



Universiteit  
Leiden  
The Netherlands

## **Differential susceptibility to an early literacy intervention**

Kooy-Hofland, V.A.C. van der

### **Citation**

Kooy-Hofland, V. A. C. van der. (2011, September 29). *Differential susceptibility to an early literacy intervention*. Retrieved from <https://hdl.handle.net/1887/17883>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/17883>

**Note:** To cite this publication please use the final published version (if applicable).

# Chapter 2

## **Effects of a brief but intensive remedial computer intervention in a sub-sample of kindergartners with early literacy delays**

This chapter will be published as: Van der Kooy-Hofland, V.A.C., Bus, A.G., & Roskos, K. (in press). Effects of a brief but intensive remedial computer intervention in a sub-sample of kindergartners with early literacy delays. *Reading and Writing*.

## **Abstract**

Living Letters is an adaptive game designed to promote children's combining of how the proper name sounds with their knowledge of how the name looks. A randomized controlled trial (RCT) was used to experimentally test whether priming for attending to the sound-symbol relationship in the proper name can reduce the risk for developing reading problems in the first two grades of primary education. A web-based computer program with more intensive practice than could be offered by teachers affords activities that prompt young children to pay attention to print as an object of investigation. The study focused on a subsample of 110 five-year-old Dutch children from 15 schools seriously delayed in code-related knowledge. Outcomes support the need for early remedial computer programs, and demonstrate that, without a brief but intensive treatment, more children from the at-risk group lack the capacity to benefit from beginning reading instruction in the early grades. With an early intervention in kindergarten, children with code-related skills delays gained about half a standard deviation on standardized tests at the end of grade 2.

## **Introduction**

Most kindergartners have a nascent awareness of the sound-symbol relationship in printed words and have begun to accrue baseline knowledge of a small assemblage of letters and sounds as is demonstrated in their invented spelling attempts (Sulzby, Barnhart, & Hieshima, 1989; Treiman & Kessler, 2003). A sub-sample of children in each kindergarten classroom is already, by 5 years of age, lacking in competencies fundamental to learn to read (Duncan et al., 2007) due to sparse experiences in the early years or inability to take advantage of their environment (Shonkoff & Phillips, 2000; Stipek & Ryan, 1997) and as a result their capacity to benefit from beginning reading instruction may be compromised (Byrne, Fielding-Barnsley, & Ashley, 2000; Duursma, Augustyn, & Zuckerman, 2008; Silva & Alves-Martins, 2002; Snider, 1995). In this line of argumentation, our study tests whether an early intervention creates a better starting position for learning to read in primary education (e.g., Byrne et al, 2000; see for meta-analytic evidence: Bus & van IJzendoorn, 1999; Ehri et al., 2001). Therefore, in addition to short-term effects of the program long-term effects are assessed after 18 months of formal reading instruction.

In contrast to many early intervention programs that target whole or small groups, this program targets only a sub-sample that is in need of an additional or more intensive program in preparation for reading instruction in primary education. We modeled, therefore, an individualized remedial computer program for young children with early literacy delays to pay attention to print as an object of investigation. In this report we present effects of this program in a sub-sample of kindergartners with code-related skills delays. Results of this educational intervention show that delayed children achieve higher gains in beginning reading instruction in the first grades of primary education when they receive computer-assisted instruction in support of the kindergarten curriculum.

## **Considerations underlying the computer program's design**

Typically, many children begin to understand how letters relate to sounds when they learn how to read and write their name or some portion thereof. Name writing is commonplace in young children's everyday life resulting in the proper name being the first word that many children learn to read and write (Levin & Bus, 2003; Levin, Both-de Vries, Aram, & Bus, 2005). Close inspection of children's emerging letter name knowledge, phonemic awareness, and invented spellings supports the hypothesis that the initial letter of the proper name serves as an early decoder illuminating how sounds relate to letters (Bloodgood, 1999; Levin & Bus, 2003; Molfese, Beswick, Molnar, & Jacobi-Vessels, 2006). Most children can name

the initial letter of the proper name earlier than other letters; most can locate the sound of the first letter in other words preceding other sounds; and most can use the first letter of the proper name first of all in their invented spellings (Both-de Vries & Bus, 2008; 2010). These results suggest that children's playful experiences with the proper name offer a framework that anchors code-related skill instruction and practice in a personally motivating context.

Years spent from ages 2 to 5 offer many opportunities to learn the name, the first letter of the name, how it sounds in words, and how word spellings can be created using this letter-sound knowledge (Levin & Aram, 2004). Replicating the home literacy environment at school to remedy the sparse print exposure for children in some families is challenging and leads to an approach that fundamentally differs from more common early interventions that practice with a range of phonemes in generic tasks (Ehri et al., 2001; Borstrom & Elbro, 1997; Byrne et al., 1998). In classrooms, there is no real match for the long stretch of early literacy experiences with cumulative effects across the preschool years, although the at-school curriculum can attempt to build in some salient features of the home environment (e.g., teachers can play games using the first letter of the name). Since at risk kindergarten children do not have the luxury of time that a print rich home literacy environment affords in the preschool years, a computer program may offer the opportunity to practice frequently with the first letter of the child's name, affording far more exposure and instruction than from the teacher alone (Heuston, 1996). Computer speech, interactions, along with interesting graphics and animation has permitted the development of programs that are highly motivating to children (Mayer & Moreno, 2002). Moreover many software programs do not require large investments in professional development for purposes of differentiating instruction (Chera & Wood, 2003).

