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

What Works for Whom?

Differential genetic effects of early literacy interventions in kindergarten

Rachel D. Plak

Cover design: Sinds 1961 Grafisch Ontwerp, Ede, The Netherlands
Printed by: Print Service Ede, Ede, The Netherlands

What Works for Whom?

Differential genetic effects of early literacy interventions in kindergarten

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Leiden
op gezag van de rector magnificus
prof. mr. C.J.J.M. Stolker,
volgens besluit College voor Promoties
te verdedigen

op donderdag 15 december 2016 klokke 10.00 uur

door

Rachel Dominique Plak
geboren te Amsterdam in 1985

Promotores

Prof. dr. A.G. Bus
Prof. dr. R. van Steensel

Promotiecommissie

Prof. dr. M. J. Bakermans-Kranenburg
Prof. dr. P.C.J. Segers (Radboud Universiteit Nijmegen)
Dr. A. ten Brummelhuis (Stichting Kennisnet)

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

Table of contents

Chapter 1

General Introduction

INTRODUCTION

There are various ways for young children to come in touch with written language at school and in their home environment. Book sharing is often considered as one of the most important activities parents and teachers can do to promote young children's early literacy skills (Bus, van IJzendoorn, & Pellegrini, 1995; Mol & Bus, 2011). In addition, literacy-related experiences in homes and schools include reading and writing of words such as the child's proper name or other people's names (Both-de Vries & Bus, 2010). More than three decades ago, Emilia Ferreiro and Ana Teberosky (1982) showed the importance of the early years for the development of literacy. Their research had a strong impact on the research community and since then a spate of articles has appeared to explain such early learning processes. As research in the field of early literacy underscores the importance of the early years for developing the foundation for future literacy, the interest in early interventions has strongly increased as well.

This dissertation explores whether computers can contribute to assisting and supporting teachers when emergent literacy skills lag behind. Since it is easier to tailor the format and content of web-based programs to individual differences than to ensure that classroom instruction meets the needs of all individual pupils, computer programs may be an attractive tool for providing additional home- and school-like experiences with literacy in order to advance early literacy skills of young children at risk. Unfortunately, there is a lack of evidence with regard to computer programs as tools to provide young children with relevant practice, even though there is increasing interest in computer programs in support of instruction in early stages of becoming literate. Moreover, the number of computers or tablets with Internet connections in Dutch kindergarten classrooms has grown from one computer per 17 pupils in 1999 to one device (computer, laptop or tablet) per four students in 2015 (Kennisnet, 2015). In recent years, there has been a gradual increase of the role that digital learning material plays in the curriculum (Blockhuis, ten Voorde, & Sluijsmans, 2014; Kennisnet, 2013). This trend is continuing and teachers are using more and more digital content. The share of digital learning materials in the curriculum was according to teachers 15% in 2007-2008, this increased in 2014-2015 to about 25%. Teachers use ICT, but do not make optimal use of the possibilities ICT can offer, for example offering students personalized learning with ICT applications (Kennisnet, 2015).

The focus in this dissertation is on two educational computer programs targeting emergent literacy skills. One program, *Living Books*, includes technology-enhanced animated books that offer extra nonverbal information that matches the story text and may thus provide guidance in understanding the text. Examples of such additions are animated illustrations, zooming, sounds, and background music (Bus, Takacs, & Kegel, 2015; Takacs, Swart, & Bus, 2015). *Living Books* are formatted as

an attractive movie-like representation of the story that closely matches the events presented in the story text. The nonverbal additions may, if well designed, replace any scaffolding provided by an adult because the additional nonverbal information directs attention to details in pictures, just as adults do while sharing a story with a child (Takacs, Swart, & Bus, 2014). Even more remarkable is that the movie-like presentation of the story, including background sounds and music, may make the books more compelling and engaging than a traditional print book (Verhallen & Bus, 2009), and that children may achieve "flow" while listening to the stories. Persons "in flow" can be described as being deeply involved in an activity that they find enjoyable and they do not have to make any effort to concentrate while involved in the activity (Kiili et al., 2012).

The second educational computer program, *Living Letters*, was composed of a series of highly structured game-like assignments in which phonemic awareness and alphabetic knowledge were trained. It is designed according to the principle that verbal responsiveness is important to stimulate young children's learning (Tamis-LeMonda, Cristofaro, Rodriguez, & Bornstein, 2006). Verbal responsiveness implies that a tutor responds to all attempts that children make to solve assignments. The program's responses are not only prompt and dependent on the child's focus of attention but they have a positive, supportive tone and add constructive information. For instance, a first error in an assignment is followed by a repetition, a second by a hint, and a third by a demonstration of the correct solution.

Even though it is likely that educational computer programs - modeled on children's daily literacy-related activities - contribute to children's literacy development, they may not do so to the extent that they have added value when compared to common experiences at home and in school. Given the usual circumstances are that children have daily literacy experiences in school, including (book) reading and writing of (the proper) name(s), educational computer programs may only imply a minor expansion in the amount of time spent on these literacy activities. It is therefore not plausible to have high expectations of effects of an educational computer program on language and literacy skills in kindergarten even when children are delayed in language and literacy skills.

On the other hand, it is quite possible that educational computer programs can have an added value to common literacy experiences due to special functionalities built in the educational computer programs that respond to children's specific needs. E.g., it is conceivable that the computer programs are especially helpful when children have problems concentrating on tasks. Some program features may help children to achieve flow, that is, children are deeply involved and do not have to make any effort to be highly concentrated (Kiili et al., 2012).

Genetic differential susceptibility

There is increasing evidence in the field of child development and pathology showing that children with a specific genetic make-up are more susceptible than others to the quality of their environment, including educational program. In a less optimal or stimulating environment they are more prone to abnormal development than their peers (e.g., Ellis, Boyce, Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2011). However, in an adaptive learning situation they may even outperform their peers (Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007; Ellis et al., 2011). Ellis and Boyce (2008) described the individual differences in genetic make-up by means of a metaphor. Similar to how dandelions are able to survive and bloom, irrespective of soil, sun, drought, or rain, some children thrive whether or not the learning environment is optimal. Dandelions are known to exist in open fields, but also between the cracks of paving stones. Orchids by contrast can, similar to more susceptible children, be considered as context-sensitive. Their survival and flourishing is intimately tied, like that of the more susceptible children, to the nurturing or neglectful qualities of the environment.

Based on ample research, both correlational and experimental, the D4 Receptor Gene, with the polymorphism DRD4 7-repeat, is widely used as a marker in relation to the child-rearing environment (e.g. Bakermans-Kranenburg & Van IJzendoorn, 2006, 2011, 2015; Ptáček, Kuželová, & Stefano, 2011). The DRD4 gene is a 48-bp variable number tandem repeat (VNTR) with on its third exon repeats varying from 2 to 11. Two alleles are most common, the 4 repeat and the 7-repeat (Chang, 1996). The greater part of the population is carrier of the 4-repeat allele, while about one third is carrier of the polymorphism 7-repeat allele. Carriers of the 7-repeat allele of the DRD4 gene, also nicknamed “the long variant”, are thought to be more susceptible to the quality of their environment than non-carriers (see for an overview Bakermans-Kranenburg & Van IJzendoorn, 2011). Carriers have lower dopamine reception efficiency, caused by a diminished anticipatory cell firing, which is related to reduced attentional and reward mechanisms (Robbins & Everitt, 1999). Although the dopamine receptor D4 is associated with (in)attention and motivation (Hsiung, Kaplan, Petryshen, Lu, & Field, 2004; Tripp & Wickens, 2008), and even with ADHD (Maher, Marazita, Ferrel, & Vanykukov, 2002) and novelty seeking (Ebstein, 1996, 1997), disappointingly there are no associations between the DRD4 gene and cognitive abilities and behavioral traits (Pappa et al., 2011). However, if the 7-repeat allele of the DRD4 gene is considered in interaction with environmental qualities it does predict outcomes like cognitive performance; e.g., four-year-old carriers of 7-repeat allele of the DRD4 gene made great progress in literacy skills when pupils were exposed to the highly responsive computer program, *Living Letters*, whilst non-carriers did not benefit from this educational computer program (Kegel, Bus, & van IJzendoorn, 2011).

Carriers of the 7-repeat allele DRD4 gene may experience problems focusing their attention, and being concentrated on an activity, but some computer activities may elicit them becoming highly engaged (Kühn et al, 2011; Koepp et al., 1998) and so enter a state of flow (Kiili et al., 2012). There is still no explanation of which aspects of computer programs are so beneficial to and result in children who are carriers of the DRD4 gene being in a state of deep concentration and learning (flow). Teachers may not be able to offer similar qualities to this particular subsample as focusing attention, explaining and scaffolding, repetitively, can be challenging to teachers. In order to stimulate the learning process, a teacher must not only give special attention to the child with attention problems in a group of 20 to 30 children, but also stay calm, remain positive and patient. Moreover, teachers are not always able to monitor the learning process closely and are therefore unable to provide the exact help the child needs at a particular moment (Van de Pol, Volman, & Beishuizen, 2011). Technology-enhanced computer programs do make it possible to offer children lagging behind personalized adaptive guidance exclusively tailored to the needs of every individual child.

AIMS AND OUTLINE OF THE PRESENT THESIS

In previous smaller scaled experiments, the effect of educational computer programs was tested and confirmed (Kegel et al., 2011; Van der Kooy-Hofland, Bus, & Roskos, 2012). The intervention effect moderated by genetic profile was based on a rather small sample, and genetically moderated intervention effects need replication (Van IJzendoorn et al., 2011) and therefore new randomized controlled trials were planned. From August 2012 to October 2012 and August 2013 to October 2013 the project was broadly advertised via e-mail, mail, social media, and phone in the Netherlands. In total 183 schools from all over the country, from urban and rural areas, participated. That is 2.7% of all Dutch schools. We, as researchers, had little involvement in the implementation of the intervention. We asked the kindergarten teachers to select eligible children in their classroom and to provide parents with brochures about the research. After parents gave written informed consent for participation of their child, we randomly assigned children to either one of the programs, *Living Books* or *Living Letters*, or the control condition. In January teachers administered the Central Institute for Test Development (Cito) Literacy Test for Kindergarten (CLT). This is a standardized literacy test (Lansink & Hemker, 2012) that is used on almost all primary schools in the Netherlands. The test consists of two versions with similar items, the first version is administered in January and the second version in June. The intervention period started in March of the senior kindergarten year and ended 12 weeks later in June. Teachers received online accounts to which they could log children in and have them work with the program to which they were assigned. Immediately following the intervention period of twelve weeks, teachers administered the CLT June. Halfway during the intervention period, buccal cell samples were collected by trained members of the research team. The samples were sent to a commercial laboratory for DNA analysis.

We also examined whether computer training by means of technology-enhanced educational programs of clearly defined academic core language and literacy skills (phoneme awareness & story comprehension and vocabulary) can reduce the risk of developing reading problems in first grade. Based on outcomes of a previous smaller scaled study, we expect that the extra input from technology-enhanced educational programs will not be equally effective for all children. For a deeper understanding of learning processes it is important to find out *who* benefits the most from these interventions. In other words, it is our aim to find moderators of the effects of technology-enhanced programs in kindergarten. In this thesis, the focus will be on a genetic characteristic, i.e., whether or not a child is a carrier of the 7-repeat allele of the DRD4 gene, which may clarify for whom technology-enhanced programs are effective and why.

In Chapter 2, a nation-wide randomized controlled trial (RCT) examines whether effects of technology-enhanced educational programs *Living Letters* and *Living Books* were moderated by the dopamine receptor D4. Effects for pupils in the senior year of kindergarten are described for carriers and non-carriers, but also for children that score in the lowest ranges of a standardized literacy test, literacy-delayed children, and children that score midrange i.e. average scoring children.

Chapter 3 looks at whether the results of the first wave could be replicated in a second wave with new participants. Furthermore, a meta-analysis was conducted to determine whether the outcomes of the Gene x Environment RCT reported in this study matched the results of previous experiments.

Educational programs can be called preventive when long-term effects are present. Since this has not yet been established for the genetic moderator DRD4, long-term effects of *Living Letters* and *Living Book*s are examined in Chapter 4.

The question of why technology-enhanced programs work for certain subsamples is further explored and discussed in Chapter 5.

Finally, the findings of the above studies are integrated and summarized in Chapter 6. Furthermore, directions for future research are discussed.

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). The hidden efficacy of interventions: Gene x environment experiments from a differential susceptibility perspective. *Annual Review of Psychology*, *66*, 381–409. doi:10.1146/annurev-psych-010814-015407

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

Blockhuis, C., ten Voorde, M., & Sluijsmans, L. (2014). Leermiddelenmonitor 13/14: Beleid, gebruik, digitalisering en ontwikkeling van leermiddelen [Policy, use and development of teaching aids]. Enschede: SLO.

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., van IJzendoorn, M. H., & Pellegrini, A. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*, 1-21. doi: 10.3102/00346543065001001.

Chang, F.M., Kidd, J.R., Livak, K.J., Pakstis, A.J., Kidd, K.K. (1996). The world-wide distribution of allele frequencies at the human dopamine D4 receptor locus. *Human Genetics*, *98*, 91–101.

Ebstein, R.P., Novick, O., Umansky, R., Priel, B., Osher, Y., Blaine, D., Bennett, E.R., Nemanov, L., Katz, M., Belmaker, R.H. (1996). Dopamine D4 receptor (D4DR) exon III polymorphism associated with the human personality trait of Novelty Seeking. *Nature Genetics*, 12, 78–80. doi:10.1038/ng0196-78

Ebstein, R.P., & Belmaker, R.H. (1997). Saga of an adventure gene: Novelty Seeking, substance abuse and the dopamine D4 receptor (D4DR) exon III repeat polymorphism. *Molecular Psychiatry*, 381–384. doi.org/10.1038/sj.mp.4000315

Ellis, B. J., & Boyce, T. W. (2008). Biological sensitivity to context. *Current Directions in Psychological Science, 17* (3), 183–187. doi: 10.1111/j.1467-8721.2008.00571.x

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.

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

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

Kennisnet (2015). Vier in balans-monitor 2015: Inzet en opbrengsten van ICT in het onderwijs. [Deployment and results of ICT in education]

Kennisnet (2013). Vier in balans monitor 2013: De laatste stand van zaken van ICT en onderwijs. [The most recent state of affairs relating to ICT and education]

Kiili, K., de Freitas, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science, 15*, 78-91. doi: 10.1016/j.procs.2012.10.060

Koepp, M.J., Gunn R.N., Lawrence A.D., Cunningham V.J., Dagher A., Jones T., et al. (1998). Evidence for striatal dopamine release during a video game. *Nature*, *393*, 266–8.

Kühn, S., Romanowski A., Schilling C., Lorenz R., Mörsen C., Seiferth N. et al. (2011). The neural basis of video gaming. *Translational Psychiatry*, *53*, doi:10.1038/tp.2011.53.

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.

Maher, B. S., Marazita, M. L., Ferrell, R. E., & Vanyukov, M. M. (2002). Dopamine system genes and attention deficit hyperactivity disorder: a meta-analysis. *Psychiatric* Genetics, *12*, 207–215. doi:10.1097/00041444-200212000-00003

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137,* 267–296.doi:10.1037/a0021890

Molloy, K., Griffiths, T.D., Chait, M., Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. *The Journal of Neuroscience,9*, *35*(49), 16046 –16054. doi:10.1523/JNEUROSCI.2931-15.2015

Pappa, I., Szekely, E., Mileva-Seitzb, V.R., Luijk M.P.C.M, Bakermans-Kranenburg, M.J. Van IJzendoorn, M.H., &Tiemeier, H. (2011). Beyond the usual suspects: a multidimensional genetic exploration of infant attachment disorganization and security. *Attachment & Human Development*,17, 3, 288–301, doi.org/10.1080/14616734.2015.1037316

Ptáček, R., Kuželová, H., & Stefano, G. B. (2011). Dopamine D4 receptor gene DRD4 and its association with psychiatric disorders. *Medical Science Monitor*, 17, 215– 220. doi: 10.12659/MSM.881925

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

Schecklmann, M., Ehlisa, A., Plichtaa, M.M, Romanosb, J., Heinea, M., Boreatti-

Hümmera, A., Jacoba, C., Fallgattera, A.J (2008). Diminished prefrontal oxygenation with normal and above-average verbal fluency performance in adult ADHD. *Journal of Psychiatric Research*, *43*(2) 2008, 98–106. doi:10.1016/j.jpsychires.2008.02.005

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2014). Can the computer replace the adult for storybook reading? A meta-analysis on the effects of multimedia stories as compared to sharing print stories with an adult. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01366

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks. A meta-analysis. *Review of Educational Research*, *85*, 4, 698–739 doi:10.3102/0034654314566989

Tamis-LeMonda, C. S., Cristofaro, T. N., Rodriguez, E. T., & Bornstein, M. H. (2006). Early language development: Social influences in the first years of life. In L. Balter & C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (2nd ed.). New York: Psychology Press.

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 de Pol, J., Volman, M., & Beishuizen, J. (2011). Patterns of contingent teaching in teacher–student interaction. *Learning and Instruction, 21,* 46-57. doi:10.1016/j.learninstruc.2009.10.004.

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., Belsky, J., Beach, S., Brody G., Dodge, K.A. et al. (2011). Gene-by-environment experiments: a new approach to finding the missing heritability. *Nature Reviews Genetics*; 12: 881–882. doi:10.1038/nrg2764-c1

Verhallen, M. J. A. J., & Bus, A. G. B. (2009). Video storybook reading as a remedy for vocabulary deficits: outcomes and processes. *Journal for Educational Research Online*, *1*, 117–141.

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 (β = .47), but not after entering executive attention in the regression model (β = .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, $\chi2$ (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; $\chi2$ = .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.

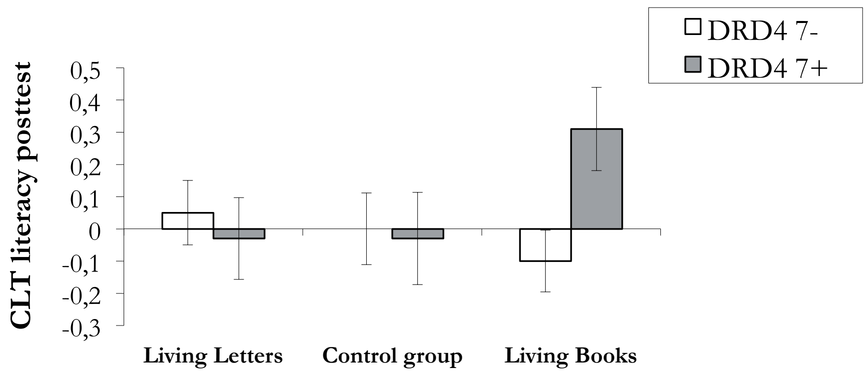


Figure 1
Standardized means and confidence intervals for CLT posttest corrected for age, father's education, and sex for carriers of DRD4 7- and DRD4 7+ scoring among the lowest 40 percent on the national CLT pre-test (N = 290). The grand mean was set to zero to give a better interpretable view of the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

Chapter 3

Brief Computer Interventions Enhance Emergent Academic Skills in Susceptible Children: A Gene-by-Environment Experiment

Published as:

Plak, R.D., Merkelbach, I., Kegel, C.A.T., Van IJzendoorn, M.H., & Bus, A.G. (2016). Brief computer interventions enhance emergent academic skills in susceptible children: A gene-byenvironment experiment. *Learning and Instruction*, 45, 1-8. doi;10.1016/j.learninstruc.2016.06.002

ABSTRACT

In this study we examined the potential of technology-enhanced educational programs for young children lagging behind in emergent literacy skills. Differential effects of technology-enhanced educational programs (*Living Letters* and *Living Books*) for poor performers were tested in a randomized controlled trial. Our previous study showed that children with a dopamine-related genetic polymorphism - DRD4 7-repeat - are more susceptible to their learning environment than children without this polymorphism, serving as a proxy for the dopamine-system related genetic pathway. In the current study, we aimed to replicate and extend these results in a sample of 583 kindergarteners from 136 schools. As predicted by the genetic differential susceptibility theory, carriers of the DRD4 7-repeat allele profited significantly from *Living Books* (d = .75), whereas non-carriers did not benefit (d = .02). *Living Letters* did not show a Gene x Environment interaction. We discuss why carriers of DRD4 7-repeat allele particularly benefit from *Living Books*.

