & Bus, 2011). This study extends this research area by testing direction of influences between RC and PE—thus, answering the question of whether RC facilitates PE, vice versa, or whether both constructs foster each other reciprocally. Moreover, we estimate

the directionality of relation in a natural experiment using a sample of twins, which allows us to account for genetic and environmental effects on both reading constructs and their relation.

RC is defined here as academic achievement on RC measures, and PE as familiarity with children’s book authors as measured by the Author Recognition Test (ART). The term PE has been generally related to the concept of voluntary reading outside the curriculum required reading (e.g., Martin-Chang & Gould, 2008). For this purpose, the current report lists terms such as voluntary reading, recreational reading, nonincidental reading, leisure or free time reading, extracurricular reading, amount of reading as synonyms for PE (National Reading Panel, 2000).

Theoretical Perspectives of the Relation Between RC and PE

Under a simple view of directionality of influences (excluding other potential confounders), the relation between RC and PE can be conceptualized in three ways. First, there could be a bidirectional relation ($RC \leftarrow \rightarrow PE$). This would mean that proficient reading comprehenders read more often, which, in turn, improves their RC. This option is in line with the abundant literature on the spiral causality or cumulative advantage and disadvantage

This research was supported, in part, by a grant from the National Institute of Child Health and Human Development (P50 HD052120). Views expressed herein are those of the authors and have neither been reviewed nor approved by the granting agencies.

Correspondence concerning this article should be addressed to Florina Erbeli, Department of Educational Psychology, 4225 Texas A&M University, College Station, TX 77843. Electronic mail may be sent to erbeli@tamu.edu.

phenomenon (Stanovich, 1986), more often known as the Matthew effect in reading. Second, it could be that extensive PE is largely a consequence of effortless RC. Good reading comprehenders are likely to seek out opportunities to read, whereas poor readers find reading an unrewarding experience, resulting in their avoidance of reading (RC → PE). This view is supported by research showing that cognitive and reading related prereading skills prior to school entry substantially predict reading achievement years later (e.g., Erbeli, Hart, & Taylor, 2018; van Bergen, De Jong, Maassen, & van der Leij, 2014). Third, it is possible that PE facilitates RC. The more a child reads and engages with print materials, the more that enhances her RC (RC ← PE). Support for this notion can be inferred from Perfetti's Verbal Efficiency Theory (Perfetti, 1985). The theory posits, among others, that higher level reading processes integral for RC, such as integrating propositions, using cognitive and meta-cognitive strategies, and activating background schemas, can be automatized through extended reading practice. However, it is also likely that the association between RC and PE is purely correlational in nature, explained merely by common genetic and/or environmental influences. This option is demonstrated in a meta-analysis by Mol and Bus (2011). The meta-analysis examined average weighted effect sizes between various reading outcomes and PE and estimated a moderate correlation of .36 between RC and PE in Grades 1 through 12. If the relation between constructs of our interest is, in fact, correlational, then such a scenario would preclude us from making conclusions about the directionality of influences and would mean that variabilities in RC and PE merely overlap.

Empirical Support for the Association Between RC and PE

In educational research, several designs are used to answer the question of directionality of influences. The gold standard to address the epitome of directionality of influence—causality—is experimental designs such as randomized control trials (RCTs). RCTs answer research questions that are linked to artificially conditional settings. As such, not all questions related to the direction of influences are feasible using such a design. The goal of this study was to estimate the directionality of influences between two important reading constructs within a natural experimental paradigm rather than artificial experimental method. One of the approaches that can in part address possible

pathways of directionality in naturally occurring situations is the longitudinal design.

One of the first studies within this realm was a study by Aarnoutse and van Leeuwe (1998). The study followed second through sixth graders to examine the development of RC and reading frequency in tandem. The authors reported that the best fitting longitudinal model was the one in which RC influenced reading frequency rather than the other way around. Such a finding was further demonstrated in another study. Leppänen, Aunola, and Nurmi (2005) designed a cross-lagged panel study and found that the extent to which children were proficient in reading fluency at the end of first grade had an effect on how likely they were to engage in out-of-school reading 1 year later. A reverse effect from out-of-school reading to reading fluency was not significant. The same type of model was tested also for the word recognition and out-of-school reading. In this model, both paths indicating directionality of influences were significant. However, the effect was stronger in magnitude for word recognition influencing out-of-school reading ($\beta = .30$) than vice versa ($\beta = .13$). In the latest study of this kind, Torppa et al. (2019) tracked Finnish students from the age of 7–16. Utilizing a random intercept cross-lagged panel model, they demonstrated that in elementary school grades poorer RC predicted less leisure reading. In contrast, in middle school active book reading was reciprocally associated with RC, with the effect of leisure reading on RC being stronger than the other way around (Torppa, Niemi, Vasalampi, Lerkkanen, Tolvanen, & Poikkeus, 2019).

Taken together, the relatively scarce existing findings from this research speak more in favor of the directionality from reading outcomes to PE rather than vice versa, even though there is one study (Torppa et al., 2019) that reported an opposite finding. While these preliminary, yet seminal studies are important first steps aiming to delineate a possible direction of influences, there is one challenge in interpreting such inferences—genetic and environmental confounding. A clear delineation of what reading construct is antecedent and what is consequent might, therefore, be harder to permit because of the putative etiological confounding that could have induced or obscured the association.

Underlying Genetic and Environmental Influences of the Relation Between RC and PE

Research readily demonstrates that both RC and PE are genetically and environmentally influenced

traits with overlapping genetic and environmental effects, giving rise to a covariation (e.g., Harlaar, Dale, & Plomin, 2007; Harlaar, Deater-Deckard, Thompson, DeThorne, & Petrill, 2011; Little, Haughbrook, & Hart, 2017; van Bergen, Snowling, de Zeeuw, van Beijsterveldt, Dolan, & Boomsma, 2018). Genetic and environmental correlational estimates, however, cannot speak to the question of directionality of influences between the constructs. Thus, a complementary approach as a natural experiment overcoming the confounding of genes and environment might be needed to assist in resolving the relation. Researchers have developed genetically informed methods, which address directionality of inferences in observational research. These methods apply cross-sectional designs combined with twin samples, serving as natural experiments. Although these methods cannot rule out all possible alternatives to conclude a strong direction of influences, they allow for an investigator to rule out many.

One of such genetically sensitive method in the twin study is called a direction-of-causation model (Heath, Kessler, Neale, Hewitt, Eaves, & Kendler, 1993). The method can best distinguish between different models of directionality if the following three conditions are met (Heath et al., 1993). First, modes of genetic and environmental influences between constructs are different. In practice that means that one construct shows a stronger influence of, say, genetic effects, whereas the other construct is determined largely by environmental factors. Second, constructs are created using multiple indicator variables as to reduce the presence of measurement error, which can bias the estimates. Third, the association between constructs can be weak as long as the first condition is met. The logic behind the direction-of-causation model is as follows. If genetic and environmental effects of the causal construct (either RC or PE in our case) were indirectly mediated through the path onto the outcome construct (again, either of the reading constructs), then this would indicate that one part of the outcome construct's variance (the common variance shared with the causal construct) can be explained from the variation in the causal construct. The other part of the outcome construct variance (the residual variance) which cannot be explained by the common part would be unique to the outcome construct (Heath et al., 1993). It should be noted that the direction-of-causation model is different from a correlational model in that it posits that the association between the constructs is fully explained by the direct influence of the causal construct on the

outcome construct, in the absence of correlational relations between the two. Influences between constructs can run both ways, which would represent a bidirectional (reciprocal) relation. Alternatively, influences run in only one direction, which would suggest a unidirectional relation between constructs. Note that in contrast with longitudinal models, the direction of influence would still be evident, regardless of the temporal order and/or lack of latency between measurements of constructs (Verhulst & Estabrook, 2012).