The *Living Letters* web-based program, developed in close collaboration among computer experts, designers, and experts in the field of education, took into account how learning about the alphabetic principle starts in literate homes and used the child's proper name as a stimulus that primes children for understanding the alphabetic principle. It is a series of adaptive games intended for kindergarten children not yet demonstrating an awareness of the sound/letter relationship in an alphabetic language. Its instructional framework is modeled on the aforementioned name writing research and emphasizes three successive skill areas: (1) recognizing the proper name in print; (2) associating the initial name letter with its sound; and (3) identifying the sound of the initial name letter in other orally presented words. The web-based software program automatically adapts to the child's proper name and provides the child with targeted instruction on sound-letter relationships modeled after parental instruction (Anderson, Boyle, & Reiser, 1985). The program registers the child's immediate responses to tailor the program to individual differences—a design advantage over traditional

classroom instruction and stand alone computer software programs (Greasser, Conley, & Olney, in press). For instance, when children produce one or more erratic responses to an assignment, the assignment is repeated one to three times, thus promoting more practice when children fall behind.

As with all computer-aided instruction, there may be drawbacks to Living Letters as an intervention for achieving code-related skill outcomes. Based on our observations, young children can indeed use a computer mouse without adult aid and can complete online educational programs independently (Bus, Verhallen, & Van der Kooy-Hofland, 2009; De Jong & Bus, 2002; 2004; Verhallen, Bus, & De Jong, 2006; Verhallen & Bus, 2010). Yet, as we have observed, the *blind eye* of computer-aided instruction can leave children to their own devices, opening the door to *free play* rather than playful engagement with the content (De Jong & Bus, 2002). Children can complete the computer assignments without seriously attempting to solve the problems they pose with the result that the potential benefits of computer-aided instruction are reduced (Kegel, Van der Kooy-Hofland, & Bus, 2009). This may explain why adaptive computer-assisted learning systems that prevent the incidence of random responses have been demonstrated to significantly increase phoneme awareness in low-performing preschoolers (Mitchell & Fox, 2001), at-risk preschoolers and kindergartners (Lonigan et al., 2003), and typical preschoolers (Foster, Erickson, Forster, Brinkman, & Torgesen, 1994) compared to a control group not utilizing any computer-assisted learning systems. We are mindful of the possible weakness of Living Letters in that children play the games by just responding randomly despite the built-in feedback loops and expect that errors are negative predictors of learning from the Living Letters intervention.

## **This study**

To test the theory that kindergarten children who have not yet begun to develop code-related skills are more at risk for developing reading problems this study tests the long-term benefits of the individualized web based program, Living Letters, in support of teacher-delivered literacy training in kindergarten. Research suggests that cueing children to sound-letter information in the proper name, particularly the initial name letter, primes the insights that (a) letters represent sounds and (b) letter-sound information can be used to make words. Once they have grasped these insights as appears from the short-term effects of the program, at risk children may benefit more fully from classroom practice in code-related skills in kindergarten as well as in the first years of reading instruction. Despite the short duration of the program (2½-3 hours), high intensity practice with the first letter of the name is expected to help kindergarten children understand

that letters in written words relate to sounds in spoken words and may enable them to make a start with invented spelling, phonemic awareness, and decoding skills (Silva & Alves-Martins, 2002). The present study is also a critical test of the universal need for early literacy interventions. Share (2008), for instance, argued that grasping basic insights in letter-sound relations prior to the start of formal reading instruction may be important in English where letter-sound relations are difficult to fathom but not so relevant in transparent languages as Dutch because letter-sound knowledge is easy to acquire. The study also probes the degree of predictability between error levels in completing program activities and code-related outcome measures, which provides an additional test of the role of the actual program as an incentive of children's learning.

The study used a randomized controlled trial (RCT) to research the instructional effects of Living Letters on a sample of at risk kindergarten children performing in the lower quartile on pre-reading assessments. Eligible children from the same kindergarten classrooms were randomly assigned to either (a) Living Letters (LL); (b) a treated control program entitled *Living Books* (LB), which focuses on story comprehension (Bus et al., 2009), or (c) a combined program consisting of Living Letters (LL) and Living Books (LB). In the present report, we only discuss the contrast between sub-sample of children at risk who were exposed to the intervention program Living Letters (alone or combined with the control program) and the control group (sub-sample of children at risk who were exposed to Living Books alone (effects of Living Books is not discussed in this paper). The target program is supplementary to an estimated hour per week of classroom practice in code-related skills by practicing letters and playing games with words and sounds per the Netherlands kindergarten curriculum. We tested the following hypotheses:

1. Participation in the computer program Living Letters alone or in combination with Living Books significantly improves at risk kindergarten children's code-related skills, including (i) letter knowledge, (ii) phonological awareness, (iii) word spelling, and (iv) decoding; a basic assumption is that practice with the initial name letter stimulates and re-organizes attention to sounds and letters in printed words toward understanding that letters relate to sounds.
2. Child gains in code-related skills accrued from participation in the computer program Living Letters are sustained beyond the kindergarten year—an assumption based on the expectation that delay in code-related skills, if not addressed in the early years, may lead to an ongoing knowledge gap, given that some code-related knowledge at the start of reading instruction enables practicing reading and spelling proficiency with success in primary school years (Juel, 1988). We opted for assessing reading and spelling skills after about two years of reading instruction because by then basic reading and

spelling skills can be reliably assessed without floor or ceiling effects (e.g., Paris, 2005).