Keywords: Early literacy intervention, Dopamine D4 receptor gene, Educational computer programs, Technology-enhanced picture storybooks, Differential susceptibility, Gene X Environment interaction

INTRODUCTION

In this study we examined the potential of educational computer programs for young children lagging behind in emergent academic skills and therefore at risk for learning problems in primary education. Building on the Simple View of Reading, Bowyer-Crane et al. (2008) proposed a two-dimensional model of early reading interventions with phonological skills positioned on one dimension, and non-phonological skills (e.g., semantics and syntax) positioned on the other. We explored the efficacy of both types of programs in kindergarten age: *Living Letters* to prevent the risk of word-level decoding difficulties, and *Living Books* to prevent the risk of reading comprehension difficulties associated with deficits in non-phonological language skills.

Average effects of technology-enhanced educational programs for young children are unfortunately rather disappointing unless subsamples are formed (e.g. Saine, Lerkkanen, Ahonen, Tolvanen, & Lyytinen, 2011). We argue that the average effects do not reflect true effects that are hidden in a subgroup of children who are more susceptible to environmental experiences, such as child rearing and school environment. Children who are thought to be susceptible to environmental factors not only catch up and perform at a level similar to other children, but they actually outperform their nonsusceptible peers. It appears from child development research that children with specific genetic or temperamental characteristics are more susceptible to the quality of the environment than others, and are at risk of a delayed development in comparison with their peers when they grow up in less stimulating families or other child rearing contexts (e.g., Ellis, Boyce, Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2011). These susceptible children easily catch up and even outperform their peers in an optimal environment. This is the For Better and For Worse principle of the Differential Susceptibility Theory (Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007; Ellis, Boyce, Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2011).

In the popular press, susceptible children have been compared to orchids - these flowers only bloom when temperature and humidity are optimal - in contrast to less susceptible species, like dandelions, which grow irrespective of the quality of the environment (Dobbs, 2009). This Differential Susceptibility Model challenges the traditional Diathesis Stress Model in which the idea is promoted that children with specific characteristics are more vulnerable to adversities or negative experiences and that others are less affected by the same experiences (Belsky & Pluess, 2013; Belsky, Bakermans-Kranenburg, & IJzendoorn, 2007). Boyce and Ellis (2005) stated that the Differential Susceptibility Model is fundamentally different from the Diathesis Stress Model in that individuals not only vary in the extent to which they are susceptible to negative experiences but also their malleability; individuals who are more malleable are thought to be more susceptible to negative and positive experiences (Belsky et al., 2007). Children, who due to specific characteristics are thought to be less susceptible, are less influenced by negative or positive environmental factors.

Correlational and experimental studies showed that children with a specific dopamine-related genetic polymorphism - dopamine D4 receptor gene, with the polymorphism DRD4 7-repeat - are thought to be more susceptible to their environment than children without this polymorphism (Bakermans-Kranenburg & Van IJzendoorn, 2006, 2011, 2015). For example, in a randomized parent training experiment, carriers of the 7-repeat variant showed less externalizing behavior when their parents were coached to enhance their sensitive caregiving, whereas with the same coaching non-carriers did not make progress (Bakermans-Kranenburg & Van IJzendoorn, 2008). Children with the 7-repeat allele of the DRD4 gene (the long variant of the DRD4 gene) have a lower dopamine reception efficiency - caused by diminished anticipatory cell firing - which is associated with reduced attentional and reward mechanisms (Robbins & Everitt, 1999). It is hypothesized that these children are more sensitive to an environment that helps to structure their activities and is supportive of reward and attentional mechanisms (Bakermans-Kranenburg and Van IJzendoorn, 2015). A crowded noisy classroom can therefore be regarded as a negative learning environment for children with the DRD4 7-repeat allele. Due to reduced attentional mechanisms, they are at risk of wandering off and failing to use the learning opportunity.

Educational computer programs may be especially helpful for supporting these children's learning processes especially when they do not just offer practice but also provide scaffolding during learning, thus helping them to stay task-focused (Kegel & Bus, 2012). *Living Letters* - a program promoting alphabetic knowledge - includes a tutor, one of the characters in the games, who responds to all attempts children make to solve the tasks. If children do not start solving a problem or their first attempt fails, the tutor encourages the child to try. After the second failure, the tutor provides a cue for finding a solution and after the third failure, the tutor models and explains the correct solution. Because of the scaffolding, children may be able to solve tasks that are just above the level that they normally can solve on their own (e.g., Kegel & Bus, 2012). *Living Books* is a technology-enhanced book reading program composed of eight different stories each repeated twice. The animated pictures, sounds, and music appearing simultaneous with the story text help to make sense of the story and 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). This format stimulates building combined verbal and nonverbal representations, also referred to as multimedia learning (Mayer, 2005). If the music is congruent to the narration, it facilitates story comprehension and learning words from the narration (Takacs, Swart, & Bus, 2015). Furthermore, additions to the story, like movie-like presentations and background music, are engaging and may therefore be helpful in staying focused

while hearing the story.

In particular, carriers of the 7-repeat allele of the DRD4 gene that has been associated with ADHD (Maher, Marazita, Ferrel, & Vanykukov, 2002) may benefit from these computer programs. Lower dopamine reception efficiency may result in being less attentive during the learning process when the learning environment is not sufficiently structured and supportive (Kegel & Bus, 2012). Because carriers of the 7-repeat variant of the DRD4 gene might be less able to focus on a task in a somewhat chaotic environment, such as in regular classrooms, and therefore are often distracted from core learning processes, they may be dependent on special tailored programs that help them to focus. Even when part of the tailored programs have overlap with the regular curriculum, carriers of the DRD4 7- repeat allele may benefit from additional educational computer programs. *Living Letters* might be very helpful to them because it includes an Intelligent Tutoring System providing consistent feedback to all of the children's responses. *Living Books* retains and guides attention with the help of a movielike presentation. It may be that children with attentional problems benefit from the movie-like presentation in the *Living Books*. Acevedo-Polakovich, Pugzles Lorch, and Richard (2007) showed that children with ADHD enjoy watching television more and have greater involvement in television-related activities when compared to typically developing children. We expected therefore to find a cross-over Gene Environment interaction, showing that carriers of the 7-repeat allele of the DRD4 gene may profit most from both computer programs and even outperform their less susceptible peers without this DRD4 7-repeat allele.

Results so far were mixed. A study by Kegel, Bus, and Van IJzendoorn (2011) was the first experiment in which genetic differential susceptibility for *Living Letters* was tested. From this study, which included typical four-year-olds, it appeared that in particular children with the 7-repeat allele of the DRD4 gene were susceptible to the educational computer intervention *Living Letters*. Carriers of the 7-repeat allele fell behind in early literacy skills in comparison to their peers without the extra input of *Living Letters*, but outperformed their peers when they did work with *Living Letters*. Plak, Kegel, and Bus (2015) tested both the educational computer programs, *Living Letters* and *Living Books*, in a group of five-year-olds delayed in early literacy skills. They could not replicate the differential effects for *Living Letters* but did find differential effects of *Living Books*. Non-carriers of the 7-repeat allele did not benefit from *Living Books* but carriers did substantially. Teachers suggested that *Living Letters* was too easy and too boring for five-year-olds, which might explain results so far. In the current study, we therefore complicated the program by dropping the easiest games.

Current study

In a large countrywide randomized controlled trial, we provided 5-year-olds, just before their transition from kindergarten to first grade, with the computer programs that trained important precursors for learning to read in first grade.We targeted children from 136 different schools whose early literacy skills were delayed or rather low according to their teacher and their scores on a national standardized literacy test. In total, 583 5-year old kindergarten children were randomly assigned to one of two early literacy computer programs, *Living Letters*, *Living Books*, or a control condition, *Clever Together* that practiced visual-spatial skills. They practiced with *Living Letters*, *Living Books* and *Clever Together* for a brief period, ranging between 160 and 220 min and over the course of about 12 weeks. The researchers assigned the children randomly to one of the three programs and provided online access but were not involved in the implementation or testing. A standardized test for literacy administered by teachers in January and June in the senior kindergarten year provided pre- and posttest scores. We tested (1) main effects of literacy programs on the standardized literacy test, (2) whether there was evidence for a Gene x Environment interaction and carriers of the 7-repeat DRD4 gene outperformed non-carriers, and (3) conducted a metaanalysis to examine whether the outcomes of the randomized GxE experiment reported in this study match the outcomes of previous experiments using the same programs and measures.

METHOD

Design

In each class one child was assigned to *Clever Together*, the control condition, and at least one child to one of the literacyrelated programs, that is *Living Letters* or *Living Books*. Children in the *Living Books* condition were offered 8 sessions with 2 books per session and children in the *Clever Together* and *Living Letters* condition ranging between 8 and 11 sessions (due to a variable number of sessions depending on the number of errors children made). Each session took about 15 minutes and children practiced once aweek. Based on the data that were stored by the program, children completed on average 33.4 out of 34 *Living Letters* games (SD = 3.09) and they "read" on average 14.7 out of 16 books (SD = 2.1).

Buccal cell samples (samples from cells from a person's cheek) were collected half way through the intervention period by trained research teammembers using a sterile swab designed for collecting buccal cells for DNA analysis (Omni Swabs, Whatman/GE Healthcare, UK). Collection took place at school. Given the large sample

and the stability of the percentage of carriers of the DRD4 gene, it was reasonable to assume that random assignment would result in a similar number of carriers of the 7-repeat allele of the DRD4 gene in each condition. Early literacy skills were tested before and after the intervention using the Central Institute for Test Development (Cito) Literacy Test for Kindergarten (CLT). In all schools this literacy test was administered group-wise in January and June (Lansink & Hemker, 2012).

Participants

A total of 136 schools participated in the experiment. From August 2013 to October 2013 information about the project was distributed via e-mail, mail, social media, and phone. In brochures and letters sent to the schools, it was emphasized that participation 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, after the intervention participating schools would receive free access to educational computer programs for kindergarten children for a period of three months (www.bereslim.nl).

Eligible children were selected between October 2013 and February 2014 by the kindergarten teachers in the 136 participating schools. Teachers were asked to select six pupils from their classroom achieving poorly in literacy. Pupils who were eligible were those, for instance, who were not yet able to write their proper name, to rhyme, to name a few letters, and to identify sounds in words. Preferably these children scored in the lowest ranges - between 0 and 59 - on the standardized literacy test CLT administered in January (Lansink & Hemker, 2012). If there were not enough children scoring below the 40th percentile, teachers also included children scoring midrange between the 40th and 60th percentile on the standardized literacy test. Dutch was required as the participants' first language. When a parent refused consent, the teacher was asked to select another eligible pupil from his or her classroom.

Due to objection to genotyping, a condition for participation, parents often refused consent, resulting in less than six lowperforming participants per classroom (average number of participants per classroom: 2.94 children). From the 607 selected pupils, data were complete for 565 pupils. About half of these pupils showed a serious lag in literacy skills (N = 307), scoring below the 40th percentile according to national norms (Lansink & Hemker, 2012). The other half (N = 258) scored in the mid-range between the 40*th* and 60*th* percentile.

Procedure

Parents of eligible children received written information about the study explaining the scientific goals and the opportunity for their child to receive extra coaching. Information about genotyping as part of the study was offered as well. Moreover, a website was available for additional information about the aim and the design of the study. Contact informationwas provided to allowparents to ask additional questions. Parents made frequent use of this opportunity. Genotypingwas a main reason for parents refusing consent for participation (roughly 25% of all children selected by the teachers) even though we guaranteed that buccal cells would be destroyed after genotyping had taken place and data would be stored anonymously.

Intervention programs

Living Letters. The intervention program *Living Letters* 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). The program adapts automatically to the child's proper namewhenavailable in the name database. If the name of the child is not available, the program uses 'mama' (mommy) - a word that is just as familiar to many young children. In all, there were 36 games from which 16 targeting name recognition, 6 games to recognize the first letter of the name, and 12 in which children are given the task of identifying pictures that start with or contain the first letter of the child's name. Feedback provided by a tutor followed every response by the child. When children answered a question incorrectly, feedback and clues were provided. After a maximum of three trials, the game ended on a positive note, irrespective of whether a correct response was given, whereupon a newgame started. Every time a child failed to fulfill an assignment, this assignment was repeated in the following two sessions. Therefore some children had more sessions than others.

Living Books, the second intervention program, was made up of eight age-appropriate digital animated storybooks. The animated pictures, sounds, and music support the meaning of the story text, which may stimulate the child's understanding of story events and language (Bus et al., 2015; Kamil, Intrator, & Kim, 2000). We assumed that when the oral narrative is accompanied by nonverbal information, and verbal and nonverbal information are simultaneously available, the narrative text would be understood and retained better than if conveyed by words alone (Bus et al., 2015). Multimedia offer optimal guidance in developing mental representations of the story and the language. Each reading of a book was interrupted four times for questions about the story and vocabulary. 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*. Each book was presented twice and four questions were included in each session.

Clever Together supports basic concepts for mathematics like practicing cardinals and visual-spatial reasoning. It includes 40 games. As in *Living Letters* and *Living Books*, a tutor provides constructive, detailed feedback for every error and every correct response. Assignments were repeated in later sessions when children made errors.

Measures

Early literacy skills

Cito Literacy Test for Kindergarten Pupils (CLT) is a standardized literacy test for kindergarten pupils that is administered in almost every Dutch school class-wise in January and June of the senior kindergarten year. Because this test is administered at almost every school in the Netherlands, the CLT test was an obvious choice. The 60-item CLT assesses vocabulary, text comprehension, rhyming, hearing the first and last word, sound blending, writing conventions (e.g. reading from left to right), and prediction of book content based on the book cover (Lansink & Hemker, 2012). The Commissie Testaangelegenheden Nederland [Committee for Test Quality in the Netherlands] evaluated the CLT as adequate.

Genetic screening for DRD4 polymorphisms

A genotype is an assortment of characteristics inherited from the parents. The genetic information about those characteristics is stored in the DNA. Genotyping is the process of determining differences in the genetic make-up (genotype) of an individual by examining the individual's DNA sequence. Biological assays are used and compared to another individual's DNA sequence. The details are explained below. PCR Amplification. The region of interest of the DRD4 gene was amplified by PCR using the following primers: a FAM-labelled primer 50-GCGACTACGTGGTCTACTCG-30, and a reverse primer 50- AGGACCCTCATGGCCTTG-30. 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 30sec 95 °C, 30sec 60 °C, 6sec 65 °C and a final extension step of 10 min 65 °C.

Analysis of PCR products for repeat number. One ml 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 allele at one or both alleles. Of the 593 participants, ten children could not be genotyped; 199 children (34%) were carriers of the 7-repeat of DRD4. The distribution of DRD4 polymorphisms was in Hardy-Weinberg equilibrium, $\chi 2$ (*df* = 1, *N* = 565) = .008, *p* = .93.

Data analysis

The posttest score of the CLT was regressed on the pretest CLT (scoring below the 40th percentile according to national norms vs. scoring midrange between the 40th and 60th percentile), *Living Letters* (contrast between the control condition *Clever Together* and *Living Letters*), *Living Books* (contrast between *Clever Together* and *Living Books*), DRD4 (7-repeat at one or both alleles vs. others), two and three-way interactions involving pretest CLT, interventions, and DRD4. Age, sex, and father's education were entered as covariates.

RESULTS

Characteristics of the sample

Table 1 presents data on children with delayed versus midrange pretest scores on the literacy test. Delayed participants had a mean age of 66.87 months (*SD* = 4.14) at pre-test, participants who scored midrange had a mean age of 67.87 months (*SD* = 4.66). Boys were overrepresented (62%) particularly in the delayed group. In the midrange group, numbers of boys and girls were equal. The mean score for father's education was 3.63 (*SD* = 1.37) on a scale ranging from 0 to 6, where 0 represents primary school and 6 represents university-level education.

Table 1: Characteristics in the conditions *Living Letters*, *Living Books*, and *Clever Together* of participants scoring below the 40th percentile (delayed) or between the 40th and 60th percentile (midrange) on the pretest

		Complete group	*Living Letters*	*Living Books*	*Clever Together*
	Male/Female	190/117	69/46	67/33	54/38
	Age months (*SD*)	66.87 (4.14)	67.17 (4.25)	66.53 (4.18)	66.87 (3.97)
	CLT pretest (*SD*)	53.40 (4.78)	53.18 (4.82)	53.91 (4.45)	53.12 (5.06)
Literacy delayed	CLT posttest (*SD*)	61.79 (7.55)	61.22 (7.11)	62.98 (8.03)	61.23 (7.49)
	Male/Female	129/129	50/43	43/51	36/35
	Age months (*SD*)	67.87 (4.66)	68.28 (4.72)	67.56 (4.79)	67.74 (4.44)
	CLT pretest (*SD*)	65.95 (6.80)	65.44 (5.53)	66.02 (6.01)	66.54 (9.00)
Midrange	CLT posttest (*SD*)	72.66 (9.58)	71.95 (9.07)	73.46 (10.92)	72.52 (8.27)

The percentage of children carrying the 7-repeat allele of DRD4 in the delayed and the midrange literacy level groups was respectively 37.5% and 31.8%, a non-significant difference, $\chi 2$ (*df* = 1, *N* = 565 = 2.00, *p* = .158. The number of children with a DRD4 7- repeat allele did not differ significantly across the three experimental conditions: *Living Letters* (37.0%), *Living Books* (33.0%), and *Clever Together* (34.4%), $\chi 2$ (*df* = 2, *N* = 565) = .744, *p* = .689.

Intervention efficacy

We tested whether or not it was necessary to allow the intercept to differ between schools and to have an interaction between intervention and school in the regression model (Twisk, 2006). The difference between the -2log likelihood of the model with a random intercept and the -2log likelihood of the model without a random intercept equaled 8.03. Following a chi-square distribution with one degree of freedom, this difference was highly significant. The difference between the -2log likelihood of the model with only a random intercept and the -2log likelihood of the model with both a random intercept and a random slope was not significant ($\chi 2$ = 1.41, *df* = 2). Therefore we applied multilevel analysis with a random intercept for schools (Luke, 2004). The intraclass correlation of.10 ([6.86/(6.86 + 62.85)], see Table 2) demonstrated that 10% of the differences in the CLT scores were attributable to school characteristics. Therefore multi-level analyses were applied to account for variation attributable to school-level characteristics using school as a random factor.

Neither the experimental conditions *Living Letters* (*Est* = -.04, *p* = .959) and *Living Books* (*Est* = .17, *p* = .838) nor DRD4 (*Est* = 1.00, *p* = .308) revealed main effects, in contrast to the dichotomized CLT pretest (*Est* = 10.88, *p* = .000). The dichotomized pretest did not show an interaction with DRD4 (*Est* = -2.16, *p* = .144). *Living Letters* (vs. control) did not show an interaction effect with the dichotomized CLT pretest (*Est*

= -1.21, p = .315) and DRD4 (*Est* = -1.82, p = .174) or a three-way interaction effect with the dichotomized CLT pretest and DRD4 (*Est* = 3.08, p = .133). *Living Books* (vs. control) did not show a two-way interaction effect with the dichotomized CLT pretest (*Est* = 1.75, p = .143), but there was a significant interaction with DRD4 (*Est* = 3.49, p = .015) and the three-way interaction including *Living Books* (vs. control), DRD4, and the dichotomized CLT pretest was found to be significant (*Est* = -6.57, p = .002).