The direction-of-causation model has been successfully applied in the psychopathological and psychiatric literature (e.g., Gillespie, Zhu, Neale, Heath, & Martin, 2003; Touloupoulou et al., 2015), but scarcely used in educational literature. More specifically, this approach has only been used once (van Bergen et al., 2018) to test the directionality of influence between a highly genetically influenced reading construct (average genetic variability was estimated at .59; Little et al., 2017) and a predominantly shared environmentally influenced PE construct (shared environmental variance estimate was .49; Harlaar et al., 2007). In the only published study of its kind, van Bergen et al. (2018) utilized a direction-of-causation model and showed that in early elementary school children reading fluency facilitated PE. The findings indicated that 16% of the PE variance was common with fluency. The amount of variance unique to the outcome construct was 84%. Specifically, genetic influences responsible for reading fluency variation dominantly influenced variation in PE. In contrast, variance unique to PE owed largely to shared environment. In sum, findings indicated that fluent readers engaged more often with print materials because they were fluent in reading. van Bergen et al.'s (2018) report is the first and fundamental step outlining the directionality of this relation in a naturally occurring setting.

The Present Study

In this study, we use the same model as van Bergen et al. (2018), but extend the findings. Here we investigate the directionality of influence between RC (rather than reading fluency) and PE, using data on late elementary, middle, and junior high schoolers (rather than early elementary schoolers) from an economically, racially, and ethnically diverse sample from Florida (rather than a homogeneous sample from the Netherlands). As such, our study further enables us to determine to what extent van Bergen et al.'s (2018) findings are

generalizable across age, economic environments, race and ethnicity, as well as different reading constructs. Based on prior reports (e.g., Aarnoutse & van Leeuwe, 1998; Leppänen et al., 2005; van Bergen et al., 2018), we hypothesized that RC would facilitate PE. We suspected that variation in RC would account for a smaller portion of the common variance with PE. We also predicted that PE would retain unique variance, which would mostly be influenced by shared environmental factors.

Method

Participants

Participants for this study are twin pairs from the Florida Twin Project on Reading, Behavior, and Environment (FTP-RBE), a cross-sequential twin study focusing on behavioral and environmental correlates of reading development, which is part of the Florida State University and Florida Center for Reading Research. FTP-RBE is a subset of a larger project Florida Twin Project on Reading. Twin pairs that form the Wave 1 of the FTP-RBE database are 414 twin pairs from the Florida Twin Project on Reading that were recruited in 2006–2010 with an additional 154 of new twin pairs that were recruited specifically for the FTP-RBE in fall 2012. Thus, the Wave 1 of the FTP-RBE database is comprised of altogether 568 twin pairs. The ascertainment method for new twins in the FTP-RBE was the same as the one used for the Florida Twin Project on Reading, which is described in detail elsewhere (Taylor, Hart, Mikolajewski, & Schatschneider, 2012). Briefly, twins were identified through a match on last name, date of birth, and school in Florida's Progress Monitoring and Reporting Network (PMRN), a statewide database of standardized achievement tests on children in schools throughout the state of Florida. Zygosity of the new twin pairs was determined by a parental five-item questionnaire obtained during intake into the FTP-RBE. It measured the physical similarities of the twins and has shown to have high correspondence to zygosity estimates from genetic markers (Lykken, Bouchard, McGue, & Tellegen, 1990).

Data used in this study came from twins from the wave 1 of the FTP-RBE database (Wave 3 data collection was finished in fall 2017), who were in fourth through ninth grades in 2012–2013. This school year was chosen because we had available data on PE measures for this school year only. The final sample included 910 twins, specifically 370 monozygotic (MZ; 194 females, 176 males) and 540

dizygotic (DZ; 200 females, 154 males, and 186 opposite sex) twins. At the assessment time point, twin pairs were on average approximately 12 years and 4 months old ($M = 12.33$, $SD = 1.41$, range = 9.72–15.70). Twins came from 240 schools throughout the U.S. state of Florida, with the number of twin pairs per school ranging from 1 to 5. The racial and ethnical breakdown was the following: 2.1% of the twins were Asian, 13.5% Black, 21.4% Hispanic, 0.9% Native American or Pacific Islander, 2.9% Mixed, 58.3% White, and 0.9% did not report race and ethnicity. The percentages reported are similar to values reported by the U.S. Census Bureau for the state of Florida. Regarding the socioeconomic status, 48.9% of the twins did not apply or were not eligible for free or reduced lunch, 40.2% were eligible for free lunch, 9.7% were eligible for reduced price lunch, and 1.2% were enrolled in a USDA approved Provision Z school.

Procedure and Measures

RC assessment data came from the PMRN fall assessment. All reading tests in the PMRN were administered by trained teachers or school staff in statewide standardized formats as part of each school's academic program. Reading grade and PE data were obtained via twin- and parent-rated questionnaires sent by mail to parents' home addresses in 2012–2013. All parents of twins provided informed consent for investigators to use their twins' PMRN data and twins provided assent to participate in the study as approved by the Florida State University Institutional Review Board.

Print Exposure

PE was measured by two tasks from the ART in the twin-rated questionnaire. Other researchers have used this measurement in prior studies (e.g., Harlaar et al., 2007; Martin-Chang & Gould, 2008). The first task measured familiarity with children's book authors. Twins were shown a series of names and asked to identify the names of people who wrote books for children by checking "Check if you know he or she is a real author." The ART includes a mixing equal number of foils and real authors (25 each). Twins who are unfamiliar with the authors would be expected to make more false positive responses. Scoring was determined by taking the proportion of author names checked and subtracting the proportion of foils checked. Range was from -1 to 1 , with 1 indicating a perfect score. The even-odd reliability

of the entire ART adjusted by the Spearman–Brown prediction formula was $r = .81$.

To account for memory which might have biased the ART scores (i.e., some individuals might have remembered the author by seeing a book of that author with their co-twin), we also calculated the primary print knowledge from the ART measurement. Researchers have used this measurement in previous research (e.g., Martin-Chang & Gould, 2008) and it served as the second measurement of the PE construct. Twins were given the following instruction: "If you checked you know he or she is a real author, then indicate one of the following, (a) I have NOT read any of the author's books, (b) I have read at least one of the author's books." This key appeared at the top of each page above the response column. Scores of primary print knowledge were obtained the following way: authors whom participants had both correctly recognized and indicated as being read were summed and divided by the total number of real authors (25). Range was from 0 to 1, with 1 being a perfect score. The even-odd reliability for the items checked as having been read adjusted by the Spearman–Brown prediction formula was $r = .81$.

Reading Comprehension

RC included scores on two standardized tests on RC from the PMRN and twin's school grade for reading from the parent-rated questionnaire. The first standardized test was The Florida Assessment for Instruction in Reading RC subtest. It's a computer administered test. Students are asked to read one to three narrative or expository passages, and answer seven to nine multiple-choice questions. The generic estimate of reliability from IRT ranges from .90 to .92 for Grades 4 to 9 (http://www.fcrr.org/fair/Technical%20manual%20-%203-12-FINAL_2012.pdf). We used standard scores, which were based on a distribution with a mean of 100 and a standard deviation of 15.