3. A low error threshold on the computer program, defined as successful completion of a majority of tasks, is required to realize optimal benefits from participation in the program although we can not exclude in advance that children learn from errors, especially if they sooner or later stop committing them, and errors may have a positive impact on outcomes. Registrations of errors can serve as level of success indicator since number of errors yields objective data about success and/or error rates that may impact program efficiency.

## **Method**

### *Participants*

The intervention sample was drawn from 15 schools in the Western part of the Netherlands. From these schools, 404 senior kindergarten children speaking Dutch as their first language and between 60 to 72 months old were screened upon kindergarten entry over a three-week period on early literacy skills in fall 2006. An estimated 12% of all pupils did not participate in the screening, due to illness or absence for other reasons or failure of parental consent. Those students scoring among the 30% lowest on the early literacy screening composite measure were selected as eligible candidates for the intervention, totaling 110 children. The 30% cut-point for the sample was based on the finding that this group demonstrated weak code-related skills. As shown in Table 1, on name writing and rhyming all students scored close to ceiling, however the lowest scoring children differed from the average or above average scoring group on three other indicators: letters ( $t = 18.13$ ,  $df = 377$ ,  $p < .001$ ), writing 'mama' [*mom*] ( $t = 15.16$ ,  $df = 326,520$ ,  $p < .001$ ), and writing other words ( $t = 13.63$ ,  $df = 274,502$ ,  $p < .001$ ). Overall the lowest scoring children did not write phonetically (i.e., they mostly produced conventional symbols, however the symbols did not represent sounds in the word) as is indicated by mean scores on the scale slightly beyond 2 whereas the highest scoring group (beyond 3) produced phonetic writing (B or BT for BOAT). Moreover, in the lowest scoring sub-sample boys were overrepresented,  $X^2=6.70$ ,  $df = 15$ ,  $p < .01$ , and mothers were lower educated,  $t = 3.42$ ,  $df = 377$ ,  $p < .001$ .

Table 1

Mean scores (and standard deviations) on gender, age, maternal education and the screening tests (including letter knowledge, rhyming and writing) for total group, highest scoring 70% and lowest scoring 30%.

|                    | <i>Max</i> | <i>Total group</i><br>(n = 404)<br>M (SD) | <i>Highest 70%</i><br>(n = 269)<br>M (SD) | <i>Lowest 30%</i><br>(n = 110)<br>M (SD) |
|--------------------|------------|-------------------------------------------|-------------------------------------------|------------------------------------------|
| Gender m / f       |            | 200 / 204                                 | 122 / 147                                 | 66 / 44                                  |
| Age                |            | 64.95 (3.59)                              | 65.12 (3.76)                              | 64.67 (3.19)                             |
| Maternal education | 8          | 5.28 (1.89)                               | 5.52 (1.84)                               | 4.80 (1.88)                              |
| Letter knowledge   | 8          | 5.62 (2.07)                               | 6.65 (1.53)                               | 3.62 (1.35)                              |
| Rhyming            | 10         | 9.07 (1.78)                               | 9.40 (1.48)                               | 8.32 (2.21)                              |
| Writing Name       | 6          | 5.63 (0.86)                               | 5.78 (0.69)                               | 5.32 (1.10)                              |
| Writing Mom        | 6          | 3.67 (1.74)                               | 4.35 (1.64)                               | 2.29 (0.96)                              |
| Writing words      | 6          | 3.04 (1.07)                               | 3.44 (0.98)                               | 2.21 (0.69)                              |

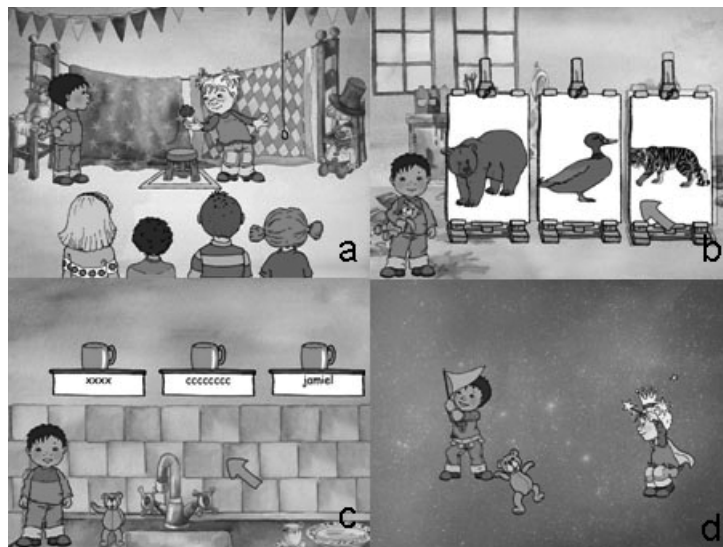
The selected sample varied from 3 to 15 children per school (17.6% - 51.7%). Eligible children were randomly assigned to Living Letters-only, Living Books-only, or Living Letters+Living Books, stratified for school and gender. No child attrition occurred during the kindergarten year; during the follow-up two years later seven children from the treatment group and five from the control group were lost due to removal ( $n = 8$ ), repeating ( $n = 3$ ) or placement in special education ( $n = 1$ ).