Table 2*: Outcomes of multilevel analysis using posttest literacy skills (CLT) language as an outcome variable (N = 565)*

	Est. (*SE*)	95% CI	t	*p*	*df*
Intercept	46.71 (5.74)	35.44 - 57.98	8.141	.000	564.98
Background					
Age	.19 (.08)	.03 - .35	2.356	.019	565.00
Sex	.43 (.71)	-.96 - 1.82	.612	.541	549.68
Father's educational level	.54 (.26)	.03 - 1.05	2.084	.038	564.60
Main Effects					
CLT pretest	10.88 (.88)	9.16 -12.61	12.396	.000	562.88
Living Letters (vs. control)	-.04 (.84)	-1.69 - 1.60	-.051	.959	546.35
Living Books (vs. control)	.17 (.82)	-1.44 - 1.77	.205	.838	523.38
DRD4 variant	1.00 (.98)	-.93 - 2.93	1.021	.308	539.99
Interaction Effects					
CLT pretest X *Living Letters*	-1.21 (1.20)	-3.57 - 1.15	-1.005	.315	556.00
CLT pretest X *Living Books*	1.75 (1.19)	-.59 - 4.09	1.47	.143	535.41
CLT pretest X DRD4 variant	-2.16 (1.48)	-5.06 - .74	-1.462	.144	541.79
DRD4 variant X *Living Letters*	-1.82 (1.34)	-4.44 - .81	-1.36	.174	557.03
DRD4 variant X *Living Books*	3.49 (1.42)	.69 - 6.29	2.45	.015	555.00
CLT pretest X *Living Letters* X DRD4 variant	3.08 (2.05)	-.94 - 7.10	1.504	.133	553.83
CLT pretest X *Living Books* X DRD4 variant	-6.57 (2.07)	-10.64 - -2.50	-3.174	.002	543.92
Random Effects	Est. (*SE*)	Wald Z		p	
Variance					
Level Child	62.85 (4.24)	14.820		.000	
Level School	6.86 (3.16)	2.172		.030	

As the results indicated that effects differed dependent on the starting level, analyses were repeated for delayed and midrange children separately. In the delayed

group (*N* = 307), there was neither a main effect for *Living Letters* (*Est* = .18, p = .804), *Living Books* (*Est* = .07, p = .924) or DRD4 (*Est* = 1.15, p = .176), but significant two-way interactions between *Living Letters* and DRD4 (*Est* = -2.29, p = .049 and between *Living Books* and DRD4 (*Est* = 3.76, p = .002). See Table 3 and Figure 1. There were no significant effects in the midrange group (*N* = 258).

Table 3: *Means and Standard Deviations for CLT post-test in the Delayed Group by Condition and DRD4 (N = 307)*

	Living Letters	*n*	*Living Books*	*n*	*Clever Together*	*n*
DRD4(7-)[b]						
Raw	61.42 (7.00)	65	61.45 (8.20)	69	61.72 (6.87)	58
Corrected[a]	61.49 (7.43)	65	61.59 (7.36)	69	61.42 (7.39)	58
DRD4(7+)[c]						
Raw	60.96 (7.30)	50	66.39 (6.54)	31	60.38 (8.48)	34
Corrected[a]	60.91 (7.37)	50	66.21 (7.37)	31	60.70 (7.40)	34
Total						
Raw	61.22 (7.11)	115	62.98 (8.03)	100	61.23 (7.49)	92
Corrected[a]	61.20 (7.42)	115	63.90 (7.98)	100	61.06 (7.61)	92

[a] Covariates appearing in the model are evaluated at the following values: gender = .38, age = 66.87, and educational level of the father = 3.61.
[b] non-susceptible.
[c] susceptible.

Of the children who received the *Living Books* intervention and were delayed at pretest and were carriers of the DRD4 7-repeat allele, 80.6% scored at least midrange on the posttest. For children assigned to the control condition this percentage did not exceed 55.6%.

Effect sizes

The overall effect size of *Living Books* in the delayed group (see Table 4), based on the corrected means and standard deviations, was rather small (Cohen's d = .36) in contrast to the effect in the sub-group of carriers of the DRD4 7-repeat allele (7+, Cohen's d = .75), which indicates a large effect. For the non-carriers of the DRD4 7-repeat allele (7-), the effect size was about zero (Cohen's d = .02). Results showed that only the group of carriers of the 7- repeat allele benefited from *Living Books*. *Living Letters* yielded a non-significant overall effect size (Cohen's d = .02), without differences between the carriers of the DRD4 7-repeat (Cohen's d = .03) and the non-carriers (Cohen's d = .01).

Table 4: *Effect Sizes and 95% Confidence Intervals for Living Letters and Living Books based on the means corrected for covariates*

Contrast	DRD4	*d*	95 % CI
Living Letters vs. Control	All	.018	-.256 - .292
	7-	.009	-.345 - .364
	7+	.028	-.407- . .464
Living Books vs. Control	All	.364	.078 - .649
	7-	.023	-.326 - .372
	7+	.746	.243 - 1.249

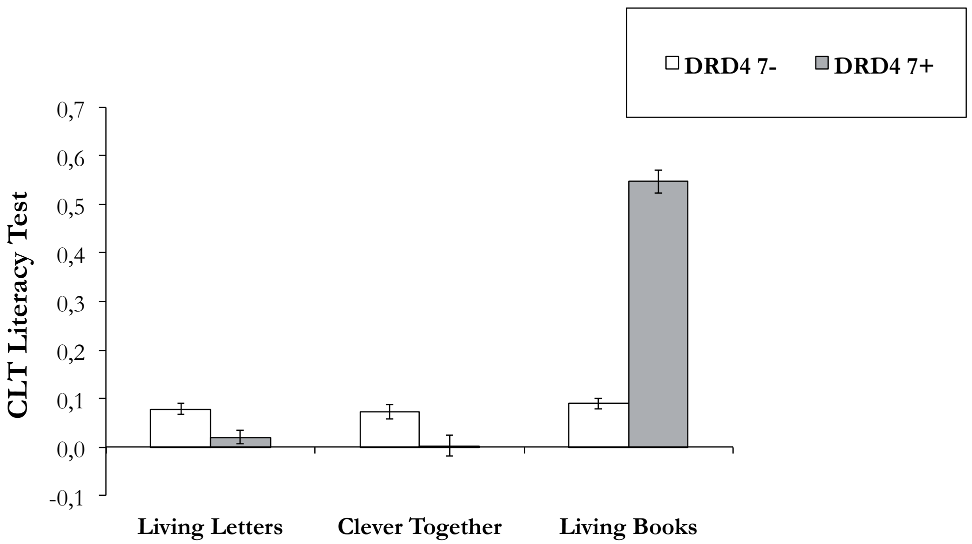


Figure 1:
Means and confidence intervals for CLT posttest corrected for father's education, age and gender for carriers of DRD4 (7-) and DRD4 (7+) scoring below the 40th percentile (delayed) on the pretest (N = 307).

Meta-analysis

To test the consistency of the current outcomes with results of prior experiments (Kegel et al., 2011; Plak et al., 2015), we conducted a meta-analysis. We included two prior experiments and the current study - in all 5 contrasts - encompassing *N* = 730 participants of which 272 were carriers of the 7-repeat allele of the DRD4 gene. The

studies revealed three contrasts between *Living Letters* and *Clever Together* and two between *Living Books* and *Clever Together*. For an overview of the contrasts included in this metaanalysis and the results, see Table 5.

Table 5: *Overview of Studies Included in the Meta-Analysis*

	Cohen's *d* overall [95%]	*n* Exp/ control	Cohen's *d* (7+) [95%]	*n* Exp/ control	Cohen's *d* (7-) [95%]	*n* Exp/ control
Kegel et al. *Living Letters*	.65 [.28, 1.02]	45/88	1.06 [.41, 1.72]	14/35	.35 [-.10, .79]	31/53
Plak et al. I *Living Letters*	.04 [-.33, .41]	97/40	-.03 [-.62, .57]	38/15[a]	.11 [-.36, .58]	59/25[a]
Plak et al. II *Living Letters*	.02 [-.32 - .36]	115/46[a]	.03 [-.52 - .58]	50/17[a]	.01 [-.43 - .45]	65/29[a]
Plak et al. I *Living Books*	.13 [-.24, .49]	103/40[a]	.39 [-.22, .99]	37/15	-.12 [-.58, .34]	66/25
Plak et al. II *Living Books*	.36 [.01 - .71]	100/46	.75 [.18 – 1.31]	31/17[a]	.02 [-.41 - .46]	69/29[a]

Notes. All values are weighted and based on the complete set unlike in earlier reports.
[a] In Plak et al., I and II, the experimental conditions (*Living Letters and Living Books*) were both compared with the same control condition. Therefore the number of children in the control condition was equally divided over the two contrasts.

Living Letters. Outcomes across three experiments are summarized in Figure 2. *Living Letters* revealed a non-significant effect both in the carriers of the 7-repeat allele (d = .33, 95% CI = -.131- .787, p = .162) and in the non-carriers (d = .14, 95% CI =-.260- .539, p = .494). The difference between carriers and non-carriers was not significant, Q(1) = .367, p = .545. The heterogeneous results of the three experiments showed a large effect in favor of the carriers of the 7-repeat allele in Kegel et al. (2011), a small effect in Plak et al. (2015) and a low effect in the current study.

Living Books. Outcomes across two experiments are summarized in Figure 2. Carriers of the 7-repeat allele were strongly affected by *Living Books* (d = .59, 95%

CI = .157- 1.105, p = .007) whereas non-carriers were not (d = -.05, 95% CI = -.363 - .267, p = .767). The difference between carriers and non-carriers was significant, Q(1) = 5.454, p = .020.

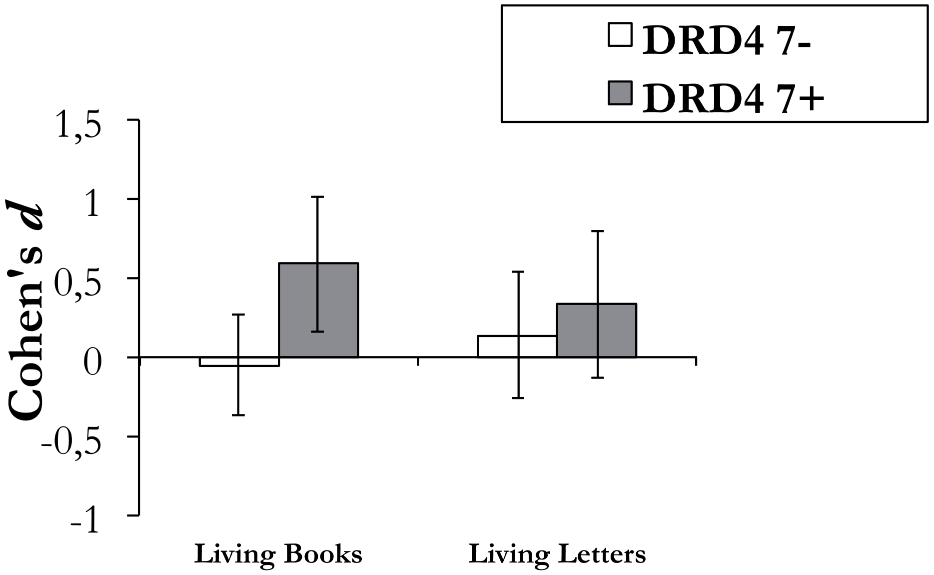


Figure 2:
Cohen's d for DRD4 (7-) and DRD4 (7+) scoring in the lowest ranges of literacy tests (N = 730) for *Living Books* and *Living Letters* across five experiments.

DISCUSSION

In the current randomized controlled trial including 565 five year- olds, we showed that two digitized programs, *Living Letters* and *Living Books* - both providing support in solving age-appropriate literacy-related tasks - failed to show effects across all subjects and programs. A subsample - slightly more than 30% of all kindergarten children - did benefit from *Living Books* as was demonstrated by a significant Gene x Environment interaction that fits the Differential Susceptibility Model (Belsky et al., 2007; Belsky & Pluess, 2009, 2013; Ellis et al., 2011; Van IJzendoorn & Bakermans-Kranenburg, 2012). As predicted by the genetic differential susceptibility hypothesis, we showed that carriers of the DRD4 7-repeat allele profited significantly from *Living Books*, whereas non-carriers did not benefit. *Living Letters*, by contrast, did not reveal a Gene x Environment interaction in the current randomized controlled trial: Carriers of the 7-repeat allele did not significantly benefit from the guidance that was offered by this program, similar to a previous trial with same age children (Plak et al., 2015) but

in contrast to a previous trial with younger children (Kegel & Bus, 2012).

Living Books guided the comprehension of the story by providing nonverbal information (animated pictures, music and sound effects) closely matched with the narration (Bus et al., 2015). This book reading intervention thus narrows the gap in language and literacy skills for poor performers who - as carriers of the DRD4 7-repeat allele - are known to be particularly susceptible to qualities of their learning environment. We hypothesize that particularly due to the multimedia in *Living Books*, including movie-like representations, sound and music, these susceptible children are more successful in understanding a story. Children with the 7-repeatallele may fail to benefit from storybook reading in a regular classroom environment because they are easily flooded with irrelevant perceptual and auditory stimuli. Just as is found in children with ADHD (see Schecklmann et al., 2008), carriers of the 7- repeat allele may become less susceptible to stimuli that are irrelevant to the task, due to numerous sources of sensory information each of which demands high levels of attention. It is even imaginable that the high attentional load of the books may result in a state of inattentional blindness and deafness (Molloy, Griffiths, Chait, & Lavie, 2015), which is known to make children insensitive to irrelevant stimuli from their surroundings, for example a crowded classroom. They may, so to speak, become hyper-focused on tasks that put a load on visual and auditory perception, resulting in outperforming their peers who are not carriers of the 7-repeat allele.

For children scoring midrange, *Living Books* may not be challenging enough since difficult words used in the stories are known by most children in kindergarten (Schaerlaekens, Kohnstamm, & Lejaegere, 1999). It is possible that children scoring midrange, who are carriers of the DRD4 7-repeat allele, would benefit from a more advanced version of *Living Books* including books with more difficult words and storylines.

There were no effects of *Living Letters*, even in the delayed group. As teachers in the Plak et al. (2015) study had complained that *Living Letters* was too tedious, we made a reduction in the easy games that mainly focused on familiarizing with the proper name. The rest of the program - learning the name of the first letter of the proper name and identifying the sound of this letter in words - remained unaltered. Despite this change, the program may not have corresponded to the level of five-year-olds even though they were delayed in literacy skills. On the other hand, *Living Letters* may not meet the criteria of a task that can keep children's attention with the 7-repeat allele focused. Even though *Living Letters* has various sources of sensory information, the pace of the program may be too slow: the push and pull between watching the general introduction, listening to specific instructions, and fulfilling the assignments may make it hard for the 7-repeat carriers to become totally engrossed in *Living Letters*.

Overall effects for carriers of the 7-repeat allele - meta analysis

To compare the results of the current randomized controlled trial with previous findings in studies that share a common design and instrumentation, we conducted a meta-analysis to show the consistency in outcomes of experiments so far, in line with guidelines of the New Statistics idea (Cumming, 2014). *Living Letters* showed an effect size of d = .33 for carriers of the 7-repeat allele, but results were not homogeneous ranging from d = 1.06 in Kegel et al.'s study to d = .03 in the current study. Younger children seem to benefit more than older children even when the older group only includes delayed children. *Living Books*, in contrast, revealed a homogeneous effect size of d = .59 for delayed carriers of the 7- repeat allele. The effect sizes for delayed non-carriers of the 7- repeat allele range from very small to non-existing. The current research result supports the results of previous studies on differential susceptibility. That is those children who are thought to be susceptible and are carriers of the 7-repeat allele benefit from this technology-enhanced educational program, while their presumably less-susceptible peers, non-carriers of the 7-repeat allele, do not.

In conclusion, it seems most plausible that *Living Letters* does not seem to fit the needs of literacy-delayed children, susceptible or non-susceptible. Please note that this does not exclude this program matching the needs of other subgroups not included in the current study (Merkelbach, Plak, van der Kooy-Hofland, Kegel, & Bus, under review). However, the program *Living Books* does support learning in carriers of the long variant of DRD4. These children may be particularly susceptible to *Living Books* because they have attentional problems: the animated stories are so engaging that they elicit a state of inattentional deafness to irrelevant stimuli from the environment (for example a noisy classroom). This may explain why carriers of the long variant benefit more from *Living Books* than they normally do from book reading while the program is not particularly effective for other children.

Limitations

We studied the role of one gene as a marker of differential susceptibility and more research is needed to obtain a good understanding of how the DRD4 gene interacts with other dopamine genes and the dopaminergic system in general. A single gene cannot, of course, be the exclusive cause of neurotransmitter levels in the brain and be responsible for a specific type of learning behavior (Kegel & Bus, 2012). The current findings suggest that the single dopamine-related gene DRD4 functions as a marker for differential susceptibility because it is a proxy for the dopaminergic system. The mechanisms that explain how the dopaminergic system interacts with the program are still unknown and need further research. A second limitation was that the qualities of *Living Letters* were insufficient to promote learning in the group of delayed learners. A strong element of this study is the number of participants. Taking into account that the current study is not correlational but a randomized controlled trial, the scale of this GxE study is substantial and well-powered (Van IJzendoorn et al., 2011). Longterm effects in the field of reading instruction are needed to show that the programs are indispensable and for whom.

Implications

We can conclude from the results of our study that brief educational computer programs can be effective specifically for a group of delayed presumably susceptible children: carriers of the DRD4 7-repeat allele. Without an additional literacy program, they lag behind in literacy skills, but they outperform their peers when they receive additional optimal instruction in a positive learning environment provided by a computer program. For their less susceptible peers, non-carriers of the 7-repeat allele, the brief additional programs have no effect probably because these children do not experience problems with on-task behavior. It is also possible that the non-carriers of the 7-repeat allele are in need of a more prolonged and intensive version of additional programs.

While the findings for *Living Letters* are mixed across experiments, *Living Books* yielded consistent positive effects. *Living Books* promotes basic language and literacy skills - comprehension and vocabulary - and may therefore best fit the needs of children lagging behind in an early stage of learning to read. The program *Living Letters* that targets basic alphabetic skills may not fit the needs of the most delayed group five-year-olds. We hypothesize that for the purpose of learning literacy skills, educational computer programs can even be more helpful than scaffolding by an adult (Takacs et al., 2015). Carriers of the 7-repeat allele can be considered as vulnerable since in a negative learning environment they may not thrive. A noisy and crowded classroom without personalized positive feedback by the teacher can be characterized as such a learning environment.

Carriers of the 7-repeat allele show their full potential when placed in a learning environment that helps them to engage in the task. Thus, the academic success of these presumably susceptible children can be enhanced if their susceptibility to the environment, for better and for worse, is acknowledged. When children, who are carriers of the 7-repeat allele, are offered a more suitable learning environment, they easily catch up with and even outperform their peers.