The second standardized test was The Florida Comprehensive Assessment Test (FCAT) reading subtest. FCAT-Reading is a measure of RC and consists of several narrative and expository passages. Students are asked to read passages and answer multiple-choice, short- or long-answer items based on passage content. Cronbach's alphas range from .85 to .90 for Grades 4 to 9. The criterion-related validities with Stanford Achievement Test Series are from .79 to .83 for Grades 4 to 9 (Florida Department of Education, 2007). We used developmental scaled scores and

they ranged from 154 to 302 (Florida Department of Education, 2014).

The third assessment of RC was a parent reported twin school grade for reading at the time of the assessment. Grades were on a scale of 1 (*A = Excellent*) to 5 (*F = Fail*). Scores from this item were recoded so that a higher rating indicated a better grade in reading.

Data Analyses

Descriptive Statistics, Correlational Analyses, and Overview of the Twin Method

Descriptive statistics and phenotypic correlations among all measured variables were calculated on raw data. Because age and sex effects can bias twin analyses, raw data on all variables for every child were corrected by residualizing on age, age-squared, and sex (McGue & Bouchard, 1984). Residualized data were subsequently standardized. All subsequent twin correlations and structural equation models were calculated on residualized and standardized data.

Twin methodology allows for an examination of the variance and covariance of etiological (genetic and environmental) influences on a particular trait (variable), or shared among traits. The assumptions under the twin method are that MZ twins share approximately 100% of their segregating genes, whereas DZ twins share, on average, 50%. Shared environmental factors are assumed to be 100% for both MZ and DZ twins if siblings are reared together. Nonshared environmental factors are not shared between the siblings for either MZ or DZ twins (Neale & Cardon, 1992). Model fitting is based on the comparison of the variance-covariance matrices in MZ and DZ twins and allows for the separation of the observed phenotypic variance into additive genetic (A), shared environmental (C), and nonshared environmental (E) components. In multivariate genetic analyses, as well as partitioning the phenotypic variance of variables, it is also the covariance between variables that is decomposed into A, C, and E influences.

Intraclass (ICC) and cross-twin cross-trait (CTCT) correlations yield preliminary information on the relative magnitudes of underlying genetic and environmental influences on variation in a particular variable and covariation of two variables. Greater ICCs for MZ relative to DZ twins indicate the presence of additive genetic effects. In contrast, ICCs for MZ and DZ twins that are similar in magnitude signify a lack of genetic effects, but shared

environmental influences. Finally, ICCs for MZ twins that are < 1.0 indicate the presence of nonshared environmental factors. The CTCT correlation is calculated by correlating a variable from one member of a pair with a different variable from his or her co-twin. The interpretation of the CTCTs follows the same logic as that of the ICCs. Descriptive statistics, phenotypic, and twin correlations were conducted in SAS 9.4 (SAS Institute Inc., Cary, NC, USA).

Structural Equation Modeling

Following the descriptive and correlational statistics, three sets of structural equation models were fitted in accordance with the aim of our study. Each subsequent model set built on the model fit of the

previous set. The first set of models entailed fitting a phenotypic two-factor measurement model as shown in the upper panel of Figure 1. To account for dependency among observations (twins clustered in pairs), we corrected standard errors and model fit statistics as proposed by Rebollo, de Moor, Dolan, and Boomsma (2006). A good model fit was evaluated using the following fit indices: chi-square statistics (χ^2), root mean squared error of approximation (RMSEA; < .08), comparative fit index (CFI; > .90), standardized root mean square residual (SRMR; < .05; Kline, 2015).

Building on the first set, in the second set of models genetic and environmental A, C, and E factors and their covariance paths were added to the two reading constructs, and to their indicators

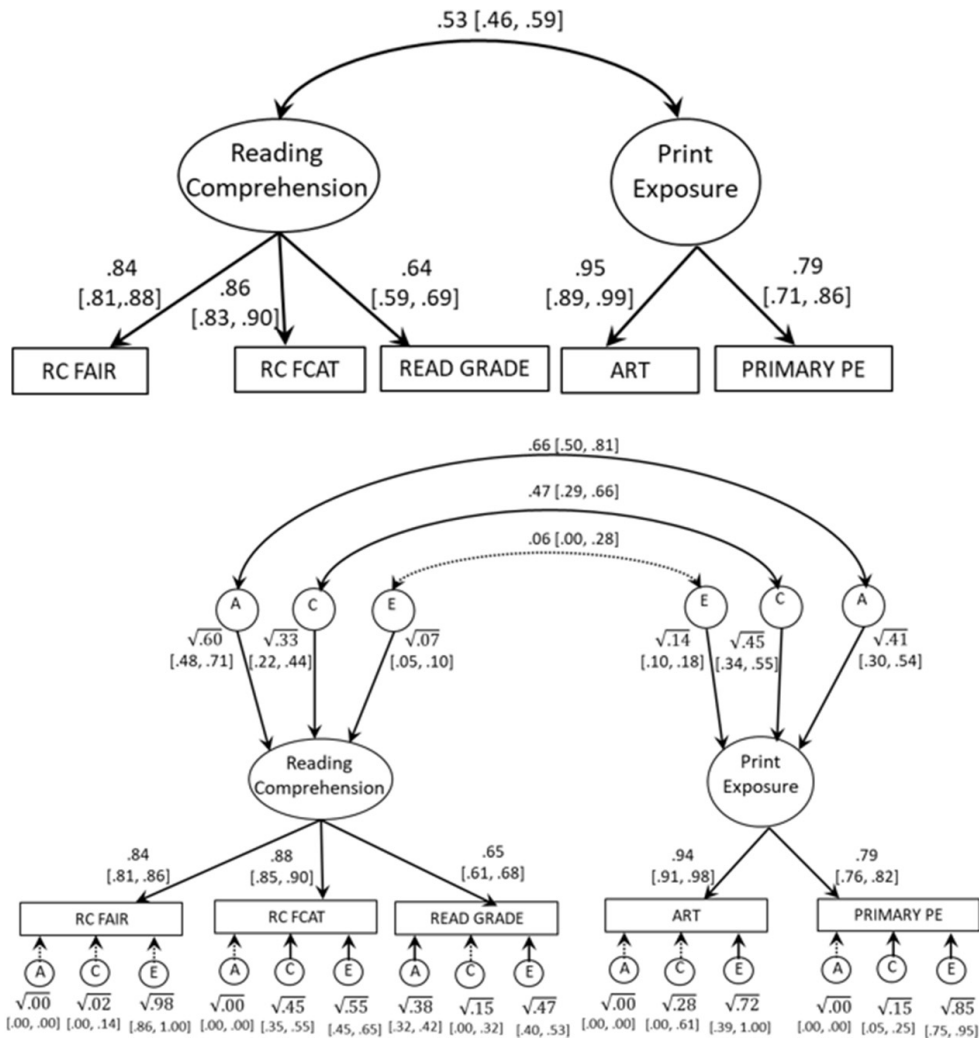


Figure 1. The upper panel shows a measurement model and the bottom panel a two-factor common pathway model, both with estimated standardized parameters. Significant paths are solid, nonsignificant paths are dotted. Amounts of variance explained are indicated under $\sqrt{\text{sign}}$. For simplification reasons, residual signs are not depicted in the figures. RC FAIR = reading comprehension The Florida Assessment for Instruction in Reading; RC FCAT = reading comprehension Florida Comprehensive Assessment Test; READ GRADE = school grade in reading; ART = Author Recognition Test; PRIMARY PE = primary print exposure.

to estimate etiological influences. This model is a two-factor common pathway model, which captures the common variance of interrelated observed indicators. We also used this model to estimate the genetic and environmental correlations between factors. The model is depicted in the bottom panel of Figure 1. The fit of a model was indicated by a χ^2 statistic and Akaike's information criterion (AIC; Akaike, 1987) estimates.