#### *Description of Treatment Conditions*

Children assigned to the intervention condition participated in a series of games of increasing difficulty in the following order:

- 22 games providing practice in recognizing the proper name (see for instance, Figure 1C); as the program adapts to the child's name tasks are unique for each child.
- 6 games focusing on recognition of the first letter of the proper name; children are asked to identify their name letter among three or more other letters; the computer pronounces the letter ("yes that is your letter, /t/").
- 12 games providing practice in identifying pictures that start or end with the first letter of the child's name. Criteria for selecting words were familiarity and transparency of words. For every child the program provides a unique selection of words attuned to the first sound in the child's name.

All sessions start with an attractive animation to explain the upcoming games (e.g., main characters, *Sim* and *Sanne*, discuss their names and discover that they begin with the same sound). Errors when solving the games are followed by increasingly supportive audio feedback in the following order: (1) repetition of the task (*Find the word that starts with the same sound as your name*); (2) a clue (*Tom starts with /t/*), and (3) demonstration of the correct solution (*You hear /t/ in tom and tent*). Apart from increasingly supportive feedback errors imply one to three repetitions of the same assignment. Tasks as well as oral feedback are adapted to the child's name. Figure 1a shows a screenshot from the instruction at the start of the last set of games; *Sanne* is the magician who finds words that start with the /s/ of *Sanne*. Figure 1b is one of the assignments in the last set of games. *Tom* has to find the word that starts the same as his name. Figure 1c shows that *bear* provides a cue when the child has not succeeded two times to find his or her name among the three alternatives. Figure 1d is a screenshot from the scene at the very end of each game. In the current study the six games focusing on recognition of the first letter of the proper name and the 12 games providing practice in identifying pictures that start or end with the first letter of the child's name were always repeated in a subsequent session, these games thus constituting two-thirds of the total computerized program.



*Figure 1.* The screenshots have been derived from four different elements of the games: the animation at the start of a new set of games (a), one of the assignments (b), bear provides a cue after an error, and (c) the scene at the very end of each game (d).

Children assigned to the control condition listened to five age-appropriate electronic books that consisted of oral narration, but no printed text, thus allowing the child to *read by listening*. In each 10-minute session, children *read* one book and responded to four follow-up questions among which two about difficult words (e.g., *What are paving stones?*) and two about story events (e.g., *Is dad happy or angry?*) by choosing one out of three pictures. Each book was repeated three times across the 15 sessions. In each repeated reading, children responded to four new questions, totaling 12 questions per book.

## Assessment measures

*Background.* As a control for factors that may influence the beneficial effects of the computer program, indicators for intelligence and SES were assessed. The Dutch version of Raven's Colored Progressive Matrices (Van Bon, 1986), a measure of nonverbal intelligence, was administered in the pre-assessment phase of the study (fall of the kindergarten year). To survey maternal education mothers ticked their highest level of education: 1 (primary school), 2 (preparatory secondary vocational education), 3 (preparatory middle-level vocational education), 4 (senior secondary vocational education), 5 (senior secondary education), 6 (pre-university education), 7 (professional higher education), and 8 (university).

*Literacy Screening to select the 30% lowest scoring children (Fall).* As a measure of invented spelling children were asked to write the *proper name* (name writing task), *mama* (mom), and four other words (e.g., *boot* [boat]). We selected words which are familiar to children (Schrooten & Vermeer, 1994) but not practiced in invented spellings. Each word was double-coded on a scale from 1 (writing-like scribbles) to 6 (conventional spelling) by trained master students (Levin & Bus, 2003). The intra-class correlation coefficient for 20 double-coded assignments was high ( $r = .99$ ). The rhyming task included 10 items asking children to select the picture among three alternatives that rhymed with a target picture. In the receptive letter knowledge task, children were asked to point to one of eight target *letters*, each presented on a card between four other letters. Alpha reliabilities for the tests were satisfactory; see Table 2. Scores on invented spelling and letter knowledge were standardized and averaged to form an early literacy composite measure to select the 30% lowest scoring children. Name writing and rhyming were not included due to ceiling effects.

*Pre/post assessments (Winter; Spring).* A more extensive set of pre/post assessments was applied to assess effects of the intervention. *Letter knowledge:* On pre-test children were asked to identify by sound or name eight high frequency letters on

a chart. To avoid placing unreasonable demands on children, only a selection of letters was given. On post-test, 14 more letters were added to the letter knowledge test. Awarding both letter sound and letter name the maximum score on the pretest was 8 and on posttest 22. *Phonological skills* (pre- and post-tested) were assessed in a 5-task series: (1) identifying among three words the one that starts with a sound different from the other two words; (2) selecting among four words two words with the same initial sound; (3) selecting from four words two words with the same final sound; (4) naming the first sound of words; and (5) naming all sounds of words. To reduce examiner bias, all picture names were pronounced by a computerized voice. All target sounds ( $n = 20$ ) were consonants; all words were monosyllabic (CVC or CVVC). Each correct response was awarded one point (maximum = 25). *Invented spelling* (pre/post-tested): Children were asked to write five randomly chosen words: *papa* (dad), *kaas* (cheese), *been* (leg), *jurk* (dress), and *duim* (thumb) that were scored on a 1-6 scale. *Word recognition* (post-tested only): Children were asked to identify the depicted target word among four printed words. The (incorrect) alternatives differed in 1, 2, or 3 letters from the target word. For instance, distracters for /*raam*/ were /*room*/, /*rat*/, and /*been*/. Correct responses were rewarded with a score of 3 (*raam*); a match of the first and last letter (*room*) with a score of 2; a match of the first letter only with a score of 1 (*rat*); and no match (*been*) with 0. *Decoding* (post-tested only): Children were tutored in decoding four vowel-consonant (VC) and four consonant-vowel-consonant (CVC) nonsense words. If children failed to pronounce the nonsense word in the first five seconds after presentation of a word, they were prompted to sound out the separate letters. If this did not elicit correct decoding, the experimenter pronounced the separate sounds and stimulated the subjects to blend the sounds. If they did not succeed, the experimenter repeated the separate sounds, blended them, and had subjects repeat the naming and blending. The list of eight words was repeated five times in different sequences. Scores per word varied from 5 (successful first attempt) to 1 (non-completion of item). Alpha reliabilities were satisfactory (see Table 2).