ACKNOWLEDGEMENTS

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

REFERENCES

Acevedo-Polakovich, I.D., Pugzles Lorch, E., & Richard, M. (2007). Comparing Television Use and Reading in Children With ADHD and Non-Referred Children Across Two Age Groups. *Media Psychology, 9*, 447-472. doi:10.1080/15213260701291387

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. (2008). Experimental evidence for differential susceptibility: dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. *Developmental Psychology, 44*, 293–300. doi: 10.1037/0012-1649.44.1.293

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). The hidden efficacy of interventions: Gene x environment experiments from a differential susceptibility perspective. *Annual Review of Psychology*, *66*, 381–409. doi:10.1146/annurev-psych-010814-015407

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. & 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

Bowyer-Crane, C., Snowling, M.J., Duff, F.J., Fieldsend, E., Carroll, J.M., Miles, J., Götz, K., Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry*, *49*, 422–432. doi:10.1111/j.1469-7610.2007.01849.x

Boyce, W.T., & Ellis B.J. (2005). Biological sensitivity to context: I. An evolutionary–developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology, 17*, 271–301. doi: 10.10170S0954579405050145

Bus, A. G., Takacs, Z. K., & Kegel, C. A. T. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. *Developmental Review*, *35*, 79-97. doi:10.1016/j.dr.2014.12.004

Cumming, G. (2014). The new statistics: Why and how. Psychological Science, *25*, 7–29. doi: 10.1177/0956797613504966

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

Dobbs D. 2009. The science of success. Atlantic 12,1–15.

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., & 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

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.

Maher, B. S., Marazita, M. L., Ferrell, R. E., & Vanyukov, M. M. (2002). Dopamine system genes and attention deficit hyperactivity disorder: a meta-analysis. *Psychiatric* Genetics, *12*, 207–215. Doi:10.1097/00041444-200212000-00003

Mayer, R. E. (2005). Principles for reducing extraneous processing in multimedia learning: coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning. New York, NY: Cambridge University Press.

Merkelbach, I., Plak, R.D., Van der Kooy-Hofland, V. A. C., & Bus, A.G.B. (under review). Mild perinatal adversities as a marker of differential susceptibility?

Molloy, K., Griffiths, T.D., Chait, M., Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. *The Journal of Neuroscience,9*, 35(49), 16046 –16054. doi:10.1523/JNEUROSCI.2931-15.2015

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

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

Saine, N. L., Lerkkanen, M. K., Ahonen, T., Tolvanen, A., & Lyytinen, H. (2011). Computer-assisted remedial reading intervention for school beginners at risk for reading disability. Child Development, 82, 1013–1028. doi:10.1111/j.1467-8624.2011.01580.x

Schaerlaekens, A., Kohnstamm, D., & Lejaegere M. (1999). Streeflijst woordenschat voor zesjarigen (3rd ed.). Lisse: Swets & Zeitlinger.

Schecklmann, M., Ehlisa, A., Plichtaa, M.M, Romanosb, J., Heinea, M., Boreatti-Hümmera, A., Jacoba, C., Fallgattera, A.J (2008). Diminished prefrontal oxygenation with normal and above-average verbal fluency performance in adult ADHD. Journal of Psychiatric Research, 43(2) 2008, 98–106. doi:10.1016/j.jpsychires.2008.02.005

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks. A meta-analysis. Review of Educational Research. doi:10.3102/0034654314566989

Twisk, J. W. R. (2006). Applied multilevel analysis. New York: Cambridge University Press.

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

Chapter 4

The Potential of Two Technology-enhanced Early Literacy Interventions to Prevent Reading Delays in First Grade

Submitted as:

Plak, R.D., Merkelbach, I., Kegel, C.A.T., & Bus, A. G. The Potential of Two Technology-enhanced Early Literacy Interventions to Prevent Reading Delays in First Grade.

ABSTRACT

The effect of educational computer programs as extra input for delayed children was examined. Children worked with the programs, a book reading program and a game-like program to stimulate phonemic awareness and alphabetic knowledge, in the last half year of kindergarten. On average children spent 2-3 hours with the programs, spread over 2-3 months. As indicators for effects we used standardized tests, halfway through and at the end of first grade. Neither of the programs caused main effects in the group of literacy-delayed children. Technology-enhanced books showed effects on text comprehension when assessed six months and a year after the intervention in a subsample with a genetic disposition to attention problems (carriers of the DRD4 gene 7-repeat allele). The digital books including animated pictures, sounds and music, seem to match specific needs of children with a genetic disposition for attention problems. We hypothesize that they elicit a state of hyperfocus, thus being more effective for these children than everyday experiences stimulating the same skills. As a substantial group (about one-third of all children) benefits from *Living Books* they seem to be indispensable in the kindergarten classroom.

Keywords: Technology-enhanced literacy interventions; digital books including animated pictures, sounds and music; Dopamine D4 receptor gene; differential susceptibility; hyperfocus; randomized controlled trial; long-term effects.

INTRODUCTION

The need for book reading is well established; book reading predicts children's later academic success (Bus, Van IJzendoorn, & Pellegrini, 1995; Scarborough & Dobrich, 1994). One aim of the current study was to test whether a new form of book exposure, the solo "reading" of technology-enhanced books, in addition to regular book reading experiences in the classroom and at home, helps to narrow gaps in a group of literacy-delayed children. The technology-enhanced books may result in children who are lagging behind in literacy skills extending their book reading experiences and thereby their emergent literacy skills. The books may also respond to specific needs and speed up literacy learning. These technology-enhanced books are known to offer guidance in understanding the story text by adding extra nonverbal information that matches the story text and thus help to concretize the text. Examples of such additions are animated illustrations, zooming, sounds, and background music (Bus, Takacs, & Kegel, 2015; Takacs, Swart, & Bus, 2015). *Living Books* are formatted as an attractive animated movie closely matching the story text. The nonverbal additions may, if well designed, replace cognitive guidance offered by an adult because the additional information directs attention to details in pictures, just as adults do while sharing a story with a child (Takacs, Swart, & Bus, 2014). The film-like presentation of the story may also make the books more compelling and engaging and thereby increase attentional arousal during the reading sessions (Verhallen & Bus, 2009). One of the aims of the current study was to test whether exposure to *Living Books*, in addition to everyday book reading experiences at home and school, contributes to children's emergent reading development and to specify which skills are in particular influenced.

From previous studies targeting children delayed in literacy, we deduced that *Living Books* could be a boost to emergent reading skills of this group (Plak, Kegel, & Bus, 2015; Plak, Merkelbach, Kegel, van IJzendoorn, & Bus, under review). From assessment with a standardized test administered immediately after an intervention with *Living Books*, we concluded that technology-enhanced books increase children's literacy-related knowledge but not that of *all* young children. We found that the books were particularly effective for five-year-old carriers of the Dopamine D4 Receptor Gene with the polymorphism DRD4 7-repeat allele – a subsample, about one-third of the total population, that is known to be vulnerable for attention problems (see for an overview Bakermans-Kranenburg & Van IJzendoorn, 2011). This recurrent finding is in line with the Differential Susceptibility Theory (Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007; Ellis et al., 2011): Some children are susceptible for specific qualities of educational input, in this instance the *Living Books*, due to their typical way of processing information.

This particular subsample of children did benefit from the technology-enhanced series of books, but not from another educational computer program that is also intended to stimulate early literacy skills. This educational computer program, *Living Letters*, was composed of a series of highly structured game-like assignments in which phonemic awareness and alphabetic knowledge were trained. Children received continuous feedback from two tutors, thus guiding them through every assignment. We neither found main effects of *Living Letters* nor effects in the subsample of carriers of the DRD4 polymorphism (Plak et al., 2015; Plak et al., 2016).

We speculate that carriers of the DRD4 7-repeat allele typically show diminished dopamine reception efficiency, often resulting in reduced attention and reward mechanisms (Robbins & Everitt, 1999). If, however, activities are satisfactory, there will be an increased release of dopamine in the ventral striatum, creating a route for learned reinforcing; that is, after being challenged and achieving a satisfactory result, dopamine becomes available thereby enabling new achievements (Koepp et al., 1998). Unlike traditional book sharing, *Living Books* may elicit this route in carriers of the DRD4-7 repeat allele and, like is reported about persons with ADHD, they may enter stages of very high levels of concentration (Maher, Marazita, Ferrel, & Vanykukov, 2002). They may show a tendency to focus very intently on things that do interest them (Schecklmann et al., 2008) - watching films or playing (video) games – and become oblivious to the world around them. They enter a state of hyperfocus blocking all other stimuli coming from the environment, and engage for hours in activities like playing games or watching movies.

In line with this argument, we hypothesize that carriers of the DRD4 polymorphism benefit more from *Living Books* than other children because this program may elicit a state of very intensive on-task behavior as only occurs in children with attention problems. The film-like presentation in *Living Books* may attract all attention and result in a state of intensive concentration. The books may even lead to diminished processing of irrelevant auditory stimuli from the surroundings resulting in a state of *inattentional deafness* – the failure to notice auditory stimuli under high visual perceptual load as in *Living Books* (Molloy, Griffiths, Chait, & Lavie, 2015).

Furthermore, it seems plausible that in a subsample of carriers of the DRD4 polymorphism, literacy development can be intensified through the program *Living Books* but not through *Living Letters*. The latter program may lack elements that engage five-year-olds' attention to the same extent as *Living Books*. Instruction films -as are characteristic of *Living Letters* and the game-like elements of the program- might be engaging to younger children, which may explain the positive effects of this program in a group of four-year-olds (Kegel, Bus, & van IJzendoorn, 2011): Carriers of the DRD4-7 repeat allele benefited from *Living Letters* while non-carriers did not. We noticed that five-year-olds had a different response to the program: they were not enthusiastic about the instruction films and game-like elements of the *Living Letters* program. Teachers complained that five-year-olds' motivation for *Living Letters*

decreased after a few sessions probably due to the repetitive character of both the instruction films and the assignments. In line with our reasoning, we therefore do not expect that playing with *Living Letters* in kindergarten will facilitate learning to read in first grade.

Living Books and *Living Letters*, each target different literacy skills; *Living Letters* focuses on inside-out skills (e.g., letter knowledge and phoneme awareness), while *Living Books* focuses on outside-in skills (e.g., vocabulary and story comprehension). Both inside-out skills and outside-in skills are fundamental to learning to read in first grade and thereafter (Whitehurst & Lonigan, 2001). In the current experiment, we expect the strongest long-term effects of *Living Books* on tests that relate to reading comprehension: vocabulary and listening comprehension. It seems less plausible that the program affects tests associated with technical reading skills: word reading and spelling. As we did not find short-term effects of *Living Letters*, neither in the complete group nor in a subsample, we do not expect long-term effects of this program on tests associated with technical reading skills.

This study

In the current study, all literacy-delayed children included participated in one of the two RCTs (Plak et al., 2015; Plak et al., 2016) in which the two computer interventions, *Living Letters* and *Living Books*, were included. The two exactly similarly designed experiments were carried out during the winter/spring semester in two consecutive years (2012-2013 and 2013-2014). Since in those experiments effects were only found in a group of literacy-delayed pupils, in the current study only literacy-delayed pupils were included. Standardized tests for literacy administered by teachers in January and June in the first grade served as posttests. We expected to find that *Living Books* would be an extra boost for the reading development of delayed children and in particular carriers of the DRD4-7 repeat allele. There may be direct effects of a technology-enhanced book reading program on children's listening comprehension and vocabulary, as well as indirect effects of extra book reading in kindergarten: positive experiences with the *Living Books* program may increase interest in stories, which may start off a snowball effect. Due to an increased interest in book reading, children elicit more reading from adults and thus become more proficient in reading skills (Mol & Bus, 2011).

We expect effects on outside-in skills like comprehension rather than on inside-out skills, such as word recognition and spelling, the technical reading skills. The current long-term study aims at proving that *Living Books* are an important supplement to the regular kindergarten curriculum because they compensate for failure to benefit from traditional book reading at home and school especially when children have a greater risk of attention problems. We hypothesize that technology-enhanced books may be indispensable for these vulnerable children. Due to a state of hyperfocus elicited by the multimedia stories, they gain more in-depth knowledge from reading technology-enhanced books than from regular book reading. They may even outperform their peers, not only on the short run but also on the long run. In this study, we present long-term effects of reading technology-enhanced books six months and a year after the intervention took place in kindergarten.

METHOD

Participants

From August 2012 to October 2013 and August 2013 to October 2013 the project was widely advertised via e-mail, mail, social media, and phone in the Netherlands. A total of 183 schools from all over the country responded, 2.7% of all Dutch schools. The participating schools are from urban as well as rural areas. Kindergarten teachers from the participating schools selected eligible kindergarten children between October 2012 and February 2014 and October 2013 and February 2014. The eligible kindergarten children had for example difficulty writing their proper name, rhyming, naming a few letters, and identifying sounds in words. Preferably pupils scored 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). Parents gave written informed consent for participation that also included consent to collect long-term data.

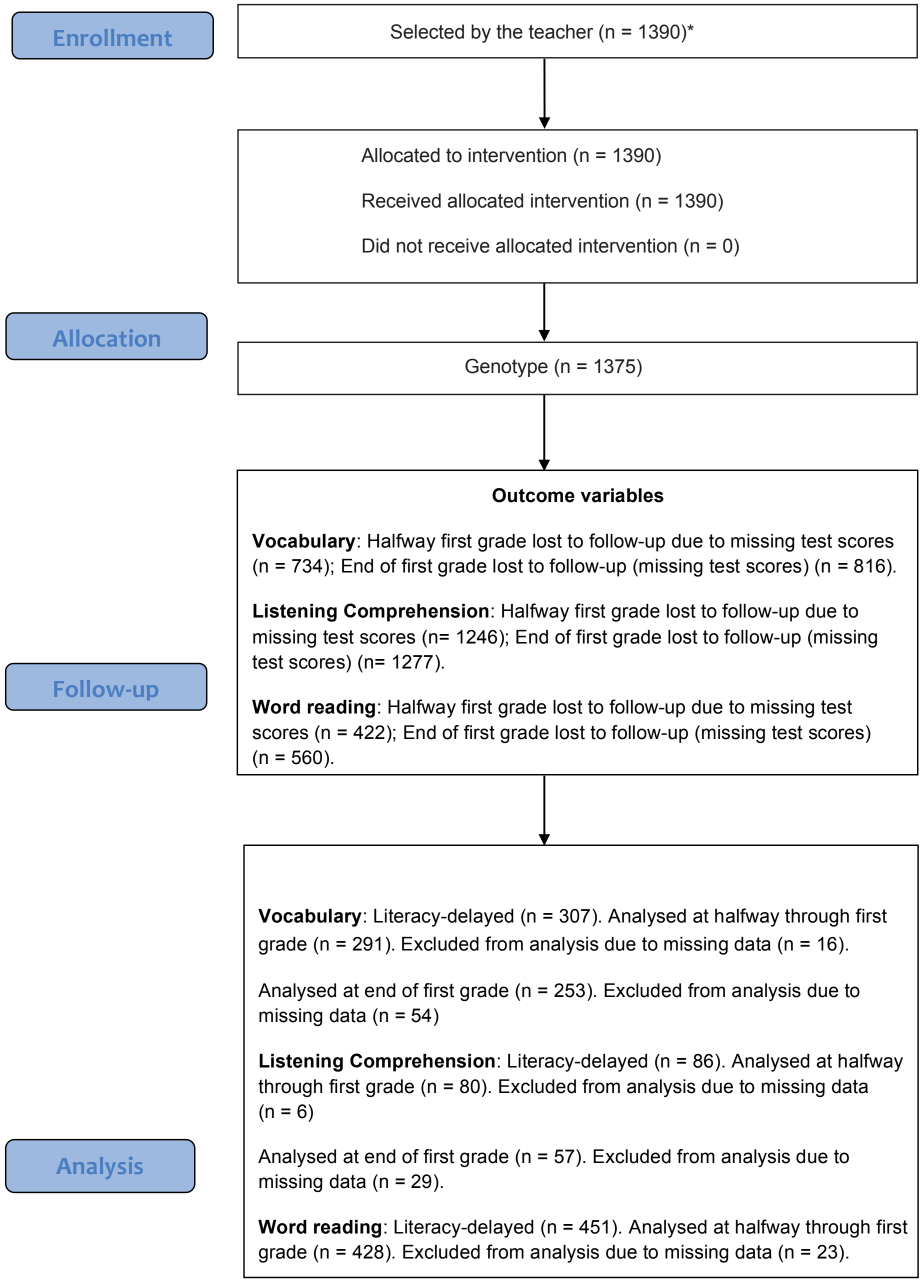
Design

After receiving informed consent to participate from parents, we randomly assigned children to one of three conditions: a control program (*Clever Together*), *Living Books* or *Living Letters*. In each class, one child was assigned to *Clever Together*, the control condition, and at least one child to one of the literacy-related programs, that is *Living Letters* or *Living Books*. The intervention, a randomized control trial, took place in the second half, between March and June, of the senior kindergarten year. Children in the *Living Books* condition were offered 8 sessions with 2 books per session and children in the *Clever Together* and *Living Letters* condition ranging between 8 and 11 sessions (the variable number of sessions depended on the number of errors children made). Each session took about 15 minutes and children practiced once a week. Based on the data that were stored by the program, children completed on average 33.62 out of 34 *Living Letters* games (*SD* = 2.50) and they "read" on average

14.8 out of 16 books (*SD* = 1.8). Immediately following the intervention period of twelve weeks, teachers administered the Central Institute for Test Development (Cito) Literacy Test for Kindergarten (CLT), a standardized literacy test (Lansink & Hemker, 2012). These short-term results are discussed in Plak et al., 2015, and Plak et al., 2016. This study aims at answering the pressing issue of whether additional computer programs in kindergarten targeting early literacy skills facilitate learning to read in first grade. To this end, we collected the results of four Cito tests - Vocabulary, Cito Listening Comprehension, Word Reading and Spelling - administered halfway through first grade and at the end of first grade, see Figure 1.
Buccal cell samples were collected halfway during the intervention period. Trained members of the research team visited the schools to collect the samples using a sterile swab designed for collecting buccal cells for DNA analysis (Omni Swabs, Whatman/GE Healthcare, UK). Subsequently the samples were sent to a commercial laboratory for DNA analysis.

Procedure

Directly after the administration of the tests, the contact persons received a request to send the scores of the literacy tests to the researchers. This was done by e-mail, phone, and mail, depending on the contact with the school. The number of times the request was repeated could range from 1 to 12 times. In the Netherlands, switching schools as a teacher is quite common, as a result of which we lost many contact persons when we returned to the schools to collect the long-term data. In those cases we contacted their colleagues, who however did not always respond to our request to send the test scores. Moreover, as Dutch schools are free to decide which tests from a nation-wide student tracking system are administered, not all schools administered all tests of interest. Most schools did administer Word Reading followed by Vocabulary and Spelling. The least popular was Listening Comprehension.



Note: *The teacher selected the pupils, therefore we do not have insight into how many students were eligible to begin with.

Figure 1
CONSORT diagram

Intervention programs

Living Books. The intervention program *Living Books*, was made up of eight age-appropriate digital animated storybooks. The animated pictures, sounds, and music support children's understanding of story events and language (Takacs et al., 2015). When the oral narrative is accompanied by nonverbal information, and both information sources are simultaneously available, the narrative text will be understood and retained better than if conveyed by words alone (Bus et al., 2015). Multimedia books offer optimal guidance in developing mental representations of the story and the language. The movie-like visual representation promotes engagement.

Each reading of a book was interrupted four times for questions about the story and vocabulary. Feedback to the child's response was adaptive, similar to *Living Letters*. The first error was followed 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 four new questions were included in each session.

Living Letters. The second intervention program *Living Letters* 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). The proper name is often the first word that young children can read and write; therefore the child's name is used to draw attention to phonemes in the spoken name and words (Van der Kooy-Hofland & Bus, 2012). The program adapts automatically to the child's proper name when available in the name database. If not, the program uses 'mama' (mommy) - a word that is just as familiar to many young children. A tutor providing feedback followed up on every response. When children answered a question incorrectly, feedback and clues were provided. More specifically, after the first error the assignment was only repeated: "Listen carefully, in which word do you hear /t/ from 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 on a positive note, irrespective of whether a correct response was given, whereupon a new game started. When children failed to give the correct answer at the first try, the assignment was repeated twice in a subsequent session, which explains why some children had a few more sessions than others. In the first games, children practiced how their name (or "mama") is written, followed by games to train the sound of the first letter of the child's name (or "mama") and thereafter by games to identify pictures that start or end with the first letter of the target name.