Finally, in the third set of models we determined the direction of influence between RC and PE. The full bidirectional direction-of-causation model (Duffy & Martin, 1994; Heath et al., 1993) is a modification of the model from the second set, such that it includes a directionality path from RC to PE and a directionality path from PE to RC ($RC \leftarrow \rightarrow PE$), whereby dropping the covariance paths between RC and PE (Figure 2). In contrast to the model from the second set, which predicts that shared sources of etiological influences account for the covariance between RC and PE, the direction-of-causation model predicts that RC and PE are each due to independent etiological influences and that the covariation between them is explained only by the direct influence of causal construct on the outcome construct. Etiological factors of PE can be thought as of representing residual variance on PE, after allowing for the influence of RC on PE; likewise, etiological factors of RC can be viewed as residual variance on RC, after allowing for an influence of PE on RC. In other words, there will be additional indirect effects of genes and environment of RC on PE, mediated through the influence of RC on PE if $RC \rightarrow PE$ is different from zero. And vice versa, there will be additional indirect effects of genes and environment of PE on RC, mediated through the influence of PE on RC if $RC \leftarrow PE$ is different from zero (Heath et al., 1993). Alternate (i.e., reduced, unidirectional) direction-of-causation models were also tested, the first specifying RC as a causal construct for PE ($RC \rightarrow PE$; Figure 3, upper panel), and the second specifying PE as a causal construct for RC ($RC \leftarrow PE$; Figure 3, bottom panel). When two correlated constructs have different modes of genetic and environmental influences, there are different expectations for the cross-covariance in alternate direction-of-causation models (Heath et al., 1993). As noted in the introduction, previous research has shown that differences in RC are mostly due to genetic factors, whereas environmental factors mostly account for differences in PE. Hence, there is an a priori expectation that the direct influence between RC and PE could be at play. If $RC \rightarrow PE$ fits data best, then this predicts

that the cross-covariance between RC and PE will be dominated by genetic influences. On the contrary, if $RC \leftarrow PE$ fits best, then this implies that the cross-covariance will be dominated by shared environmental influences (Heath et al., 1993).

Again, the fit of a direction-to-causation model was indicated by a χ^2 statistic estimate. Reduced models were compared to the full direction-to-causation model using a χ^2 difference test. A significant χ^2 difference test would indicate that the reduced model not be accepted over the full model, whereas a nonsignificant χ^2 difference test would allude for the reduced model be accepted over the full model. Structural equation models were fit in Mplus 8.1 using maximum likelihood estimator with robust standard errors (Muthén & Muthén, 2012). Significance of parameter estimates was based on the 95% confidence intervals not including zero.

Power Analysis

Following the Muthén and Muthén (2002) approach, a post-hoc Monte Carlo simulation study from the observed data was conducted in Mplus 8.1 to determine whether our sample size of MZ and DZ twins was large enough to reject the hypotheses that important effects in the model (i.e., parameters indicating direction of influence between RC and PE) were zero. Specifically, we studied the values of parameter and standard error biases, coverage, and power. The values of parameter and standard error biases for the parameter for which power is being assessed should not exceed 10% and 5%, respectively. The recommended values for coverage are between 0.91 and 0.98. As of power, the value of 0.80 was used because it is a commonly accepted value for sufficient power (Muthén & Muthén, 2002). The alpha level was set at .05. An alternative, yet additional, approach to the power question was determining the width of the confidence intervals around the parameter estimates. According to Greenland (2012) and Greenland et al. (2016), comparison of models (hypotheses) together with confidence intervals should be presented as additional and more easily interpreted information to determine post-hoc power, rather than only calculating power using data from the study.

Results

Descriptive Statistics and Correlational Analyses

Descriptive statistics, phenotypic, ICC, and CTCT correlations for all measured variables are

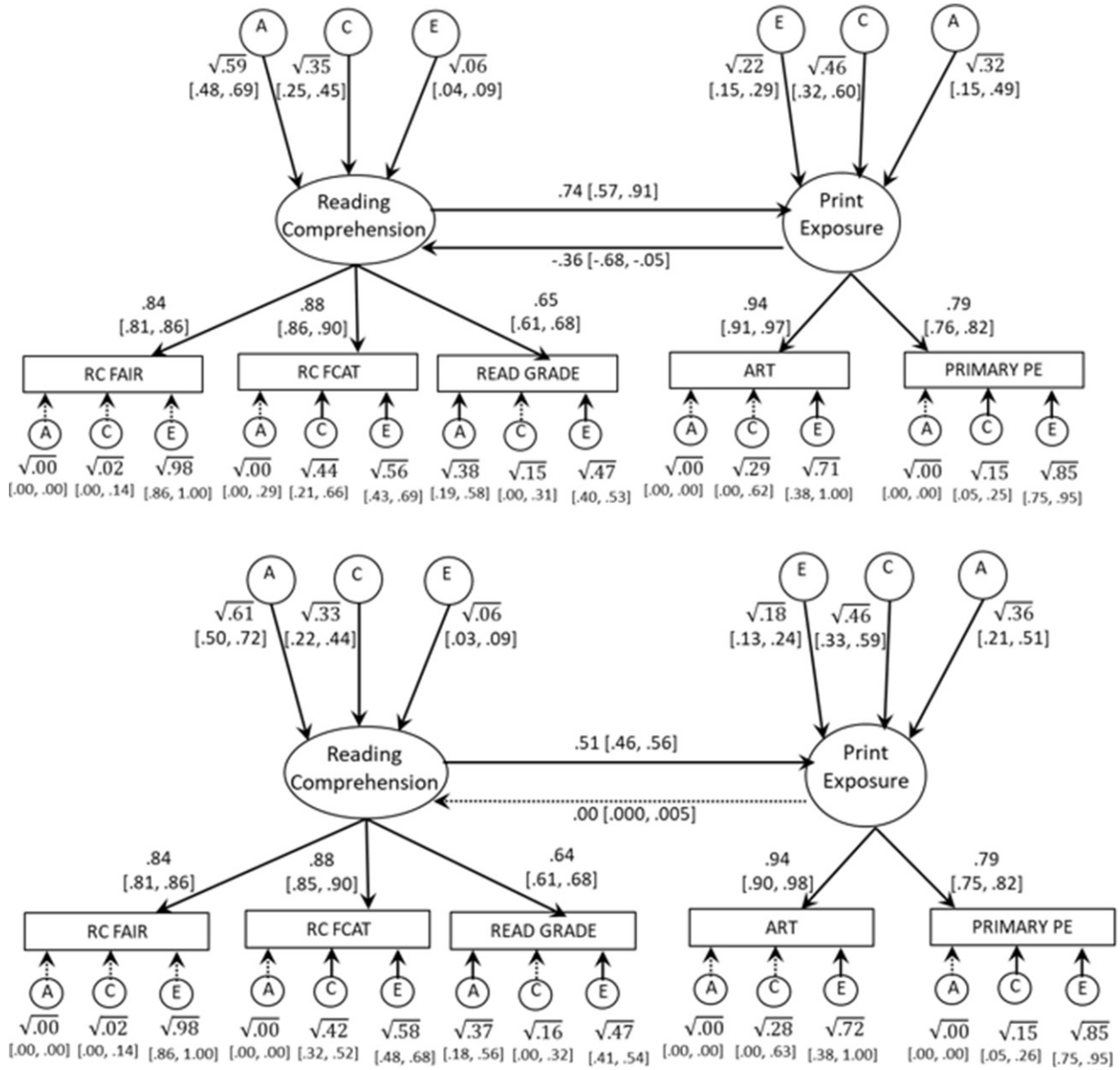


Figure 2. The upper panel shows a full bidirectional direction-to-causation model with no bound paths between constructs, and the bottom panel shows a full bidirectional direction-to-causation model with paths running between constructs bound to take only values larger than zero. Both panels indicate estimated standardized parameters. Significant paths are solid, nonsignificant paths are dotted. Amounts of variance explained are indicated under $\sqrt{\text{sign}}$. For simplification reasons, residual signs are not depicted in the figure. RC FAIR = reading comprehension The Florida Assessment for Instruction in Reading; RC FCAT = reading comprehension Florida Comprehensive Assessment Test; READ GRADE = school grade in reading; ART = Author Recognition Test; PRIMARY PE = primary print exposure.

