INTERNATIONAL LITERACY ASSOCIATION READING RESEARCH QUARTERLY

How Are Practice and Performance Related? Development of Reading From Age 5 to 15

Elsje van Bergen
Vrije Universiteit Amsterdam, The Netherlands

Kati Vasalampi
Minna Torppa
University of Jyväskylä, Finland

ABSTRACT

Does reading a lot lead to better reading skills, or does reading a lot follow from high initial reading skills? The authors present a longitudinal study of how much children choose to read and how well they decode and comprehend texts. This is the first study to examine the codevelopment of print exposure with both fluency and comprehension throughout childhood using autocorrelations. Print exposure was operationalized as children's amount of independent reading for pleasure. Two hundred children were followed from age 5 to age 15. Print exposure was assessed at ages 5, 7, 8, 9, and 13. Prereading skills were tested at age 5 and reading skills at ages 7, 8, 9, 14, and 15 (the latter with the Programme for International Student Assessment [PISA]). Before children learned to read (i.e., age 5), prereading skills and print exposure were not linked. Path analyses showed that children's print exposure and reading skills reciprocally influence each other. During the early school years, the effects run from reading fluency to comprehension and print exposure, so from skills to amount. The effect of accumulated practice only emerged in adolescence. Reading fluency, comprehension, and print exposure were all important predictors of age 15 PISA reading comprehension. These findings were largely confirmed by post hoc models with random intercepts. Because foundational reading skills predicted changes in later reading comprehension and print exposure, the authors speculate that intervening decoding difficulties may positively impact exposure to and comprehension of texts. How much children read seems to matter most after the shift from learning to read to reading to learn.

Individuals differ greatly in their performance in sports, games, music, and education, and hardly surprising, those who practice more perform better (Macnamara, Hambrick, & Oswald, 2014; Macnamara, Moreau, & Hambrick, 2016; Mosing, Madison, Pedersen, Kuja-Halkola, & Ullen, 2014). The big question that psychologists have sought to answer is whether differences in performance level are merely the result of differences in accumulated practice, or the other way around, in other words, whether initial success stimulates and failure discourages practice. In this study, we focused on education and asked, Do children who achieve above average in school do so because they have practiced a lot, or do children with a high initial skill level engage more in activities in which they develop those skills even further? We investigated the direction of effect between skill and amount of practice for the key academic skill, reading, which lays the foundation for the remainder of one's educational path. Within reading practice and reading skills, we considered whether direction of effect is specific to the subskill (i.e., reading fluency, reading comprehension) or developmental stage (i.e., developmental in nature).

INTERNATIONAL LITERACY ASSOCIATION

Reading Research Quarterly, 0(0)
pp. 1–20 | doi:10.1002/rrq.309

Previous studies have demonstrated that avid readers are also better readers (Anderson, Wilson, & Fielding, 1988; Cunningham & Stanovich, 1997). How much children read is often referred to as print exposure (Cunningham & Stanovich, 1990). Mol and Bus (2011) meta-analyzed 99 studies (total $N = 7{,}669$) that reported correlations between (precursors of) reading ability and print exposure. The researchers found that the relation between print exposure and reading skills becomes even more evident in time: from modest correlations in kindergarten to strong correlations in higher education. The associations among reading interest, print exposure, and reading skills are stronger for reading for pleasure rather than for reading for school (Schiefele, Schaffner, Möller, & Wigfield, 2012). In addition, the association between print exposure and reading skills is stronger for reading books than reading online (McGeown, Duncan, Griffiths, & Stothard, 2015; McGeown, Osborne, Warhurst, Norgate, & Duncan, 2016; Pfost, Dörfler, & Artelt, 2013; Torppa et al., 2019). Hence, we focused on the independent reading of printed material for pleasure.

The typical theory to explain the link between print exposure and reading skills is that interest in reading leads children to read more, which in turn makes them better readers (Becker, McElvany, & Kortenbruck, 2010; Schiefele et al., 2012). Indeed, exposure to words is needed to acquire orthographic knowledge of the words, which in turn speeds up the reading of those words (Share, 1999). Subsequently, effortless decoding may free up cognitive resources to enhance comprehension (Mol & Bus, 2011). In addition, frequent print exposure may support other essential components of efficient reading, such as general verbal skills (Florit & Cain, 2011; Mol & Bus, 2011) and vocabulary size (Cain & Oakhill, 2011). Thus, children who are interested in reading and therefore read for pleasure get more practice in both basic decoding skills (i.e., word-reading accuracy, fluency) and higher order reading skills (i.e., comprehension of texts), and as a result, their reading skills might improve more than those of children who are not interested in reading. In longitudinal research, this would lead to an effect of early print exposure on later reading skills.

The problem in previous studies is that most have been cross-sectional, and the direction of causality has been assumed rather than tested. The association between print exposure and reading skills could arise from the effects running from skills to exposure. Indeed, experiencing progress and competence in reading increases children's motivation to read (Becker et al., 2010), which may lead to an effect of early reading skills on later print exposure. In line with this hypothesis, Cunningham and Stanovich (1997) found a predictive association between first-grade reading skills and 11th-grade print exposure in a 10-year longitudinal study. However, the researchers only tested the direction from skills to exposure. To test what comes first, or causal predominance, we need to take a developmental approach and measure all of the constructs at each timepoint. Consider constructs X and Y, which show longitudinal stability and are correlated at any single timepoint. If at time 1 we only measured X, and at time 2 only Y, we might jump to the conclusion that X influences Y; yet, had we measured Y_{time1} and X_{time2}, we might have come to the opposite conclusion, that Y influences X. In sum, we need to measure both X and Y at both timepoints. This is because it is crucial to correct for the autoregressive effect of Y to evaluate the impact of X on the development or growth of Y, and vice versa for the effect of Y on X.

To our knowledge, there are only four longitudinal studies on the development of reading skills and print exposure that corrected for autoregressive effects. First, Aarnoutse and van Leeuwe (1998) assessed children's print exposure, reading pleasure, and comprehension annually between grades 2 and 6. The researchers found that reading pleasure and print exposure developed in conjunction, but largely independently from reading comprehension. Second, Leppänen, Aunola, and Nurmi (2005) assessed students in grades 1 and 2 on print exposure and reading skills (i.e., accuracy, fluency, comprehension). Apart from one small effect of print exposure on word recognition, cross-lagged effects went from reading skills to print exposure. Third, Harlaar, Deater-Deckard, Thompson, DeThorne, and Petrill (2011) measured children at ages 10 and 11 on print exposure and reading skills (i.e., a composite of accuracy and comprehension). The cross-lagged effect from skills to exposure was statistically significant ($\beta = 0.28$), whereas the reverse effect was absent ($\beta = 0.00$). Finally, Torppa et al. (2019) modeled a random intercept cross-lagged panel model (CLPM) with print exposure and reading skills (i.e., fluency, comprehension) from grade 1 to grade 9. The researchers assessed print exposure of different types of reading material, of which only the reading of books was associated with reading skills. Within-person paths in the early grades ran from both types of reading skills to print exposure. However, from grade 4 onward, the association between reading comprehension and print exposure was reciprocal. In sum, all four studies found how well children read to be more stable over time than how much they read. Together, there seems to be more evidence for reading skills affecting the development of print exposure than vice versa during the early grades. However, studies focusing on the later grades are still rare, and their findings have been inconsistent.

Two recent studies took a different approach to studying causality. Van Bergen et al. (2018) studied causality between reading skills (stressing reading fluency) and print exposure in a very large cross-sectional sample of 7.5-year-old Dutch twins. The researchers took advantage of the genetically sensitive nature of the data to infer the direction of causation. In line with the four longitudinal studies mentioned earlier, van Bergen et al. found evidence for a causal influence of reading skills on print exposure and no evidence at all for the reverse. This study was replicated and extended in a U.S. sample by Erbeli, van Bergen, and Hart (2019). They used the same method but studied reading comprehension (rather than basic reading skills) in older children (approximately 12 years of age). Although the picture was less clear than in the study of 7.5-year-olds (van Bergen et al., 2018), Erbeli et al. also found the most support for an effect of skills on print exposure. That is, it seemed that the extent to which children chose reading activities reflected their reading skills at least partly.

The picture is still far from complete, as the direction and strength of effects may well depend on the developmental stage and the type of reading skill. There is ample evidence (e.g., Florit & Cain, 2011) to separate basic decoding skills from higher order reading skills in their association with print exposure. Basic skills precede and are a prerequisite for higher order reading skills. Skills other than decoding that are needed for comprehending texts are language and cognitive skills that help construct a representation of a text, such as syntax, vocabulary, background knowledge, inference making, comprehension monitoring, and working memory (Oakhill, Berenhaus, & Cain, 2015). Basic reading skills have partly distinct underlying cognitive factors. Hence, children can be impaired in just one reading domain (Catts & Weismer, 2006; Torppa, Tolvanen, 2007). Parallel to the development of reading skills, the reading circuit in the brain develops (Ozernov-Palchik & Gaab, 2016), the pattern of cognitive underpinnings alters slightly (e.g., Vaessen & Blomert, 2010), the emphasis in teaching moves gradually from decoding to comprehension, and the reading material often changes. Accordingly, the relations between reading abilities and reading habits may very well change during and after primary school. One of the proposed theories to explain the print exposure–reading ability relation (Becker et al., 2010; Guthrie, Wigfield, Metsala, & Cox, 1999) states that more exposure to texts increases reading fluency or efficacy. More efficient reading in turn frees up cognitive resources, which can then be employed for higher order information processing, leading to better text comprehension (verbal efficiency theory; e.g., Perfetti, 1985). In our final path model (see Figure 4), we tested this proposed path (i.e., print exposure → later reading fluency → still later reading comprehension).

Furthermore, the link between reading ability and amount may start to develop even prior to primary school. Morgan, Fuchs, Compton, Cordray, and Fuchs (2008) speculated that those who lag behind on prereading skills at the preschool stage may not be interested in books and, hence, may not benefit from early intervention efforts. As of yet, to our knowledge, there has been no study that has tackled this hypothesis.

In the current study, we tracked children's development over a 10-year time span from the two years prior to school entry (age 5 years) all the way to the end of lower secondary school, with reading comprehension from the Programme for International Student Assessment (PISA) at age 15 as the ultimate outcome measure. We modeled the associations between reading skills and print exposure, taking autoregressive effects into account. Additionally, we compared the print exposure of preschool children who did and did not lag behind on prereading skills. Such an investigation gave us the opportunity to address the question of the direction of effect between print exposure and reading skills. In brief, we aimed to determine whether reading skills predict print exposure, vice versa, or both (i.e., reciprocal influences). Within this broader aim, we were interested in possible differences in these relations across skills (i.e., fluency, comprehension) and across development (from age 5 to 15).