Clever Together supports basic concepts for mathematics like practicing cardinals and visual-spatial reasoning. It includes 40 games. As in *Living Letters* and *Living Books*, a tutor provides constructive, detailed feedback for every error and every correct response. Assignments were repeated in later sessions when children made errors.

Literacy measures

Cito Literacy Test for Kindergarten Pupils [Taal voor Kleuters, groep 2]. To identify the literacy-delayed pupils we used the Cito Literacy Test for Kindergarten Pupils (CLT), a standardized literacy test for kindergarten pupils. The 60-itemed CLT is administered group-wise in January ($\alpha = .89$) of the senior kindergarten year. Vocabulary, text comprehension, rhyming, hearing the first and last word, sound blending, writing conventions (e.g. reading from left to right), and prediction of book content based on the book cover are assessed (Lansink & Hemker, 2012). The Committee for Test Quality in the Netherlands [Commissie Testaangelegenheden Nederland] evaluated the CLT as adequate.

Cito Vocabulary, posttest, first grade [Woordenschat, groep 3]. With Cito Vocabulary for the first grade, the level of receptive vocabulary can be determined. The tests for the first grade consist of 50 assignments in January and at the end of the school year (June). The tests are administered group-wise in class. The students receive classroom instruction and a number of exercises on paper. In first grade the teacher provides the questions orally. For example: “Lena looks at the *procession*, where do you see a *procession*?” The students make a choice out of three pictures. Cronbach’s alpha is .81 for grade 1 January and .83 for grade 1 June (Van Berkel et al., 2010). Based on the criteria of the Committee for Test Quality in the Netherlands [Commissie Testaangelegenheden Nederland] the reliability of the Cito Vocabulary test is considered good.

Cito Listening Comprehension, posttest, first grade [Begrijpend Luisteren, groep 3]. The test determines the level of listening comprehension. The test measures proficiency in giving meaning to spoken language (Krom, Ouborg, & Kamphuis, 2001). For the first grade two versions are available, one for halfway through first grade and one for the end of first grade. Different stories are played from a CD. After each story, children answer multiple-choice questions about the story presented in a booklet. Answers take the form of pictures since some students still have trouble reading when they are in first grade. Based on the criteria of the Committee for Test Quality in the Netherlands [Commissie Testaangelegenheden Nederland] the reliability of the Cito Listening Comprehension test is considered adequate.

Cito Word Reading, posttest, first grade [Drie-Minuten-Toets, groep 3]. The Three-Minute Test (TMT) measures the word reading speed of a child and is administered individually by the teacher. The student is given one minute to read aloud

as many words from a reading chart. The words must be read by column, from top to bottom. While reading, the administrator scores which words are read incorrectly. In general, when pronunciation does not match the spelling this is considered to be an error. Stress differences are not included. Spelling out words (/ s / - / a / - / p /) is incorrect, except if spelling is followed by / sap [juice]/. When a pupil spontaneously corrects him or herself this is considered correct. Raw test scores are the number of correct words. This score can be converted into a standard score, which can be converted, in turn, in a skill level. Cronbach’s alpha is calculated for the combination of TMT cards, categorized in level 1 (kindergarten junior year and senior year) and level 2 (kindergarten junior year, kindergarten senior year and first grade). For the first grade January test Cronbach’s alpha is .964 (level 1, level 2 is not applicable), for the first grade June test Cronbach’s alpha is .967 for level 1 and .971 for level 2. Based on the criteria of the Committee for Test Quality in the Netherlands [Commissie Testaangelegenheden Nederland] the reliability can be considered as adequate (Krom, Jongen, Verhelst, Kamphuis, & Kleintjes, 2010).

Cito Spelling, posttest First Grade [Spelling voor groep 3]. With the Cito Spelling test a child’s ability to spell words can be established. This test is administered group-wise. During the administration of the Cito Spelling in the first grade January (M3) and June (E3) tests, words from different levels of complexity are dictated (for example words with 1 to 3 syllables; words with an intermediate sound that is not written, et cetera).

Genetic screening for DRD4 polymorphisms

An assortment of characteristics inherited from the parents is a genotype. Genotyping is the process of determining differences in the genetic make-up (genotype) of an individual by examining the individual’s DNA sequence. Biological assays are used and compared to another individual’s DNA sequence. The details are explained below.

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 100ng genomic DNA template, 10pmol 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 30µl using the following cycling conditions: initial denaturation step of 10 min at 95 ºC, followed by 27 cycles of 30 sec 95 ºC, 30 sec 60 ºC, 60 sec 65 ºC and a final extension step of 10 min 65 ºC.

Analysis of PCR products for repeat number. One µl of PCR product was mixed with 0.3µl LIZ-500 size standard (Applied Biosystems) and 11.7µl 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 allele at one or both alleles. Of the 593 participants, ten children could not be genotyped; 199 children (34%) were carriers of the 7-repeat allele of DRD4. The distribution of DRD4 polymorphisms was in Hardy-Weinberg equilibrium, $\chi 2$ (df = 1, N = 428) = .069, p = .793.

Data Analysis

In the literacy-delayed group (scoring below the 40th percentile according to national norms), we regressed the posttest scores including vocabulary, listening comprehension, word reading, and spelling on Vocabulary, Listening Comprehension and Word Reading, *Living Letters* (contrast between the control condition *Clever Together* and *Living Letters*), *Living Books* (contrast between *Clever Together* and *Living Books*), DRD4 (7-repeat allele vs. others), two-way interactions involving interventions, and DRD4. Age, sex, and father's education were entered as covariates.

RESULTS

Characteristics of the sample

Table 1 shows the characteristics of the 428 participants. Participants at pretest (halfway through senior kindergarten year) had a mean age of 67.00 (SD = 4.30). The mean score for father's education was 3.64 (SD = 1.69) on a scale ranging from 0 – 6,with 0 representing primary school and 6 representing university-level education. Of the participants 57% was male (n = 244) and 43% female (n = 184). Across the three experimental conditions *Living Letters* (88 males/64 females), *Living Books* (88 males/64 females) and *Clever Together* (68 males/56 females), there was no significant difference between the number of children with a DRD4 7-repeat allele: *Living Letters* (35.5%), *Living Books* (35.5%), and *Clever Together* (29.0%), χ^2 (df = 2, N = 428) = .939, p = .62.

Table 1: *Characteristics of the experimental conditions Living Letters, Living Books, and Clever Together of participants scoring below 40th percentile (literacy-delayed)*

	Complete group	*Living Letters*	*Living Books*	*Clever Together*
Age months (SD)	67.00 (4.30)	67.35 (4.33)	66.85 (4.35)	66.75 (4.21)
CLT pretest (SD)	53.82 (4.38)	53.74 (4.89)	54.34 (3.89)	53.29 (4.24)
Male/Female	244/184	88/64	88/64	68/56
DRD4 7+/7-R (%)	37.5/64.3	38.2/61.8	32.9/67.1	36.3/63.7

We first checked per dependent measure whether scores of missing cases differed from the scores included in the analyses. The group with missing data for Word Reading January/June and Spelling January/June differed significantly from the group with complete data on the pretest. The scores of the group with complete data were somewhat higher than scores in the missing group. For the complete group the mean for Vocabulary/ January was 60.11(SD = 8.42) vs. 58.10 (SD = 9.51) for the missing group. The complete group had a mean of 59.99 (SD = 8.01) for Vocabulary/ June, while the missing group had a mean of 58.81 (SD = 9.85). The complete group mean for Spelling/ January was 60.45 (SD=8.81), and for the missing group 58.55 (SD = 8.69). Finally, the complete group mean for Spelling/ June was 60.41 (SD = 8.48), while for the missing group it was 58.82 (SD = 8.98). Since we expected the effects of the programs to be stronger the more children were delayed, the outcomes of the current experiment may therefore underestimate the effects of the programs. Furthermore, the group with missing scores was somewhat younger.

Multilevel analysis

To test whether or not it is necessary to allow the intercepts and slopes to differ between schools (Twisk, 2006), we tested the difference between the -2log likelihood of the model with a random intercept and the -2log likelihood of the model without a random intercept and the difference between the -2log likelihood of the model with a random intercept and slope and the -2log likelihood of the model with only a random intercept, for every outcome variable. Adding a random intercept, the differences were highly significant (chi-square distributions with one degree of freedom $\geq$ 20.02. However, the differences between the -2log likelihood of the model with a random intercept and the -2log likelihood of the model with both a random intercept and a random slope were not significant ($\chi^2 \leq 5.99$). Therefore a multilevel analysis with a random intercept for schools was applied (Luke, 2004).

The variance in scores halfway through and at the end of first grade attributable to school characteristics was calculated by dividing random variance related to school by all random variance (variance related to school and child). As may be expected, intraclass correlations were higher for Vocabulary January/June (27%/ 31%) and Listening Comprehension January/June (26%/ 25%) than for Word Reading January/June (8%/ 4%) and Spelling January/June (17%/ 15%); see Table 2 for estimates and significance levels. Apparently, home and school input is more diverse for comprehension skills resulting in higher intraclass correlations.

Table 2: *Outcomes of Multilevel Analysis Using Posttest Vocabulary, Comprehension, Word Reading, and Spelling Halfway of First Grade as an Outcome Variable of participants scoring below 40st percentile (literacy-delayed)*

	Outside-in skills				Inside-out skills			
Fixed effects	Vocabulary Jan	Vocabulary June	Comprehension Jan	Comprehension June	Word reading Jan	Word Reading June	Spelling Jan	Spelling June
Intercept	46.79 (12.33)*	72.65 (14.81)*	62.60 (12.36)*	53.96 (17.77)*	31.57 (7.72)*	54.00 (12.51)*	126.11 (6.08)*	121.88 (5.88)*
Background								
Age	-.32 (.18)**	-.49 (.21)*	-.38 (.18)*	-.08 (.26)	-.20 (.11)**	-.40 (.18)*	-.31 (.09)*	-.17 (.08)*
Sex	-1.56 (1.50)	-.03 (1.82)	.25 (1.59)	-.79 (2.09)	1.19 (.97)	.83 (1.58)	.55 (.75)	1.10 (.71)
Fathers educational level	.15 (.41)	.13 (.48)	.15 (.40)	-.71 (.54)	-.28 (.29)	.18 (.43)	-.10 (.21)	.13 (.19)
Main Effects								
Living Letters (vs. control)	1.20 (1.35)	1.99 (1.62)	-.47 (1.38)	-1.71 (1.93)	-1.36 (.84)	-2.59 (1.38)**	-1.17 (.68)**	-.16 (.65)
Living Books (vs. control)	-1.69 (1.27)	-3.34 (1.51)*	-1.47 (1.30)	.79 (1.69)	.00 (.81)	-.74 (1.32)	.22 (.63)	.18 (.60)
DRD4 variant	4.07 (1.56)*	3.80 (1.91)*	4.63 (1.61)*	7.22 (2.34)*	.11 (1.00)	1.24 (1.63)	-.18 (.77)	.28 (.73)
Interaction Effects								
DRD4 variant X *Living Letters*	-2.20 (2.18)	-2.16 (2.64)	-2.37 (2.35)	.68 (3.10)	.29 (1.38)	1.77 (2.28)	1.20 (1.10)	-.23 (1.04)
DRD4 variant X *Living Books*	4.87 (2.23)*	5.58 (2.70)*	5.08 (2.39)*	5.27 (3.31)	.51 (1.40)	.75 (2.27)	.82 (1.07)	.74 (1.00)
Random Effects				Est. (SE) Wald Z				
Variance								
Level Child	138.03 (12.95) 10.66*	173.31 (17.43) 9.94*	41.33 (7.37) 5.61*	54.97 (11.91) 4.62*	90.90 (6.82) 13.33*	209.90 (16.58) 12.66*	37.34 (3.50) 10.67*	28.80 (2.91) 9.91*
Level School	52.25 (15.02) 3.48*	78.53 (22.39) 3.51*	14.69 (9.30) 1.58	18.09 (14.13) 1.28	7.48 (3.87) 1.93**	8.58 (6.79) 1.26	7.79 (3.38) 2.31*	5.09 (2.51) 2.03*

Note: * $p<.05$, ** $p<.10$

Outside-in skills

Outcomes were similar for vocabulary and listening comprehension. There were no main effects of the intervention programs with one exception: *Living Books* had a significantly lower score on Vocabulary in June (Est = -3.34, p = .028). There were main effects of DRD4 indicating that on all tests the DRD4-7R subsample outperformed their peers (Est≥ 3.80, p = .028), probably due to the strong effects of *Living Books* in this subsample. With one exception (Listening Comprehension in June) we found significant interactions between genotyping and *Living Books* (Est ≥ 4.87). Overall, carriers of the 7-repeat allele of the DRD4 gene benefited from *Living Books* while the non-carriers did not. Effect sizes in the sub-sample of carriers ranged from .35 (Listening Comprehension/January) to 1.38 (Listening Comprehension/June), which indicates that effect sizes were moderate to large. The non-carriers tended to benefit more from *Clever Together*, as is indicated by negative *d*'s (-.29 for Vocabulary/ January; -.40 for Vocabulary/June; -.36 for Listening Comprehension/January). *Living Letters* did not reveal any differential effect.

Does the intervention foster a reduction in the number of low achievers particularly in carriers of the 7-repeat allele of the DRD4 gene? Among carriers were fewer children that scored in the lowest ranges, as compared to the non-carriers. For Vocabulary/January scores of non-carriers and carriers in the lowest ranges were 38.2% and 32.4%, respectively, for Vocabulary/June 33.1% and 26.3%,and for Listening Comprehension/January 9.9% and 6.5%.

Inside-out skills

Living Letters approached significant effects on word reading in June (Est = -2.59, p = .062) and spelling in January (Est. = -1.17, p = .084). The negative estimates indicate, contrary to what we expected, higher scores for the control condition, *Clever Together*. No other main effects of intervention programs or DRD4 were found. We also found no interactions between DRD4 and intervention programs.

Table 3: *Adjusted Means and Standard Deviations for Vocabulary Halfway through First Grade and End of First Grade, and Listening Comprehension Halfway through First Grade and End of First Grade in the Delayed Group by Condition and DRD4*

		Total group	*n*	*Living Letters*	*n*	*Living Books*	*n*	*Clever Together*	*n*
Jan Vocabulary	DRD4 (7-)	23.81(14.27)	189	24.52 (14.28)	57	21.67 (9.75)	67	25.23 (14.27)	65
	DRD4 (7+)	28.53 (14.32)	102	28.14 (14.24)	39	31.68 (14.22)	33	25.76 (14.22)	30
	Total	26.17 (14.98)	291	26.33 (14.48)	96	26.68 (15.13)	100	25.49 (15.33)	95
June Vocabulary	DRD4 (7-)	39.12 (16.49)	167	41.11 (16.51)	49	34.81 (16.43)	58	41.42 (16.46)	60
	DRD4 (7+)	43.38 (16.54)	86	44.26 (16.42)	33	46.19 (16.40)	28	39.70 (16.42)	25
	Total	41.25 (17.40)	253	42.68 (16.73)	82	40.50 (17.51)	86	40.56 (18.00)	85
Jan Listening Comprehension	DRD4 (7-)	37.67 (7.91)	52	37.21 (7.99)	17	36.49 (7.89)	19	39.30 (7.872)	16
	DRD4 (7+)	42.45 (8.05)	28	40.30 (7.95)	9	44.92 (8.05)	8	42.11 (7.88)	11
	Total	40.06 (8.27)	80	38.76 (8.30)	26	40.71 (8.66)	27	40.71 (8.00)	27
June Listening Comprehension	DRD4 (7-)	44.28 (9.41)	39	43.78 (9.34)	12	44.71 (9.34)	16	44.35 (9.38)	11
	DRD4 (7+)	51.50 (9.68)	18	51.00 (9.31)	8	58.20 (9.36)	5	45.30 (9.28)	5
	Total	47.89 (10.15)	57	47.39 (9.45)	20	51.45 (10.82)	21	44.82 (9.93)	16
Jan Word Reading	DRD4 (7-)	17.80 (10.13)	275	16.27 (10.11)	94	17.84 (10.06)	102	19.28 (10.06)	79
	DRD4 (7+)	17.89 (10.14)	153	16.93 (10.06)	58	18.33 (10.09)	50	18.40 (10.08)	45
	Total	17.84 (10.55)	428	16.60 (10.37)	152	18.09 (10.71)	152	18.84 (10.47)	124
June Word Reading	DRD4 (7-)	28.67 (15.06)	235	26.01 (15.08)	77	27.97 (15.01)	88	32.02 (15.01)	70
	DRD4 (7+)	29.840 (15.05)	128	29.27 (15.03)	47	29.71 (14.99)	43	30.55 (15.02)	38
	Total	29.25 (15.76)	363	27.64 (15.49)	124	28.84 (15.98)	131	31.28 (15.71)	108
Jan Spelling	DRD4 (7-)	105.11 (6.81)	198	104.02 (6.86)	60	105.25 (6.78)	75	106.07 (6.78)	63
	DRD4 (7+)	104.76 (6.78)	114	104.79 (6.77)	38	106.09 (6.78)	39	103.39 (6.78)	37
	Total	104.93 (7.03)	312	104.41 (6.95)	98	105.67 (7.13)	114	104.73 (7.02)	100
June Spelling	DRD4 (7-)	111.82 (5.94)	167	111.74 (5.97)	48	111.89 (5.91)	65	111.84 (5.91)	54
	DRD4 (7+)	112.01 (5.94)	98	111.75 (5.91)	34	113.07 (5.91)	36	111.22 (5.90)	28
	Total	111.92 (6.14)	265	111.75 (5.99)	82	112.48 (6.16)	101	111.53 (6.22)	82

DISCUSSION

In this randomized controlled trial we proved that a short technology-enhanced educational computer program can be important, as appears from literacy skills, not just directly after the intervention but also in the long term. In first grade, a *subgroup* of children still benefits from being exposed to a set of *Living Books* in kindergarten even though the intervention was rather brief and did not encompass more than 2 to 3 hours. An effect across the board was not found for *Living Books*, but this digital book program *was* effective in a subsample, with effect sizes ranging from .35 to .42.

These long-term results showed that *Living Books* is particularly effective for developing outside-in skills. Effects were found on vocabulary (halfway through first grade and at the end of first grade) and on listening comprehension, specifically halfway through first grade. The effect on listening comprehension was substantial a year after the intervention was implemented, but the small number of participants might have played a part in the failure to find a significant effect for this point of measurement. On inside-out skills, no effects of *Living Books* were found, which makes sense since the books did not include print or other incentives for the development of such skills. Therefore it was not likely that children would develop inside-out skills. In contrast to *Living Books, Living Letters* did not show significant effects on inside-out or outside-in skills. Considering the earlier lack of findings for *Living Letters* (see Plak et al., 2015 and Plak et al., 2016), it may come as no surprise that no effects were found six months or a year after the intervention was implemented.

Findings suggest that *Living Books* are not effective for all children but that a specific group of children, carriers of the DRD4-7 repeat allele, *does* benefit from being exposed to a series of animated books in addition to traditional book sharing at home and school (Plak et al., 2015; Plak et al., 2016) and that those effects are still measurable six months and a year later. When carriers of the 7-repeat allele of the DRD4 gene are exposed to *Living Books*, even when it is only for a brief period, they catch up and – as has been proven in the current study – even outperform their non-carrier peers on vocabulary and listening comprehension. Non-carriers do not benefit from *Living Books* probably because they maximally benefit from daily book reading sessions, in school and at home, and the eight extra *Living Books*, even when those books include nonverbal support for story comprehension, represent a relatively small expansion of the normal book reading diet. For the majority of kindergarten children, *Living Books* is more of the same and therefore does not result in any identifiable effects. The close to significant negative effects suggest that in this group the control program, *Clever Together*, tends to provide a more unique contribution to listening comprehension and vocabulary as compared to *Living Books*.

