What works for whom? Differential genetic effects of early literacy interventions in kindergarten

Plak, R.D.

Author: Plak, R.D.
Title: What works for whom? Differential genetic effects of early literacy interventions in kindergarten
Issue Date: 2016-12-15

Chapter 2

Genetic Differential Susceptibility in Literacy Delayed Children: A Randomized Controlled Trial on Emergent Literacy in Kindergarten

Published as:

Plak, R.D., Kegel, C.A.T., & Bus, A. G. (2015). Genetic differential susceptibility in literacy delayed children: A randomized controlled trial on emergent literacy in kindergarten. *Development and Psychopathology*, 27, 69–79. doi:10.1017/S0954579414001308

ABSTRACT

In this randomized controlled trial 528 five-years-old kindergarten children participated, of whom 290 were delayed in literacy skills as they belonged to the lowest 40 percent of a national standard literacy test. We tested the hypothesis that some children are more susceptible to school-entry educational interventions than their peers due to their genetic make-up, and thus whether DRD4 moderated intervention effects. Children were randomly assigned to a control condition or one of two interventions involving computer programs tailored to the literacy needs of delayed pupils: *Living Letters* for alphabetic knowledge, and *Living Books* for text comprehension. Effects of *Living Books* met the criteria of differential susceptibility. For carriers of the DRD4 7-repeat allele (about one-third of the delayed group), the *Living Books* program was an important addition to the common core curriculum in kindergarten (effect size d = .43), whereas the program did not affect the other children (d = -.13). Findings for *Living Letters* did not fulfill the statistical criteria for differential susceptibility. Implications of differential susceptibility for education and regarding the crucial question 'what works for whom?' are discussed.

INTRODUCTION

Children starting school with limited emergent literacy skills are at risk for encountering difficulties in reading throughout school and being classified as (pseudo) dyslectic in later years (Stanovich, 1986). Intervention programs to ensure timely development of key reading precursors for all at-risk children are currently the gold standard (Snow, Burns, & Griffin, 1998), yet compensatory educational programs that aim to improve school-entry literacy skills seem to have only modest effects on children's development (see for instance NELP, 2008). Despite half a century of research into preemptive measures in kindergarten, few attempts have been made to understand the moderate efficacy of programs promoting school-entry literacy skills.

Educational programs may affect some children's literacy substantially but evaluations focused on average or across-the-board effects may underestimate the impact of programs on such children. For instance, the overall effect of an extensive, nation-wide intervention stimulating parent-child verbal interaction in the first year after birth on language development at 15 months was small (d = .05), but the effect was moderately high (d = .46) in a sub-sample of temperamentally highly reactive children (Van den Berg & Bus, 2014). A reactive temperament proved a serious risk factor for language development but an asset when parents increased verbal parent-child interaction as stimulated by the intervention.

In our research program *What Works for Whom* we seek to shed light on the hidden efficacy of kindergarten programs to enhance early literacy. Thus far the dominating theory has been that kindergarten children with risk factors such as poor regulatory skills are less able to benefit from their less than optimal "natural" environment at home and in school (Justice, Chow, Capellini, Flanigan, & Colton, 2003). In accordance with the differential susceptibility model we expect that specific sub-groups of children, defined by their genetic make-up, may be more susceptible than their peers to the environment (Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007; Belsky & Pluess, 2009, 2013; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2011); Van IJzendoorn et al., 2011). Although they lag behind without additional support, they outperform their peers when they receive optimal instruction (Kegel, Bus, & Van IJzendoorn, 2011). In the intervention experiment reported herein, we test whether young children whom we assume to be more susceptible to the environment because of their genetic make-up respond better than their putatively less susceptible peers to early interventions promoting important precursors of literacy.

In line with a series of genetic differential susceptibility studies by Bakermans-Kranenburg and Van IJzendoorn (2006, 2011, in press), the focus in the current inquiry is on a dopamine-related genetic polymorphism as moderator of intervention effects. Bakermans-Kranenburg and Van IJzendoorn (2006), for example, found that maternal sensitivity observed when children were 10 months of age predicted externalizing problems more than two years later, but only for carriers of the 7-repeat dopamine receptor D4 (DRD4) allele. The dopamine receptor D4 may also be a relevant moderator of effects of educational programs as it is associated with attention and motivation (Hsiung, Kaplan, Petryshen, Lu, & Field, 2004; Tripp & Wickens, 2008). Here we present results from a randomized controlled trial examining genetic moderation of the effects of two early literacy interventions in children with delayed literacy development.

Differential susceptibility

In developmental psychopathology differential susceptibility studies are a major challenge to the traditional diathesis-stress model (Belsky et al., 2007; Belsky & Pluess, 2009, 2013; Ellis et al., 2011). Children susceptible to adversity not only catch up and achieve at a level similar to other children when a program compensates for their vulnerabilities, but they actually outperform peers lacking the putative 'vulnerable' constitution under optimized learning conditions. In general, evidence is accumulating that specific neurobiological markers of high reactivity to the environment, whether measured at the emotional, behavioral, or biological level, affect how children respond to negative *and* positive environments (Belsky et al., 2007; Belsky & Pluess, 2009, 2013; Ellis et al., 2011; Van IJzendoorn et al., 2011). Most of the evidence, however, originates from developmental and psychopathological studies. A first test of differential susceptibility in the educational domain was an investigation with *Living Letters* – a computer program to promote basic alphabetic knowledge. Narrowing gaps in phonological skills at an early stage is important considering that the risk of word-level decoding difficulties in reading is often carried by phonological deficits (Goswami & Bryant, 1990; Hulme, Bowyer-Crane, Carroll, Duff, & Snowling, 2012).

In a group of four-year-olds who did not yet understand the alphabetic principle (i.e., that letters relate to sounds in spoken words), we tested whether the dopamine D4 receptor gene (DRD4) moderated children's susceptibility for input from the learning environment. The experiment provided more support for differential susceptibility than for diathesis-stress. With *Living Letters* the group expected to be most susceptible to the environment – carriers of the long variant of DRD4 – performed at the lowest level without intervention and highest with intervention, thereby demonstrating their high reactivity to input from the environment. Especially notable was that the *Living Letters* group scored almost one standard deviation higher than a control group similar in genotypic susceptibility (Kegel et al., 2011). The group that was considered less susceptible to environmental support also benefited from *Living Letters*, but the effect size was only modest. These findings corroborate the theory that the dopamine genotype indeed functions as a susceptibility marker in the domain of early literacy acquisition.