Method

Participants

Our final sample for the current article comprised 200 children from the Jyväskylä Longitudinal Study of Dyslexia (JLD). The JLD is a prospective study of 222 children who participated from 1993 to 2012 (see Lyytinen, Erskine, Hämäläinen, Torppa, & Ronimus, 2015). The participating families were invited from maternity clinics in central Finland between 1993 and 1996. More than 9,000 families responded to the first questionnaires of interest to participate in the study. After questionnaire screening, parental interviews, and assessments of the parents' reading, spelling, and cognitive skills, families with and without familial risk for dyslexia were invited. Originally, 222 families participated in the study, which followed the development of the yet unborn child. Out of the 222 children, 22 were excluded from the current study because they did not have any data on reading skills. Of these 22, 20 had dropped out before school age, and the remaining two and were omitted because they only had data on print exposure until age 7 or 9 (besides having no data on reading skills). The remaining 200 were included in the analyses, 107 of whom had high family risk for dyslexia (i.e., parental dyslexia) and 93 of whom had low family risk (i.e., both parents had typical reading skills). For the assessment of parental reading skills, see Leinonen et al. (2001). Fifty-three of the high-risk and 52 of the low-risk children were boys.

The educational levels of the families of high and low family risk were matched. On a scale ranging from 1 (*comprehensive school education without any vocational education*) to 7 (*master's or doctoral degree*), the mothers and fathers of the study participants had an average level of education of 4.34 (standard deviation [*SD*] = 1.43) and

3.74 (*SD* = 1.33), respectively. All of the children spoke Finnish as their native language (and were tested in Finnish) and had no severe mental, physical, or sensory impairments. None of the children had a standard score below 80 on performance and verbal IQ assessed in grade 2 (via the third edition of the Wechsler Intelligence Scale for Children; Wechsler, 1991). All of the children attended mainstream public schools following the national curriculum. In Finland, children enter school and begin formal reading education in August of the year they turn 7.

Previous studies on the JLD sample have shown that the group at high family risk performs lower than the group at low family risk on reading (related) skills. This was also reflected in the finding that dyslexia is approximately four times more prevalent in JLD's high-risk group as compared with the low-risk group. Having said that, the collapsed data of the groups with and without family risk showed a normal distribution (no evidence of bimodality). This is in line with family risk being multifactorial and continuous (van Bergen, van der Leij, & de Jong, 2014). Despite mean differences, the pattern among variables (i.e., the variance–covariance matrices) could still be highly similar in the two groups. We investigated this by testing the similarity of the path models (see the Statistical Analyses section). Regarding the home literacy environment, the JLD groups at high and low family risk did not differ statistically significantly in terms of shared reading, library visits, or number of books at home (Torppa, Poikkeus, 2007).

Classmates of the Participants

To investigate whether the longitudinal sample described earlier was representative of the population, we compared their reading level with that of their classmates. The longitudinal sample's and their classmates' reading skill were assessed using group-administered tasks in grades 1, 2, 3, 7, and 9. These group-administered tasks assess silent reading and are described in Appendix A. The participants of the longitudinal sample attended many different schools. Hence, a large number of children were tested: approximately 1,500 classmates per wave. Despite the large power, none of the group comparisons were statistically significant (see Appendix B).

Ethical approval for the JLD was obtained from the Research Ethics Committee of the Central Finland Health Care District (protocol number 66/2004).

Measures

The children's prereading skills were assessed at age 5, and their reading skills were assessed in grades 1, 2, 3, 8, and 9. At the time of our assessments, the participants were on average 5.50 years old, 7.91 years old (May, grade 1), 8.98 years old (June, grade 2), 9.86 years old (April, grade 3), 14.36 years old (November, grade 8), and 15.90 years old (May, grade 9). Questionnaires on print exposure were sent to the parents around the time of their child's 5th, 7th, 8th, 9th, and 13th birthdays. In the current article, we report on the children's reading fluency, comprehension, and print exposure at all timepoints when they were assessed. Due to lack of funding, children were not assessed during grades 4, 5, and 6.

Prereading Skills

At age 5, children's letter knowledge and phonological awareness were individually tested. Phonological awareness was measured with four tasks: first-phoneme identification, first-phoneme production, segment identification, and synthesis. A composite score for phonological awareness was calculated by averaging the *z*-scored scores (sums of correct answers) of the four tasks. *Z*-scores were calculated based on the low-risk group's distribution. Letter knowledge was assessed by asking the children to identify, one by one, 29 lowercase letters in the Finnish alphabet. The measure for prereading skills was the mean of the *z*-scored phonological awareness composite and the *z*-scored letter knowledge measure (number of letters named correctly). Cronbach's alpha for the prereading skills measure was .77. Next, we give the full descriptions of the four phonological awareness tasks.

In the first-phoneme identification task, children were shown four pictures of objects, and the computer named each object. The sound of an initial phoneme was presented, and children were asked to select the picture of the word that starts with that phoneme (e.g., "In the beginning of which word do you hear *k*?"). There were two practice items and nine test items.

In the first-phoneme production task, children were shown a picture of an object and asked to articulate the first sound (phoneme or letter name) of the object. There were two practice items and eight test items.

In the segment identification task, three pictures of objects were presented on a computer screen, and each object was named by the computer (e.g., *koira* [dog], *kissa* [cat], *kukko* [cock]). Children were asked to identify on a touch screen the picture that contained a specified subword-level unit (syllable or phoneme) within the target (e.g., the *koi* in the word *koira*). The size of the segment to be identified varied from one to four phonemes and came from the beginning, end, or middle part of the word. There were three practice items and 14 test items.

In the synthesis task, children were presented with segments (syllables or phonemes) by the computer, each separated by 750 milliseconds. Children were asked to blend the segments to produce the resulting word (e.g., *per-ho-nen* [butterfly], *m-u-n-a* [egg]). The items were three to nine phonemes long, each divided into three or four segments. Three items required synthesis at the level of syllables (e.g., *per-ho-nen*), five items required synthesis at the level of syllables and phonemes (e.g., *tuo-l-i* [chair]), and four items required synthesis at the level of phonemes (e.g., *m-u-n-a*). There were two practice items and 12 test items.

Reading Fluency

Participants were given a printout of a grade-level-appropriate text and were asked to read the text aloud as quickly and accurately as they could. The total time to read the text was measured. This measure was converted to the number of words read correctly per minute, which was the score used in the analyses. Our measure of oral reading fluency thus emphasized accurate and automatic word recognition (see Fuchs, Fuchs, Hosp, & Jenkins, 2001; Kuhn, Schwanenflugel, & Meisinger, 2010). The task was administered by trained research assistants. Testing took place in grade 1 (May), grade 2 (June), grade 3 (April), and grade 8 (November). In grade 1, the text consisted of 19 sentences (122 words/859 letters); in grade 2, the text consisted of 19 sentences (124 words/877 letters); in grade 3, the text consisted of 18 sentences (189 words/1,154 letters); and in grade 8, the text consisted of 16 sentences (207 words/1,591 letters).

The correlations between the two waves with a one-year gap were .83–.85 (see Table 2), indicating good reliability. Validity of the text-reading fluency task was also good, as indicated by the following correlations of (oral) text-reading fluency with (oral) word list–reading fluency, the group-administered silent reading tests, and the teacher's evaluation of reading skills (for descriptions, see Appendix A): In grade 1, r = .87 with the silent word-reading fluency task and .77 with the teacher's evaluation; in grade 2, r = .90 with word list–reading fluency, .74 with a composite of the silent word-reading fluency task and the word chain task, and .74 with the teacher's evaluation; in grade 3, r = .88 with word list–reading fluency, .60 with the silent word-reading fluency task, and .65 with the teacher's evaluation; in grade 8, r = .70 with word list–reading fluency; and in grade 9, r = .74 with a composite of the silent sentence-reading fluency task and the word chain task. The text-reading fluency task was chosen for the current study because it is a fluency measure that is closest to natural reading and because it was assessed at all timepoints.

Reading Comprehension

Reading comprehension was assessed in the participants' classrooms in grades 2, 3, and 9. In grade 2 (April) and grade 3 (April), we used the nationally normed reading test battery (Ala-Asteen Lukutesti [The comprehensive school reading test]; Lindeman, 2000). The children silently read an informational text (about gymnastics in grade 2 and photography in grade 3) and then answered 11 multiple-choice questions (with four answer options) and one question in which they had to arrange five statements in the correct sequence based on the information gathered from the text. All of the questions could be answered using the text, so no background knowledge was required. The text contained 114 words in grade 2 and 139 words in grade 3. Children were given as much time as they needed to complete the test. Lindeman (2000) reported the Kuder–Richardson reliability coefficients of .80 in grade 2 and .75 in grade 3. The score was the number of correct answers, ranging from 0 to 12.

In grade 9 (May), reading comprehension was assessed with the PISA reading test. The test used in this study consisted of the link items, which are used repeatedly in each PISA cycle to ensure the comparability of the measurement (OECD, 2010, 2013). All of the students in our sample took this test as part of the current study, not as a part of the PISA assessments. In the booklet, there were eight different reading materials that the students were asked to read before answering several questions. The reading materials included texts, tables, graphs, and figures. There were 15 multiple-choice questions (with varying answer options) and 16 questions that required written responses. There were three types of questions: 12 required students to access and retrieve information, 12 to integrate and interpret information, and seven to reflect and evaluate information. Students had 60 minutes to complete the task. The score used in the analysis was a mean of standardized scores (mean = 0, SD = 1) for each type of question. The Cronbach's alpha reliability coefficient for the total score in the current sample was .80.

Print Exposure

Print exposure was assessed via a parental questionnaire at ages 5, 7, 8, 9, and 13. The following six items were rated on a 5-point Likert-type scale at ages 5, 7, 8, and 9:

1. How often does your child look at/read books or magazines independently? (1 = *never*; 5 = *many times per day*)
2. What is the typical duration of your child's independent reading episode? (1 = *5 minutes*; 5 = *more than 45 minutes*)
3. How long does your child read per day independently? (1 = *5 minutes*; 5 = *more than 45 minutes*)
4. How often does your child read children's books? (1 = *never*; 5 = *every day*)
5. How often does your child read comics? (1 = *never*; 5 = *every day*)
6. How interested is your child in book reading? (1 = *not at all interested*; 5 = *very interested*)

At age 5, none of the children could yet read texts, but the word *reading* in the original Finnish items included both text reading and looking at picture books. At age 13, there were four items: two for the mother and two for the father. The following two questions were rated on a 5-point Likert-type scale:

1. How often does your child read books? (1 = *every day*; 5 = *never*)
2. How often does your child read magazines or comics? (1 = *every day*; 5 = *never*)

The measure for print exposure at each age was the mean of the items (range = 0–5). The Cronbach's alpha reliability coefficients were .75 at age 7, .81 at age 8, .85 at age 9, and .72 at age 13. Regarding validity, at age 13, participants were also asked about their print exposure. Self-report and parent-report correlated as .77, showing good validity.

Statistical Analyses

We originally planned to fit only CLPMs. We report these models in Figures 1–4. Additionally, we fitted post hoc RI-CLPMs (see Figures 5 and 6).

Planned Analyses

All models reported were fitted with a maximum likelihood estimator in M*plus* 7.3 (Muthén & Muthén, 2012). The goodness of fit of the estimated models was evaluated using five indicators: chi-square test, comparative fit index (CFI), Tucker–Lewis index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Good model fit is indicated by a small, preferably statistically nonsignificant chi-square, CFI and TLI greater than 0.95, RMSEA less than 0.06, and SRMR less than 0.08 (Hu & Bentler, 1999).