Nonverbal information in *Living Books* such as animations, background sounds and music increase children's comprehension, especially when children are at risk for language delays (Takacs, Swart, & Bus, 2015). *Living Books* are designed in such a way that the same information is presented simultaneously through words and nonverbal information thus enabling that *multimedia learning* can take place ; (Mayer, 2005). Maybe due to the film-like presentations, *Living Books* may incite a state of deep concentration or hyperfocus in this specific group. That is, they focus on what is presented in the computer program, thereby blocking irrelevant stimuli coming from the environment. Actually, findings corroborate the hypothesis that digital storybooks are much more effective for children with a genetic disposition to attentional problems than regular book sharing with adults. Where a teacher or a parent does not have the means to incite deep engagement and thereby high achievement, the technology-enhanced computer program *Living Books* does.

So far we do not have direct evidence for the hypothesized differences in processing information. Further testing of responses to the sessions is needed. Are they, as a state of hyperfocus implies, indeed "deaf " (Molloy et al., 2015) to sounds coming from the environment while they "read" the books? *Living Books* meet the criteria of a task with a high visual perceptual load, making it probable that when children -who are highly susceptible to their learning environment- work with the program, experience "deafness"; they enter a state of hyperfocus blocking irrelevant sounds from their surroundings, for example sounds from a noisy classroom.

Furthermore, we do not have information to explain the enduring effects. We hypothesize that a brief period of exposure to multimedia books raises these children's interest in books, which may explain why they still outperform their peers in vocabulary and listening skills a year after exposure to the books. However, this hypothesis awaits further testing. The program *Living Letters* is not a source of additional input for carriers of the 7-repeat allele of the DRD4 gene, probably because it lacks characteristics that elicit a state of hyperfocus.

Limitations

The longitudinal design of the current randomized controlled trial has several advantages, such as the opportunity to determine the stability of the effects of technology-enhanced educational computer programs and contributing to the strengths of the current study. A disadvantage of the longitudinal character is the dependence on the willingness of the participants to participate over longer periods of time. In this study, we relied on the schools to send the scores of the various CLT tests administrated in first grade. Because of large turnovers of staff - typical of elementary schools in the Netherlands - the researchers often lost the contact person within the school who was responsible for sending the scores. Data were missing relatively often from schools with many children from families with low incomes and children from various ethnic backgrounds, probably due to organizational and social problems within these schools. From the comparison between groups with complete and missing data, it seems that in so far there are differences, the difference is mostly found in the pretest score. The children with missing data had a somewhat lower pretest score, and thus belong to the sample that benefits most from the intervention. However, even without those low-scoring pupils, the effects were substantial. Moreover, the findings are in line with the short-term outcomes in two previous RCTs (Plak et al., 2015; Plak et al., 2016).

Summary and implications

In the current longitudinal study, we proved that a brief technology-enhanced program can be vital for a subgroup. *Living Books* gave a boost to story and word comprehension in a subgroup encompassing about one-third of all children. Carriers of the 7-repeat allele of the DRD4 gene fall behind in vocabulary and listening comprehension when they are not exposed to the books, but catch up and outperform non-carriers when they had a chance to work with *Living Books* in kindergarten, even six months and a year after the intervention. The non-carriers did not benefit from an extra program, in addition to the traditional book reading experiences, because they benefit from regular book reading experiences and *Living Books* does not constitute a substantial addition. This suggests that computer programs with particular features, even when they are only temporarily available for a short period of time, can be vital incentives for children with attention problems. Another educational program, *Living Letters*, as a result of the lack of characteristics that engage children with attention problems, showed no effects; neither carriers nor non-carriers of the 7-repeat allele of the DRD4 gene benefited from *Living Letters*. We suspect that this learning environment lacked elements that would make it more stimulating for the susceptible group in the current study because it lacks elements that would make this learning environment more stimulating than what children encounter in their everyday environment.

The results of the current study show that with a brief technology-enhanced early intervention in literacy education good long-term results can be achieved. Some children have a hard time keeping up with their peers, but when exposed to an enriched learning environment as offered by *Living Books*, even for a short period of time (intervention duration ranged between 160 and 220 minutes for a period of 12 weeks), they seize the opportunity and benefit from the extra guidance that they are offered, even a year later. The results support the hypothesis that technology-enhanced programs that include film-like components, and probably also game-like elements, can elicit a state of hyperfocus in subgroups that strengthen their learning attempts and results in outperforming their peers. Traditional forms of education do not offer teachers many means for getting children, who are predisposed to

attention problems, to hyperfocus, whilst programs like *Living Books* and maybe also technology-enhanced computer programs that include *gamification* - the use of video game elements in non-gaming systems to improve user engagement (Deterding, Sicart, Nacke, O'Hara, & Dixon, 2011) – can respond to the needs of susceptible children by providing a highly engaging learning environment.

REFERENCES

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

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

Bus, A. G., Takacs, Z. K., & Kegel, C. A. T. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. *Developmental Review, 35,* 79-97. doi:10.1016/j.dr.2014.12.004

Bus, A. G., van IJzendoorn, M. H., & Pellegrini, A. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65*, 1-21. doi: 10.3102/00346543065001001.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification: using game-design elements in non-gaming contexts. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, 2425-2428. doi.org/10.1145/1979742.1979575.

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

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

Krom, R.S. H., Ouborg, M.J., & Kamphuis, F.H. (2001). Wetenschappelijke verantwoording van de toetsseries Luisteren 1, 2 en 3 [Scientific justification of listening comprehension tests for kindergarten and grade 1]. Arnhem, The Netherlands: Cito.

Koepp, M.J., Gunn R.N., Lawrence A.D., Cunningham V.J., Dagher A., Jones T., Brooks, C., Bench, J. & Grasby, P.M. (1998). Evidence for striatal dopamine release during a video game. *Nature*, *393*, 266-268. doi:10.1038/30498

Krom, R., Jongen, I., Verhelst, N., Kamphuis, F., & Kleintjes, F. (2010). Wetenschappelijke verantwoording DMT en AVI [Scientific justification of DMT and AVI]. Arnhem, The Netherlands: Cito.

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.

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

Maher, B. S., Marazita, M. L., Ferrell, R. E., & Vanyukov, M. M. (2002). Dopamine system genes and attention deficit hyperactivity disorder: a meta-analysis. *Psychiatric Genetics*, *12*, 207–215. doi:10.1097/00041444-200212000-00003

Mayer, R.E. (2005). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R.E. Mayer (Ed.). *The Cambridge handbook of multimedia learning* (pp.183-200). New York, NY: Cambridge University Press.

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin, 137,* 267–296.doi:10.1037/a0021890

Molloy, K., Griffiths, T.D., Chait, M., Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. *The Journal of Neuroscience,9*, *35*(49), 16046 –16054. doi:10.1523/JNEUROSCI.2931-15.2015

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

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

Scarborough, H. S., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review*, *14*, 245–302. doi:10.1006/drev.1994.1010

Schecklmann, M., Ehlisa, A., Plichtaa, M.M, Romanosb, J., Heinea, M., Boreatti-Hümmera, A., Jacoba, C., Fallgattera, A.J (2008). Diminished prefrontal oxygenation with normal and above-average verbal fluency performance in adult ADHD. *Journal of Psychiatric Research*, *43*(2) 2008, 98–106. doi:10.1016/j.jpsychires.2008.02.005

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2014). Can the computer replace the adult for storybook reading? A meta-analysis on the effects of multimedia stories as compared to sharing print stories with an adult. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.01366

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks. A meta-analysis. *Review of Educational Research*, *85*, 4, 698–739 doi:10.3102/0034654314566989

Twisk, J.W.R. (2006). *Applied Multilevel Analysis.* Cambridge, England: Cambridge University Press.

Van Berkel, S. Hilte, M., Engelen, R., Kamphuis, F., Kleintjes, F., & Krom, R. (2010). Woordenschat groep 3 t/m 5 wetenschappelijke verantwoording. [Scientific justification of word knowledge test for first grade to third grade]. Arnhem, The Netherlands: Cito.

Van der Kooy-Hofland, V. A. C., & Bus, A. G. (2012). Long-term Effects of a Computer Intervention to Narrow the Code-related Knowledge Gap in Young Children. Thesis, University of Leiden, Leiden.

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

Verhallen, M. J. A. J., & Bus, A. G. B. (2009). Video storybook reading as a remedy for vocabulary deficits: outcomes and processes. *Journal for Educational Research Online*, *1*, 117–141.

Whitehurst, G. J., & Lonigan, C. J. (2001). Emergent literacy: Development from prereaders to readers. In S. B. Neuman & D. K. Dickinson (Eds.), Handbook of early literacy research (pp. 11–29). New York: Guilford Press.

4

Chapter 5

Brief Report: Better Focus, Better Performance: Account of On-task Child Behavior while Working with Online Educational Programs

Submitted as:

Plak, R.D., Merkelbach, I., Kegel, C.A.T., & Bus, A. G. Brief Report: Better Focus, Better Performance: Account of On-task Child Behavior while Working with Online Educational Programs.

ABSTRACT

Carriers of the 7-repeat allele of the DRD4 gene - a gene that is related to efficiency in dopamine production - benefit from extra-curricular animated movie-like books alongside regular print book sharing. This is in contrast to non-carriers for whom movie-like books provide no extra reinforcement. It is hypothesized that the movie-like books are more engaging for carriers of this genetic marker than for non-carriers, whilst other literacy-related computer programs may not produce such differences. The number of errors in computer assignments recorded while playing with the computer programs, so-called big data, were analyzed and used as indicators of children's engagement. Results corroborate the hypothesis that carriers of the 7-repeat allele of the DRD4 gene were more engaged in digital books than non-carriers, but these differences did not occur while interacting with another program without movie-like features.

Keywords: big data, engagement, the 7-repeat allele of the DRD4 gene, educational computer programs, movie-like books, process data.

INTRODUCTION

Large-scaled randomized controlled trials have shown that *Living Books* – digital storybooks with movie-like elements, including animated pictures, background sounds and music – can be particularly effective for literacy-delayed kindergarten pupils. Findings show that, due to such a program, the gap in early literacy skills between midrange and delayed children narrows, even when the additional program is of short duration. The strongest effects were found when children were carriers of the 7-repeat allele of the DRD4 gene, a gene that is involved in dopamine production in the brain (see Plak, Kegel, & Bus, 2015; Plak, Merkelbach, Kegel, van IJzendoorn, & Bus, 2016). Children with this genetic characteristic easily catch up and even outperform their peers when they are exposed to *Living Books* (e.g. Ellis, Boyce, Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2011; Kegel & Bus, 2013; Plak et al., 2015; Plak et al., 2016, Plak et al., submitted). By way of explaining the effects of *Living Books*, it was hypothesized that the movie-like features of the digital books are particularly engaging for carriers of this dopamine related gene polymorphism 7-repeat allele of the DRD4 gene. Due to the movie-like presentations, stories as offered by *Living Books*, may, probably just like other engaging digital game-based activities (see Kühn et al, 2011; Koepp et al., 1998), bring carriers of the DRD4 7-repeat allele in a state of deep engagement or "flow". This might explain why these children learn intensely and catch up with their peers when exposed briefly (in total 2 to 3 hours) to movie-like animated books. Due to a state of flow, they may not only understand the target stories better and learn more difficult words used in these stories, but may also become more motivated for storybook reading, whether digital or printed.

Children that are carriers have lower dopamine reception efficiency, caused by diminished anticipatory cell firing, which is associated with reduced attention and reward mechanisms (Robbins & Everitt, 1999). In particular these children benefit from the dopamine boost released by activities that are particularly engaging: satisfaction after being immersed in the books releases dopamine, thus enabling a self-reinforcing route and thereby gradually better achievements. Their focus while listening to the stories, may be so intense that they even become oblivious to the world around them, resulting in a state called *inattentional deafness* (Molloy, Griffiths, Chait, & Lavie, 2015). That is, while "reading" *Living Books,* children may not notice irrelevant stimuli from their environment. Flow is a state of complete absorption or engagement in an activity: nothing else seems to matter because he or she is so involved with the goal-driven activity. In game-based learning, the flow framework may have particular value. Because the child is captivated by playing the game, time seems to fly and the experience is so rewarding that the player wants to experience it again and again (Kiili, de Freitas, Arnab, & Lainema, 2012). We expect that carriers of the 7-repeat allele of the DRD4 gene may be so concentrated while working with *Living Books* that comprehension increases and therefore the number of errors in on-task assignments decreases.

The repeated finding in random controlled trials that movie-like books are particularly supportive of a subsample's learning led to a post hoc formulated hypothesis about children attaining a state of flow when they "read" *Living Books*. To test this hypothesis we analyzed children's responses to built-in multiple-choice questions while interacting with educational computer programs. Data for two educational programs were available: the book reading program, *Living Books*, and a phonemic awareness program, *Living Letters*. Arguing that carriers of the 7-repeat allele of the DRD4 gene are more engaged when they are exposed to the movie-like animated books, and may even attain a state of flow, we expected in the *Living Books* group that carriers would make fewer errors in the multiple choice-questions than non-carriers. On the other hand, we did not expect similar differences in number of errors in the multiple-choice questions in *Living Letters* that may not engage carriers more than non-carriers (Plak et al., 2015; Plak et al., 2016; Plak et al., submitted). In contrast to *Living Books*, *Living Letters* lacks characteristics that may elicit a state of intense engagement, such as a movie-like presentation or gamification may do. Educational games should stretch a player's capacity to its limits and encourage effort to overcome challenges (Kiili et al, 2012) and we assume that this is not the case for carriers of the 7-repeat allele of the DRD4 gene interacting with *Living Letters*.

In sum, it was expected that (1) carriers of the DRD4-7repeat allele would be more successful in answering the questions embedded in the *Living Books* stories than non-carriers, (2) effects would be only present in a literacy-delayed group (under the 40th percentile of a standardized literacy test) but not in a literacy-midrange group - this is because the books are thought to have a better match with the needs of less literacy-advanced children - and (3) there would be no differences between carriers and non-carriers in on-task behavior while working with a less engaging program, *Living Letters*, in the literacy-delayed or in the midrange group.

METHOD

Design

In the current study the focus was on process data that were collected as part of a randomized control trial testing effects of two educational computer programs, *Living Letters* and *Living Books*. In both programs, children's responses to built-in multiple-choice questions were sent to and stored on the server of Bereslim, the provider of both programs. In the current study, the number of errors in answering the questions throughout all sessions was the dependent variable, and predictors were

child characteristics: (1) literacy-delayed children - a group of children scoring under the 40th percentile on a standardized literacy test - versus children scoring midrange on literacy, and (2) carriers of the 7-repeat allele of the DRD4 gene versus non-carriers.

Participants

The early literacy project was advertised nationwide from August 2012 to October 2013 and August 2013 to October 2013 via phone, social media, and (e-) mail. Between October 2012 and February 2013 and October 2013 and February 2014, teachers from the 183 participating schools, (from both urban and rural areas) selected eligible kindergarten children to participate in the study. If the parents of the selected children gave informed consent, children were randomly assigned to the experimental conditions, *Living Books* or *Living Letters*.

Procedure

Participants in this study worked online with one of the two educational computer programs, once a week for 15 minutes over a period of about 12 weeks. Based on the process data that were stored by the program (big data), children completed on average 33.62 out of 34 *Living Letters* games (SD = 2.50) and they "read" on average 14.8 out of 16 books (SD = 1.8).

Educational computer programs

The program *Living Books*, was made up of eight age-appropriate digital animated storybooks. The animated pictures, sounds, and music may support children's understanding of story events and language (Bus, Takacs, & Kegel, 2015). Each book was interrupted four times for questions about the story and vocabulary. If children had difficulty with the questions and therefore failed to answer the question correctly, they received feedback and guidance in order to find the correct answer. The first error was followed 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!"). When all the answers were incorrect and the maximum amount of time was taken to answer each question, children spent a maximum of 2 minutes (approximately 25% of the time) answering questions. So for most of the time, children were listening to the story. Each book was presented twice, i.e. two series of 8 books, and in each session four new questions were included.

The program *Living Letters* offered 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). The program adapted automatically to the child's proper name when available in the name database. If the name of the child was not available, the program used 'mama' (mommy) - a word that is just as familiar as their own name to many young children. Feedback provided by a tutor followed up on every response of the child. The program *Living Letters* included five sets of assignments targeting name recognition, targeting recognition of the target word "mama", recognizing the first letter of the name among other letters, and identifying pictures that start with or contain the first letter of the child's name. When children answered a question incorrectly, feedback and clues were provided. After a maximum of three trials, the game ended on a positive note, irrespective of whether a correct response was given, whereupon a new game started. When children failed to give the correct answer at first try, the assignment was repeated twice in subsequent sessions; therefore some children had a few more sessions than others.

Measures

The Central Institute for Test Development (Cito) Literacy Test for Kindergarten (CLT). CLT is a standardized literacy test for kindergarten pupils that is administered group-wise in January (α = .89) and June (α = .87) of senior kindergarten year used in almost all primary schools in the Netherlands. The items assess vocabulary, listening comprehension, rhyming, hearing the first and last word, sound blending, writing conventions (e.g. reading from left to right), and prediction of book content based on the book cover (Lansink & Hemker, 2012).

DNA analysis. Halfway during the intervention period, buccal cell samples were collected by trained members of the research team that visited the schools to collect the samples using a sterile swab designed for collecting buccal cells for DNA analysis (Omni Swabs, Whatman/GE Healthcare, UK). The samples were sent to a commercial laboratory for DNA analysis.

On-task behavior. Children's scores on the multiple-choice questions in *Living Letters* and *Living Books* were stored and made available to the researchers. The responses were formatted as Incorrect, Correct or No response. E.g., when a child gave the correct response at the first try, the format was C (first response Correct). When no response was given at the first try, but then the correct answer was given, the format was NC (No response at the first try → second response Correct). If the child needed feedback twice (made an error twice) before giving the correct answer, the format was IIC (first response Incorrect → second response Incorrect → third response Correct) and when the child needed feedback twice (made an error twice) before giving the incorrect answer, the format was III (first response Incorrect → second response Incorrect → third response Incorrect). This information was used to determine whether or not the player needed to be given the same task again in the next session but also as indicator of the number of errors. Analyzing the data, each multiple-choice question was scored based on the number of incorrect attempts, varying from 0 to 3:

0 - the child answered the question correctly the first time,
1 - the child needed feedback once before answering the question correctly,
2 - the child needed feedback twice before answering the question correctly,
3 - the child failed to answer the question correctly after receiving feedback twice.

For *Living Books* we calculated per book (16 in total) the average score on four questions. For *Living Letters* average scores per set of assignments, five in total, each including 36 unique questions, were calculated.

Genetic screening for DRD4 polymorphisms. Genotyping is the process of determining differences in the genetic make-up (genotype) of an individual by examining the individual's DNA sequence. Biological assays are used and compared to another individual's DNA sequence. 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 100ng genomic DNA template, 10pmol 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 ºC, followed by 27 cycles of 30 sec at 95 ºC, 30 sec at 60 ºC, 60 sec at 65 ºC and a final extension step of 10 min at 65 ºC.

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 an 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 allele at one or both alleles. Of the 593 participants ten children could not be genotyped; 199 children (34%) were carriers of the 7-repeat of DRD4. The distribution of DRD4 polymorphisms was in Hardy-Weinberg equilibrium, $\chi 2$ (df = 1, N = 428) = .069, p = .793.