Neurobiological markers in the cognitive domain

There are some good reasons for including the DRD4 genotype in experiments with early literacy interventions. Transmission of electric signals, especially in the prefrontal lobe monitoring impulses from the limbic system, may be less efficient in carriers of the long variant of the DRD4 genotype and, consequently, children may be easily distracted by irrelevant elements in the learning environment, with poor achievement as a result (Robbins & Everitt, 1999). Direct support for this hypothesis comes from a longitudinal study in which we assessed, apart from the dopamine-related genotype, executive attention when children were four years of age along with their alphabetic skills after three months in kindergarten and in first grade (Kegel & Bus, 2013). Carriers of the long variant of the DRD4 polymorphism gene benefited less from reading instruction in kindergarten and first grade than their peers. Moreover, executive attention measured using Stroop-like tasks, digit span forward, and digit span backward, fully mediated the link between the dopamine DRD4 gene and alphabetic skills. DRD4 was a significant predictor of alphabetic skills at four months in first grade ($\beta = .47$), but not after entering executive attention in the regression model ($\beta = .16$). These findings clearly suggest that carriers of the risk genotype demonstrate lower levels of executive attention than their peers and may, as a result, have benefited less from instruction in kindergarten and first grade.

But how can it be explained that carriers of the DRD4 gene with 7 repeats (7+) proves to have great learning potential as outcomes of the *Living Letters* experiment indicate (Kegel et al., 2011)? There is evidence that the performance feedback to children's responses might have been an important promotive mechanism in *Living Letters* for the highly reactive children (Howe, Beach, & Brody, 2010). When a program includes elements that mobilize children's attention for solving the tasks by providing intensive, closely monitored, and individualized scaffolding it may, especially in the case of highly susceptible children, stimulate high reactivity to the problems to be solved, thereby turning the putative "risk" group into the most successful group who actually benefit more than -and thus outperform- their peers (Belsky et al., 2007; Obradović, Bush, Stamperdahl, Adler, & Boyce, 2010).

Indispensable elements of an optimal early literacy intervention

As a direct test of this hypothesis, two versions of *Living Letters*, the complete and an abbreviated version, were contrasted with each other and with a control group in a randomized controlled trial. In both *Living Letters* versions, instruction and assignments were exactly the same, but in the cut-down version there was no computer tutor - that is, an animated character that comments on the child's responses to the tasks - who provides intensive, closely monitored and individualized scaffolding. For instance, finding the first letter of the name among four other letters or selecting the picture that starts with the same sound as the child's proper name were included in the complete version followed by feedback from a tutor when children made errors. In the abbreviated version, however, children did not receive feedback. With the help of technology these small variations in a program (i.e. the presence of a computer tutor providing feedback vs. no tutor) can be implemented with high fidelity.

The experiment demonstrated that the computer tutor makes the difference between underachievement and high achievement in carriers of the susceptibility genotype. That is, DRD4-long-allele carrying four-year-olds benefited most from *Living Letters* when the computer tutor continuously corrected and confirmed children's responses (Kegel & Bus, 2012). Apparently, not the assignments and instructions in the program but continuous performance feedback canalizes the learning capacities of these children in particular (Kegel et al., 2011). The computer tutor enables them to make optimal use of their cognitive abilities while carrying out computer assignments. High reactivity to an often over-stimulating learning environment leads to distraction and inefficient use of learning opportunities, whereas this same reactivity may at the same time make children highly responsive to a program that continuously stimulates, structures, and regulates their learning behavior by providing positive performance feedback. The program may thus improve children's latent potential to solve tasks and to acquire new skills.

Current study

By failing to consider the differential susceptibility of children, educators and policy makers may easily overlook the potential impact of literacy intervention programs (e.g., Van den Berg & Bus, 2014).Thus, in the current study we tested whether an average effect across all participants may mask the effectiveness of early literacy intervention programs. When rather modest or absent intervention effects in the total group are juxtaposed with strong effects for a susceptible group of children the efficacy of the program may be (strongly) underestimated. Differential susceptibility theory offers a vital heuristic in designing studies that aim at evaluating educational programs to improve school entry skills of the most susceptible children who are delayed in literacy.

We target a group of five-year-old kindergarten children delayed in literacy skills who score in the lowest 40 percent of a national standard literacy test. We aim, first, at replicating and extending earlier findings for *Living Letters* in an older age group. Second, we test whether genetic differential susceptibility could be found for another computerized intervention - *Living Books* - carried out within the same time frame and

based on the same principles of immediate positive performance feedback. Similar to *Living Letters*, *Living Books* includes a tutor who coaches the learning process by providing feedback but addresses less time-constrained literacy skills than phonemic awareness that is mostly reached within a brief period of rapid growth (Paris, 2005). The children read digital storybooks and during each reading answer questions about story events and difficult words in the text. Story reading is a vital precursor of learning to read in first grade because in storybooks children become familiarized with complex phrasing and sophisticated vocabulary as is common in text.

METHOD

Participants

A total of 90 schools responded to our request to participate in the experiment. In brochures and letters sent to the schools they were offered both a chance to provide extra guidance to pupils with literacy delays and an opportunity to experience how to implement technology-based programs in their teaching. Furthermore, participating schools would receive free access to educational computer programs for kindergarten children during three months after the intervention was completed (www.bereslim.nl). Information about the project was distributed via e-mail, mail, social media, and phone from August 2012 to October 2012. The schools willing to participate were from all parts of the Netherlands.

Eligible children were selected between October 2012 and February 2013 by the kindergarten teachers in the 90 participating schools. Teachers were asked to select six pupils lagging behind in literacy skills per kindergarten classroom. The eligible pupils should, for instance, not yet be able to write their proper name, to rhyme, to name a few letters, and to identify sounds in words. As a guideline the eligible children preferably would score in the lowest 40 percent -between 0 and 59- on a standardized literacy test (i.e., the Central Institute for Test development [Centraal Instituut voor Toetsontwikkeling] (Cito) Literacy Test for Kindergarten Pupils, CLT) administered at most Dutch schools (Lansink & Hemker, 2012). The CLT administered in January 2013 was used to check whether the teachers had correctly selected the literacy-delayed children. Dutch was required as the participants' first language. When a parent refused consent, the teacher was asked to select another pupil from her classroom. In 40 schools the number of participants was somewhat lower than six because too few pupils were eligible for the intervention or too many parents refused consent (M = 3.18 pupils per classroom, SD = 1.74). Eight schools (with 92 children) were not included because these schools did not test their pupils with the standardized CLT test in the kindergarten year preceding the first grade. Due to

incidental missing scores, 42 children were lost.
Teachers complained that parents of children who were most in need of the intervention often refused consent. As a result, only slightly more than half of the 528 selected pupils scored in the lowest CLT 40 percent at pre-test and thus making up the delayed-literacy group (Lansink & Hemker, 2012). In most schools about half of the selected children met this criterion. The other half of the children selected by the teachers scored in the mid- range of the CLT, between 60 and 64 (Lansink & Hemker, 2012). We included these typically developing children in the first round of analyses although our primary focus was on the efficacy of the interventions for the delayed pupils. Only the delayed group (*n* = 290) was included in the statistical tests of genetic differential susceptibility. Table 1 presents numbers per condition and level (children with delayed versus typical literacy). Participants had a mean age of 66.84 months (*SD* = 4.35) at pre-test. The mean score for father's education was 3.97 (*SD* = 1.93) on a scale ranging from 0 - 6, where 0 represents primary school and 6 represents university-level education.