*Post-post measures* were conducted after 18 months of reading instruction, as follows: *Word reading fluency* was tested by administering the Een-Minuuut-Test (one minute test), a standardized test to determine how many words from a list can be read during one minute (Brus & Voeten, 1973). *Klepel*, a standardized nonsense word reading test, assesses how many words are read accurately in 2 minutes (Van den Bos, Lutje Spelberg, Scheepstra, & de Vries, 1994). *Spelling*: For fifteen dictated words, well-known, with two or three syllables (including more complex letter-sound rules), the correctly spelled words were scored. Schools did not allow us to use a standardized test for spelling to avoid interference with regular progress monitoring. *Compound measure for reading and spelling*: The measures

Table 2

Mean scores (and standard deviations) on age, gender, maternal education, nonverbal intelligence, screening tests, pre-tests, and post-tests for intervention group (LL) and control (LB).

| Measure                             | $\alpha$         | Timing      | LL <sup>c</sup><br><i>n</i> = 73 | LB <sup>d</sup><br><i>n</i> = 37 |
|-------------------------------------|------------------|-------------|----------------------------------|----------------------------------|
| Gender M / F                        | -                |             | 45 / 28                          | 21 / 16                          |
| Age at the screening in months      | -                | Fall K      | 64.89(3.28)                      | 64.41(3.24)                      |
| Maternal education (max = 8)        | -                | Winter K    | 4.79 (1.98)                      | 4.81 (1.70)                      |
| Nonverbal intelligence              | .92 <sup>a</sup> | Winter K    | 4.97 (1.84)                      | 5.18 (1.94)                      |
| Number of errors in games           | -                |             | 13.16 (7.31)                     | --                               |
| Letter knowledge (max = 8)          | .73 <sup>b</sup> | Fall K      | 3.66 (1.32)                      | 3.54 (1.43)                      |
| Rhyming (max = 10)                  | .82 <sup>b</sup> | Fall K      | 8.55 (2.03)                      | 7.86 (2.49)                      |
| Writing Proper Name (max = 6)       | -                | Fall K      | 5.37 (1.12)                      | 5.22 (1.06)                      |
| Writing Mom (max = 6)               | -                | Fall K      | 2.24 (.83)                       | 2.39 (1.18)                      |
| Writing words (max = 6)             | .87 <sup>b</sup> | Fall K      | 2.27 (.67)                       | 2.08 (.73)                       |
| Letter knowledge (max = 8)          | .66 <sup>b</sup> | Winter K    | 3.63 (2.10)                      | 3.03 (1.80)                      |
| Letter Knowledge (max = 22)         | .90 <sup>b</sup> | Spring K    | 10.93 (6.00)                     | 9.46 (4.43)                      |
| Phonological skills (max = 25)      | .81 <sup>b</sup> | Winter K    | 8.05 (5.05)                      | 7.35 (4.22)                      |
| Phonological skills (max = 25)      | .87 <sup>b</sup> | Spring K    | 13.58 (5.89)                     | 11.11 (5.69)                     |
| Invented spelling (max = 6)         | .85 <sup>b</sup> | Winter K    | 2.45 (.65)                       | 2.15 (1.03)                      |
| Invented spelling (max = 6)         | .79 <sup>b</sup> | Spring K    | 3.27 (.81)                       | 2.91 (.63)                       |
| Word Recognition (max = 45)         | .80 <sup>b</sup> | Spring K    | 28.42 (7.58)                     | 26.32 (5.76)                     |
| Decoding (max = 40)                 | .98 <sup>b</sup> | Spring K    | 25.10 (8.09)                     | 22.26 (5.17)                     |
| Word reading fluency <sup>c</sup>   |                  | End Grade 2 | 44.38 (15.47)                    | 39.66 (13.23)                    |
| Klepel <sup>c</sup>                 |                  | End Grade 2 | 40.29 (18.03)                    | 32.81 (17.16)                    |
| Spelling <sup>c</sup>               | .85 <sup>b</sup> | End Grade 2 | 6.00 (3.83)                      | 3.94 (2.84)                      |
| Reading and Spelling (factor score) |                  | End Grade 2 | .15 (1.03)                       | -.31 (.86)                       |
| Reading Comprehension               | .85 <sup>b</sup> | End Grade 2 | 10.41 (4.18)                     | 8.45 (4.09)                      |

Notes. <sup>a</sup> Guttman's split-half reliability. <sup>b</sup> Cronbach's alpha. <sup>c</sup> (*n*=66). <sup>d</sup> (*n*=32).

for word reading, non word reading, and spelling showed high correlations (> .55). Principal Component Analysis revealed one component explaining 85% of variance with test loadings ranging from .87 to .95. This new variable was normally distributed. *Reading comprehension* was assessed with a cloze test. This test on paper consists of a 142 words text with 21 words removed, where the child is asked to write the missing words. Children completed the test on their own.