Statistical analyses

Repeated measures ANOVAs were conducted on number of errors as a within-subjects variable and results of genotyping as a between-subjects factor.

For *Living Letters*, repeated measures were average scores for each of five sets of assignments (name recognition, recognition of "mama", recognition of the first letter of the name, and identifying pictures that contain the first letter or a middle letter of the child's name) and for *Living Books* repeated measures were average scores for each of the 16 books. Analyses for both *Living Books* and *Living Letters* were carried out for the literacy-delayed group and the literacy-midrange group separately. DRD4 was entered as a between-subject variable, age, sex, and father's education were entered as covariates.

RESULTS

There were 661 participants with a mean age of 67.19 months (*SD* = 4.55) of which 57% was male (*n* = 374). The mean score on father's education was for *Living Books* 3.84 (*SD* = 1.62) and for *Living Letters* 3.80 (*SD* = 1.70) on a scale ranging from 0 – 6; 0 represents primary school and 6 university-level education. There were no significant differences in father's education between *Living Letters* and *Living Books,* χ2 (*df* = 7, *N* = 661) = 2.069, *p* = .956.

Between the two experimental conditions, *Living Letters* and *Living Books*, there were no significant differences in the percentage of children with a DRD4 7-repeat allele, 34.9% and 33.3%, respectively; χ^2 (*df* = 1, *N* = 661) = .984, *p* = 1.00. Comparing *Living Letters* and *Living Books* no significant differences were found for age, *t* (659) = .21, *p* = .212, or gender, χ^2 (*df* = 1, *N* = 661) = .344, *p* = .558.

Within the *Living Books* condition, the number of errors between literacy-midrange and literacy-delayed children differed, *t* (203.46) = 4.30, *p* = .000; literacy-midrange children made fewer errors (*M* = 1.37, *SD* = .17) than literacy-delayed children (*M* = 1.48, *SD* = .22). Similarly, within the *Living Letters* condition there was a significant difference in the number of errors between literacy-midrange and literacy-delayed children, *t* (430.71) = 3.70, *p* = .000. Children scoring midrange made fewer errors (*M* = 1.56, *SD* = .31) than literacy-delayed children (*M* = 1.68, *SD* = .38).

Table 1: *Characteristics of Living Books and Living Letters for the total group, the literacy-delayed group and the midrange group.*

		Total n = 220	Literacy-delayed n = 108	Midrange n = 112
Living Books	Age in months (SD)	67.14 (4.60)	66.78 (4.34)	67.48 (4.82)
	CLT pretest (SD)	60.09 (8.49)	53.45 (4.69)	66.48 (6.07)
	Father's education	3.84 (1.62)	3.78 (1.82)	3.90 (1.40)
	Male/Female	128/92	62/46	66/46
	DRD4 7+/7-R	143/77	70/38	73/39
	Number of errors in mc-questions	1.42 (.21)	1.48 (.22)	1.37 (.17)
		Total n = 441	Literacy-delayed n = 227	Midrange n = 214
Living Letters	Age in months (SD)	67.22 (4.53)	66.87 (4.36)	67.59 (4.67)
	CLT pretest (SD)	59.51 (8.52)	53.14 (4.97)	66.28 (5.87)
	Father's education	3.80 (1.70)	3.65 (1.77)	3.97 (1.60)
	Male/Female	246/195	135/92	111/103
	DRD4 7+/7-R	287/154	137/90	150/64
	Number of errors in assignment sets	1.62 (.35)	1.68 (.38)	1.56 (.31)

Living Books

Overall. The significant Maulchy's test for both literacy-delayed and literacy-midrange - χ^2 (119) = 194.30, *p* = .000 and χ^2 (119) = 244.69, *p* = .000, respectively - showed that the assumption of sphericity was violated. Therefore degrees of freedom were corrected with Greenhouse-Geisser estimates of sphericity (ε = .794 and ε = .752, respectively). These estimates were closer to 1 than the lower limit of ε = .067 for both literacy-delayed and literacy-midrange, indicating that there was limited deviation from sphericity (Field, 2009).

Literacy-delayed vs. literacy-midrange. In the literacy-delayed group, the number of errors did not differ across books, *F* (2.39, 253.73) = 1.00, *p* = .452, whilst this did differ in the literacy-midrange group (*F* (2.58, 107.00) = 2.88, *p* = .044), indicating that, for this group, the books were not of a similar level of difficulty. The interaction *Living Books* and DRD4 , and the interaction *Living Books* and the covariates did not show significant effects for the literacy-delayed group or the literacy-midrange group. DRD4 did not cause a significant between-subjects effect in the literacy-midrange group (*F* (1, 107) = .050, *p* = .823), while in the literacy-delayed group DRD4 caused a significant between-subjects effect (*F* (1, 103) = 4.53, *p* =

.036). In the group of literacy-delayed children, carriers of the DRD4 gene made fewer errors (M = 1.42, SD = .22, n = 38) than non-carriers (M = 1.52, SD = .22, n = 70).

Living Letters

Overall. The assumption of sphericity was violated, as is indicated by a significant Maulchy's test for *Living Letters* for both the literacy-delayed and literacy-midrange groups $\chi^2(9)$ = .756, p = .000 and $\chi^2(9)$ = .774, p = .000, respectively - but the deviation from sphericity was limited; Greenhouse-Geisser estimates, for both literacy-groups, ε = .874 and ε = .884, respectively, were closer to 1 than the lower limit of ε = .250 for both groups.

Literacy-delayed vs. literacy-midrange. In the literacy-delayed group the effect of the set of assignments approached significance, F (3.54, 222.30) = 2.26, p = .069, but the number of errors did differ significantly in the literacy-midrange group, F (3.50, 205.95) = 2.71, p = .036, indicating that the sets of assignments differed in difficulty. For literacy-delayed children, all sessions were of such a consistent level of difficulty that little variation between the sessions was found, in contrast to literacy-midrange children; overall they made fewer errors (M = 1.55, SD = .34) than literacy-delayed children (M = 1.68, SD = .38), and the level of difficulty of the program seemed to be a better fit for the literacy-midrange group. The interaction *Living Letters* with DRD4 and the covariates did not show significant effects for the literacy-delayed group or the literacy-midrange group. For both the literacy-delayed group (F (1, 222) = 1.24, p = .267), and for the literacy-midrange group DRD4 (F (1, 209) = .679, p = .411), DRD4 was not significant.

DISCUSSION

Based on the results of two large-scaled RCTs it is possible to conclude that movie-like books may be particularly successful in engaging the attention of children who are carriers of the 7-repeat allele of the DRD4 gene. The results of the data, stored while children interacted with the educational computer programs, are in line with the post hoc formulated hypothesis that programs that include movie-like presentations with nonverbal additions such as background music, and zooming may be particularly engaging for carriers of the 7-repeat allele of the DRD4 gene as these additions direct attention to details in pictures (Bus et al., 2015; Takacs, Swart, & Bus, 2015) Due to the movie-like presentation of stories in *Living Books* that closely match the narratives, carriers of the 7-repeat allele may achieve "flow" while "reading" the stories; they are so engaged and involved that concentrating on the stories goes effortlessly (Kiili et al., 2012).

The program *Living Books* elicited higher levels of concentration and engagement in carriers than non-carriers as can be derived from the significantly lower number of errors in the built-in multiple choice-questions. Another program, *Living Letters*, which also included multiple choice-questions, revealed no differences between carriers and non-carriers in the number of errors. This finding supports the conclusion that movie-like features, typical for *Living Books*, stimulate engagement.

However, we cannot disregard the possibility that differences in number of errors may also mean that carriers of the 7-repeat allele of the DRD4 gene have more knowledge about storylines than non-carriers and are more capable of figuring out the meaning of difficult words used in the stories by means of the context. This is, however, not a very likely explanation when we take into account that carriers had a similar baseline level on a standardized language and literacy test as non-carriers.

When children scored midrange on a standardized literacy test, no differences between carriers and non-carriers were found in the number of errors in multiple-choice questions while reading the books. This may indicate that the *Living Books* program, as used in the current study, was not particularly effective for children at a more advanced level, probably because the selection of books was attuned to low performers and rather simple for the midrange group as is indicated by the error rate; children scoring midrange made significantly fewer errors than their literacy-delayed peers.

The current findings offer further support to the hypothesis that movie-like books may be an indispensable element of the kindergarten curriculum for a substantial group of pupils. For carriers of the 7-repeat allele of the DRD4 gene, about one third of all children, programs such as *Living Books* are more engaging and therefore are a valuable addition to regular book reading sessions (Bus et al., 2015; Takacs, Swart, & Bus, 2015). A movie-like presentation may boost their dopamine production, as a result of which carriers of the 7-repeat allele of the DRD4 gene may enter a state of flow (Kiili et al, 2012) and make substantially fewer errors in questions about the storyline and difficult words used in the story than non-carriers. During traditional book reading, children who are carriers of the 7-repeat allele of the DRD4 gene may experience difficulty staying attentive, whilst they are extremely focused when they interact with *Living Books*, consequently showing more progress in literacy skills with this program than their peers.

REFERENCES

Bus, A. G., Takacs, Z. K., & Kegel, C. A. T. (2015). Affordances and limitations of electronic storybooks for young children's emergent literacy. Developmental Review, 35, 79-97. doi:10.1016/j.dr.2014.12.004

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

Field, A.P. (2009). Discovering Statistics Using SPSS. London: Sage.

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

Kiili, K., de Freitas, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. Procedia Computer Science, 15, 78-91. doi: 10.1016/j.procs.2012.10.060

Koepp M.J., Gunn R.N., Lawrence A.D., Cunningham V.J., Dagher A., Jones T., et al. (1998) Evidence for striatal dopamine release during a video game. Nature, 393, 266–8.

Kühn S., Romanowski A., Schilling C., Lorenz R., Mörsen C., Seiferth N. et al. (2011). The neural basis of video gaming. Translational Psychiatry, 53, doi:10.1038/tp.2011.53.

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.

Molloy, K., Griffiths, T.D., Chait, M., Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. The Journal of Neuroscience, 9, 35(49), 16046 –16054. doi:10.1523/JNEUROSCI.2931-15.2015

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

Plak, R.D., Merkelbach, I., Kegel, C.A.T., Van IJzendoorn, M. H., & Bus, A.G. (2016). Brief computer interventions enhance emergent academic skills in susceptible children: A gene-by-environment experiment. Learning and Instruction, 45, 1-8.doi;10.1016/j.learninstruc.2016.06.002

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

Takacs, Z. K., Swart, E. K., & Bus, A. G. (2015). Benefits and pitfalls of multimedia and interactive features in technology-enhanced storybooks. A meta-analysis. Review of Educational Research, 85, 4, 698–739 doi:10.3102/0034654314566989

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

Chapter 6

General Discussion

Until recently, an educational intervention was considered to be successful if main effects could be established (e.g. Fukkink, Jilink, & Oostdam, 2015). There is however evidence that not every child benefits to the same extent from educational interventions. From the large-scale experiment conducted for this study it appears that effects of additional educational computer programs do not become manifest when the focus is on the complete group, but effects are shown in subsamples.

Two large randomized controlled trials were carried out to examine the effect of educational computer programs. In the first wave - school year 2012-2013 - 90 schools from all over the Netherlands participated in the experiment. Children were randomly assigned to one of the experimental computer programs, *Living Letters* or *Living Books*, or to a control program that does not stimulate early literacy skills (Plak, Kegel, & Bus, 2015). In the second wave - school year 2013-2014 - the randomized controlled trial was expanded by an additional 93 schools (Plak, Merkelbach, Kegel, Van IJzendoorn, & Bus, 2016). The intention was to replicate a Gene x Environment interaction effect, which had been found in the first wave (Plak et al., 2015). To test long-term effects of the educational computer programs in kindergarten, first grade test scores were collected and analyzed (Chapter 4). Big data were used to ascertain differences in on-task behavioral data while working with the experimental technology-enhanced programs *Living Letters* and *Living Books* (Chapter 5). The aim was to show that some children attain "flow" while making the assignments, whereas other do not.

In Kegel, Bus, & Van IJzendoorn (2011) four-year olds who were carriers of the 7-repeat allele of the DRD4 gene benefitted from exposure to *Living Letters*. We were unable to replicate this finding in experiments in which we targeted carriers who were one year older; no effects were found for five-year-old carriers of the 7-repeat allele of the DRD4 gene when exposed to *Living Letters*. It is possible that the five-year-old participants in the current experiments did not benefit from this educational computer program because the tasks were not challenging and exciting enough for their age. In line with this argument, teachers complained that *Living Letters* was too tedious for their students, thus suggesting that *Living Letters* was somewhat boring for most five-year-olds. If the level of the educational program is not tailored exactly right, it is not surprising that no effects were found. It is also possible that the pace of the program is too slow: the push and pull between watching the general introduction, listening to specific instructions, and completing the assignments may make it hard for children to become totally engrossed in *Living Letters*. On the other hand, in parallel studies that focused on another vulnerable subsample, five-year-olds born preterm, *Living Letters* did have added value (Merkelbach, Plak, Kegel, & Bus, under review; Van der Kooy-Hofland, Bus, & Roskos, 2012). These incompatible results for different subsamples, carriers of the 7-repeat allele of the DRD4 gene and children born preterm, emphasize that program characteristics are important, but that fine-tuning matching program characteristics and child characteristics is indispensable.

The educational computer program *Living Books* was not effective for the sample as a whole, but it was for carriers of the 7-repeat allele of the DRD4 gene. *Living Books* had added value only in this particular subsample compared to the control condition in which children had only the daily book sharing experiences. As most children had book-sharing experiences at home and in school, typically several times a day, it is not surprising that the majority did not benefit from a brief intervention with *Living Books*. The program provides extra books, but they may not add much to most children's daily dose of reading, i.e., it hardly matters whether children received the book reading program or not. In spite of the majority not benefiting from *Living Books*, a minority - carriers of the 7-repeat allele of the DRD4 gene - received a boost as a result of the intervention. In this subsample, a brief intervention with animated storybooks contributed substantially more to literacy skills despite of a wealth of traditional book sharing that children experienced.

The effects found were typical for literacy-delayed children; that is, children that scored in the ranges (≤40 percent) of a nationally implemented standardized literacy test for children in kindergarten (Central Institute for Test Development (Cito) Literacy Test for Kindergarten). Effects were not found for children scoring midrange on this test or for carriers of the 7-repeat allele of the DRD4 gene effects, indicating that *Living Books* as used in this large-scale experiment is only effective for those children who struggle to acquire literacy skills. The books that were used in *Living Books* had simple plots, were easy to follow and all contained repeating phrases that create rhythm and structure, as often found in storybooks for very young children. It may be that if more advanced books were to be included in *Living Books*, midrange-scoring children could also benefit from this educational computer program.

The effect of educational computer programs persevere; in first grade, carriers of the 7-repeat allele of the DRD4 gene still outperform their peers in outside-in skills like vocabulary and story comprehension. These long-term effects should probably not only be attributed to direct effects of *Living Books* on those skills. What is more plausible is an indirect snowball effect; the experience with *Living Books* has increased pleasure in book reading for carriers, as a result of which reading motivation in young children grows and they continue to read and ask for books.

Toward a theory explaining the effects of *Living Books*

The positive response that carriers of the 7-repeat allele of the DRD4 show to the *Living Books* may support the hypothesis that educational computer programs that include movie-like components can strengthen these children's learning. *Living Books* are designed in such a way that movie-like presentations of the story - including background music - are constantly available. *Living Books* may thus create a maximally stimulating learning environment for children who may have problems focusing on activities, but who at the same time have a tendency to get carried away by activities that do interest them. Actually the results of analyzing children's online behavior corroborate the hypothesis that the movie-like presentations in *Living Books* may incite a state of deep concentration or "flow" in this specific subsample.

Carriers of the 7-repeat allele of the DRD4 gene may typically show diminished dopamine reception efficiency, often resulting in reduced attention and reward mechanisms (Robbins & Everitt, 1999). If, however, activities are satisfactory, there will be an increased release of dopamine in the ventral striatum, creating a route for learned reinforcement; that is, after being challenged and achieving a satisfactory result, dopamine becomes available thereby enabling new achievements (Koepp et al., 1998). *Living Books* may elicit this route in carriers of the 7-repeat allele of the DRD4 gene and, as have been reported about persons with ADHD, they may then enter stages of very high levels of concentration (Schecklmann et al, 2008).

In other words, it is plausible that carriers of the DRD4 7-repeat allele reach a state of hyperfocus. That is, they focus on what is presented in the computer program, thereby blocking all other stimuli coming from the environment, and engage in activities. Things like playing games or watching movies that do really interest them elicit such a state. In this state it may be hard for them to "shift gear": to stop the task at hand and take up boring but necessary tasks. However a tendency to become immersed in activities that do interest them may turn out to be a good thing when it is used to reach optimal performance. With *Living Books,* carriers of the DRD4 allele even can outperform non-carriers because these books, due to the engaging film-like elements, afford a state of hyperfocus, which supports more intense learning than occurs in non-carriers (Plak et al., 2015; Plak et al., 2016).

It is possible that carriers of the 7-repeat allele of the DRD4 gene, while playing with *Living Books* and watching movie-like images, experience flow that often leads to *inattentional deafness* - not noticing irrelevant stimuli from their environment (Molloy, Griffiths, Chait, & Lavie, 2015) like their somewhat noisy and chaotic classroom. As a result, their attention is solely focused on *Living Books*. Nothing else may matter because the child is deeply involved in reading a story (Kiili, de Freitas, Arnab, & Lainema, 2012). Direct evidence for the state of deep concentration is not yet available although we did find some proof from analyzing children's responses to questions that were built in the stories. The carriers of the 7-repeat allele of the DRD4 gene made significantly fewer errors in those questions than non-carriers, thus indicating that they were much more concentrated on and involved in the task than non-carriers.
Similarly, game-like elements in educational computer programs may elicit flow in subsamples of children. *Gamification* may therefore be a potential beneficial addition to educational programs for carriers of the DRD4 gene. Gamification is the use of video game elements in non-gaming systems to improve user engagement (Deterding,

Sicart, Nacke, O'Hara, & Dixon, 2011). Gamification can take many forms and is used in corporate settings, marketing and education. When playing games, children's level of engagement - a challenge in education - may increase (Van den Boer, 2013). Flow is especially important when involved in game-like activities; a child is highly engaged and tunes out irrelevant stimuli from their surroundings. Because such children are so captivated by playing, and the experience was so rewarding, the children want to experience it again and again (Kiili et al., 2012).

Recommendations for future research

Since it remains unclear whether movie-like images contribute to the effect of *Living Books*, a comparison should be made in which carriers of the 7-repeat allele of the DRD4 gene are presented with digital animated books or with static books.

Furthermore, the focus in the current thesis was solely on the DRD4 gene. In further research however, other candidate genes or pathways should be included in experiments in order to expand our knowledge of markers of differential effects on learning. In addition to this, educational computer programs that include different functionalities than the programs included in this study should be included in future research.

In order to identify children who are carriers of the 7-repeat allele of the DRD4 gene, we used genotyping in the current thesis. For mere research purposes, this is defensible, but it is rather laborious and charged with ethical issues. By including the genetic marker, we were able to prove that *Living Books* are an indispensable element of the kindergarten curriculum. Teachers should make the books part of their daily activities, even though they are unable to identify the children for whom the books are most beneficial. In this study we noticed that for teachers it is hard to believe that the books *do* facilitate learning in subsamples and free reading time is a necessary element of the curriculum, as is the case in higher grades. Integrating *Living Books* in the curriculum is only possible when teachers recognize that free reading time adds to the shared book reading activities.