Procedure and design

Parents of eligible children received written information about the study explaining the scientific goals and the opportunity for their child to receive extra coaching. They also received information about genotyping to be part of the research. Moreover, a website was available for additional information about aim and design of the research. Contact information was provided to allow parents to ask additional questions. Parents made frequent use of this opportunity. Genotyping was a main reason for parents to refuse consent for participation (about 25%).
The children were randomly assigned to one of the three conditions: *Living Letters*, *Living Books*, and a control condition consisting of playing *Clever Together.* At least one child in each class was assigned to an intervention condition (*Living Letters* or *Living Books*). Twice a week for 15 minutes per session the participants engaged a computer program on their own. Children in the *Living Books* condition were involved in 16 sessions and in Clever Together and *Living Letters* in a variable number of sessions, averaging 15. The more errors children made the more sessions. About half way through the intervention period, buccal cell samples were collected by trained research team members using a sterile swab specifically designed for collecting buccal cells for DNA analysis (Omni Swabs, Whatman/GE Healthcare, UK). The samples were stored at -20ºC directly after collection. Literacy skills were tested before and after the intervention using the Cito standardized literacy test CLT (Lansink & Hemker, 2012). Children were group-wise examined by their teacher.

Intervention programs

Living Letters promotes understanding of the alphabetic principle, the notion that letters in print relate to sounds in spoken words. The program offers a framework that anchors instruction and practice in a personally motivating context of activities using children's own proper name (Van der Kooy-Hofland, Bus, & Roskos, 2012). This approach is based on a series of studies showing that most children can name the initial letter of the own proper name earlier than other letters (Levin, Both-De Vries, Aram, & Bus, 2005) and that the sound of this letter is the first one that children can identify in spoken words and use correctly in spelling (Both-De Vries & Bus, 2008, 2010). The program adapted automatically to the child's proper name when it was available in the data base; 240 common Dutch names were obtainable. When the name was not available in the data base or irregularly spelled, the word "mama" [mommy] was used in its place, as this is a well-known name (Both-De Vries & Bus, 2010). Dutch is rather regularly spelled and most names can be used to highlight the alphabetic principle that letters in print relate to sounds in spoken words. In a less regularly spelled language like English more names might not be usable to illustrate the alphabetic principle.

In the first 20 games, children practiced how their name (or "mama") is written, followed by 10 games to train the sound of the first letter of the child's name (or "mama") and thereafter by 10 games to identify pictures that start or end with the first letter of the target name. Each session began with animations of two preschoolers (called "Sim" and "Sanne") who announced a new game and demonstrated how to play the games. Feedback provided by Sim's teddy bear followed up on every response of the child. When children produced one or more erratic responses to an assignment, the assignment was repeated one to three times, thus promoting additional practice when children performed poorly. After each additional error children received more clues to solve the assignment. More specifically, after the first error the assignment was only repeated: "Listen carefully, in which word do you hear /t/ of Tom?" After the second error children received a clue: "How does your teacher write your name?" If the child failed to give the correct answer after the third attempt, the solution was demonstrated together with a spoken explanation by the digital tutor. After a maximum of three trials the game ended with a positive note, irrespective of whether a correct response was given, whereupon a new game started. When children failed to give the correct answer, the assignment was repeated twice in subsequent sessions which explains why some children had more sessions than others.

Living Books was made up of eight age-appropriate digital animated storybooks. The animated pictures, sounds, and music support the meaning of the story text and thus enable the child to understand story events and language even when the oral text is difficult for the child (Bus, Takacs, & Kegel, 2015; Kamil, Intrator, & Kim, 2000). Each reading of a book was interrupted four times to ask questions about the story (e.g. “Eventually Little Mouse found a house. Whose house do you think it is?”) and about word knowledge (e.g. “Little Mouse peeked inside. On which picture do you see her peeking?). If the child’s response was incorrect the question was repeated maximally three times and feedback was adapted to the child’s response -- similar to *Living Letters* (see above). The first error was followed up by a repetition of the question, the second by a clue (“Peeking is secretly watching. Where do you see Little Mouse peeking?”), and the third by demonstrating the correct response together with a spoken explanation (“Of course, this house is Little Mouse’s own house!”). Each book was presented twice and in each session four questions were included. During each session the child ‘reads’ one book for 10 minutes. In contrast to the more adaptive program *Living Letters*, assignments were not repeated in the next session when children made errors.

Clever Together, also a computer program, does not target story comprehension or code-related skills. It includes 40 hide and seek games. For example, the child is told that one of the main characters is hidden behind a yellow object. As in *Living Letters* and *Living Books*, a tutor provides constructive, detailed feedback for every error and every correct response (“Good job, you found Sanne behind the yellow tractor.”). The first error is followed up by a repetition of the question (“Where again would Sanne hide?”), and a second error by clues. Assignments were repeated in future sessions when children made errors. *Measures*

Early literacy skills

CITO Literacy Test for Kindergarten Pupils (CLT) is a group administered standardized literacy test for kindergarten pupils, given in January (α = .89) and June (α = .87) of each year. The 60-item CLT concern word knowledge, critical listening, rhyming, hearing the first and last word, sound blending, writing orientation, and prediction of book content based on the book cover (Lansink & Hemker, 2012). Commissie Testaangelegenheden Nederland [Committee for Tests in the Netherlands] evaluates the reliability and validity of the CLT, judging it adequate. According to the CLT manual, pupils with CLT scores lower than 40 percent are considered delayed in their literacy development. The pre-test CLT score was coded as delayed (n = 290) for children scoring in the lowest 40 percent according to national norms (0) or as typical literacy level for children (n = 238) scoring above 40 percent (1). At post-test we used the full range of scores on CTL.

Genetic screening for DRD4 polymorphisms

DNA isolation. Buccal swabs were incubated in lysis buffer (100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8, 0.1 mg/ml proteinase K, and 0.5% w/v SDS) until further processing. Genomic DNA was isolated using the Chemagic buccal swab kit on a Chemagen Module I workstation (Chemagen Biopolymer-Technologie AG, Baesweiler, Germany).