provided in Table 1. Phenotypic correlations between RC as well as those between PE measures were moderate. Correlations across different construct measures were, however, weak to moderate. The magnitudes were similar to those reported in the van Bergen et al. (2018) study. Table 1 also shows twin correlations. ICCs were consistently higher for MZ than DZ twins for all measures,

indicating presence of some genetic influences on the variation in these measures. ICCs for PE measures were close to each other in magnitude and, as expected, alluded to shared environmental influences accounting for differences in these traits. CTCTs also alluded to genetic and shared environmental influences on the covariance between reading measures.

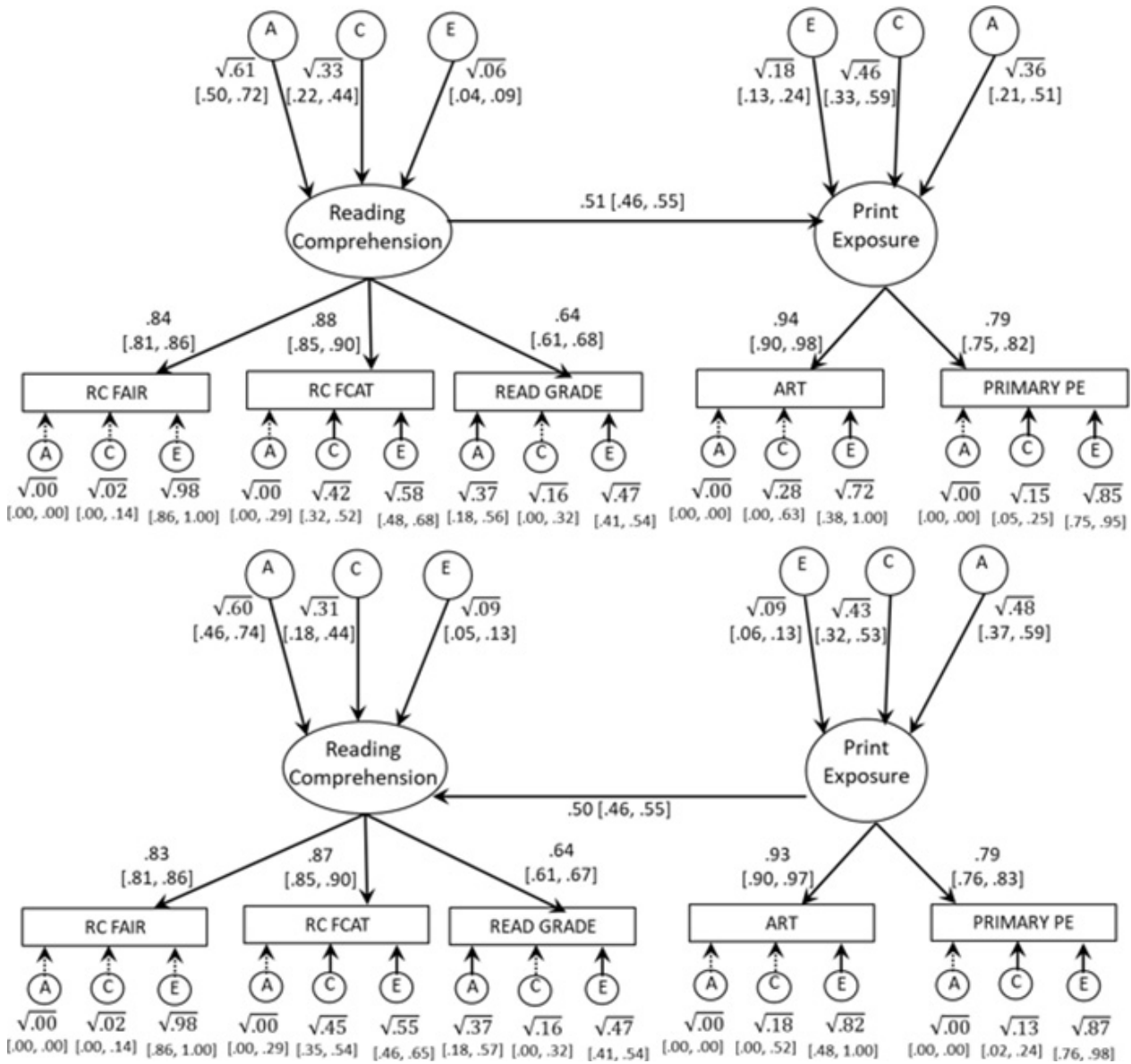


Figure 3. The upper panel shows a reduced direction-to-causation model with the unidirectional pathway running from reading comprehension to print exposure (the most parsimonious model to describe our data), and the bottom panel shows a reduced direction-to-causation model with the unidirectional pathway running from print exposure to reading comprehension. Both panels indicate estimated standardized parameters. Significant paths are solid, nonsignificant paths are dotted. Amounts of variance explained are indicated under $\sqrt{\text{sign}}$. For simplification reasons, residual signs are not depicted in the figures. RC FAIR = reading comprehension The Florida Assessment for Instruction in Reading; RC FCAT = reading comprehension Florida Comprehensive Assessment Test; READ GRADE = school grade in reading; ART = Author Recognition Test; PRIMARY PE = primary print exposure.

Structural Equation Modeling

Measurement Model

A two-factor measurement model fitted the data well: $\chi^2(df = 4, N = 910) = 5.087, p = .28, RMSEA = .019 (.000, .060), CFI = 1.000, SRMR = .010$. Standardized path estimates are indicated in the upper panel of Figure 1. This measurement model served

as a basis for the two-factor common pathway model.

