Improving early phoneme skills with a computer program: Differential effects of regulatory skills

Cornelia A.T. Kegel, Verna A.C. van der Kooy-Hofland, Adriana G. Bus *

Department of Education and Child Studies, Leiden University, The Netherlands

A R T I C L E   I N F O
Article history:
Received 17 October 2008
Received in revised form 26 June 2009
Accepted 5 July 2009

Keywords:
Regulatory skills
Phoneme skills
Computer training program
Kindergarten children at risk

A B S T R A C T

Research findings: The study focused on 90 five-year-olds from fifteen Dutch schools. The children scored among the 30% lowest on literacy tests. Half were randomly assigned to a phonological skills program on the computer, the other half to a book program. Both programs consisted of 15 ten-minute sessions. During the phonological skills program children’s mouse behavior was registered every tenth of a second. Intelligence, phoneme skills, and regulatory skills were tested. Children scoring average on regulatory skills benefited from teacher-free encounters with the phonological skills program, children scoring low or high did not. Typically, the lowest-scoring children showed more meaningless mouse activity and more random clicking. Practice or policy: Computer programs can be used to stimulate early phoneme skills of poorly performing kindergarten children, but not for all children. Children with poor regulatory skills did not benefit from the intervention program.

© 2009 Elsevier Inc. All rights reserved.
The study addressed the following questions:

1. Is a computer program intended to fix attention on how written words relate to spoken words effective for children performing poorly on early literacy tests?
2. Do regulatory skills explain differential effects of the computer intervention program beyond verbal and nonverbal intelligence?
3. Are children’s regulatory skills related to their computer behavior during the computer games?

1. Method

1.1. Participants

We selected 90 children out of 404 children from 15 schools. All selected children (a) spoke Dutch as their first language, (b) were 60 to 72 months old, and (c) were among the 30% with the lowest scores on screening tests for early literacy: a letter test, a rhyming test, name writing, and a word dictation test. Eligible pupils were randomly assigned by the main researcher to a condition, with the restriction that boys and girls, and children from the same school, were distributed about equally across the two conditions. Intervention and control groups were similar in age, gender, verbal (Peabody Picture Vocabulary Test; Schlichting, 2005), and nonverbal intelligence (Raven’s Coloured Progressive Matrices; Van Bon, 1986). Groups differed marginally in parental education, t(88) = −1.94, p < .056 (Table 1). Children were very capable of using a mouse to operate the educational software, because computers were in use in the participating kindergarten classes.

1.2. Programs

1.2.1. Living Letters

The Internet program Living Letters, recently made available for schools and parents via subscription (www.Bereslim.nl), is aimed at practicing phoneme skills. The program uses the spelling of a familiar word like the child’s name (Levin, Shatil-Carmon, & Asif-Rave, 2006) to draw attention to phonemes in spoken words (Bus & Van IJzendoorn, 1999; Ehri et al., 2001). The program uses the child’s proper name unless the spelling is inconsistent with Dutch orthography (e.g., ‘mama’; Fig. 1a). The first 20 games practicing the proper name and ‘mama’ (e.g., find your name; Fig. 1a and b) are followed by 10 games on the sound of the first letter (‘which one is the /m/ of mama?’; Fig. 1c), and 10 games to identify pictures that start or end with the first letter of the child’s name or ‘mama’ (Fig. 1d). Each of the last 20 games is played twice in two consecutive sessions, thus constituting two-thirds of the program. All 15 sessions start with an attractive animation using two characters who explain the upcoming games; for instance, the two characters, Sim and Sanne, discover that their names start with the same sound. Errors when solving the games are followed by increasingly supportive feedback. First the task is repeated, next a clue is given (“Do you remember how the teacher writes your name?”), and lastly the correct solution is demonstrated.

1.2.2. Living Books

The control group were given the Internet program Living Books (www.Bereslim.nl), made up of five age-appropriate picture story-books on the computer. As text is orally available, children can “read” individually. Per session, the children “read” one book followed by questions. Each book was repeated three times.