PCR Amplification. The region of interest of the DRD4 gene was amplified by PCR using the following primers: a FAM-labelled primer 5’- GCGACTACGTGGTCTACTCG -3’, and a reverse primer 5’- AGGACCCTCATGGCCTTG -3’. Typical PCR reactions contained between 10 and 100 ng genomic DNA template, 10 pmol of forward and reverse primer. PCR was carried out in the presence of 7.5% DMSO, 5x buffer supplied with the enzyme and with 1.25U of LongAmp *Taq* DNA Polymerase (NEB) in a total volume of 30 µl using the following cycling conditions: initial denaturation step of 10 min at 95 ˚C, followed by 27 cycles of 30 sec 95 ˚C, 30 sec 60 ˚C, 60 sec 65 ˚C and a final extension step of 10 min 65 ˚C.

Analysis of PCR products for repeat number. One µl of PCR product was mixed with 0.3 µl LIZ-500 size standard (Applied Biosystems) and 11.7 µl formamide (Applied Biosystems) and run on a AB 3730 genetic analyser set up for fragment analyses with 50 cm capillaries. Results were analysed using GeneMarker software (Softgenetics). The genetic variable was coded as 0 or 1 for absence or presence, respectively, of a 7-repeat at one or both alleles. Of the 528 participants one child could not be genotyped; one hundred eighty-nine children (36%) were carriers of the long variant of DRD4 - the susceptible group. Three-hundred thirty-nine participants (64%) belonged to the less susceptible group because they did not carry the 7-repeat. The distribution of DRD4 polymorphisms was in Hardy-Weinberg equilibrium, χ^2 (df = 1, N = 528) = .01, p = .926.

RESULTS

We tested whether the program only proved effective for the lowest scoring 40 percent, as predicted, as teachers had broadened the sample by also including mid-range-scoring children. Included in the analysis were data on child sex, age in months, father’s education, child gene polymorphism (DRD4), the experimental or control condition to which the child was randomly assigned, and the child’s literacy level on the standardized CLT test before and after the intervention had taken place. The percentage of putatively susceptible children - carrying the 7-repeat allele of DRD4

- in the delayed and typical literacy level groups was 37% and 34%, respectively, the difference being non-significant, $\chi 2$ (df = 1, N = 528) = .59, p > .05. The number of children with a DRD4 7-repeat also did not differ significantly across the three experimental conditions: *Living Letters* (35.7%), *Living Books* (35.1%), and Clever Together (36.7%), the latter being the Control group; $\chi 2$ = .99. The sample was almost equally divided on sex (46% female).

Intervention efficacy

The post-test CLT was regressed on the following predictor terms: pre-test CLT (delayed versus mid-term), the contrasts between control group and *Living Letters* and control group and *Living Books*, DRD4 (carrier of one or two 7-repeat alleles versus others), and two- and three-way interactions involving pre-test CLT, interventions, and DRD4. The two group interventions were effect-coded by creating variables for the contrast between Control group and *Living Letters*, and Control group and *Living Books* (Cohen, Cohen, West, & Aiken, 2003). The child's sex, age in months, and father's education were entered as covariates. Since the assignment to the conditions was random, inclusion of covariates is not required to correct for any baseline differences, especially as the child's sex and age and father's education did not vary across the different groups (see Table 1). Inclusion of covariates, however, does reduce unexplained outcome variance and thereby increases power (Van Breukelen & Van Dijk, 2007).

Table 1. *Descriptives for the Complete Group and the Conditions Living Letters, Living Books and Clever Together*

	Complete Group (n = 528)	*Living Letters (n = 196)*	*Living Books (n = 174)*	*Clever Together (n = 158)*	*F-value (2, 526)*	*p-value*
Male/female	287/241	109/87	93/81	85/73	.13	.88
Age in months	66.83 (4.35)	66.70 (4.46)	66.97 (4.26)	66.86 (4.33)	.19	.83
Father's education	3.97 (1.93)	3.95 (1.99)	3.91 (1.89)	4.05 (1.91)	.24	.79
CLT pre	58.96 (8.67)	58.48 (8.75)	58.22 (8.18)	59.13 (9.07)	1.01	.37
Lowest 40%/ mid-range	290/238	101/95	107/67	82/76	-	-
CLT post	66.44 (9.95)	66.24 (9.35)	66.02 (10.34)	67.15 (10.25)	.59	.56

As the intraclass correlation coefficient was substantial we applied multi-level analysis using mixed models in SPSS in order to account for variation attributable to school-level characteristics (Luke, 2004). The intraclass correlation of [7.68/(7.68 + 61.30)] .11, demonstrated that 11% of the differences in the CLT scores was attributable to school characteristics (see random effects in Table 2).

Table 2. *Predicting CLT Posttest from CLT Pre-test, Living Letters, Living Books, and DRD4 with Age, Sex, and Father's Education as Covariates*

Measure	Estimate (*SE*)	95% CI	*t*	*p*-value	*df*
Fixed effects					
Intercept	54.21 (5.67)	43.07 -65.34	9.56	.00	527.47
Background					
Age	.09 (.08)	-.07 -.26	1.12	.26	526.60
Sex	.59 (.71)	-.80 -1.98	.84	.40	493.72
Father's education	.22 (.19)	-.15 -.59	1.16	.25	527.56
Main effects					
CLT pre	10.52 (.94)	8.66 -12.37	11.16	.00	505.73
Living Letters	.81 (.84)	.36 -3.81	.96	.34	496.93
Living Books	-.76 (.82)	-2.36 -.83	-.94	.35	493.36
DRD4	.96 (.99)	-.98 -2.91	.97	.33	479.64
Interaction effects					
CLT pre * *Living Letters*	-3.07 (1.21)	-5.45 -.68	-2.53	.01	510.24
CLT pre * *Living Books*	1.37 (1.26)	-1.11 -3.85	1.09	.28	501.75
CLT pre * DRD4	-2.49 (1.50)	-5.43 -.46	-1.66	.10	517.22
Living Letters * DRD4	-2.18 (1.38)	-4.89 -.52	-1.59	.11	514.85
Living Books * DRD4	3.01 (1.36)	.34 -5.69	2.21	.03	508.99
CLT pre * *Living Letters* * DRD4	5.80 (2.04)	1.80 -9.80	2.85	.005	505.69
CLT pre * *Living Books* * DRD4	-4.99 (2.13)	-9.18 - -.80	-2.34	.02	508.45

	Estimate (SE)	Wald Z	*p*-value
Random effects			
Level Child	61.30 (4.08)	15.02	.00
Level School	7.68 (3.22)	2.615	.01

Note. N = 528

The regression analysis revealed a significant main effect for pre-test CLT literacy level, significant two-way interactions between *Living Letters* and pre-test CLT literacy level and between *Living Books* and DRD4, and significant three-way interactions between the programs, pre-test CLT, and DRD4 (see Table 2). There was no significant main effects of *Living Books*, *Living Letters* or DRD4 on post-test CLT literacy. To address Keller's (2014) concerns regarding covariate interaction inclusion in gene x environment studies, we repeated the above analysis with the inclusion of the interactions of each of the three covariates (the child's sex, age, and father's education) with each of the four main variables (CLT literacy level, *Living Books*, *Living Letters*, and DRD4). The main effect of CLT literacy level was no longer significant, but the two-way interactions between *Living Letters* and pre-test CLT literacy level and between *Living Books* and DRD4, and the three-way interactions between the programs, pre-test CLT, and DRD4 remained significant. Thus, we restrict reporting here to these significant interactions.