Two-Factor Common Pathway Model

A full two-factor common pathway model with standardized estimates and percentages of variances explained is shown in the bottom panel of

Table 1
Descriptive Statistics, Phenotypic, Intra-class, and Cross-Twin, Cross-Trait Correlations

Measured variables	M (SD)	N	Min	Max	Skew	1.	2.	3.	4.	1.	2.	3.	4.	5.
1. RC FAIR	100.15 (13.83)	624	66.00	145.00	0.46	1				.60** / .52**	.49**	.33**	.27**	.26**
2. RC FCAT	236.13 (25.92)	894	155.00	296.00	-0.32	.69**	1			.64**	.80** / .60**	.35**	.23**	.24**
3. READ GRADE	4.34 (0.81)	898	1.00	5.00	-1.08	.55**	.55**	1		.38**	.48**	.69** / .44**	.25**	.24**
4. ART	0.25 (0.18)	758	-0.28	1.00	1.08	.44**	.39**	.33**	1	.41**	.42**	.28**	.75** / .66**	.54**
5. PRIMARY PE	0.18 (0.15)	758	0.00	1.00	1.04	.33**	.36**	.27**	.74**	.21**	.35**	.23**	.56**	.52** / .51**

Note. Phenotypic correlations are presented in Columns 7–10. Intra-class and cross-twin, cross-trait correlations are in Columns 11–15. Intraclass correlations are on the diagonal; the first estimate is for monozygotic (MZ) twins, the second for dizygotic (DZ) twins. Cross-twin, cross-trait correlations are off the diagonal; the estimates for MZ twins are below the diagonal and for DZ twins above the diagonal. N = number of twins; Min = minimum; Max = maximum; Skew = skewness; RC FAIR = reading comprehension The Florida Assessment for Instruction in Reading; RC FCAT = reading comprehension Florida Comprehensive Assessment Test; READ GRADE = a parent reported twin school grade in reading; ART = Author Recognition Test; PRIMARY PE = primary print exposure.
** $p < .0001$.

Figure 1. The model fitted the data well: $\chi^2(df = 98, N = 455 \text{ pairs}) = 106.877, p = .25, AIC = 16,249.606$. As expected, the RC construct was mostly explained by genetic influences (60%). Shared environment contributed 33% to the RC variance. On the other hand, the variability in PE construct was explained to almost identical extent by genetic (41%) and shared environmental (45%) influences. The genetic influences between reading constructs correlated strongly (.66). This correlation was higher than that of shared environment (.47).

Direction-of-Causation Model

Finally, we fitted the third set of models attempting to determine the direction of influence between RC and PE. The full bidirectional direction-to-causation model, as shown in the upper panel of Figure 2, fitted the data only slightly better than the two-factor common pathway model ($\chi^2(df = 99, N = 455 \text{ pairs}) = 107.830, p = .26, AIC = 16,248.559, \Delta AIC = -1.047$). Given the negative parameter estimate of -0.36 on the path $RC \leftarrow PE$ (Figure 2, upper panel), which appeared to be theoretically and practically counterintuitive, we re-ran the bidirectional model with the paths running between our constructs bound to only take values larger than zero. The fit of the model with bound parameter was as follows $\chi^2(df = 99, N = 455 \text{ pairs}) = 111.411, p = .18, AIC = 16,252.140, \Delta AIC = 2.534$. The model is depicted in the bottom panel of Figure 2.

Next, we fitted two reduced direction-of-causation models. In the first one, we dropped the path from PE to RC. This model, depicted in the upper panel of Figure 3, fitted the data well ($\chi^2(df = 100, N = 455 \text{ pairs}) = 111.559, p = .19, AIC = 16,250.288$). It did not fit significantly worse than the full bidirectional direction-of-causation model ($\Delta\chi^2(1) = 3.729, p = .053$). In addition, it did not fit significantly worse than the bound full bidirectional model ($\Delta\chi^2(1) = 0.148, p = .700$).

In the second reduced model, we dropped the path from RC to PE. This model, indicated in bottom panel of Figure 3, did not fit the data as well as the full bidirectional model ($\chi^2(df = 100, N = 455 \text{ pairs}) = 124.643, p = .05, AIC = 16,263.372$). In addition, it fit significantly worse than the full bidirectional direction-of-causation model ($\Delta\chi^2(1) = 16.813, p < .001$), as well as the bound full bidirectional model ($\Delta\chi^2(1) = 13.232, p < .001$), indicating that the drop of the path from RC to PE construct resulted in deterioration of fit. Hence, of the two reduced direction-of-causation models, the former

was determined as a final model of choice (the upper panel of Figure 3).

As of presentation of power, a Monte Carlo simulation study (Muthén & Muthén, 2002) showed that both reduced direction-of-causation models were at power (0.80) to reject the hypotheses that the unidirectional pathway $RC \rightarrow PE$ and $RC \leftarrow PE$, respectively, was zero. As recommended by Muthén and Muthén (2002), the values of parameter error bias were below 10%. Specifically, they were 0.678% and 0.120% for each unidirectional pathway, respectively. In addition, the values of standard error bias did not exceed 5%. Specifically, they were 4.794% and 1.592%, respectively. Coverage was 1.000 in both cases. Estimation of the Monte Carlo simulation for the full bidirectional direction-to-causation model failed to converge. Nonetheless, Greenland et al.'s (2016) recommended that presentation of power should not obviate the need to provide direct tests of alternative models with interval estimates. Our results on model comparisons alluded to the reduced, unidirectional model $RC \rightarrow PE$ being accepted as our final model over other alternative models.

Consistent with our hypothesis, the direction-of-causation model indicating an influence from RC to PE (Figure 3, upper panel) appeared to be the most parsimonious model to describe our data. As indicated in the upper panel of Figure 3, there were indirect effects of genetic (61%) and shared environmental (33%) influences of RC on PE mediated through the pathway indicating directionality of influence. Squaring the path estimate gives an estimated common variance shared with RC, $(.51) \times (.51) = .26$ (or 26%). The residual variance ($100\% - 26\% = 74\%$) was unique to the outcome construct PE. Decomposing it, it owed largely to shared environmental factors (46%) and less to genetic factors (36%).

Discussion

One of the interesting questions in the reading literature is whether reading skills facilitate PE or vice versa. Prior work showed that reading fluency fostered PE when children are beginning to learn to read (e.g., van Bergen et al., 2018). The aim of this study was to test the directionality of influence between RC and PE in early adolescence when adolescents are reading to learn. We also estimated genetic and environmental influences of this relation. Our results replicated van Bergen et al.'s (2018) findings, suggesting that the extent to which

adolescents chose to embrace opportunities to read, in part, reflected their proficiency in reading.

As hypothesized, there appears to be a direction of influence from reading achievement to PE. That means that adolescents who had a more sophisticated set of skills and knowledge of reading took increasing initiatives in engaging with reading, whereas those lacking such skills rather avoided such practices (i.e., genetic niche-picking or active gene-environment correlation; Johnson, 2007). This relation was driven mainly by genetic and to a lesser extent by shared environmental influences of RC. Although this model cannot identify the exact sources of these etiological influences, we can surmise that at this developmental stage genetic and shared environmental influences of early reading skills (e.g., reading fluency) coupled with the novel genetic and environmental influences for RC in adolescence (Erbeli et al., 2018) might be reflected in this relation. Higher-level reading skills, such as inference, comprehension monitoring, and knowledge and use of story structure, might have contributed to differences in adolescent RC. These higher-level reading skills might be mediating, through the pathway indicating directionality of influence, individual differences in adolescents' amount of voluntary reading. Previous research has shown higher-level reading skills emerged as distinct predictors of RC in sixth grade, even after controlling for autoregressive effects of comprehension (Oakhill & Cain, 2012), giving support to our assumption.

Our results mirror previous findings (van Bergen et al., 2018), indicating that the direction of influence from reading to PE might be generalizable across ages, samples, and reading constructs. Nonetheless, we need to take care not to draw black or white conclusions and overreach in interpretation of the findings. Improvements in model fitting between the correlational model, full bidirectional direction-of-causation models (not bound and bound), and our final model were marginal at best. As such, we cannot completely reject the hypothesis that the covariation between RC and PE was due to shared genetic and environmental factors, with no direction of influences, or that it arose via a reciprocal effect between our two reading constructs. Taking that into consideration though, in the bound bidirectional model the path from PE to RC was zero and statistically nonsignificant (see bottom panel of Figure 2), also mirroring findings from prior research (e.g., van Bergen et al., 2018). This further bolstered our confidence in concluding that if there is, indeed, evidence of a potential direction of influence between RC and PE, it is likely

demonstrated as reading being an antecedent and PE a consequent construct.