*Error registration in the target program.* To determine the frequency of child appeals for feedback while engaged in the intervention program, the position and location of the mouse onscreen were recorded every 10th second. How successful children were at solving the computer assignments immediately or after one or more repetitions can be derived from these registrations.

## **Procedure**

The training regimes, consisting of one weekly session, held over a period of 15 weeks, were incorporated into the kindergarten curriculum. Children receiving one program spent 10-15 minutes a week on the intervention while those assigned to the condition that combined the control and intervention program spent an estimated 15-30 minutes per week playing computer games. Sessions occurred during the morning either in classroom or computer room conditional upon the school routines. Children wore headphones to reduce noise and distractibility. University students at the master's level were present but did not provide any guidance while children solved the computer tasks even when children asked for support. It was the students' task to prevent and solve technical problems. They logged children in on the website and provided supervision and assistance to ensure that children could complete all sessions. When the supervisor entered the child's name and the system had identified the child, the correct game appeared on screen and the system was programmed in a way that the session automatically discontinued after four games. An off-site helpdesk was available for emergencies. Failed or missed sessions were repeated within one week.

In fall (screening), one month before the 15-week intervention (winter 2006), directly after the intervention (spring 2007), and after 18 months of instruction (April 2009) master's level university students, blind to treatment, tested the children. Assessments were delivered in a fixed order to all participants. Examiners were extensively trained in administration procedures. With the exception of the invented spelling task, pre/post assessment was videotaped.

Table 3  
Correlations among pre- and posttest measures and number of errors in computer tasks

|                               | Time        | 1 <sup>a</sup> | 2 <sup>a</sup> | 3 <sup>a</sup> | 4 <sup>a</sup> | 5 <sup>a</sup> | 6 <sup>a</sup> | 7 <sup>a</sup> | 8 <sup>a</sup> | 9 <sup>b</sup> | 10 <sup>b</sup> | 11 <sup>c</sup> |
|-------------------------------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|-----------------|
| 1. Letter knowledge           | Winter K    | -              | .78**          | .42**          | .56**          | .20*           | .51**          | .58**          | .58**          | .34**          | .41**           | -.42**          |
| 2. Letter Knowledge           | Spring K    |                | -              | .44**          | .65**          | .28**          | .63**          | .65**          | .78**          | .38**          | .45**           | -.58**          |
| 3. Phonological skills        | Winter K    |                |                | -              | .50**          | .27**          | .36**          | .47**          | .59**          | .13            | .29**           | -.41**          |
| 4. Phonological skills        | Spring K    |                |                |                | -              | .22*           | .54**          | .59**          | .80**          | .29**          | .46**           | -.46**          |
| 5. Invented spelling          | Winter K    |                |                |                |                | -              | .34**          | .24*           | .24*           | .25            | .28**           | -.21            |
| 6. Invented spelling          | Spring K    |                |                |                |                |                | -              | .68**          | .69**          | .30**          | .44**           | -.40**          |
| 7. Word recognition           | Spring K    |                |                |                |                |                |                | -              | .69**          | .37**          | .48**           | -.39**          |
| 8. Decoding                   | Spring K    |                |                |                |                |                |                |                | -              | .29**          | .47**           | -.57**          |
| 9. Reading & spelling         | End Grade 2 |                |                |                |                |                |                |                |                | -              | .63**           | -.42**          |
| 10. Reading Comprehension     | End Grade 2 |                |                |                |                |                |                |                |                |                | -               | -.32**          |
| 11. Number of errors in tasks | Spring K    |                |                |                |                |                |                |                |                |                |                 | -               |

Notes. \*\*  $p < .01$ . \*  $p < .05$ . <sup>a</sup>  $n = 110$ . <sup>b</sup>  $n = 98$ . <sup>c</sup>  $n = 73$ .

## **Results**

This study examined the effects of the computer-based intervention Living Letters on at risk kindergarten children's code-related skills within the context of the kindergarten curriculum. Although the correlations between dependent measures were rather high, as can be seen in Table 3, outcome measures were analyzed separately to determine those skills specifically influenced by the program and those that were not. Because intra-class correlations (ICC) ranging from .03 to .24 indicated interdependence of observations within schools, we adapted standard errors within schools with the Huber-White sandwich estimator (cf. Hatcher et al., 2006; Miles, 2006). We used Complex Samples Analyses (General Linear Model) to test the contrast intervention versus control group while we controlled for differences associated with gender, maternal education, nonverbal intelligence, and, when available, pre-intervention levels of performance on the same task.

*Hypothesis 1: The treatment group outperformed the control group in the acquisition of code-related skills in kindergarten.* Pre-test early literacy skills were a significant covariate when assessed whereas maternal education level and gender caused significant effects on invented spelling, word recognition, and decoding and non-verbal intelligence on phonological skills. Table 4 reports effects of the treatment on all dependent measures after controlling for the covariates. A Cohen's  $d$  that equals 1.0 represents a difference of 1  $SD$  between treatment and control group and is equivalent to a strong effect size if  $d$  equals .8, moderate if  $d$  equals .5, and small if  $d$  equals .2 (Cohen, 1988). As is shown in Table 4, effects sizes were small to moderate. They were strongest for word recognition ( $d = .48$ ), followed by phonological awareness ( $d = .47$ ), invented spelling ( $d = .46$ ), and decoding ( $d = .39$ ). For three out of five assessments, namely those tapping invented spelling, phonological awareness skills, and word recognition, the difference between treatment and control group reached significance; however, for decoding the effect was marginally significant ( $p < .07$ ) and for letter knowledge non-significant.