The significant *Living Letters* x pre-test CLT interaction and the three-way interaction between the programs, pre-test CLT, and DRD4 indicated that *Living Letters* had a negative effect on the typically developing pupils and in particular in the not at risk group; see Table 2. See Table 3 for means and standard deviations for the CLT post-test.

Table 3. *Means and Standard Deviations for CLT post-test in the Delayed Group by Condition and DRD4*

	Complete	*n*	Living Letters	*n*	Living Books	*n*	Clever Together	*n*
DRD4 (7-)	61.64(8.60)	167	63.48(9.57)	56	59.95(7.81)	62	61.67(8.10)	49
DRD4 (7+)	62.00(7.11)	90	63.27(7.22)	37	62.58(7.07)	31	59.05(6.42)	22
Total	61.77(8.10)	257	63.40(8.67)	93	60.83(7.64)	93	60.86(7.67)	71

7- = low-susceptible. 7+ = high-susceptible.

Regressing the post-test CLT on *Living Letters* revealed non-significant main and interaction effects in the delayed group. The typical children in the *Living Letters* condition scored about 2.5 points (estimate = -2.42, *SE* = .92) lower than the Control group on the post-test CLT, which is a significant difference. Cohen's *d* was .20. A significant interaction *Living Letters* x DRD4 indicated that differences between *Living Letters* and control group were largest in the group without risk gene.

Gene x Intervention Interaction

The effect of *Living Books*, however, depended both on pre-test CLT level and DRD4, as revealed by the significant three-way interaction involving pre-test CLT x DRD4 x *Living Books*. The overall effects size of *Living Books* was low: Cohen's *d* was .15 (Table 5).

Table 5. *Cohen's d's and r's for Two Contrasts within DRD4 Groups in the Delayed Sample*

Contrast	DRD4	*d*	*r*
LivingLetters vs. control	all	.05	.03
	7-	.10	.05
	7+	-.04	-.02
LivingBooks vs. control	all	.06	.03
	7-	-.13	-.07
	7+	.40	.20

Notes. Cohen's $d = M_1 - M_2 / s_{pooled}$ where $s_{pooled} = \sqrt{[(s_1^2 + s_2^2) / 2]}$
r_{Yl} = d / √(d² + 4)

However, for the delayed children who were also carriers of a DRD4 7-repeat, evidence of a strong effect emerged from *Living Books* (Cohen's *d* = .43), but this was decidedly not the case for the children who did not carry the 7-repeat allele (Cohen's *d* = -.13); see Table 5 and Figure 1.

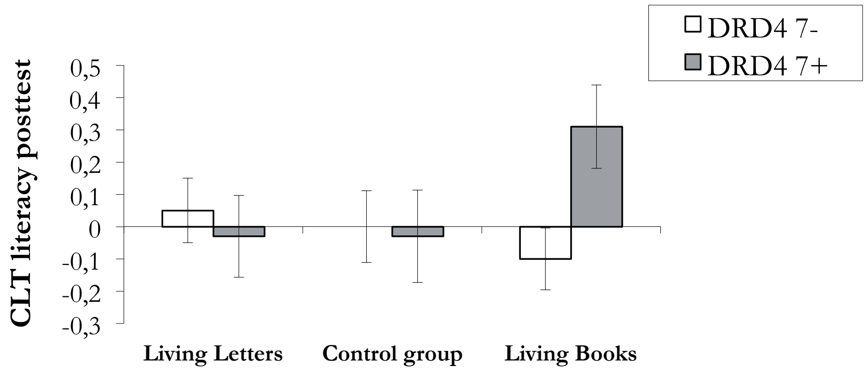

Figure 1
Standardized means and confidence intervals for CLT posttest corrected for age, father's education, and sex for carriers of DRD4 7- and DRD4 7+ scoring among the lowest 40 percent on the national CLT pre-test (N = 290). The grand mean was set to zero to give a better interpretable view of the differences between the groups.

Regressing the *Living Books* intervention on post-test CLT in the delayed and non-7R group yielded a non-significant estimate of -.80 (SE = .81). In the delayed but high-susceptibility group (i.e., carriers of the 7-repeat allele), however, the *Living Books* intervention group scored significantly higher than the Control group ($p < .015$). The estimate of 2.20 (SE = .88) means that the *Living Books* group scored on average more than 2 points higher on the post-test CLT. Results support the differential susceptibility hypothesis that only the genetically susceptible group benefited from the *Living Books* intervention.

DISCUSSION

The majority of children, by virtue of being immersed in a literate society, acquire emergent literacy concepts and skills relatively effortlessly during the course of early childhood (Ferreiro & Teberosky, 1982). For many children, the basis for emergent literacy is acquired within the period preceding formal literacy instruction, from birth to about six years of age. The subplot in this story, however, is equally important: An unacceptably large number of children are, at school entry, already lacking in competencies fundamental to their school success; they lack cognitive multipliers to engage in intensive practice of literacy once they are exposed to formal instruction in first grade, and at risk of being classified as (pseudo)dyslectic in later years (Stanovich, 1986).

In the current randomized controlled trial we tested literacy interventions that may narrow gaps in school entry skills. They are designed in a way that they can be used in addition to the regular curriculum because children can practice on the computer on their own. Both *Living Letters* and *Living Books* appeared to be effective interventions for pupils who are delayed in literacy according to a standardized Dutch test that is applied nation-wide twice during the year preceding first grade. Children scoring in the mid-range of the test, in contrast, did not benefit from the computer programs. This is an understandable outcome given that both programs train elementary literacy skills: Basic alphabetic knowledge and simple story comprehension. Thus, these programs designed for use with delayed or at-risk pupils are not effective for typically developing children scoring above the lowest 40 percent of literacy skills.