Beyond the common variance shared with RC, variability in PE also retained a considerable amount of unique variance. The residual variance was less due to genetics (36%) and more due to shared environment (46%). That was expected and consistent with prior research findings (e.g., Harlaar et al., 2007). The genetic portion indicated that not all genetic influences that contributed to reading in one context (comprehending texts as part of a formal assessment battery) contributed to reading in another context (recreational after school reading). More likely, however, is that at least some genetic influences on PE reflected adolescent-driven behaviors through gene–environment correlation (e.g., van Bergen, van der Leij, & de Jong, 2014). For example, adolescents partly selected and shaped their own recreational reading environments based on their relevant genetic background. They would take part in reading clubs, ask for books as presents, and take other initiatives in honing their reading related skills. Another possibility is that parents and teachers exposed avid readers to books more often, talked to them about reading materials, and provided texts and other reading opportunities that engaged adolescents in intellectually stimulating ways. All these examples mirror gene–environment correlations that tend to be reflected as genetic influences in models such as the one we applied here.

As of the shared environmental factor of PE, it likely reflected aspects of home and school environments, or, more broadly, pedagogical, institutional, and/or socioeconomic factors. For example, parental influences toward creating environments of rich literacy related activities (Sénéchal & Cornell, 1993) might have instilled adolescents' interest for reading activities. Similarly, the tone set by classroom reading instruction and school policy on leisurely reading activities could have influenced adolescents' engagement with books and their reading aspirations (Merga, 2015). Schools with consistent reading instruction and financial and human resources could have been investing time and resources in stimulating and emphasizing the value of independent reading for all students (Merga, 2015).

Regardless of the mechanism uniquely influencing PE variability, results indicated that reading proficiency was, in part, the driver of voluntary reading. Such a result must not be interpreted to represent one-way evidence that interventions aimed at targeting only RC will work to remediate reading deficits. We cannot exclude the possibility

that interventions stimulating PE activities might transfer and help ameliorate reading difficulties. Share (1995, 1999), for example, demonstrated that the more frequent a child has been exposed to a particular word, the more likely that word will be in her sight word vocabulary. In addition, there is evidence from extensive intervention research showing that exposure to texts with new concepts, such as texts targeting content areas knowledge, indicated gains on content RC as well as standardized RC tests (e.g., Swanson, Wanzek, Vaughn, Roberts, & Fall, 2015; Vaughn et al., 2013). Targeting different comprehension strategies (e.g., inferring and activating background knowledge) via exposure to different texts also led to gains on RC tests (e.g., Clarke, Snowling, Truelove, & Hulme, 2010). As such, we can only speculate on the meaning of our findings for intervention studies.

Moreover, as the foregoing makes it clear, the nature of our study speaks to individual differences and directionality of influences as they exist in a particular population at a particular time in a naturally occurring setting (Plomin & Haworth, 2010, refer to this as a *what is* situation). Intervention research, such as RCTs, on the other hand, focuses on differences between intervention and control group in order to understand *what could be* and prove causality of the intervention in an artificial experimental setting (Plomin & Haworth, 2010). Notwithstanding these precautions, our results help justify implementing interventions, in particular the ones targeting RC and possibly monitoring those adolescents over time to examine whether immediate impact on RC proficiency would be expected to also boost adolescents' engagement with print materials. Such research has not been conducted thus far, however, it may be of major significance in particular for poor reading comprehenders.

There were limitations to consider when interpreting the present findings. First, ratings of PE came only from adolescents themselves and this may not present a complete picture of their reading practices. Ideally, ratings would have been obtained from multiple informants across different contexts (e.g., parents and teachers). Second, the ART measure is a proximate indicator of voluntary reading, but not a direct measure. As such, familiarity with children's book authors does not necessarily mean that adolescents actually read the books. Nonetheless, this aspect was, in part, controlled for by accounting for twins' print knowledge gained through personal reading. Third, our study sampled pre- and early adolescents across a fairly wide age range (9–15 year-olds). This may have obscured

possible developmental variations in the etiological associations between RC and PE. Fourth, the magnitude of genetic and environmental variance of the reading constructs, as well as their covariance, depends on the nature of the environment from which a population is recruited. In this regard, it is notable that the present results were indicative of the economically, racially, and ethnically diverse population of twins in Florida. Nonetheless, it appears that the main finding of our study replicated that from the Dutch homogeneous sample.

Despite these caveats, the findings from this study extend the literature regarding the question of the covariation of RC and PE at a time when proficient RC is key to learning various content areas in school. Our results aligned with those of prior studies (e.g., van Bergen et al., 2018). Specifically, it appears that reading proficiency in adolescence to some extent directly affected how often adolescents chose to read after school. This may have implications for interventions of reading difficulties in that it underscores the importance of individual differences in their manifestation (e.g., Connor, Morrison, Fishman, Crowe, Al Otaiba, & Schatschneider, 2013). If proficient RC, or lack thereof, is partly responsible for more or less engagement in PE activities, then parents and school-based practitioners may need to be called upon to help find activities targeted toward giving an adolescent the best chance to improve RC and, at the same time, engage in more literacy-related practices.

References

- Aarnoutse, C., & van Leeuwe, J. (1998). Relation between reading comprehension, vocabulary, reading pleasure, and reading frequency. *Educational Research and Evaluation, 4*, 143–166. <https://doi.org/10.1076/edre.4.2.143.6960>
- Akaike, H. (1987). Factor analysis and AIC. *Psychometrika, 52*, 317–332. <https://doi.org/10.1007/BF02294359>
- Clarke, P. J., Snowling, M. J., Truelove, E., & Hulme, C. (2010). Ameliorating children's reading-comprehension difficulties: A randomized controlled trial. *Psychological Science, 21*, 1106–1116. <https://doi.org/10.1177/0956797610375449>
- Connor, C. M., Morrison, F. J., Fishman, B., Crowe, E. C., Al Otaiba, S., & Schatschneider, C. (2013). A longitudinal cluster-randomized controlled study on the accumulating effects of individualized literacy instruction on students' reading from first through third grade. *Psychological Science, 24*, 1408–1419. <https://doi.org/10.1177/0956797612472204>
- DeWalt, D. A., Berkman, N. D., Sheridan, S., Lohr, K. N., & Pignone, M. P. (2004). Literacy and health outcomes: A systematic review of the literature. *Journal of General Internal Medicine, 19*, 1228–1239. <https://doi.org/10.1111/j.1525-1497.2004.40153.x>
- 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. <https://doi.org/10.1002/gepi.1370110606>
- Erbeli, F., Hart, S. A., & Taylor, J. (2018). Longitudinal associations among reading related skills and reading comprehension: A twin study. *Child Development, 89*, e480–e493. <https://doi.org/10.1111/cdev.12853>
- Florida Department of Education. (2007). *Assessment and accountability briefing book*. Tallahassee, FL: Author. Retrieved from <http://www.fldoe.org/core/fileparse.php/7478/urlt/briefingbook07web.pdf>
- Florida Department of Education. (2014). *2014 FCAT 2.0*. Tallahassee, FL: Author. Retrieved from http://www.fldoe.org/core/fileparse.php/5668/urlt/0066933-2014fcatt20_media.pdf
- Gillespie, N. A., Zhu, G., Neale, M. C., Heath, A. C., & Martin, N. G. (2003). Direction of causation modeling between cross-sectional measures of parenting and psychological distress in female twins. *Behavior Genetics, 33*, 383–396. <https://doi.org/10.1023/A:1025365325016>
- Greenland, S. (2012). Nonsignificance plus high power does not imply support for the null over the alternative. *Annals of Epidemiology, 22*, 364–368. <https://doi.org/10.1016/j.annepidem.2012.02.007>
- Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology, 31*, 337–350. <https://doi.org/10.1007/s10654-016-0149-3>
- Harlaar, N., Dale, P. S., & Plomin, R. (2007). Reading exposure: A (largely) environmental risk factor with environmentally-mediated effects on reading performance in the primary school years. *Journal of Child Psychology and Psychiatry, 48*, 1192–1199. <https://doi.org/10.1111/j.1469-7610.2007.01798.x>
- Harlaar, N., Deater-Deckard, K., Thompson, L. A., DeThorne, L. S., & Petrill, S. A. (2011). Associations between reading achievement and independent reading in early elementary school: A genetically informative cross-lagged study. *Child Development, 82*, 2123–2137. <https://doi.org/10.1111/j.1467-8624.2011.01658.x>
- 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. <https://doi.org/10.1007/BF01067552>
- Johnson, W. (2007). Genetic and environmental influences on behavior: Capturing all the interplay. *Psychological Review, 114*, 423–440. <https://doi.org/10.1037/0033-295X.114.2.423>
- Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling*. New York, NY: Guilford.