Table 4

*Intervention effects on short term (spring K) and long term (Grade 2), after controlling for background variables and pretest if available.*

| Measure               | Timing   | Estimate (SE) | 95%CI B  | <i>t</i>          | <i>p</i> | Cohen's <i>d</i> <sup>c</sup> |
|-----------------------|----------|---------------|----------|-------------------|----------|-------------------------------|
| Letter knowledge      | Spring K | .02 (.06)     | -.11,.16 | .38 <sup>a</sup>  | .71      | .01                           |
| Phonological skills   | Spring K | .18 (.08)     | .02,.38  | 2.42 <sup>a</sup> | .03      | .47                           |
| Writing               | Spring K | .18 (.08)     | .02,.35  | 2.38 <sup>a</sup> | .03      | .46                           |
| Word recognition      | Spring K | .16 (.07)     | .02,.30  | 2.50 <sup>a</sup> | .03      | .48                           |
| Decoding              | Spring K | .21 (.10)     | -.01,.42 | 2.01 <sup>a</sup> | .07      | .39                           |
| Reading and Spelling  | Grade 2  | .22 (.09)     | .03,.42  | 2.43 <sup>b</sup> | .03      | .50                           |
| Reading Comprehension | Grade 2  | .23 (.08)     | .06,.40  | 2.89 <sup>b</sup> | .01      | .53                           |

*Notes.* <sup>a</sup> *n* = 110, *df* = 14. <sup>b</sup> *n* = 98, *df* = 14. <sup>c</sup> For calculating Cohen's *d* we used the formula  $2t/\sqrt{n-2}$  (Thalheimer & Cook, 2002).

*Hypothesis 2: Gains in code-related skills after the computer treatment sustained beyond the kindergarten year.* The treatment effect was tested for reading and spelling and comprehension after 18 months of reading instruction while controlling for differences associated with gender, maternal education, and nonverbal intelligence. For both reading and spelling, and comprehension skills, the difference between treatment and control group reached significance. Effect sizes (*d*'s) for reading and spelling, and for comprehension were moderate equaling .50 and .53, respectively (see Table 4).

*Hypothesis 3: Low error rate predicted code-related skill acquisition in the treatment group.* On average the 73 children in the treatment group made errors in 8 out of 40 games. In particular the last set of games - identifying words with the sound of the first letter of the proper name - revealed more errors than the games that included recognition of the written form of the name and recognition of the first letter of the proper name between alternatives. The percentage of errors in the last set (48%) was much higher than the percentage in the first two sets (10%). Table 3 shows that correlations with pre-tested scores were moderate; those with post-tested scores ranged from moderate for invented spelling, phonological skills, and word recognition to strong for letter knowledge and decoding; and those with post-post tested scores were moderate as well. After estimating robust standard errors for the dependent measures Complex Samples Regression Analyses were carried out with gender, maternal education, nonverbal intelligence, and number of errors as covariates. As more errors were made children's skills improved less.

As Table 5 shows, the number of errors explained substantial variance, ranging from 7% for phonological awareness to 30% for decoding at post-tests and from 13% for reading and spelling to 21% for comprehension at post-post tests.

Table 5

*Effects of computer behavior on short term (End-K) and long term (Grade 2) posttest scores, expressed as estimates (SE) and increase in explained variance of the model ( $\Delta R^2$ ).*

| Measure             | Timing  | Estimate (SE) | 95%CI B   | p-value | R <sup>2</sup> | $\Delta R^2$ |
|---------------------|---------|---------------|-----------|---------|----------------|--------------|
| Letter knowledge    | End-K   | -.33 (.08)    | -.50,-.17 | .001    | .619           | .076         |
| Phonological skills | End-K   | -.32 (.12)    | -.57,-.06 | .02     | .323           | .066         |
| Writing             | End-K   | -.35 (.10)    | -.57,-.13 | .004    | .177           | .114         |
| Word recognition    | End-K   | -.37 (.10)    | -.59,-.16 | .002    | .124           | .176         |
| Decoding            | End-K   | -.62 (.10)    | -.84,-.41 | .000    | .104           | .300         |
| Reading & Spelling  | Grade 2 | -.44 (.12)    | -.70,-.18 | .003    | .037           | .206         |
| Comprehension       | Grade 2 | -.28 (.15)    | -.61,.04  | .085    | .080           | .126         |

*Note.* R<sup>2</sup> = explained variance of a model including background variables (gender, maternal education, nonverbal intelligence) and pretest if available.

## Discussion

This is one of the few randomized controlled trials testing effects of a remedial computer program at the kindergarten age. In a child sub-sample with code-related skills delays, treated children seemed to thrive on the remedial program. They outperformed an equally delayed treated control group in phonological skills, invented spelling, and decoding of words and nonsense words directly after the intervention period. Children's skill gains are considerable—on the order of a half a standard deviation on average. Effect sizes (Cohen's *d*'s), ranging from .39 for decoding to .48 for word recognition, approximate those reported in an earlier meta analysis of phonemic awareness training programs (overall *d* = .44); see Bus & Van IJzendoorn (1999). More importantly, this study is one of the first showing that an effective and efficient intervention targeted toward gaps in early literacy skills in kindergarten-age is essential to the developmental success of children in the first two grades of primary education. The group at-risk who received the remedial computer program at kindergarten-age scored on average half a standard deviation higher on reading and spelling tests and on reading comprehension after about two years of beginning reading instruction. The program seems to narrow in noticeable way the skills gap by about 8%.