Because of the reading fluency differences between the high– and low–family risk groups, (i.e., CLPMs) we first ran models separately for the two groups. We built multigroup models and tested the equality of all paths in the two groups, using the chi-square difference test. The final model fit did not deteriorate statistically significantly after setting all paths equal between the two groups, $\Delta\chi^2(37) = 51.08$, $p = .06$. This suggests that the models for the two study groups did not differ statistically significantly. This is in line with other longitudinal family risk studies that found similar relations among variables in groups of children with and without family risk (Hulme, Nash, Gooch, Lervåg, & Snowling, 2015; van Bergen, de Jong, et al., 2014). Therefore, we report only models (depicted in Figures 1–4) on the full sample. The final model that we report (see Figure 4) includes all three traits (i.e., reading fluency, reading comprehension, print exposure) and includes age 5 (prereading skills and print exposure). The first three models each include a pair of traits from grade 1 onward. We report these models as well in the interest of thoroughness and to be able to compare our findings with those in the literature.

In setting up the models, we included all stability (i.e., autoregressive) paths and all cross-lagged paths between measurements in subsequent timepoints. At the final stage, we examined modification indexes provided by M*plus*, and we added paths to the models that were aligned with theory and were suggested to improve the model fit. This led to the inclusion of two paths: an extra print exposure stability path (grade 1 to grade 3) and a path from grade 3 print exposure to grade 9 reading comprehension.

Of particular interest in these path models are the cross-lagged paths. Note that these models control for autoregressors, meaning that they control for the prior level of the trait being predicted. Hence, the autoregressor predicts the stable portion of the trait, and the cross-lagged path predicts part of the change portion of the trait. Cross-lagged effects in different directions should not be compared if the different constructs vary widely in their reliability. However, this was not the case in the current study, as the reliabilities only ranged from .72 to .85. The literature mentioned in the introduction found, for all three traits, longitudinal within-trait correlations (e.g., from reading comprehension at time 1 to reading comprehension at time 2) that are neither close to zero nor close to unity, justifying modeling stability (with autoregressors) and change (with cross-lagged paths and residuals). The CLPMs do not focus on stability and change within persons but rather describe stability and change in individual differences (Selig & Little, 2012). Statistically significant cross-lagged paths suggest systematic effects on change over time. Therefore, we think that even small path estimates are of importance.

Post Hoc Analyses

Finally, we fitted post hoc models to address the concern that the traditional CLPM estimates mix between-person variance (stable differences between individuals across time) and within-person variance (fluctuations around the stable level at each timepoint; e.g., Berry & Willoughby, 2017; Curran, Howard, Bainter, Lane, & McGinley, 2014; Hamaker, Kuiper, & Grasman, 2015). Therefore, we applied models based on the RI-CLPM, as suggested by Hamaker et al. (2015). In RI-CLPMs, the stable differences between individuals over time are estimated separately from the within-person changes from timepoint to timepoint. Cross-lagged paths are estimated between the within-person factors at each timepoint. In these analyses, we focused only on the models that answered our main questions: the associations of print exposure with each of the reading skills (fluency and comprehension).

Results

This section starts with two short subsections devoted to the descriptive statistics and associations prior to school entry, respectively. This is followed by subsections on longitudinal modeling.

Descriptive Statistics

For all variables, Table 1 reports the means, standard deviations, *N*s, and comparisons between the low– and high–family risk groups. Table 2 shows the correlations among all variables. Data for all variables were approximately normally distributed. The proportion of missing values ranged between 1% and 13% per measure, except for grade 7 print exposure, where 25% was missing according to Little's missing completely at random test, $\chi^2(405) = 392.17, p = .67$.

There were no statistically significant differences (with alpha set at .05) between the groups differing in risk status (see Appendix C) in reading comprehension or print exposure (effect sizes = −0.26–0.30). However, the group with high family risk read less fluently on average (effect sizes = 0.48–0.61) and had weaker prereading skills (effect size = 0.58).

TABLE 1
Descriptive Statistics

Timepoint and measure	*N*	Mean	Standard deviation	Minimum	Maximum
Prereading skills					
Age 5	194	−0.26	0.93	−2.23	1.65
Reading fluency (words read correctly per minute)					
Grade 1	189	35.45	22.62	2.67	103.33
Grade 2	196	61.28	25.92	6.10	135.27
Grade 3	192	73.30	25.71	16.55	144.62
Grade 8	182	87.58	18.09	28.53	128.40
Reading comprehension					
Grade 2	170	8.94	2.77	0.00	12.00
Grade 3	179	9.88	1.70	3.00	12.00
Grade 9	159	0.06	0.92	−2.83	1.49
Print exposure					
Age 5	200	3.31	0.72	1.33	5.00
Grade 1	181	3.06	0.65	1.50	4.33
Grade 2	180	3.33	0.66	1.67	5.00
Grade 3	175	3.31	0.72	1.50	4.33
Grade 7	163	2.74	0.67	1.00	4.50

Note. The score on prereading skills is the mean of *z*-scores for phonological awareness and letter knowledge.

Prereading Skills and Print Exposure

To address the question of whether exposure to print and literacy development are linked already prior to school entry, we examined whether print exposure and prereading skills were already associated at age 5. Prereading skills and early print exposure were not statistically significantly correlated, $r = .06, p = .393$. In addition, the print exposure of the children with low prereading skills (≥ 1 *SD* below the mean of the low-risk group) did not differ from that of the other children, $t(192) = 0.31, p = .760$. The conclusions remained the same when we conducted the analyses separately for the low- and high-risk groups.

CLPMs

To recapitulate from the introduction, we aimed to investigate the direction of effects over time (i.e., reading skills → print exposure, print exposure → reading skills, and/or reciprocal influences). More specifically, we asked whether these effects might be developmentally different and reading skill dependent. The path models are depicted in Figures 1–4; the path estimates given are standardized. Here, we describe the three bivariate models, followed by the final trivariate model, which includes reading fluency, reading comprehension, and print exposure. For each endogenous variable, the figures include the amount of explained variance (*R*-squared). Squaring a path estimate gives the path's contribution to *R*-squared. For example, in Figure 4, the prereading skills explained 32% ($0.563^2 = 0.32$) of the variance in grade 1 reading fluency. To be precise, the prereading skills also explained variance in grade 1 reading fluency via age 5 print exposure, but this was negligible given the very small path estimates (0.08 and −0.01).

Developmental Model for Reading Fluency and Reading Comprehension

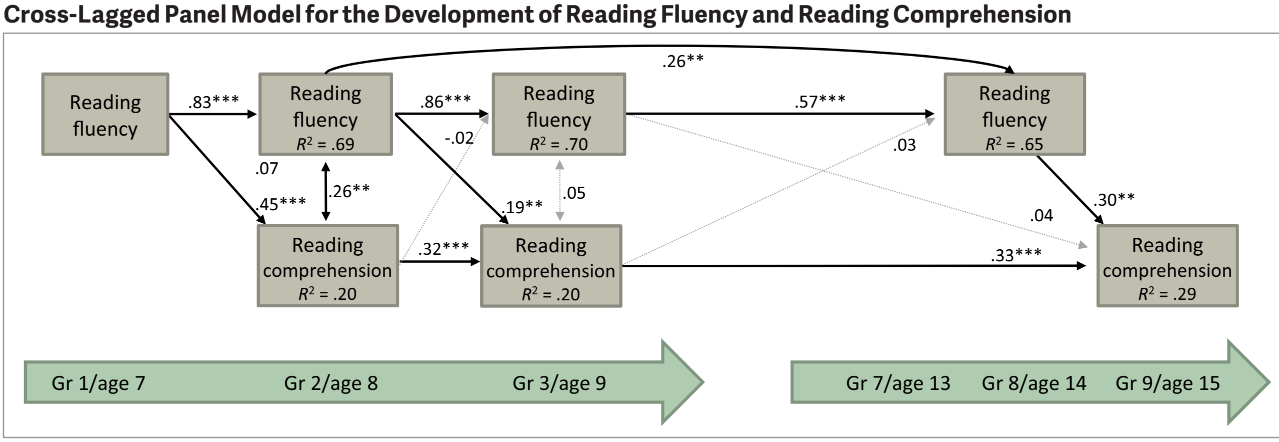
Although the interrelation between the two aspects of reading skills is not the focus of the current article, we included the path model of their codevelopment for completeness. The model (see Figure 1) showed strong stability for reading fluency and lower stability for reading comprehension. Reading fluency predicted reading comprehension over and above the autoregressive effect, but reading comprehension did not predict later reading fluency. Regarding the variance in reading fluency, 65–70% at different timepoints was predicted by the model, whereas 20–29% of the variance in reading comprehension was predicted by the model. The model fit to the data was excellent, $\chi^2(7) = 7.10, p = .42$; CFI = 1.00; TLI = 1.00; RMSEA = 0.01, 90% confidence interval (CI) [0.00, [0.09]; SRMR = 0.02.

TABLE 2
Correlations Among All Variables

Timepoint and measure	1	2	3	4	5	6	7	8	9	10	11	12	13
Prereading skills													
1. Age 5	—												
Reading fluency													
2. Grade 1	.57***	—											
3. Grade 2	.56***	.83**	—										
4. Grade 3	.45***	.73**	.85**	—									
5. Grade 8	.43***	.62**	.76**	.79**	—								
Reading comprehension													
6. Grade 2	.49***	.45**	.54**	.44**	.42**	—							
7. Grade 3	.34***	.30**	.38**	.36**	.36**	.43**	—						
8. Grade 9	.45***	.30**	.40**	.40**	.42**	.41**	.45**	—					
Print exposure													
9. Age 5	.06	.07	.07	.01	.06	.24**	.05	.20*	—				
10. Grade 1	.21**	.30**	.29**	.24**	.08	.22**	.01	.28**	.48***	—			
11. Grade 2	.12	.29**	.26**	.20**	.11	.24**	.12	.26**	.52***	.69**	—		
12. Grade 3	.26**	.41**	.45**	.36**	.19*	.36**	.23**	.37***	.45***	.68**	.73**	—	
13. Grade 7	.10	.20*	.16	.14	.13	.11	.17	.23**	.24**	.31**	.36**	.45**	—

*p < .05, two-tailed. **p < .01, two-tailed. ***p < .001, two-tailed.