- Leppänen, U., Aunola, K., & Nurmi, J. E. (2005). Beginning readers' reading performance and reading habits. *Journal of Research in Reading, 28*, 383–399. <https://doi.org/10.1111/j.1467-9817.2005.00281.x>
- Little, C. W., Haughbrook, R., & Hart, S. A. (2017). Cross-study differences in the etiology of reading comprehension: A meta-analytical review of twin studies. *Behavior Genetics, 47*, 52–76. <https://doi.org/10.1007/s10519-016-9810-6>
- Lykken, D. T., Bouchard, Jr., T., McGue, M., & Tellegen, A. (1990). The Minnesota Twin Family Registry: Some initial findings. *Acta Geneticae Medicae et Gemellologiae, 39*, 35–70. <https://doi.org/10.1017/S0001566000005572>
- Martin-Chang, S. L., & Gould, O. N. (2008). Revisiting print exposure: Exploring differential links to vocabulary, comprehension and reading rate. *Journal of Research in Reading, 31*, 273–284. <https://doi.org/10.1111/j.1467-9817.2008.00371.x>
- McGue, M., & Bouchard, T. J. (1984). Adjustment of twin data for the effects of age and sex. *Behavior Genetics, 14*, 325–343. <https://doi.org/10.1007/BF01080045>
- McLaughlin, M. J., Speirs, K. E., & Shenassa, E. D. (2014). Reading disability and adult attained education and income: Evidence from a 30-year longitudinal study of a population-based sample. *Journal of Learning Disabilities, 47*, 375–386. <https://doi.org/10.1177/0022219412458323>
- Merga, M. K. (2015). Access to books in the home and adolescent engagement in recreational book reading: Considerations for secondary school educators. *English in Education, 49*, 197–214. <https://doi.org/10.1111/eie.12071>
- Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137*, 267–296. <https://doi.org/10.1037/a0021890>
- Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling, 9*, 599–620. https://doi.org/10.1207/S15328007SEM0904_8
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus. The comprehensive modeling program for applied researchers: User's guide* (7th ed.). Los Angeles, CA: Author.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction* (NIH Pub. 00-4769). Washington, DC: National Institute of Child Health and Human Development.
- Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht, The Netherlands: Springer Science.
- Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading, 16*, 91–121. <https://doi.org/10.1080/10888438.2010.529219>
- Perfetti, C. A. (1985). *Reading ability*. New York, NY: Oxford University Press.
- Plomin, R., & Haworth, C. M. (2010). Genetics and intervention research. *Perspectives on Psychological Science, 5*, 557–563. <https://doi.org/10.1177/1745691610383513>
- Rebollo, I., De Moor, M. H., Dolan, C. V., & Boomsma, D. I. (2006). Phenotypic factor analysis of family data: Correction of the bias due to dependency. *Twin Research and Human Genetics, 9*, 367–376. <https://doi.org/10.1375/183242706777591326>
- Sénéchal, M., & Cornell, E. H. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly, 28*, 360–374. <https://doi.org/10.2307/747933>
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition, 55*, 151–218. [https://doi.org/10.1016/0010-0277\(94\)00645-2](https://doi.org/10.1016/0010-0277(94)00645-2)
- 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. <https://doi.org/10.1006/jecp.1998.2481>
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407. <https://doi.org/10.1598/RRQ.21.4.1>
- Swanson, E., Wanzek, J., Vaughn, S., Roberts, G., & Fall, A. M. (2015). Improving reading comprehension and social studies knowledge among middle school students with disabilities. *Exceptional Children, 81*, 426–442. <https://doi.org/10.1177/0014402914563704>
- Taylor, J. E., Hart, S. A., Mikolajewski, A. J., & Schatschneider, C. (2012). An update on the Florida State Twin Registry. *Twin Research and Human Genetics, 16*, 471–475. <https://doi.org/10.1017/thg.2012.74>
- Torppa, M., Niemi, P., Vasalampi, K., Lerkkanen, M. K., Tolvanen, A., & Poikkeus, A. M. (2019). Leisure reading (but not any kind) and reading comprehension support each other—A longitudinal study across grades 1 and 9. *Child Development, 90*, 1324–1341. <https://doi.org/10.1111/cdev.13241>
- Toulopoulou, T., Van Haren, N., Zhang, X., Sham, P. C., Cherny, S. S., Campbell, D. D., . . . Brouwer, R. (2015). Reciprocal causation models of cognitive vs volumetric cerebral intermediate phenotypes for schizophrenia in a pan-European twin cohort. *Molecular Psychiatry, 20*, 1386–1396. <https://doi.org/10.1038/mp.2014.152>
- van Bergen, E., De Jong, P. F., Maassen, B., & van der Leij, A. (2014). The effect of parents' literacy skills and children's preliterate skills on the risk of dyslexia. *Journal of Abnormal Child Psychology, 42*, 1187–1200. <https://doi.org/10.1007/s10802-014-9858-9>
- van Bergen, E., Snowling, M. J., de Zeeuw, E. L., van Beijsterveldt, C. E., Dolan, C. V., & Boomsma, D. I. (2018). Why do children read more? The influence of reading ability on voluntary reading practices. *Journal of Child Psychology and Psychiatry, 59*, 1205–1214. <https://doi.org/10.1111/jcpp.12910>
- van Bergen, E., van der Leij, A., & De Jong, P. F. (2014). The intergenerational multiple deficit model and the case of dyslexia. *Frontiers of Human Neuroscience, 8*, 346. <https://doi.org/10.3389/fnhum.2014.00346>

Vaughn, S., Swanson, E. A., Roberts, G., Wanzek, J., Stillman-Spisak, S. J., Solis, M., & Simmons, D. (2013). Improving reading comprehension and social studies knowledge in middle school. *Reading Research Quarterly, 48*, 77–93. <https://doi.org/10.1002/rrq.039>

Verhulst, B., & Estabrook, R. (2012). Using genetic information to test causal relationships in cross-sectional data. *Journal of Theoretical Politics, 24*, 328–344. <https://doi.org/10.1177/0951629812439348>