To account for effects comparable to those in more comprehensive and lengthier interventions, we argue that by combining children's understanding of how the proper name sounds with knowledge of how the name looks the computer program acts as a catalyst that stimulates the young child's attention to the grapheme-phoneme relationship in printed words. Once children have accrued baseline knowledge of a small assemblage of letters and sounds they benefit more from the kindergarten curriculum in so far it focuses on training of phonological skills, and they show progress on tests that encompass a set of code-related skills beyond those targeted by the computer program. Yet, experimental and control children were on par in letter knowledge after the intervention period which may indicate that baseline knowledge as promoted by the computer program, is not required for learning associations between letters and sounds/names (National Early Literacy Panel, 2008).

As a result of a program that helps children to combine their understanding of how the name looks with knowledge of how the name sounds in kindergarten-age the treated group seems better prepared to benefit from the reading curriculum in the first two grades of primary education whereas the non-treated group is more vulnerable to a phenomenon referred to as the Matthew effect (Stanovich, 1986). That is, early delays persist until the end of second grade when no interventions are carried out to stem achievement gaps (Juel, 1988). An appropriate program may help to avoid the consequences of these early delays in the learn-to-read process. Compared to children with similar code-related skills delays in kindergarten-age the treated children scored half a standard deviation higher on standardized reading and spelling tests at the end of grade 2, thus narrowing the achievement gap. To account for the long-term effects of a brief intervention, we argue that Living Letters creates a vital *opportunity to learn* how reading works, how letters and words work, and how literacy learning can be applied to new situations (Castles, Coltheart, Wilson, Valpied, & Wedgwood, 2009).

Apart from the computer treatment, there were no differences between treatment and control condition that might explain the treatment group's advantage at the end of grade two; the treated control group and treatment groups were taught by the same teachers and children in both conditions received the same general classroom instruction, in kindergarten and beyond, for about the same amount of time. We can also be confident that the computer program was applied with high implementation fidelity. Due to the storage of children's mouse behavior during the computer sessions the principal researcher immediately noticed when children had skipped one or more sessions. Moreover, the presence of a researcher during the sessions served to monitor that children used the materials as intended and on each occasion completed all aspects of a given activity.

Not all children participating in the treatment, however, realized the full benefits of the software program; those with high task completion on first

attempts outperformed their peers after the training and in second grade. This result is another indicator that the computer program itself and not its side effects explain the short- and long-term effects. This finding makes it less plausible that experimental children outperformed control children because their teachers saw children practice these skills on the computer and therefore decided to offer more practice in code-related skills in classroom. It becomes also less plausible that children assigned to the treatment program practiced more on other occasions stimulated by the computer program.

So far we can only theorize about the causes of these differential effects of the program. The present findings refute the hypothesis that poor prior literacy skills cause more errors and thereby lower performance on outcome measures after the intervention. The relation between errors and outcome measures still exists after controlling for pre-tested skills. It may be more plausible to assume that in particular a sub-sample of children with under-developed regulatory skills profit less from computer programs as described here compared to same-age children with appropriate regulatory skills because they are unable to stay attentive and focused without continuous adult support (Kegel et al., 2009).

## **Limitations**

The current study is, of course, not without limitations. The first is that we assumed that the name is the best pathway through which children develop code-related knowledge, without testing this assumption. Because the name provides surface perceptual features that help children discover sound-symbol relationships between the first letter of the name and its sound in the spoken name, we gave the name a very special function in the program (Ferreiro & Teberosky, 1982). However, the critical test of comparing a program that uses the name as a starting point for code-related knowledge with an alternative program using other words remains to be done.

Second, we did not compare individualized early computer interventions with additional classroom instruction provided by a teacher. The computer format has pivotal qualities that might make it superior to teacher-led interventions (Vernadakis, Avgerinos, Tsitskari, & Zachopoulou, 2005). For instance, the computer program individualizes opportunity to learn and practice around someone's own proper name, the most familiar word to young children. Furthermore, the program not only enables more individual tuition than teachers can provide, it also enables fine-tuning of instruction to the child's skills and it provides a better match for the long stretch of early literacy experiences in literate homes than teachers could realize given instructional demands. Another plus-point of a computer program is that it allows for individual variations, such as

skipping tasks or an increased number of repetitions. However, to prove that a well-founded computer program could make a substantial contribution to the learning environment, other experiments are warranted.

Lastly, when children make relatively many errors in the program's assignments, they benefit less from playing the games. However, a sound explanation for making comparatively many errors in computer assignments resulting in differential effects of the computer program is so far missing.

## **Conclusions**

This study supports the need for remedial programs at the kindergarten-age, and demonstrates that a brief, intensive treatment can improve the capacity to benefit from beginning reading instruction in the early grades (Raudenbush, 2009). Unlike previous long-term studies (Henning, McIntosh, Arnott, & Dodd, 2010), the evidence obtained here indicates that pre-emptive measures in kindergarten can effectively interrupt a potentially downward spiral at an early stage in the learn-to-read process (Stanovich, 1986) although not all eligible children may be susceptible to the intervention to the same extent (Kegel, Bus, & Van IJzendoorn, 2011). A personalized remedial computer intervention with an individualized and consistent feedback loop seems better equipped than whole or small group interventions to prevent reading problems (Byrne et al., 2000; Henning et al., 2010). But this hypothesis awaits further research.