FIGURE 1
Cross-Lagged Panel Model for the Development of Reading Fluency and Reading Comprehension

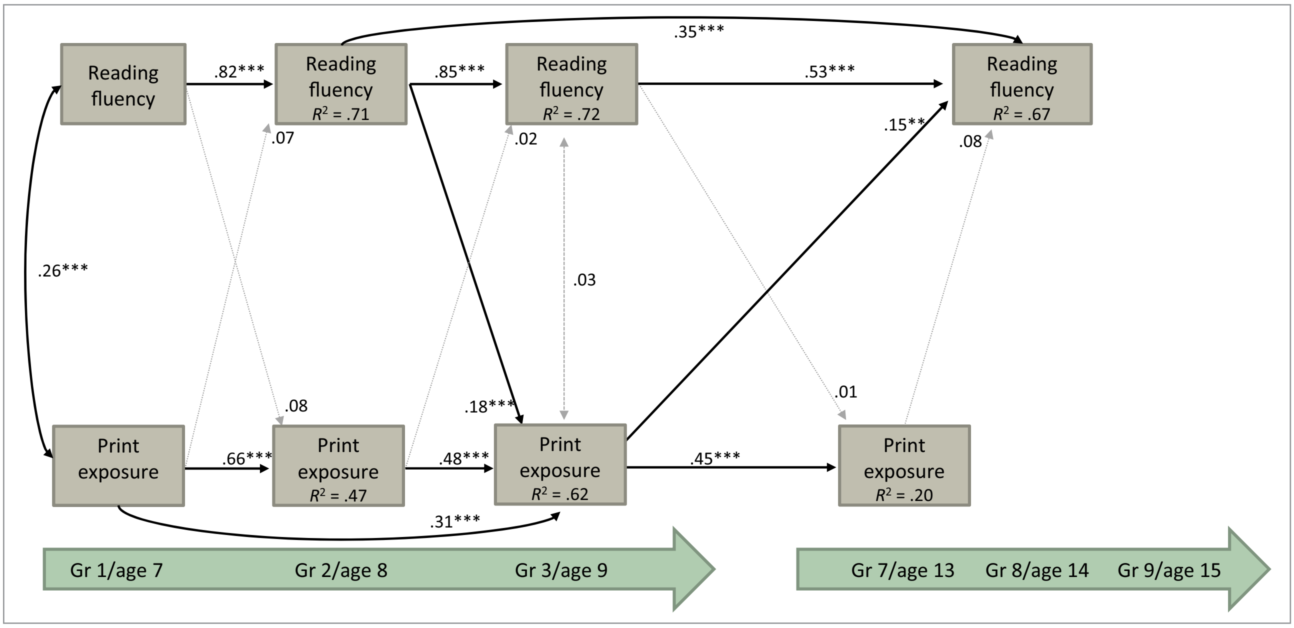


Note. Gr = grade. All paths represent standardized estimates (βs). The color figure can be viewed in the online version of this article at http://ila.onlinelibrary.wiley.com.
**p < .01. ***p < .001.

Developmental Model for Reading Fluency and Print Exposure

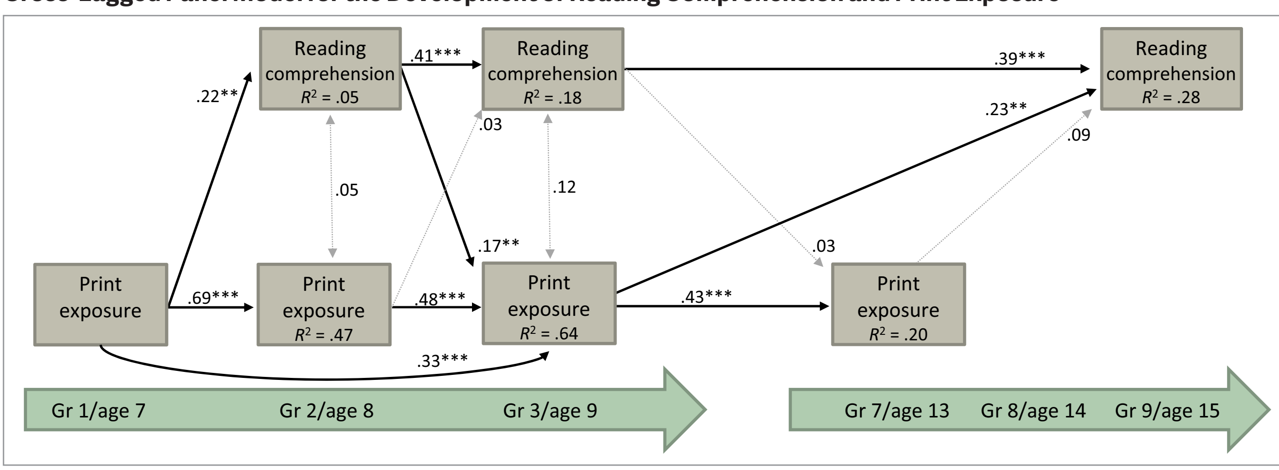
The path model for reading fluency and print exposure (see Figure 2) showed reciprocal links between the measures across time. Grade 2 reading fluency predicted grade 3 print exposure, and grade 3 print exposure predicted grade 8 reading fluency. There was a modest correlation (.26) between reading fluency and print exposure in grade 1. The model fit to the data was excellent, $\chi^2(10)$ = 6.29, p = .79; CFI = 1.00; TLI = 1.00; RMSEA = 0.00, 90% CI [0.00, 0.05]; SRMR = 0.01.

FIGURE 2
Cross-Lagged Panel Model for the Development of Reading Fluency and Print Exposure



Note. Gr = grade. All paths represent standardized estimates (ßs). The color figure can be viewed in the online version of this article at http://ila.onlinelibrary.wiley.com.
$^{**}p < .01$. $^{***}p < .001$.

FIGURE 3
Cross-Lagged Panel Model for the Development of Reading Comprehension and Print Exposure



Note. Gr = grade. All paths represent standardized estimates (ßs). The color figure can be viewed in the online version of this article at http://ila.onlinelibrary.wiley.com.
$^{**}p < .01$. $^{***}p < .001$.

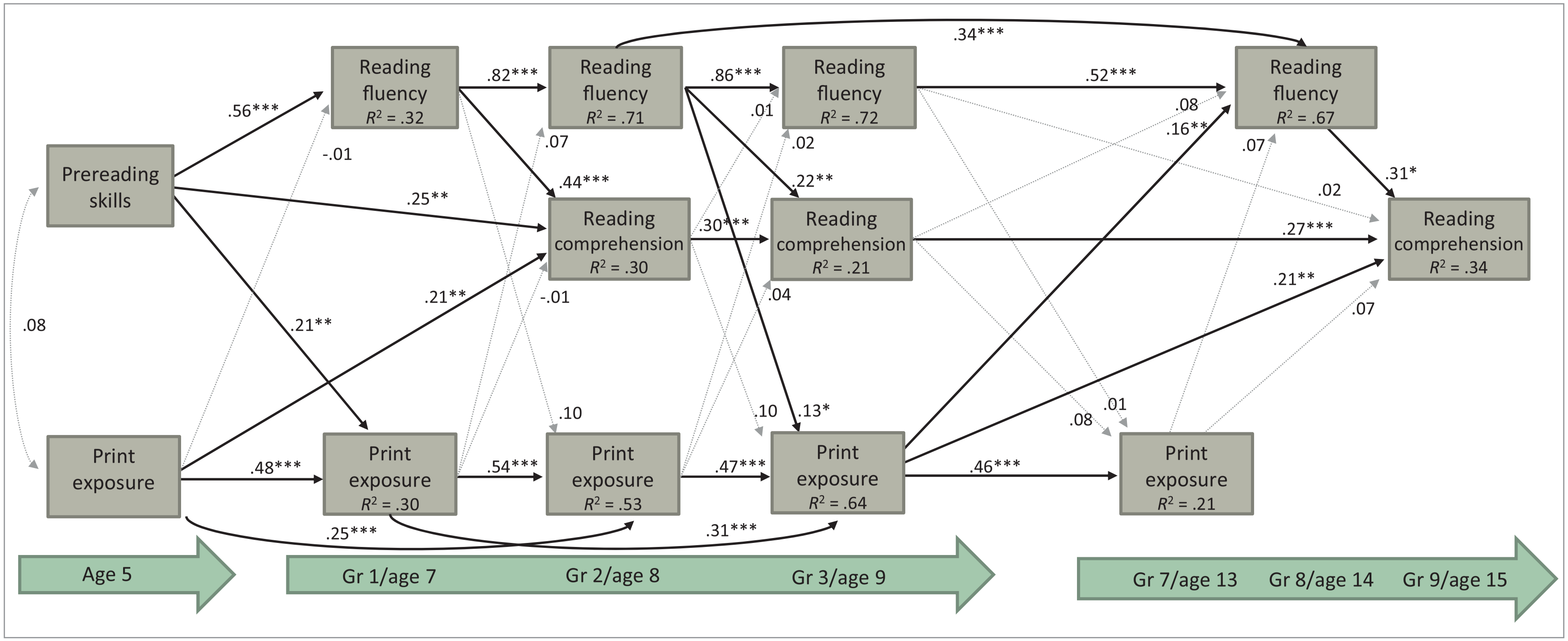
Developmental Model for Reading Comprehension and Print Exposure

The path model for reading comprehension and print exposure (see Figure 3) showed reciprocal links between the measures across time. Grade 1 print exposure predicted grade 2 reading comprehension, which in turn predicted print exposure in grade 3. Finally, grade 3 print exposure predicted grade 9 reading comprehension. The model fit to the data was acceptable, $\chi^2(7) = 14.23$, $p = .05$; CFI = 0.98; TLI = 0.94; RMSEA = 0.08, 90% CI [0.01, 0.13]; SRMR = 0.04.

Developmental Model for Reading Fluency, Reading Comprehension, and Print Exposure

The final model (see Figure 4) included all reading fluency, reading comprehension, and print exposure measures from

FIGURE 4
Cross-Lagged Panel Model for the Development of Prereading Skills (Letter Knowledge and Phonological Awareness), Reading Fluency, Reading Comprehension, and Print Exposure



Note. Gr = grade. All paths represent standardized estimates (βs). The color figure can be viewed in the online version of this article at http://ila.onlinelibrary.wiley.com.
$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

age 5 onward. In addition to the paths depicted in Figure 4, we included error covariances among reading comprehension, reading fluency, and print exposure within grades 2 and 3 in the model. Two of these error covariates were statistically significant: the one between reading fluency and print exposure in grade 1 (standardized estimate 0.18) and the one between reading fluency and reading comprehension in grade 2 (standardized estimate 0.25). Model fit to the data was good, $\chi^2(34) = 50.75$, $p = .03$; CFI = 0.99; TLI = 0.97; RMSEA = 0.05, 90% CI [0.02, 0.08]; SRMR = 0.03. As expected, prereading skills (phonological awareness and letter knowledge) predicted later reading skills, especially reading fluency. Print exposure at age 5 was unrelated to prereading skills but predicted reading comprehension three years later, in grade 2. Further variance in grade 2 reading comprehension was explained by grade 1 reading fluency. Note that these effects were not over and beyond an autoregressor, as reading comprehension was not measured earlier in time.

There were five statistically significant cross-lagged paths after taking autoregressors into account. To begin with, prereading skills at age 5 predicted changes in print exposure in grade 1. Next, grade 2 reading fluency predicted growth in grade 3 reading comprehension and print exposure. Subsequently, grade 3 print exposure predicted growth in grade 8 reading fluency and grade 9 reading comprehension (PISA).

Finally, reading comprehension in grade 9 (PISA) was predicted by all three constructs: earlier reading comprehension, reading fluency, and print exposure. Note that for print exposure, the statistically significant contribution came from the grade 3 measure, not the grade 7 measure.

RI-CLPMs (Post Hoc)

We modeled the RI-CLPMs for print exposure and reading fluency (see Figure 5) and print exposure and reading comprehension (see Figure 6) post hoc. The variance of the observed variables is disaggregated into the timepoint-specific within-person variances (shown in the middle circles) and the between-person variance (captured by the random intercepts in the ovals at the top and bottom), which indicates trait-like stability, or a child's overall level. The models fitted well, reading fluency model: $\chi^2(9) = 9.96$, $p = .35$; CFI = 1.00; RMSEA = 0.00, 90% CI [0.00, 0.08]; SRMR = 0.02; reading comprehension model: $\chi^2(1) = 0.82$, $p = .36$; CFI = 1.00; RMSEA = 0.00, 90% CI [0.00, 0.18]; SRMR = 0.01. In both models, the random intercepts were only weakly linked (0.10 and 0.08, respectively), so most of the trait associations were at the within-person level.