Finally, in further research the focus should be on specifying underlying mechanisms that explain the efficacy of technology-enhanced books for subsamples, like the hypothesis that some children attain a state of flow when they read the books.

CONCLUSION

The current findings have far-reaching consequences for kindergarten schools and how to use educational computer programs. In this study it was possible to prove that, for a substantial subsample of kindergartners, the *Living Books* program was more instructive than regular book reading experiences. For one third of the children, the educational program *Living Books* has an added value when compared to conventional reading situations. For those children, the program is indispensable and therefore should be part of the curriculum. The program should not be seen as a bonus for children as is often the case with computer programs. Reading independently with *Living Books* should be implemented in their daily routine. With this program only used as little as once or twice a week, a substantial proportion of the students experience increasing success since the extra guidance in these books is crucial for them. In the current experiments, *Living Books* enabled literacy-delayed children who are carriers of the 7-repeat allele of the DRD4 gene to show their full potential. A selection of more complex books may result in similar effects in more literacy-advanced subsamples. For non-carriers, the *Living Books* program does not contribute to their learning, although the program is not harmful to their learning process. For them, the program is only an extension of reading moments, but it does not add value compared with the regular book reading experiences. The finding that for some children a computer program is more effective than an approach that involves teacher-student interaction is surprising, and for many educators, counterintuitive.

REFERENCES

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., & Dixon, D. (2011). Gamification: using game-design elements in non-gaming contexts. *Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems*, 2425-2428.

Fukkink, R., Jilink, L., & Oostdam, R. (2015). Met een blik op de toekomst : een meta-analyse van de effecten van VVE op de ontwikkeling van kinderen in Nederland. [With a view to the future: a meta-analysis of the effects of pre-and early school education on the development of children in the Netherlands] *Kenniscentrum Onderwijs en Opvoeding.*

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

Kiili, K., de Freitas, S., Arnab, S., & Lainema, T. (2012). The design principles for flow experience in educational games. *Procedia Computer Science, 15*, 78-91. doi: 10.1016/j.procs.2012.10.060

Koepp, M.J., Gunn R.N., Lawrence A.D., Cunningham V.J., Dagher A., Jones T., Brooks, C., Bench, J. & Grasby, P.M. (1998). Evidence for striatal dopamine release during a video game. *Nature*, *393*, 266-268. doi:10.1038/30498

Merkelbach, I., Plak, R.D., Kegel, C.A.T., & Bus, A.G.B. (under review). Mild Perinatal Adversities as a Marker of Differential Susceptibility?

Molloy, K., Griffiths, T.D., Chait, M., Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. *The Journal of Neuroscience,9*, *35*(49), 16046 –16054. doi:10.1523/JNEUROSCI.2931-15.2015

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

Plak, R.D., Merkelbach, I., Kegel, C.A.T., Van IJzendoorn, M. H., & Bus, A.G. (2016). Brief computer interventions enhance emergent academic skills in susceptible children: A gene-by-environment experiment. *Learning and Instruction*, *45*, 1-8.doi;10.1016/j.learninstruc.2016.06.002

Plak, R.D., Merkelbach, I., Kegel, C.A.T., & Bus, A. G. (submitted). The Potential of Two Technology-enhanced Early Literacy Interventions to Prevent Reading Delays in First Grade.

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

Schecklmann, M., Ehlisa, A., Plichtaa, M.M, Romanosb, J., Heinea, M., Boreatti-Hümmera, A., Jacoba, C., Fallgattera, A.J (2008). Diminished prefrontal oxygenation with normal and above-average verbal fluency performance in adult ADHD. *Journal of Psychiatric Research*, *43*(2) 2008, 98–106. doi:10.1016/j.jpsychires.2008.02.005

Van den Boer, P. (2013). Introductie in gamification [Introduction in gamification]. www.pietvandenboer.nl

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

6

Nederlandse samenvatting (Dutch summary)[1]

[1]Delen hiervan zijn gepubliceerd als Plak, R.D., Kegel, C.A.T., & Bus, A. G. (2014). Ontluikende geletterdheid stimuleren met digitale prentenboeken: verschillen tussen kleuters. 4W, 3, 22-29.

Een opmerkelijke bevinding in recent orthopedagogisch onderzoek is dat niet ieder kind even gevoelig is voor interventies. Uit dit onderzoek is gebleken dat genetische verschillen, en met name dopamine-gerelateerde genen, kunnen voorspellen wie wel van interventies profiteren, en wie niet. De onderzoekers beschrijven diverse studies waaruit blijkt dat dragers van een specifieke variant op het dopamine-D4-receptor-gen – dit zijn kinderen met aanleg voor een minder efficiënte dopaminehuishouding - extra gevoelig zijn voor programma's waarbij zij positieve feedback krijgen. Dat verklaart bijvoorbeeld waarom een interventie waarbij ouders leerden om hun kind positief te disciplineren in plaats van te straffen, alleen effectief was voor dragers van de variant van het dopamine-D4-receptor-gen. In lijn met deze opmerkelijke resultaten is in het project "*Wat Werkt voor Wie*" onderzocht of ook sprake is van verschil in gevoeligheid voor cognitieve interventies op het gebied van ontluikende geletterdheid en of dopamine-gerelateerde genen geschikt zijn om interventie-gevoelige kinderen van minder gevoelige te onderscheiden.

Het staat vast dat dopamine - een neurotransmitter - een belangrijke rol speelt bij leren en het D4 receptor gen met het polymorfisme DRD4 7-repeat allel zou geschikt kunnen zijn als marker bij het opsporen van differentiële effecten. Dragers van het dopamine-D4-receptor-gen, ook wel de "lange variant" genoemd, hebben een minder efficiënte dopaminehuishouding wat een negatieve invloed heeft op hun aandacht.

In hoofdstuk 2 testen we de hypothese dat een deel van de kinderen, dragers van het DRD4 7-repeat-allel, meer ontvankelijk is voor digitale onderwijsinterventies; een programma gericht op het leren van letters en het herkennen van klanken in woorden, en een voorleesprogramma, digitale prentenboeken, die wordt aangeboden naast het gewone voorlezen thuis en op school. Deze digitale prentenboeken bieden meer begeleiding dan gewone statische prentenboeken, waardoor gemakkelijk afgeleide kinderen zich er wellicht beter op kunnen concentreren. Tegelijkertijd met het voorlezen verschijnen geanimeerde beelden op het scherm met bijpassende achtergrondgeluiden en muziek. Deze extra informatie helpt kinderen niet alleen om tekst aan non-verbale informatie zoals de prenten te koppelen, maar helpt gemakkelijk afleidbare kinderen mogelijk ook om hun aandacht erbij houden.

In een grootschalig landelijke experiment op 82 scholen werden kleuters uit groep 2 die zeer laag scoorden op de CITO Taal voor Kleuters (< 40ste percentiel), door de onderzoekers willekeurig toegewezen aan één van drie programma's: het programma *Letters in Beweging*, een programma gericht op het verbeteren van foneembewustzijn, het programma *Bereslimme Boeken*, digitale prentenboeken met animaties en achtergrond muziek die gericht zijn op het verbeteren van tekstbegrip en woordenschat, of aan het controleprogramma *SamenSlim*, een programma dat niet gericht is op taal, maar op rekenen. In totaal deden 528 kinderen mee aan

het experiment. De onderzoekers waren niet geïnvolveerd in de implementatie: de programma's werden beschikbaar gemaakt via een website waarop leerkrachten de deelnemende kinderen in konden loggen. De programma's registreerden tot waar kinderen gekomen waren en startten bij elke nieuwe sessie automatisch op het juiste punt. Wel controleerden de onderzoekers via de automatische registraties van de programma's of kinderen in het juiste tempo met de programma's werkten en als dit niet het geval was stimuleerden zij de leerkrachten om de kinderen op de afgesproken momenten met de programma's te laten werken; twee keer per week gedurende twee maanden. Voorafgaand aan en na afloop van de interventies, namen leerkrachten de *Cito Taal voor Kleuters* af, een toets uit het leerlingvolgsysteem Cito dat door de meeste scholen in Nederland wordt gebruikt.

Van de *Bereslimme Boeken* profiteerden de dragers van het DRD4 7-repeat-allel, een derde van de kleuters en gingen meer dan niet-dragers vooruit op de *Cito Taal voor Kleuters*. Voor de niet-dragers gold dat de boeken niets extra's bijdroegen aan de ontwikkeling; ze scoorden even hoog met of zonder *Bereslimme Boeken*; met andere woorden de digitale boeken voegden niets toe aan de ervaringen die kinderen al hebben met voorlezen. Voor deze groep achterblijvers zijn de *Bereslimme Boeken* kennelijk een cruciale toevoeging aan het curriculum in de kleuterklas. We vonden geen effecten voor kinderen die gemiddeld scoorden op de *Cito* Taal voor Kleuters. Dit lijkt erop te wijzen dat de voor deze studie geselecteerde *Bereslimme Boeken* vooral geschikt waren voor de zwakste leerlingen.

Replicatie en overzicht van de resultaten

In hoofdstuk 3 wordt een replicatieonderzoek in een nieuwe grotere groep kinderen gepresenteerd; 583 kleuters uit groep 2, afkomstig van 136 scholen deden mee met dit experiment.

Opnieuw bleek dat met *Bereslimme Boeken* achterblijvers die drager zijn van het DRD4 7-repeat-allel, meer vooruit gingen op de *Cito Taal voor Kleuters* dan niet-dragers. Alleen voor dragers van het DRD4 7-repeat-allel waren de digitale prentenboeken een extra stimulans, in vergelijking tot het gewone voorlezen. Het programma *Letters in Beweging* had daarentegen in geen van beide groepen extra effect naast de alledaagse ervaringen met letters en klanken in woorden.

De resultaten van deze onderzoeken zijn door middel van een meta-analyse vergeleken met resultaten uit eerdere studies waarin dezelfde digitale leesinterventies werden gebruikt. Uit deze meta-analyse blijkt dat de effecten voor het programma *Letters in Beweging* heterogeen zijn; effectgroottes lopen uiteen van d = 1.06 tot d = .03 vermoedelijk door variatie in leeftijd. Jongere kinderen profiteren meer van *Letters in Beweging* dan oudere kinderen. De inhoud van dit programma is wellicht geschikter voor jongere kinderen en hun belevingswereld. De effecten van *Bereslimme Boeken* zijn homogeen, *d* = .59. Voor niet-dragers is het effect zeer klein.

De *Bereslimme Boeken* lijken het meest geschikt voor kinderen die moeite hebben hun aandacht bij taken te houden. De bewegende beelden in *Bereslimme Boeken* gecombineerd met de achtergrondmuziek, zijn zo aantrekkelijk voor deze kinderen dat ze alle aandacht opeisen. Wellicht resulteert dit erin dat ze doof en blind worden voor stimuli uit de omgeving. In de literatuur wordt *inattentional deafness* als mogelijk effect genoemd. Het resultaat is dat de kinderen in een staat van opperste concentratie verkeren: *hyperfocus* of *flow*.

Langetermijneffecten van digitale leesinterventies

De bevinding dat dragers van het 7-repeat-allel van het DRD4-gen baat hebben bij de *Bereslimme Boeken* wordt niet alleen direct na de interventie gevonden (hoofdstuk 2 en 3), maar ook een jaar na de interventie (hoofdstuk 4). De kinderen uit hierboven beschreven experimenten zijn gevolgd tot halverwege groep vier. Eindmaten waren de *Cito*-taaltoetsen die in groep 3 en 4 worden afgenomen - Woordenschat, Begrijpend Luisteren (*outside-in* vaardigheden) en Spelling en de Drie Minuten Test (*inside-out* vaardigheden).

De *Bereslimme Boeken* bleken ook op lange termijn een boost te geven aan de geletterdheid van de dragers van het DRD4 7-repeat-allel. De positieve effecten betroffen met name de *outside-in* vaardigheden (woordenschat en verhaalbegrip), en niet *inside-out* vaardigheden (decoderen). In groep 3 profiteerden achterblijvers die drager zijn er nog steeds van dat ze een jaar eerder met de *Bereslimme Boeken* hebben gewerkt; ze presteerden zelfs beter dan hun achterblijvende leeftijdsgenoten.

Gedrag tijdens het werken met de programma's

Zowel *Letters in Beweging* als de *Bereslimme Boeken* bevatten meerkeuzevragen waarop kinderen antwoord moeten geven tijdens het werken met het programma. Elk respons op deze vragen werd geregistreerd en opgeslagen waardoor het gedrag van kinderen *tijdens* het werken met de programma's gemonitord kon worden. In beide programma's verscheen vier keer een adaptieve tutor die de leerling een vraag stelde over letters of klanken (*Letters in Beweging*) of over het verhaal (*Bereslimme Boeken*). Het kind antwoordde door een keuze te maken uit enkele antwoordmogelijkheden.

Als kinderen - zoals we vermoeden - onder invloed van *Bereslimme Boeken* in een *flow* raken, mag worden verwacht dat ze geconcentreerder zijn en daardoor minder fouten maken. Deze hypothese werd bevestigd voor de *Bereslimme Boeken*; dragers van het DRD4 7-repeat-allel met een taalachterstand maakten minder fouten dan niet-dragers. Bij *Letters in Beweging* werd dit effect niet gevonden.

Aanbevelingen

Traditioneel voorlezen is voor de meeste kinderen een effectieve manier om taalvaardigheden te stimuleren, maar niet voor alle kinderen; voor sommige kinderen is meer nodig. Deze studie heeft aangetoond dat een deel van de kleuters pas leert als de boeken filmachtige beelden en muziek bevatten. Informatie- en communicatietechnologie (ICT) biedt in deze gevallen een essentiële toevoeging aan verhalen. Voor deze kleuters is de implementatie van digitale prentenboeken in het curriculum van het kleuteronderwijs cruciaal. Uit deze studies blijkt dat het van groot belang is rekening te houden met moderatoren als effecten van programma's getest worden. Wordt dit genegeerd en wordt enkel naar de groep als geheel gekeken, dan is de kans aanwezig dat effecten van interventies worden gemist.

De marker DRD4 7-repeat-allel kan worden gezien als risicokenmerk, maar ook als kenmerk van leerbaarheid. Dragers van het DRD4 7-repeat-allel van het 7-repeat-allel van het Dopamine-D4-receptor-gen hebben een groter risico om achter te lopen op school wanneer niet wordt voorzien in hun behoefte. Aan de andere kant zijn juist zij degene die het meest vooruit gaan wanneer zij op de juiste wijze worden benaderd.

Voor niet-dragers is er in de huidige studies geen bewijs gevonden dat *Letters in Beweging* of de *Bereslimme Boeken* extra bijdragen aan taalvaardigheden; evenmin ondervinden niet-dragers nadelen van de digitale boeken. Daarom zouden de *Bereslimme Boeken* die voor dragers zeer gunstige effecten opleveren door alle kinderen in de klas gebruikt kunnen worden; een derde van de kinderen zal baat hebben bij het programma, maar alle kinderen zullen plezier beleven aan de geanimeerde verhalen.

Curriculum Vitae

Rachel Dominique Plak werd op 12 mei 1985 geboren te Amsterdam, Nederland. In 2003 behaalde zij haar eindexamen aan het Sint Ignatiusgymnasium te Amsterdam. Aansluitend begon zij aan de studie Communicatiewetenschap aan de Universiteit van Amsterdam. Tijdens deze studie werd haar interesse voor de ontwikkeling van kinderen gewekt en besloot zij de studie Psychologie aan dezelfde universiteit te volgen. Vanaf 2008 werkte zij voor een regionale televisieomroep, onder andere op de redactie voor een schooltelevisieprogramma. In 2009 behaalde zij haar bachelor Communicatiewetenschap en deed zij ervaring op als behandeld psycholoog bij de afdeling Kinder- en Jeugdpsychiatrie van het Academisch Medisch Centrum Amsterdam waar zij kinderen met somatoforme en emotionele stoornissen behandelde. Kort nadat zij in 2010 begon als klinisch ontwikkelingspsychologe bij de Geheime Tuin, een kinder- en jeugafdeling van Molemann Clinics, behaalde zij haar Master of Science in de Klinische Ontwikkelingspsychologie. In 2012 startte zij als promovenda bij de afdeling Orthopedagogiek bij de Universiteit van Leiden onder begeleiding van prof. dr. A.G. Bus waar zij onderzoek verrichtte naar differentiële ontvankelijkheid voor de effecten van educatieve computerprogramma's. De resultaten van het promotieonderzoek zijn in dit proefschrift beschreven.

Dankwoord

Mijn dank gaat uit naar Stichting Kennisnet voor het toekennen van subsidie aan mijn promotor Adriana Bus teneinde dit onderzoek mogelijk te maken.

Alle deelnemende scholen, leerkrachten, ouders en leerlingen, hartelijk bedankt. Zonder jullie medewerking had dit onderzoek niet uitgevoerd kunnen worden.

Ik prijs mijzelf gelukkig met zoveel fijne mensen om mij heen die mij de afgelopen jaren hebben geholpen en bijgestaan. Daarom wil ik de volgende mensen bedanken:

Thijs, Elise en Heleen, bedankt voor het altijd willen meedenken met mijn onderzoek. Wat hebben we het leuk gehad. Bedankt voor de leuke tijden op de borrels, weekendjes weg en congressen. Jullie hebben mijn promotietraject er heel veel leuker op gemaakt.

Marga, Verna en Marianne, fijne collega's, ik ga jullie missen. Bedankt voor álles.

An, Niek, Mar, Yaan, Liek en Callie, bedankt voor het luisteren naar mijn (soms eindeloze) verhalen over mijn werk en het bieden van de nodige afleiding. Jullie zijn geweldig. Anniek, wat fijn dat je me als paranimf bijstaat tijdens de verdediging!

Inge, paranimf en partner in crime, hoe had ik dit allemaal zonder jou kunnen doen? Bedankt voor de samenwerking, voor het lachen tot ik buikpijn had, voor de etentjes, voor de discussies over allerhande zaken, voor het maar naar je hoeven kijken en dan snappen wat ik bedoel, en voor het koelhouden van mijn hoofd.

Klaas en Nini, bedankt voor jullie adviezen en betrokkenheid. Ik heb geboft met zulke lieve schoonouders.

Mam, je bent mijn grootste fan. Bedankt voor je enthousiasme en steun. Hoe groot sommige uitdagingen ook waren, jij bleef altijd in mij geloven.

Lieve Daniël, je was de afgelopen jaren mijn persoonlijke ICT-helpdesk, degene die het beste uit me naar boven haalde en bovenal mijn steun en toeverlaat. Het waren soms roerige tijden, maar jij sleepte me er altijd doorheen. Bedankt voor het luisteren, het meeleven én het meelachen. Ik kijk uit naar ons verdere leven samen met onze lieve Boris.

List of Publications

Plak, R.D., Kegel, C.A.T., & Bus, A. G. (2014). Ontluikende geletterdheid stimuleren met digitale prentenboeken: verschillen tussen kleuters [Encourage early literacy with digital story books: differences between children]. *4W*, 3, 22-29.

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

Plak, R.D., Merkelbach, I., Kegel, C.A.T., Van IJzendoorn, M. H., & Bus, A.G. (2016). Brief computer interventions enhance emergent academic skills in susceptible children: A gene-by-environment experiment. *Learning and Instruction*, *45*, 1-8. doi;10.1016/j.learninstruc.2016.06.002

Plak, R.D., Merkelbach, I., Kegel, C.A.T., & Bus, A. G. (submitted). The Potential of Two Technology-enhanced Early Literacy Interventions to Prevent Reading Delays in First Grade.

Plak, R.D., Merkelbach, I., Kegel, C.A.T., & Bus, A. G. (submitted). Brief Report: Better Focus, Better Performance: Account of On-task Child Behavior while Working with Online Educational Programs.