The current research shows evidence of genetic differential susceptibility for *Living Books*. Not all delayed pupils are affected by this computer intervention to promote early literacy skills. Differential effects of interventions are generally framed in dual-risk terms or diathesis-stress. Due to genetic characteristics (i.e. so-called ‘risk’ genotypes), some individuals need additional input to catch up and develop precursors

for literacy whereas other individuals without these ‘risk’ genes are not in need of a special program. For *Living Books* we found strong evidence for an alternative model to the diathesis-stress model: differential susceptibility. This model is based on the assumption that some of the children are not particularly susceptible to environmental input and hardly benefit from an intervention in addition to regular experiences with literacy. A susceptible group, in contrast, clearly responds in a positive way to the intervention: They lag behind without a special program but outperform their peers when receiving additional input which takes into account their reactive and easily distracted attention.

As the plot in Figure 1 shows, the 7-repeat polymorphism of DRD4 moderates the effects of *Living Books*. A high-susceptible group – carriers of the 7-repeat polymorphism of DRD4 – benefits from *Living Books* (d = .43) while the low-susceptible group does not (d = -.13). Inspection of Figure 1 reveals that the high-susceptible group does better than the low susceptible group in the case of *Living Books*, but that the reverse patterns is not seen for the control group. In other words, the high-susceptible manifest the “for-better” -but not for-worse pattern. This is plausible result as we selected the lowest achieving pupils.

In other words, about two-thirds of the delayed children do not benefit from additional book reading experiences beyond regular book readings in school and at home. The one-third susceptible children - carriers of a genetic marker of genetic differential susceptibility - learn substantially more when they receive *Living Books* with prompt and personalized performance feedback canalizing their attention and motivation toward the tasks at hand.

The current findings for *Living Letters*, in contrast, do not meet the criteria for differential susceptibility. Children lagging behind in literacy did not benefit from this program. Results of the current research do not meet the statistical criteria for genetic differential susceptibility as we found in a previous study of younger children with the same program (Kegel et al., 2011). *Living Letters* might have been less appropriate to reveal differential effects in an older group of delayed children as most children may acquire the target skills in this program within a brief period of rapid growth (Paris, 2005). Even the most delayed five-years old pupil may easily reach a high level on the most difficult task in *Living Letters* - identifying the first letter of the proper name or mama as the last or middle sound in words - and score at ceiling on target skills after playing the games in *Living Letters*. Had we included more advanced phonemic skills to *Living Letters* we might have found more variation in effects between low- and high-susceptible children similar to findings in a younger group (Kegel et al., 2011).

Current results underline the importance of identifying sub-samples of genetically high-susceptible pupils in education. An emergent corpus of work has shown the value of early interventions for supporting literacy achievements in young at-risk children (e.g., Lonigan, Farver, Phillips, & Clancy-Menchetti, 2011). However, these experiments have rarely taken into account genetic differences or other markers that may moderate program effects. As appears from this study and previous ones (Kegel et al., 2011; Van den Berg & Bus, 2014), it may even happen that a program's effect may not become manifest when the focus is on the complete, undivided group of children and the crucial question about 'what works for whom' is not asked. Differential susceptibility theory implies, *a priori,* that markers of differential effectiveness to be tested as moderators in educational interventions.

Genetically high-susceptible children may benefit from extra computer-based instruction due to continuous feedback to their responses that teacher are not able to provide in over-crowded classrooms. From previous research comes strong evidence supporting the hypothesis that continuous feedback to children's reactions built into the literacy program is an effective mechanism especially for genetically high-susceptible children such as carriers of the DRD4 polymorphism. Feedback may help these children to stay attentive despite of distractors and to avoid responding randomly, which proved to be the case in an earlier investigation in which we compared an abbreviated version of *Living Letters* in which feedback was omitted with the complete version of the program (Kegel et al., 2011; Kegel & Bus, 2012). In sum, feedback as part of *Living Books* may explain why high-susceptible children benefit more from this program than from similar daily book reading experiences within the regular kindergarten curriculum.

Implications and Future Directions

The current account of variation in effects of early intervention programs challenges the traditional double-risk or diathesis stress model in education and highlights the need for a paradigm shift towards differential susceptibility (Belsky, Jonassaint, Pluess, Stanton, Brummet, & Williams, 2009; Belsky & Pluess, 2009, 2013). The fact that only some children proved susceptible to treatment may explain why Aptitude Treatment Interaction (ATI) failed as an explanation of differential outcomes of instruction (Cronbach & Snow, 1977). The ATI model, popular in the seventies of the last century, is based on the assumption that all children have different susceptibilities and need instruction attuned to their susceptibilities. Our findings in particular with *Living Books* indicate that a sub-group of children identified on an *a priori* basis using a specific genetic marker are especially susceptible to this program. Special literacy

programs can profoundly affect some children's literacy, but average or across-the-board effects will often mis-estimate the impact of a program, underestimating it for some and overestimating it for others. Focus on genetically more susceptible sub-samples is needed to demonstrate the power of early literacy programs. As we found for *Living Books* effect sizes for high-susceptible children may be much higher than effects for the total group. Our findings thus contrast with the received ATI model to address the question what works best for whom in education and account for an alternative model, differential susceptibility, with a theoretical basis in evolutionary theory and neurobiology, and with more clear-cut hypotheses about relevant markers. It is therefore imperative to include markers of differential susceptibility as moderators in experimental designs to make correct estimates of the importance of intervention programs to improve early literacy. Armed with specific differential susceptibility hypotheses about neurobiological or behavioral markers as moderators of program effects, researchers can shed new light on previously hidden efficacy of programs that were reported not or only moderately effective (Van IJzendoorn & Bakermans-Kranenburg, 2012; Bakermans-Kranenburg & Van IJzendoorn, 2015 in press).

Furthermore, neurobiological markers that predict differential outcomes of early literacy programs may typically cause high reactivity to the environment, for better and for worse (Belsky et al., 2007; Belsky & Pluess, 2009, 2013; Ellis et al., 2011). Carriers of the DRD4 polymorphism may lag more behind under bad learning conditions but outperform the low-susceptible children when they receive optimal additional input. Thus, susceptibility markers are doubled edged, serving as a risk factor for academic skills under negative learning conditions but as a potential asset and promotive factor under optimal conditions. This new and exciting idea has potentially far-reaching implications for early academic education. It should be noted that genetic measures may reach beyond traditional boundaries of behavioral measures in showing reactivity and predicting which children are likely to make good progress (Kegel & Bus, 2012; Kegel, Van der Kooy-Hofland, & Bus, 2009; Wasserman & Drucker Wasserman, 2012). Ultimately, a thorough understanding of how genetic mechanisms regulate children's susceptibilities to environmental influences should provide a solid foundation for shaping programs to maximally benefit children. It is likely that in due course DRD4 will be shown to be a sensitive index of an underlying genetic pathway modulating dopamine production and re-uptake and that more easily observed endophenotypic correlates will be found that represent this pathway.