Regarding reading fluency, the models with and without random intercept differed in two respects (compare Figures 2 and 5). First, cross-lagged pathways between grades 1 and 2 reading fluency and print exposure became stronger and statistically significant. This suggests that if students were active readers in grade 1 in comparison with their own overall print exposure level across time, their reading fluency developed quickly between grades 1 and 2. Similarly, if students were fast readers in grade 1 in comparison with their overall level across time, they also became more active readers between grades 1 and 2. Second, the path estimate from grade 3 print exposure to grade 8 reading fluency decreased from 0.15 (see Figure 2) to 0.09 (see Figure 5) and was no longer statistically significant. What remained the same across models was the statistically significant path from grade 2 reading fluency to grade 3 print exposure.

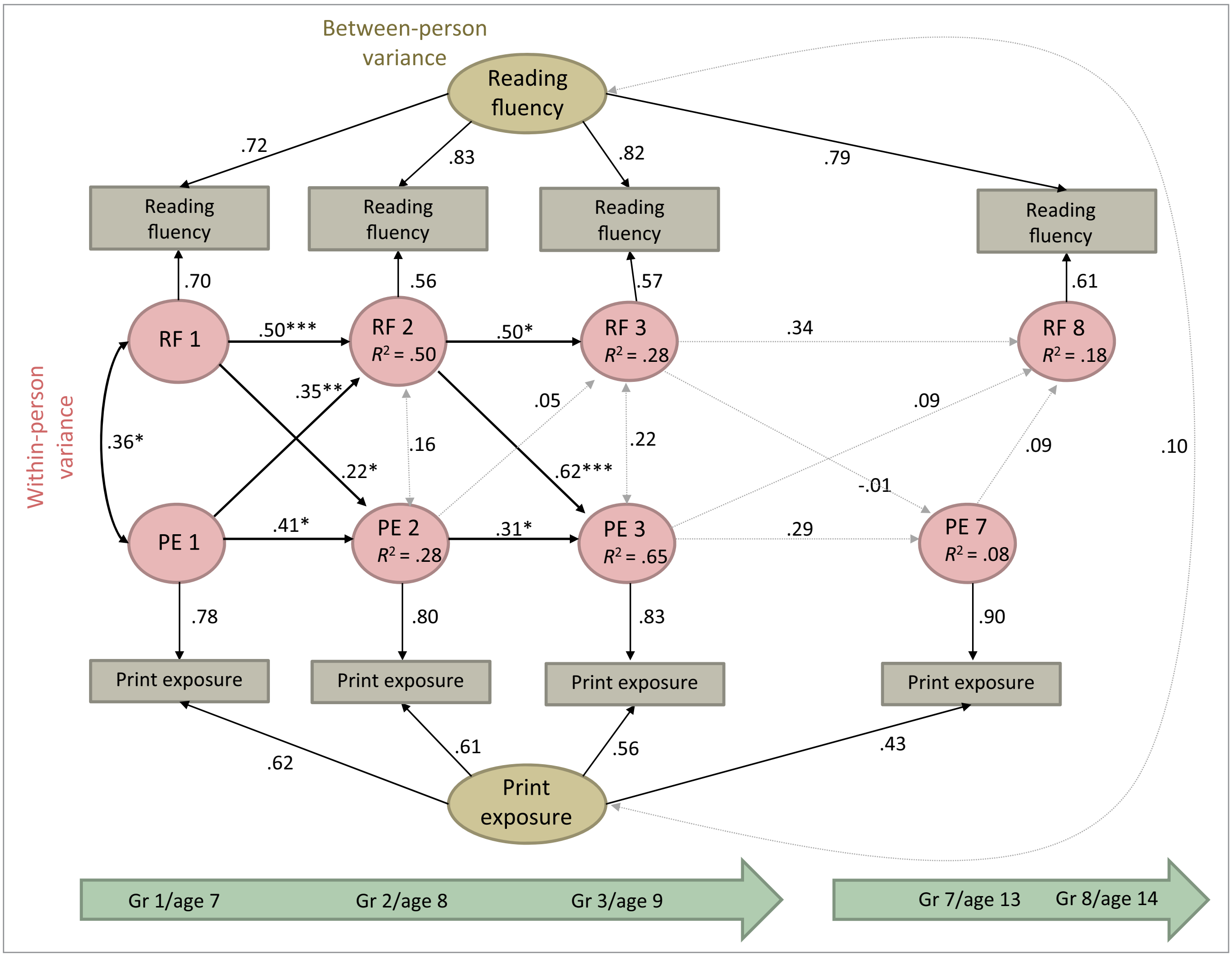
Regarding reading comprehension, the same two cross-lagged paths emerged as statistically significant in the models with and without random intercept (compare Figures 3 and 6): grade 2 comprehension to grade 3 print exposure, and grade 3 print exposure to grade 9 reading comprehension.

Discussion

We set out to investigate whether young children who read a lot become good readers as a result of all the practice or whether children who initially pick up reading skills easily go on to read a lot. Accordingly, we followed the progress of children from age 5 (preschool) to 15 (ninth grade) in terms of how much they read and two aspects of how well they read. We modeled the links within and between constructs over time, as depicted in Figure 4. After accounting for autoregressive effects, a few effects were weak but statistically significant. The direction of causation ran during the early grades mainly from foundational reading skills (i.e., prereading skills, reading fluency) to later reading comprehension and print exposure, meaning that children's early reading fluency levels predict future change in how much they read and how well they comprehend what they read. In contrast, after grade 3, print exposure predicted later growth in reading skills, particularly comprehension. Thus, we found that the association between print exposure and reading skills is developmental in nature.

Although previous work has shown that reading ability and amount of reading are linked during the school years, we showed for the first time that the link between ability and amount only kicks in when formal reading instruction starts. Preschoolers scoring low on prereading skills were not less interested in reading activities than their peers were. Also, the concurrent correlation between print exposure and skills was absent in preschool (0.06) yet present in grade 1 (0.30). Note that none of the children could yet read at age 5, so print exposure here assessed how often

FIGURE 5
Random Intercept Cross-Lagged Panel Model for the Development of Reading Fluency and Print Exposure

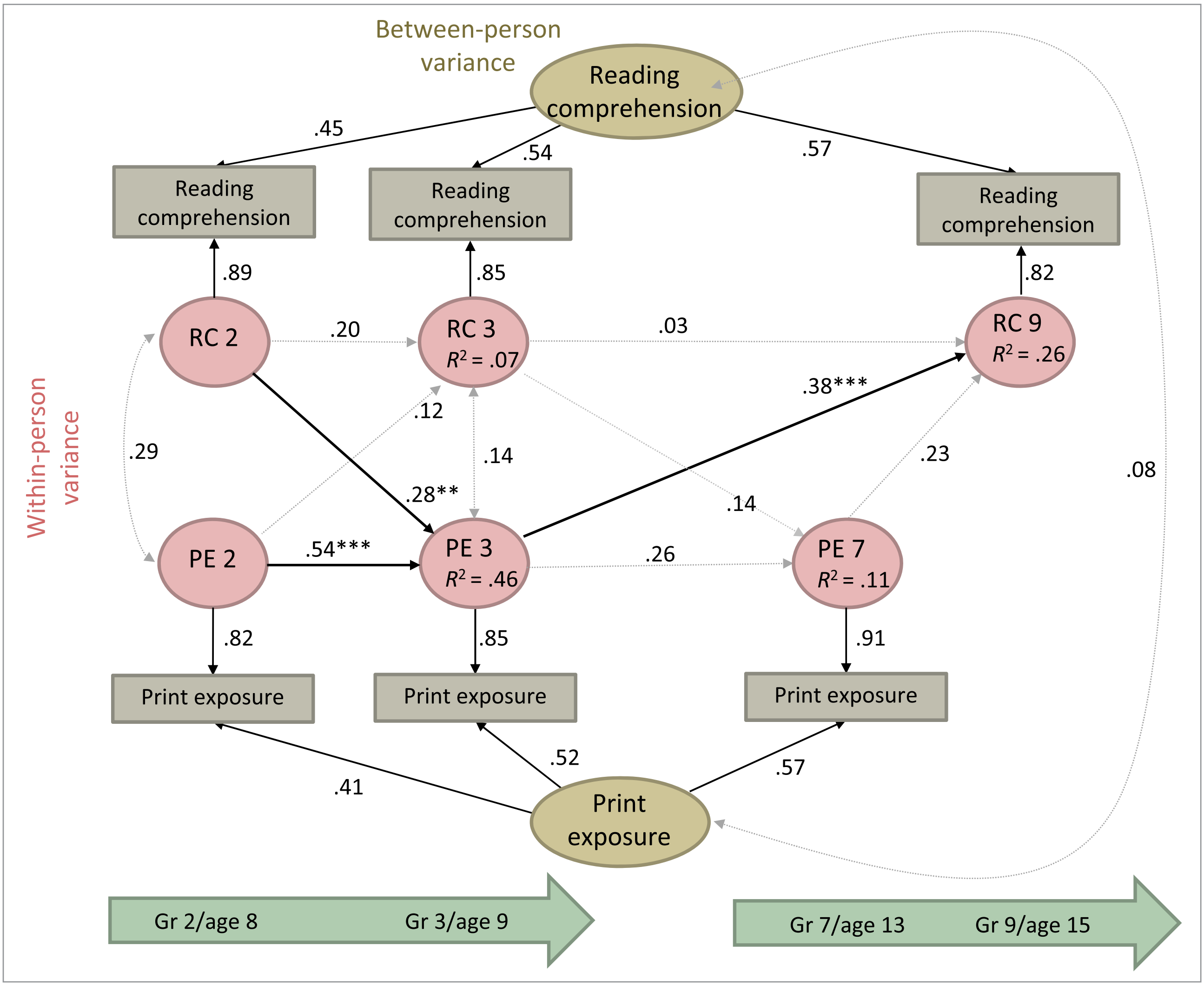


Note. Gr = grade; PE 1 = print exposure in grade 1, etc.; RF 1 = reading fluency in grade 1, etc. The between-subject part (the top and bottom ovals) captures children's general level. The within-subject part (the circles in the middle) shows the within-person associations. All paths represent standardized estimates (ßs). The color figure can be viewed in the online version of this article at http://ila.onlinelibrary.wiley.com.
*p < .05. **p < .01. ***p < .001.

prereaders amused themselves with comics and picture books. Even though this was not related to prereading skills, it predicted reading comprehension three years later. We cannot conclude from the current study whether that longitudinal link is (partly) causal or due to a third factor, like their common link with oral language skills.

Regarding the direction of effects in the early school years, prereading skills at age 5 and both aspects of reading skills in grade 2 predicted whether children would go on to read more or less in the following year (see Figures 2 and 3; in line with the findings of Leppänen et al., 2005). In the full model (see Figure 4), the effects that remained statistically significant ran from age 5 prereading skills to subsequent growth in print exposure and from grade 2 reading fluency to subsequent growth in comprehension skills and print exposure, so generally, skills → exposure. The findings that supported the predictive association from early reading skills to print exposure were also confirmed in the post hoc RI-CLPM for both reading comprehension and reading fluency. However, in the RI-CLPM, we found one path that ran in the opposite direction (i.e., grade 1 exposure to grade 2 fluency). Our finding of skills → exposure during the early grades is in line with the findings of the six studies discussed in the introduction (Aarnoutse & van Leeuwe, 1998; Erbeli et al., 2019; Harlaar et al., 2011; Leppänen et al., 2005; Torppa et al., 2019; van Bergen et al., 2018). Be aware that all of these studies were conducted in alphabetic writing systems. It remains to be studied how reading skills and print exposure relate in nonalphabetic writing systems with a large set of characters to be learned. Our finding of skills → exposure is also in line with the literature on intrinsic reading motivation, which is linked to print exposure. Becker et al. (2010) showed that early intrinsic motivation

FIGURE 6
Random Intercept Cross-Lagged Panel Model for the Development of Reading Comprehension and Print Exposure



Note. Gr = grade; PE 2 = print exposure in grade 2, etc.; RC 2 = reading comprehension in grade 2, etc. The between-subject part (the top and bottom ovals) captures children's general level. The within-subject part (the circles in the middle) shows the within-person associations. All paths represent standardized estimates (βs). The color figure can be viewed in the online version of this article at http://ila.onlinelibrary.wiley.com.
**p < .01. ***p < .001.

did not predict later reading comprehension over and beyond early reading comprehension. This led them to conclude that children do not struggle with reading because they lack motivation; rather, children lack motivation because they struggle. Struggling readers do not experience enjoyment or competence from reading.

Our study was less suited to test leading and lagging effects after grade 3. Yet, the exceptions from the early trend of ability to print exposure were the paths from print exposure in grade 3 to fluency and comprehension in grades 8 and 9. From a statistical viewpoint, however, keep in mind that autoregressors for reading skills in grades 8 and 9 were less strong due to the large time gap. Hence, there was more scope for print exposure to explain variance. Nevertheless, it is striking that how much children read had a small but statistically significant impact on the development of basic reading skills 5 years later. Nevertheless, this predictive path from reading fluency to later print exposure did not reach statistical significance in the model with random intercept. In contrast, the RI-CLPM confirmed the predictive path from comprehension to later exposure (0.38; see Figure 6). This result is in line with the findings of Torppa et al. (2019), who showed that from grade 3 onward, frequent book reading predicted growth in reading comprehension. In both Torppa et al.'s study and the current study, most of the print exposure → reading skills effects only appeared from grade 3 onward. This fits with grade 3 being a turning point in reading education, when the curriculum switches from learning to read to reading to learn. After having mastered the technique of reading, children can choose reading material according to their own interests. In other words, at that age, motivation (i.e., the choice to read) can bring about significant individual differences in the frequency and complexity of the reading experiences. For example, reading a Harry Potter book can provide as many as 250,000 words of reading practice and, at the same time, could also improve one's vocabulary and understanding of contextual information, which can later on help the reader infer the meaning of previously unknown words (Cain & Oakhill, 2011; Nagy, Anderson, & Herman, 1987). Our finding also fits with Nation's (2017) lexical legacy hypothesis, in which she states that print exposure fosters word-level reading skills because knowledge of words (including their phonology, orthography, semantics, and frequency) accumulates by encountering words in meaningful and diverse texts. However, returning to the mechanism mentioned in the introduction (Becker et al., 2010; Guthrie et al., 1999), in the path model spanning age 5 to age 15 (see Figure 4), we did not see evidence for a mechanism in which increased print exposure early on leads to more efficient reading, which in turn leads to better comprehension.

Moving to the outcome measure at age 15—PISA reading comprehension—earlier reading comprehension, fluency, and print exposure all contributed statistically significantly and about equally, together explaining one third of individual PISA differences. Children who have read a lot may comprehend texts better simply because of having accumulated more practice but also because of having better grammatical understanding and a larger vocabulary, both important for reading comprehension (Hulme & Snowling, 2011). It is remarkable that basic reading skills also contributed uniquely to PISA scores. This may be because readers who struggle with word identification (i.e., readers with dyslexia) have fewer cognitive resources left to devote to extracting meaning from text (e.g., Miller et al., 2013; Perfetti, 1985). Also, readers with dyslexia often have attention problems (Boada, Willcutt, & Pennington, 2012), which may lower reading comprehension performance. What would likely aid in explaining the PISA reading variance are measures of listening comprehension, working memory, inference-making skills, and comprehension monitoring, which could overshadow the effects of word-level reading (Cain, Oakhill, & Bryant, 2004; Eklund, Torppa, Sulkunen, Niemi, & Ahonen, 2018; Torppa et al., 2016).

The term *print exposure* may suggest that the child is exposed to a certain reading environment in a passive way. However, especially as children age, they actively choose whether to engage in reading. Although the current study was not genetically sensitive, a large body of genetically sensitive research has demonstrated that reading skills are highly heritable (de Zeeuw, de Geus, & Boomsma, 2015; Olson, Keenan, Byrne, & Samuelsson, 2014). As far as we know, the heritability of print exposure has not been studied in prereaders. However, a few groups studied print exposure in primary school children, and here print exposure also seems to be genetically influenced, partly by the same genes that influence reading skills (Erbeli et al., 2019; Harlaar et al., 2011; Harlaar, Trzaskowski, Dale, & Plomin, 2014; van Bergen et al., 2018). More general and contrary to common belief, motivation for school subjects is 40% heritable, with no influences from the family environment (Kovas et al., 2015). Thus, the literacy environment that children seek out is partly driven by their genotype. This was recently demonstrated by van Bergen et al. (2018) and replicated by Erbeli et al. (2019): How much children chose to read for themselves (i.e., print exposure) was causally influenced by their reading skills, and one third to one half of the individual differences in print exposure were due to genetic differences. These genetic influences acted partly through the causal path from reading skills and partly through unmeasured traits. In general, a heritable inclination to select environmental niches is called genetic niche picking, or active gene–environment correlation, and contributes to the observed association between environmental exposure and skill level.

Study Characteristics and Limitations

Our findings should be interpreted in the light of the characteristics of this study. First, tracing developmental changes over a 10-year period allowed us to establish time precedence. This is an important first step in finding support for causality, but it is not a sufficient condition. Our design was especially strong during grades 2 and 3, when all three constructs were measured at each timepoint. In grade 1, students could not yet read with sufficient comprehension to warrant testing reading comprehension. In secondary school, the constructs were measured at different timepoints, hindering comparison of the direction of effects. Our data set would have been even stronger had we had data on all of the constructs in every grade.

Second, the current sample was not randomly selected from the population, which potentially poses a threat to the generalizability of the findings. However, the fitted models did not differ statistically significantly between the high– and low–family risk groups, indicating that the pattern of effect was similar. Furthermore, the full sample's reading skills did not differ from that of their approximately 1,500 classmates. Although we could not test whether our sample differed from the general population in print exposure and its relation with reading skills, it is reassuring that the sample's reading levels were representative.

Third, print exposure is commonly measured using either questionnaires (as in our study) or checklists, such as the title recognition test, in which a participant checks the books that he or she is familiar with in a list of titles that includes foils (Cunningham & Stanovich, 1990). Neither of these measurements would be as valid as directly observing children's reading behavior 24-7, but this is not feasible. Checklists circumvent bias due to social desirability. However, they assess familiarity with books, thus undesirably also tapping memory and linguistic skills and only measuring familiarity with the books included in the list (typically best sellers), whereas a specific child may only read, say, horse books (Torppa et al., 2019). Questionnaires directly ask about reading volume and time but might be biased by social desirability. We were not interested in absolute levels of print exposure (which parents might exaggerate) but in associations between print exposure and reading skills. As a result, our use of parental questionnaires would render our results less reliable only if the degree to which parents over- or underreport their children's print exposure depends on the children's reading skills.

Mol and Bus (2011) meta-analyzed the correlation between reading ability and print exposure, including only studies with checklist measures. They found a Fisher z for reading comprehension of 0.38 and for word reading of 0.40. Our correlations are slightly lower (concurrently: .23–.36; see Table 2) but within the confidence intervals of Mol and Bus's estimates. Another reassuring finding is that the direction-of-causation study of Erbeli et al. (2019) replicated that of van Bergen et al. (2018). That is, both found skills → print exposure, with Erbeli et al. using checklists and van Bergen et al. using parent reports. Also, an advantage of a longitudinal study of this scope is that by using parent reports, we could use the same instrument to measure print exposure from prereading to adolescence, which would not have been possible with checklists. Having said that, we encourage the field to continue research on identifying trustworthy measures of print exposure and to replicate our findings using other indicators of children's amount of voluntary reading.

Finally, we note that reading fluency is very stable over time (see Figure 4 and Verhoeven & van Leeuwe, 2009), more so than reading comprehension and amount (see also Betjemann et al., 2008; Harlaar et al., 2011). The stability of all constructs would have been higher had we used multiple measures, especially when used as indicators of latent variables. However, using latent variables, or adding another interesting trait such as language skills, would have increased the complexity of the model and hence decreased the subjects-to-measures or subjects-to-parameters ratio below recommended ratios (Kline, 2005). In particular, reading comprehension skills in grades 2 and 3 would have been better tapped by the use of multiple texts. Nevertheless, the medium grade 2–grade 3 stability of .43 was mirrored in an independent and large Finnish sample (Torppa et al., 2019; $r = .48$). Multiple texts were used in grade 9, when students could handle a long test duration. Reading fluency, in contrast, was very stable, which renders it difficult for other variables to contribute to explaining developmental changes. This is also seen in our path models: In the light of being conservative by correcting for autoregressors, the cross-lagged paths that reached statistical significance were small yet meaningful.

Given that the traits in our study differed in stability, one might argue that the CLPMs that we employed are less suitable than recently developed random intercept models, in which within- and between-person effects are separated (Berry & Willoughby, 2017; Hamaker et al., 2015). Applying these models to subsets of the variables (see Figures 5 and 6) largely yielded converging findings. However, we were not able to include all of the variables in these models, as one needs all constructs measured at each of at least three occasions. Also, for the interpretation of the between-person factor, measurement invariance is preferred. This is a question not only of identical items but also of qualitatively assessing the same domain across time. For any skill assessed over an extended period in child development, this is questionable. In our study, reading comprehension is particularly problematic from this point of view, as it leans more heavily on decoding skills in the early stages of reading acquisition and more on linguistic comprehension in skilled readers (Florit & Cain, 2011; Torppa et al., 2016).

Conclusion

We found that early reading fluency predicts not only later fluency but also later comprehension and amount of reading. Therefore, it seems that children with good basic reading skills may enter in an upward spiral, whereas children with deficient basic reading skills may enter into a downward spiral. This upward or downward spiral will magnify differential exposure to print between struggling and proficient readers. To counteract this Matthew effect (Stanovich, 1986), it may be that early reading intervention for poor readers pays off, in terms of not only improving skill level but also developing positive reading habits. Morgan et al. (2008) tested this idea and boosted the reading skills of 15 poor readers in first grade. However, the study size and/or intervention effect was not large enough to demonstrate a transfer effect on reading practices. A more powerful (i.e., longer and larger) study of this kind is required to test causality. Apart from the effects of reading skills on practice, we found that how much children read in middle childhood affects reading skills later on. It seems that the habit of engaging in regular reading activities stems from being interested in reading. Hence, it is important that caregivers and teachers offer various types of reading materials that fit the child's interest. Nevertheless, practice does not always make perfect. The absence of feedback (as in independent reading) limits the positive effect of practice (Hattie & Timperley, 2007; Reitsma, 1988). This aligns with findings from a large randomized controlled trial in which the intervention group received books matched with their level and interests over the summer vacation. The intervention indeed increased children's print exposure, but this did not translate into improved reading skills (Kim, 2007).