Lastly, successful literacy intervention programs that change the odds for children may not just intensify experiences with relevant tasks as Justice and colleagues (2003) advocate but additionally provide support regarding how to approach tasks. There is evidence that an emphasis on performance feedback while solving problems is especially important. Correcting how children approach tasks – realized by continuous performance feedback to children's responses in the programs which were the focus of this report - may be especially effective for highly susceptible children but not for all learners (e.g., Bodrova & Leong, 2006). When children are highly susceptible to the environment, the mainstream classroom environment may be an obviously unsatisfactory, distracting, and chaotic environment; overcrowded early literacy settings are likely to challenge these students much more than their more sturdy peers. They may, however, outperform their classmates when a (computer) program succeeds in mobilizing and channeling children's high reactivity by providing intensive, closely monitored, and individualized scaffolding.

Programs such as *Tools of the Mind* may therefore be good candidates to support the learning of highly susceptible children (Bodrova & Leong, 2006). So far research does not demonstrate strong effects for this literacy intervention for preschool and kindergarten, children; and we suspect this is because relevant research informed by differential-susceptibility thinking has not yet been conducted (Barnett et al., 2008).

Afterword

An obvious practical implication of the current finding that children carrying the 7-repeat DRD4 allele especially benefit from *Living Books* may involve screening of pupils in search of an optimal fit between the program and individual characteristics. Increasing knowledge of factors that determine susceptibility for instruction may provide concrete guidance in identifying (*a priori*) subsets of pupils that are especially susceptible to specific instructional mechanisms. Practitioners and policymakers will thus obtain more realistic estimates of the effectiveness of preventive and curative efforts. It is therefore an important area for future investigation to further specify genetic and behavioral characteristics of children who need intensive, closely monitored, and individualized practice as in *Living Books*—and who can especially benefit from them.

However, as long as realistic estimates of the effectiveness of preventive or curative programs cannot be made by practitioners, it seems prudent to address school entry skills of all kindergarten children who are delayed in these skills, even though some learn as much when they are exposed to the regular curriculum with additional treatment compared to the regular curriculum only. Given the promising outcome that susceptible pupils benefit most from an additional computer program beyond formal reading instruction, it seems important to present such extra programs to all delayed five-year-olds especially because these computerized programs are very cost-effective and fun to do. It should be noted as well that there are no indications for

negative effects of the intervention among children not carrying the 7-repeat allele. As yet it seems therefore most in line with the idea of No Child Left Behind to include all children with delayed early literacy development in the intervention.

An alternative implication may be blaming the susceptible children and trying to change them to better cope with adverse environments (Ellis et al., 2011). The differential susceptibility model does not promote blaming or making the vulnerable more durable. On the contrary, the model provides a new perspective on how to support susceptible children in need with an emphasis on a better fit between individual characteristics and environmental input.

The quite modest or even absent effects of programs in the majority of pupils is a source of concern for researchers and educators. Of course, it is possible that children who belong to the less-susceptible group are simply nonresponsive to any intervention. Until that is found to be the case, it is probably best to presume that programs that are tailored to other child characteristics of learning may speed up the acquisition of literacy skills among *seemingly* low-susceptible children.

ACKNOWLEDGMENTS

This study was supported by a grant from the Dutch organization Kennisnet to Adriana G. Bus.

REFERENCES

Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2006). Gene–environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in pre- schoolers. *Developmental Psychobiology, 48*, 406-409. doi: 10.1002/dev.20152

Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2011). Differential susceptibility to rearing environment depending on dopamine-related genes: New evidence and a meta-analysis. *Development and Psychopathology*, *23,* 39–52. doi:10.1017/S0954579410000635

Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2015, in press). The hidden efficacy of interventions: Gene x environment experiments from a differential susceptibility perspective. *Annual Review of Psychology*.

Barnett, W., Jung, K., Yarosz, D., Thomas, J., Hornbeck, A., Stechuk, R., & Burns, S. (2008). Educational effects of the Tools of the Mind curriculum: A randomized trial. *Early Childhood Research Quarterly, 23*, 299–313. doi: 10.1016/j.ecresq.2008.03.001

Belsky, J., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2007). For better and for worse. Differential susceptibility to environmental influences. *Current Directions in Psychological Science*, *16,* 300-304. doi: 10.1111/j.1467-8721.2007.00525.x

Belsky, J., Jonassaint, C., Pluess, M., Stanton, M., Brummet, B., & Williams, R. (2009). Vulnerability genes or plasticity genes? *Molecular Psychiatry*, *14,* 746-754. doi: 10.1038/mp.2009.44

Belsky, J. & Pluess, M. (2009). Beyond diathesis-stress: differential susceptibility to environmental influences. *Psychological Bulletin, 135*, 885-908. doi: 10.1037/a0017376

Belsky, J., & Pluess, M. (2013). Beyond risk, resilience and dysregulation: Phenotypic plasticity and human development. *Development and Psychopathology, 25*, 1243-1261. doi: 10.1017/S095457941300059X

Belsky, J., Pluess, M., & Widaman, K. F. (2013). Confirmatory and competitive evaluation of alternative gene-environment interaction hypotheses. *Journal of Child Psychology and Psychiatry, 54*, 1135-1143. doi:10.1111/jcpp.12075

Bodrova, E., & Leong, D. J. (2006). Vygotskian perspectives on teaching and learning early literacy. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 2, pp.243-256). New York: Guilford Press.

Both-De Vries, A. C., & Bus, A. G. (2008). Name writing: A first step to phonetic writing? Does the name have a special role in understanding the symbolic function of writing? *Literacy Teaching and Learning*, *12*, 37-55. Retrieved from http://www.readingrecovery.org/rrcna/journals/ltl/index.asp

Both-De Vries, A. C., & Bus, A. G. (2010). The proper name as starting point for basic reading skills. *Reading and Writing*, *23*, 173-187. doi: 10.1007/s11145-008-9158-2

Bus, A. G., Takacs, Z. K., & Kegel, C. A. T. (in press). Affordances and limitations of electronic storybooks for young children's literacy: Consequences for engineering apps. *Developmental Review.*

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Commissie Testaangelegenheden Nederland [Committee Test Affaires Netherlands]. Retrieved May 2014, from http://www.boomtestuitgevers.nl/producten/onderwijs/**cotan**_documentatie.

Cronbach, L. J., & Snow, R. E. (1977). *Aptitudes and instructional methods*. New York: Irvington.

Ellis, B. J., Boyce, W. T., Belsky, J., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2011). Differential susceptibility to the environment: An evolutionary-neurodevelopmental theory. *Development and Psychopathology*, *23*, 7–28. doi:10.1017/S0954579410000611

Ferreiro, E., & Teberosky, A. (1982). *Literacy before schooling*. Exeter, NH: Heinemann.