Returning to the overarching question of the practice–skill relation, a large meta-analysis calculated that accumulated deliberate practice accounts for approximately 20% of variance in games, music, and sports; only 4% in education; and <1% in professions (Macnamara et al., 2014). This was based on simply squaring the correlation between skill level and retrospective estimations of amount of practice. In our data, this would yield an effect of ($\sim 0.25^2 =$) approximately 6% of reading practice on reading skills, in line with the meta-analysis for education. However, again, the direction of effect is assumed rather than tested. Long-term prospective studies, like the current one, can shed light on the direction of effects between practice and performance as they develop over time. Such studies can establish causal precedence (a cause must precede the consequence), an important aspect of causal inference. Causal inference relies in the end on support from a variety of empirical research designs that, together with strong theory, build an argument in favor of a causal relation (Hulme & Snowling, 2009; Selig & Little, 2012).

To conclude, we showed in a prospective study that followed children from age 5 to 15 that children's reading exposure and skills keep reciprocally affecting each other throughout development, with slightly stronger effects from skills to practice than vice versa in the early grades. Mastering effortless skilled reading early on seems to foster the development of comprehension skills and to initiate a lifelong habit of reading; subsequently, the accumulation of many hours of reading may make children better readers.

NOTES

Van Bergen's work was supported by a Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) Rubicon Fellowship (grant 446-12-005), a NWO Veni Fellowship (grant 451-15-017), and a NWO Gravitation–funded Consortium on Individual Development grant (024.001.003). Vasalampi's work was supported by Academy of Finland grants (299506 and 323773). Torppa's work was supported by an Academy of Finland Research Fellowship (grant 2762392). We offer thanks to all of the parents and children in the JLD study.

REFERENCES

Aarnoutse, C., & van Leeuwe, J. (1998). Relation between reading comprehension, vocabulary, reading pleasure, and reading frequency. *Educational Research and Evaluation*, *4*(2), 143–166. https://doi.org/10.1076/edre.4.2.143.6960

Anderson, R.C., Wilson, P.T., & Fielding, L.G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, *23*(3), 285–303. https://doi.org/10.1598/RRQ.23.3.2

Becker, M., McElvany, N., & Kortenbruck, M. (2010). Intrinsic and extrinsic reading motivation as predictors of reading literacy: A longitudinal study. *Journal of Educational Psychology*, *102*(4), 773–785. https://doi.org/10.1037/a0020084

Berry, D., & Willoughby, M.T. (2017). On the practical interpretability of cross-lagged panel models: Rethinking a developmental workhorse. *Child Development*, *88*(4), 1186–1206. https://doi.org/10.1111/cdev.12660

Betjemann, R.S., Willcutt, E.G., Olson, R.K., Keenan, J.M., DeFries, J.C., & Wadsworth, S.J. (2008). Word reading and reading comprehension: Stability, overlap and independence. *Reading and Writing*, *21*(5), 539–558. https://doi.org/10.1007/s11145-007-9076-8

Boada, R., Willcutt, E.G., & Pennington, B.F. (2012). Understanding the comorbidity between dyslexia and attention-deficit/hyperactivity disorder. *Topics in Language Disorders*, *32*(3), 264–284. https://doi.org/10.1097/TLD.0b013e31826203ac

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers: Reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities*, *44*(5), 431–443. https://doi.org/10.1177/0022219411410042

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31–42. https://doi.org/10.1037/0022-0663.96.1.31

Catts, H.W., & Weismer, S.E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278–293. https://doi.org/10.1044/1092-4388(2006/023)

Cunningham, A.E., & Stanovich, K.E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology*, *82*(4), 733–740. https://doi.org/10.1037/0022-0663.82.4.733

Cunningham, A.E., & Stanovich, K.E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945. https://doi.org/10.1037/0012-1649.33.6.934

Curran, P.J., Howard, A.L., Bainter, S.A., Lane, S.T., & McGinley, J.S. (2014). The separation of between-person and within-person components of individual change over time: A latent curve model with structured residuals. *Journal of Consulting and Clinical Psychology*, *82*(5), 879–894. https://doi.org/10.1037/a0035297

de Zeeuw, E.L., de Geus, E.J.C., & Boomsma, D.I. (2015). Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends in Neuroscience and Education*, *4*(3), 69–76. https://doi.org/10.1016/j.tine.2015.06.001

Eklund, K., Torppa, M., Sulkunen, S., Niemi, P., & Ahonen, T. (2018). Early cognitive predictors of PISA reading in children with and without family risk for dyslexia. *Learning and Individual Differences*, *64*, 94–103. https://doi.org/10.1016/j.lindif.2018.04.012

Erbeli, F., van Bergen, E., & Hart, S.A. (2019). Unraveling the relation between reading comprehension and print exposure. *Child Development*. Advance online publication. https://doi.org/10.1111/cdev.13339

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*(4), 553–576. https://doi.org/10.1007/s10648-011-9175-6

Fuchs, L.S., Fuchs, D., Hosp, M.K., & Jenkins, J.R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 239–256. https://doi.org/10.1207/S1532799XSSR0503_3

Guthrie, J.T., Wigfield, A., Metsala, J.L., & Cox, K.E. (1999). Motivational and cognitive predictors of text comprehension and reading amount. *Scientific Studies of Reading*, *3*(3), 231–256. https://doi.org/10.1207/s1532799xssr0303_3

Hamaker, E.L., Kuiper, R.M., & Grasman, R.P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

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*(6), 2123–2137. https://doi.org/10.1111/j.1467-8624.2011.01658.x

Harlaar, N., Trzaskowski, M., Dale, P.S., & Plomin, R. (2014). Word reading fluency: Role of genome-wide single-nucleotide polymorphisms in developmental stability and correlations with print exposure. *Child Development*, *85*(3), 1190–1205. https://doi.org/10.1111/cdev.12207

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Häyrinen, T., Serenius-Sirve, S., & Korkman, M. (1999). *Lukilasse*. Helsinki, Finland: Psykologien.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hulme, C., Nash, H.M., Gooch, D., Lervåg, A., & Snowling, M.J. (2015). The foundations of literacy development in children at familial risk of dyslexia. *Psychological Science*, *26*(12), 1877–1886. https://doi.org/10.1177/0956797615603702

Hulme, C., & Snowling, M.J. (2009). *Developmental disorders of language learning and cognition*. Chichester, UK: Wiley-Blackwell.

Hulme, C., & Snowling, M.J. (2011). Children's reading comprehension difficulties: Nature, causes, and treatments. *Current Directions in Psychological Science*, *20*(3), 139–142. https://doi.org/10.1177/0963721411408673

Kim, J.S. (2007). The effects of a voluntary summer reading intervention on reading activities and reading achievement. *Journal of Educational Psychology*, *99*(3), 505–515. https://doi.org/10.1037/0022-0663.99.3.505

Kline, R.B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.

Kovas, Y., Garon-Carrier, G., Boivin, M., Petrill, S.A., Plomin, R., Malykh, S.B., … Vitaro, F. (2015). Why children differ in motivation to learn: Insights from over 13,000 twins from 6 countries. *Personality and Individual Differences*, *80*, 51–63. https://doi.org/10.1016/j.paid.2015.02.006

Kuhn, M.R., Schwanenflugel, P.J., & Meisinger, E.B. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, *45*(2), 230–251. https://doi.org/10.1598/RRQ.45.2.4

Leinonen, S., Müller, K., Leppänen, P.H.T., Aro, M., Ahonen, T., & Lyytinen, H. (2001). Heterogeneity in adult dyslexic readers: Relating processing skills to the speed and accuracy of oral text reading. *Reading and Writing*, *14*(3/4), 265–296. https://doi.org/10.1023/A:1011117620895

Leppänen, U., Aunola, K., & Nurmi, J.-E. (2005). Beginning readers' reading performance and reading habits. *Journal of Research in Reading*, *28*(4), 383–399. https://doi.org/10.1111/j.1467-9817.2005.00281.x

Lindeman, J. (2000). *Ala-asteen lukutesti: Käyttäjän käsikirja* [The comprehensive school reading test: User's manual]. Turku, Finland: Oppimistutkimuksen keskus [Center for Learning Research], University of Turku.

Lyytinen, H., Erskine, J., Hämäläinen, J., Torppa, M., & Ronimus, M. (2015). Dyslexia—early identification and prevention: Highlights from the Jyväskylä Longitudinal Study of Dyslexia. *Current Developmental Disorders Reports*, *2*(4), 330–338. https://doi.org/10.1007/s40474-015-0067-1

Macnamara, B.N., Hambrick, D.Z., & Oswald, F.L. (2014). Deliberate practice and performance in music, games, sports, education, and professions: A meta-analysis. *Psychological Science*, *25*(8), 1608–1618. https://doi.org/10.1177/0956797614535810

Macnamara, B.N., Moreau, D., & Hambrick, D.Z. (2016). The relationship between deliberate practice and performance in sports: A meta-analysis. *Perspectives on Psychological Science*, *11*(3), 333–350. https://doi.org/10.1177/1745691616635591

McGeown, S.P., Duncan, L.G., Griffiths, Y.M., & Stothard, S.E. (2015). Exploring the relationship between adolescent's reading skills, reading motivation and reading habits. *Reading and Writing*, *28*(4), 545–569. https://doi.org/10.1007/s11145-014-9537-9

McGeown, S.P., Osborne, C., Warhurst, A., Norgate, R., & Duncan, L.G. (2016). Understanding children's reading activities: Reading motivation, skill and child characteristics as predictors. *Journal of Research in Reading*, *39*(1), 109–125. https://doi.org/10.1111/1467-9817.12060

Miller, A.C., Keenan, J.M., Betjemann, R.S., Willcutt, E.G., Pennington, B.F., & Olson, R.K. (2013). Reading comprehension in children with ADHD: Cognitive underpinnings of the centrality deficit. *Journal of Abnormal Child Psychology*, *41*(3), 473–483. https://doi.org/10.1007/s10802-012-9686-8

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*(2), 267–296. https://doi.org/10.1037/a0021890

Morgan, P.L., Fuchs, D., Compton, D.L., Cordray, D.S., & Fuchs, L.S. (2008). Does early reading failure decrease children's reading motivation? *Journal of Learning Disabilities*, *41*(5), 387–404. https://doi.org/10.1177/0022219408321112

Mosing, M.A., Madison, G., Pedersen, N.L., Kuja-Halkola, R., & Ullen, F. (2014). Practice does not make perfect: No causal effect of music practice on music ability. *Psychological Science*, *25*(9), 1795–1803. https://doi.org/10.1177/0956797614541990

Muthén, L.K., & Muthén, B.O. (2012). *Mplus statistical analysis with latent variables: User's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.

Nagy, W.E., Anderson, R.C., & Herman, P.A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, *24*(2), 237–270. https://doi.org/10.3102/00028312024002237

Nation, K. (2017). Nurturing a lexical legacy: Reading experience is critical for the development of word reading skill. *NPJ Science of Learning*, *2*, article 3. https://doi.org/10.1038/s41539-017-0004-7

Oakhill, J.V., Berenhaus, M.S., & Cain, K. (2015). Children's reading comprehension and comprehension difficulties. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 344–360). New York, NY: Oxford University Press.