Goswami, U., & Bryant, P. E. (1990). *Phonological skills and learning to read*. Hillsdale, NJ: Erlbaum.

Howe, G. W., Beach, S. R. H., & Brody, G. H. (2010). Microtrial methods for translating gene-environment dynamics into preventive interventions. *Prevention Science*, *11*, 343-354. doi: 10.1007/s11121-010-0177-2

Hsiung, G. Y. R., Kaplan, B. J., Petryshen, T. L., Lu, S., & Field, L. L. (2004). A dyslexia susceptibility locus (DYX7) linked to dopamine D4 receptor (DRD4) region on chromosome 11p15.5. *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics)*, 125b, 112-119. doi: 10.1002/ajmg.b.20082

Hulme, C., Bowyer-Crane, C., Carroll, J. M., Duff, F. J., & Snowling, M. (2012). The causal role of phoneme awareness and letter-sound knowledge in learning to read: Combining intervention studies with mediation analysis. *Psychological Science*, *23*, 572–577. doi: 10.1177/0956797611435921

Justice, L. M., Chow, S. M., Capellini, C., Flanigan, K., & Colton, S. (2003). Emergent literacy intervention for vulnerable preschoolers: Relative effects of two approaches. *American Journal of Speech-Language Pathology, 12*, 320-332. doi: 10.1044/1058-0360(2003/078)

Kamil, M. L., Intrator, S. M., & Kim, H. S. (2000). The effects of other technologies on literacy and literacy learning. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research:* Vol. 3. (pp. 771-778). Mahwah, NJ: Erlbaum.

Kegel, C. A. T., & Bus, A. G. (2012). Online tutoring as a pivotal quality of web-based early literacy programs. *Journal of Educational Psychology*, *104*, 182-192. doi: 10.1037/a0025849

Kegel, C. A. T., & Bus, A. G. (2013). Links between DRD4, executive attention, and alphabetic skills in a nonclinical sample. *Journal of Child Psychology and Psychiatry, 54*, 305-312. doi: 10.1111/j.1469-7610.2012.02604.x

Kegel, C. A. T., Bus, A. G., & Van IJzendoorn, M. H. (2011). Differential susceptibility in early literacy instruction through computer games: The role of the dopamine D4 receptor gene (DRD4). *Mind, Brain, and Education*, *5*, 71–78. doi: 10.1111/j.1751-228X.2011.01112.x

Kegel, C. A. T., Van der Kooy-Hofland, V. A. C., & Bus, A. G. (2009). Improving early phoneme skills with a computer program: Differential effects of regulatory skills. *Learning and Individual Differences, 19*, 549–554. doi: 10.1016/j.lindif.2009.07.002

Keller, M. C. (2014). Gene x environment interaction studies have not properly controlled for potential confounders: The problem and the (simple) solution. *Biological Psychiatry, 75*, 18-42. doi: 10.1016/j.biopsych.2013.09.006

Lansink, N., & Hemker, B. (2012). *Wetenschappelijke verantwoording van de toetsen Taal voor Kleuters voor groep 1 en 2 uit het Cito volgsysteem primair onderwijs* [Scientific justification of literacy tests for the kindergarten years]. Arnhem, The Netherlands: Cito.

Levin, I., Both-De Vries, A. C., Aram, D., & Bus, A. G. (2005). Writing starts with own name writing: From scribbling to conventional spelling in Israeli and Dutch children. *Applied Psycholinguistics*, *26*, 463–477. doi:10.1017/S0142716405050253

Lonigan, C. J., Farver, J. M., Phillips, B. M., & Clancy-Menchetti, J. (2011). Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models. *Reading and Writing, 24*, 305–337. doi:10.1007/s11145-009-9214-6

Luke, D. A. (2004). *Multilevel Modeling*. Thousand Oaks, CA: Sage.

National Center for Family Literacy (2008). *Developing early literacy: Report of the National Early Literacy Panel: A scientific synthesis of early literacy development and implications for intervention.* Washington, DC: National Institute for Literacy.

Obradović, J., Bush, N. R., Stamperdahl, J., Adler, N. E., & Boyce, W. T. (2010). Biological sensitivity to context: The interactive effects of stress reactivity and family adversity on socio-emotional behavior and school readiness. *Child Development, 81*, 270-289. doi: 10.1111/j.1467-8624.2009.01394.xParis, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly, 40*, 184–202. doi: 10.1598/RRQ.40.2.3

Robbins, T. W., & Everitt, B. J. (1999). Drug addiction: Bad habits add up. *Nature*, *398*, 567-570. doi:10.1038/19208

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.) (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407. doi: 10.1598/RRQ.21.4.1

Tripp, G., & Wickens, J. R. (2008). Dopamine transfer deficit: A neurobiological theory of altered reinforcement mechanisms in ADHD. *Child Psychology and Psychiatry, 49,* 691–704. doi: 10.1111/j.1469-7610.2007.01851.x

Van Breukelen, G. J. P., & Van Dijk, K. R. A. (2007). Use of covariates in randomized controlled trials. *Journal of the International Neuropsychological Society,* 13, 903–904. doi: 10.10170S1355617707071147

Van den Berg, H., & Bus, A. G. (2014, under review). Beneficial effects of BookStart in temperamentally highly reactive infants.

Van der Kooy-Hofland, V. A. C., Bus, A. G., & Roskos, K. A. (2012). Effects of a

brief but intensive remedial computer intervention in a sub-sample of kindergartners with early literacy delays. *Reading and Writing, 25*, 1479–1497.doi: 10.1007/s11145-011-9328-5

Van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2012). Differential susceptibility experiments: Going beyond correlational evidence: Comment on beyond mental health, differential susceptibility articles. *Developmental Psychology*, *48*, 769-774. doi: 10.1037/a0027536

Van IJzendoorn, M. H. & Bakermans-Kranenburg, M. J. (2015, this issue). Genetic differential susceptibility on trial: Meta-analytic support from randomized controlled experiments. *Development and Psychopathology*.

Van IJzendoorn, M., Bakermans-Kranenburg, M., Belsky, J., Beach, S., Brody, G., Dodge, K., Greenberg, M., Posner, M., & Scott, S. (2011). Gene-by-environment experiments: A new approach to find missing heritability. *Nature Reviews Genetics*, *12*, 881. doi:10.1038/nrg2764-c1

Wasserman, T., & Drucker Wasserman, L. (2012). The sensitivity and specificity of neuropsychological tests in the diagnosis of attention deficit hyperactivity diagnosis. *Applied Neuropsychology: Child, 1*, 90-99. doi: 10.1080/21622965.2012