OECD. (2010). *PISA 2009 results: Learning trends: Changes in student performance since 2000* (Vol. 5). Paris, France: Author.

OECD. (2013). *PISA 2012 results: What students know and can do: Student performance in mathematics, reading and science* (Vol. 1). Paris, France: Author.

Olson, R.K., Keenan, J.M., Byrne, B., & Samuelsson, S. (2014). Why do children differ in their development of reading and related skills? *Scientific Studies of Reading*, *18*(1), 38–54. https://doi.org/10.1080/10888438.2013.800521

Ozernov-Palchik, O., & Gaab, N. (2016). Tackling the 'dyslexia paradox': Reading brain and behavior for early markers of developmental dyslexia. *WIREs Cognitive Science*, *7*(2), 156–176. https://doi.org/10.1002/wcs.1383

Perfetti, C.A. (1985). *Reading ability*. New York, NY: Oxford University Press.

Pfost, M., Dörfler, T., & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, *26*, 89–102. https://doi.org/10.1016/j.lindif.2013.04.008

Reitsma, P. (1988). Reading practice for beginners: Effects of guided reading, reading-while-listening, and independent reading with computer-based speech feedback. *Reading Research Quarterly*, *23*(2), 219–235. https://doi.org/10.2307/747803
Schiefele, U., Schaffner, E., Möller, J., & Wigfield, A. (2012). Dimensions of reading motivation and their relation to reading behavior and competence. *Reading Research Quarterly*, *47*(4), 427–463. https://doi.org/10.1002/RRQ.030
Selig, J.P., & Little, T.D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T.D. Little, & N.A. Card (Eds.), *Handbook of developmental research methods* (pp. 265–278). New York, NY: Guilford.
Share, D.L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 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*(4), 360–407. https://doi.org/10.1598/RRQ.21.4.1
Torppa, M., Georgiou, G.K., Lerkkanen, M.-K., Niemi, P., Poikkeus, A.-M., & Nurmi, J.-E. (2016). Examining the simple view of reading in a transparent orthography: A longitudinal study from kindergarten to grade 3. *Merrill-Palmer Quarterly*, *62*(2), 179–206. https://doi.org/10.13110/merrpalmquar1982.62.2.0179
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*. Advance online publication. https://doi.org/10.1111/cdev.13241
Torppa, M., Poikkeus, A.-M., Laakso, M.-L., Tolvanen, A., Leskinen, E., Leppänen, P.H.T., … Lyytinen, H. (2007). Modeling the early paths of phonological awareness and factors supporting its development in children with and without familial risk of dyslexia. *Scientific Studies of Reading*, *11*(2), 73–103. https://doi.org/10.1080/10888430709336554
Torppa, M., Tolvanen, A., Poikkeus, A.-M., Eklund, K., Lerkkanen, M.-K., Leskinen, E., & Lyytinen, H. (2007). Reading development subtypes and their early characteristics. *Annals of Dyslexia*, *57*(1), 3–32. https://doi.org/10.1007/s11881-007-0003-0
Vaessen, A., & Blomert, L. (2010). Long-term cognitive dynamics of fluent reading development. *Journal of Experimental Child Psychology*, *105*(3), 213–231. https://doi.org/10.1016/j.jecp.2009.11.005
van Bergen, E., de Jong, P.F., Maassen, B., Krikhaar, E., Plakas, A., & van der Leij, A. (2014). IQ of four-year-olds who go on to develop dyslexia. *Journal of Learning Disabilities*, *47*(5), 475–484. https://doi.org/10.1177/0022219413479673
van Bergen, E., Snowling, M.J., de Zeeuw, E.L., van Beijsterveldt, C.E.M., 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*(11), 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 in Human Neuroscience*, *8*, article 346. https://doi.org/10.3389/fnhum.2014.00346
Verhoeven, L., & van Leeuwe, J. (2009). Modeling the growth of word-decoding skills: Evidence from Dutch. *Scientific Studies of Reading*, *13*(3), 205–223. https://doi.org/10.1080/10888430902851356
Wechsler, D. (1991). *WISC-III: Wechsler Intelligence Scale for Children: Manual* (3rd ed.). San Antonio, TX: Psychological.

Submitted March 18, 2019
Final revision received February 3, 2020
Accepted February 4, 2020

ELSJE VAN BERGEN (corresponding author) is an assistant professor in the Department of Biological Psychology at Vrije Universiteit Amsterdam, The Netherlands; email e.van.bergen@vu.nl; Twitter @drElsje. When she started working on this article, she was a postdoctoral research fellow in the Department of Experimental Psychology at the University of Oxford, UK. Her research focuses on how and why children differ in academic achievement, particularly reading achievement, and she uses multidisciplinary methodologies from psychology, education, and genetics.

KATI VASALAMPI is an Academy Research Fellow in the Department of Psychology at the University of Jyväskylä, Finland; email kati.vasalampi@jyu.fi. Her research focuses on the role of learning motivation and well-being in different interpersonal environments and their associations with learning outcomes.

MINNA TORPPA is an associate professor in the Department of Teacher Education at the University of Jyväskylä, Finland; email minna.p.torppa@jyu.fi; Twitter @minnatorppa. Her research focuses on the development of reading, spelling, and math and on the identification of supportive factors in the environment that can mitigate the cascading negative associations among learning difficulties, motivation, and well-being.

APPENDIX A

Descriptions of the Additional Reading Tests

Oral Word-Reading Fluency Task (Grades 2, 3, and 8)

In the Lukilasse nationally standardized reading test (Häyrinen, Serenius-Sirve, & Korkman, 1999), participants had two minutes (grades 2 and 3) or one minute (grade 8) to read aloud as many words as possible from a 90-item (grade 2) or 105-item (grades 3 and 8) word list. The score for (oral) word-reading fluency was the number of correctly read words within the allotted time. The inter-rater reliability was .99. Second, the longitudinal sample and their classmates were assessed on group-administered reading ability tests to investigate representativeness of the longitudinal sample (see the Method section). All group-administered tasks were paper-and-pencil tasks and assessed silent reading. In addition, teachers assessed children's reading skills.

Silent Word-Reading Fluency Task (Grades 1, 2, and 3)

This is a subtest of the nationally normed reading test battery (Ala-Asteen Lukutesti; Lindeman, 2000). The 80 items consisted of a picture with four phonologically similar words. The child's task is to silently read the four words and draw a line connecting the picture with the matching word. The score was the number of correct answers within the allotted time (five minutes in grade 1 and two minutes in grades 2 and 3). Lindeman (2000) reported the Kuder–Richardson reliability coefficients as .97 in grade 1 and .82 in grade 2.

Silent Word Chain Task (Grades 2 and 9)

This is a timed test with 10 rows of word chains. Each row has four to six words that are joined together. The child has to separate the words with pencil strokes. The score was the number of correct responses (maximum = 40) within the time limit (1.5 minutes).

Silent Sentence-Reading Fluency Task (Grade 9)

Sentence-reading fluency was measured with a sentence verification task. Students were given a list of short statements and were asked to circle "correct" or "incorrect." The statements were short, and verification required minimal comprehension (e.g., "A cat is an animal"). The score was the number of correct answers within three minutes.

Teacher's Evaluation of Reading Skills (Grades 1, 2, and 3)

Teachers were asked to evaluate students' reading skills (with respect to students of similar age) on a 5-point Likert-type scale ranging from 1 (*very poor progress in reading*) to 5 (*excellent progress in reading*).

APPENDIX B

Comparison of the Follow-Up Sample and Their Classmates on the Group-Administered Reading Tests

	Follow-up sample			Classmates			
Timepoint and measure	*N*	*M*	*SD*	*N*	*M*	*SD*	$t(df)$
Grade 1							
Word-reading fluency	191	43.35	19.70	1,361	43.93	17.35	$t(233.20) = 0.39, p = .70$
Teacher's evaluation	168	3.54	1.21	1,324	3.65	1.14	$t(1{,}490) = 1.19, p = .23$
Grade 2							
Word-reading fluency	183	28.56	9.58	1,370	29.89	8.65	$t(223.45) = 1.93, p = .08$
Word chains	180	13.33	6.83	1,358	14.03	7.13	$t(1{,}536) = 1.23, p = .22$
Teacher's evaluation	150	3.59	1.20	1,179	3.62	1.09	$t(181.70) = 0.30, p = .76$
Grade 3							
Word-reading fluency	190	35.83	10.58	2,566	35.69	9.316	$t(211.26) = 0.17, p = .87$
Teacher's evaluation	169	3.57	1.23	2,450	3.65	1.123	$t(2{,}617) = 0.85, p = .40$
Grade 9							
Word chains	165	64.74	18.90	1,539	64.54	17.24	$t(1{,}702) = 0.14, p = .89$
Sentence-reading fluency	156	35.88	9.45	1,508	35.37	8.43	$t(1{,}662) = 0.70, p = .48$

Note. *df* = degrees of freedom; *M* = mean; *SD* = standard deviation.

APPENDIX C

Descriptive Statistics and Comparisons Between the High— and Low—Family Risk Groups

	High family risk			Low family risk				
Timepoint and measure	*N*	*M*	*SD*	*N*	*M*	*SD*	*F*	Effect size (*d*)
Prereading skills								
Age 5	102	−0.51	0.99	92	0.01	0.78	$F(1, 192) = 15.99$, $p < .001$	0.58
Reading fluency (words read correctly per minute)								
Grade 1	99	29.93	20.46	90	41.53	23.43	$F(1, 187) = 13.22$, $p < .001$	0.53
Grade 2	107	54.69	24.95	89	69.19	24.96	$F(1, 194) = 16.40$, $p < .001$	0.58
Grade 3	101	67.65	25.88	91	79.57	24.15	$F(1, 190) = 10.83$, $p < .001$	0.48
Grade 8	101	80.19	18.33	81	90.35	14.64	$F(1, 180) = 16.45$, $p < .001$	0.61
Reading comprehension								
Grade 2	97	8.67	3.11	73	9.29	2.20	$F(1, 168) = 2.09$, $p = .150$	0.23
Grade 3	101	9.69	1.77	78	10.13	1.59	$F(1, 177) = 2.90$, $p = .090$	0.26
Grade 9	88	21.73	6.82	71	23.76	6.76	F (1,157) = 3.47, p=.067	0.30
Print exposure								
Age 5	107	3.41	0.74	93	3.20	0.68	$F(1, 198) = 4.38$, $p = .038$	−0.30
Grade 1	91	3.04	0.64	90	3.08	0.66	$F(1, 179) = 0.16$, $p = .693$	0.06
Grade 2	93	3.34	0.64	87	3.31	0.67	$F(1, 178) = 0.12$, $p = .730$	−0.05
Grade 3	91	3.26	0.74	84	3.37	0.70	$F(1, 173) = 1.03$, $p = .311$	0.15
Grade 7	81	2.77	0.80	67	2.57	0.73	$F(1, 146) = 2.48$, $p = .143$	−0.26

Note. *M* = mean; *SD* = standard deviation. Effect sizes were estimated with Cohen's *d* using pooled standard deviations.