1.3. Procedure

Fifteen sessions of approximately 10 min took place during morning hours at school over a four-month period (February-May). Children sat alone at the computer screen in their classroom or the computer room, with a headset on. Researchers logged children in on the Internet site and made sure they completed all fifteen sessions. A helpdesk was available for emergencies. After entering the child’s name, the correct games appeared automatically and the session was discontinued after four games. When a child received all available feedback a game took at most 90 s extra. Mouse behavior was written to the data store of the provider and saved.

With written consent from parents three assessments took place: a pretest, a mid-test (for testing regulatory skills) after about 8 weeks, and a posttest immediately after the intervention. Each session took approximately 25 min. Testing took place in a room where only the child and examiner were present. In most cases the four Master’s students who did the testing were blind to group allocation. The order of the tests was always the same, except for regulatory skills: computer and paper tests were counterbalanced, as were the tests within the two clusters. All sessions were videotaped and afterwards scored by Master’s students blind for group allocation.

1.4. Measures

1.4.1. Parental education

Parents ticked their highest level of education on a scale ranging from primary education to university (1–8).

1.4.2. Intelligence

To test verbal and nonverbal intelligence we used Dutch versions of the Peabody Picture Vocabulary Test (PPVT; Schlichting, 2005) and Raven’s Coloured Progressive Matrices (RCPM; Van Bon, 1986).

1.4.3. Phoneme skills

1.4.3.1. Word dictation. Five dictated words (i.e., ‘papa’ (daddy), ‘sim’ (boy’s name), ‘been’ (leg), ‘jurrk’ (dress), ‘duim’ (thumb)) were assigned one of the following codes (Levin & Bus, 2003):

- (0) drawing-like scribble
- (1) writing-like scribbles, but not similar to conventional symbols
- (2) conventional symbols not representing sounds in the word
- (3) one phonetic letter
- (4) two or more phonetic letters
- (5) invented spelling (readable but not spelled correctly)
- (6) conventional spelling

The intraclass correlation coefficients for all double-coded words were high (r’s > .99). The scores on the words were averaged (α’s > .84) for pre- and posttest.

Table 1

<table>
<thead>
<tr>
<th>Gender</th>
<th>Living Letters</th>
<th>Control group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n = 45)</td>
<td>(n = 45)</td>
</tr>
<tr>
<td>Male/Female</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Age in months</td>
<td>64.67</td>
<td>3.25</td>
</tr>
<tr>
<td>Parental education (max = 8)</td>
<td>5.41</td>
<td>2.15</td>
</tr>
<tr>
<td>PPVT* (raw scores)</td>
<td>81.36</td>
<td>12.75</td>
</tr>
<tr>
<td>RCPM* (raw scores)</td>
<td>16.09</td>
<td>3.45</td>
</tr>
<tr>
<td>Letter knowledge (max = 10)</td>
<td>3.53</td>
<td>1.36</td>
</tr>
<tr>
<td>Rhyming (max = 12)</td>
<td>9.98</td>
<td>2.19</td>
</tr>
<tr>
<td>Writing mama (max = 6)</td>
<td>2.30</td>
<td>.93</td>
</tr>
<tr>
<td>Writing words (max = 6)</td>
<td>2.24</td>
<td>.81</td>
</tr>
</tbody>
</table>

* PPVT = Peabody Picture Vocabulary Test.
* RCPM = Raven’s Coloured Progressive Matrices.
1.4.3.2. Phoneme identification. Children identified the first sounds of five CVC or CVCC words, the last sounds of five CVC or CCVC words, and named all three or four phonemes of five CVC, CCVC or CVCC words. Cronbach's alphas for pre- and posttest equaled .93 and .92, respectively.

1.4.3.3. Aggregate measure. Principal component analysis (PCA) resulted in one component explaining 63% and 73% of pre- and posttest, respectively. Component loadings ranged from .79 to .86. The distributions of the variables were normal. A higher score indicates better phoneme skills.

1.4.4. Regulatory skills

1.4.4.1. Stroop-like task (dogs). Following the Stroop paradigm, children had to switch rules by responding with an opposite, i.e., saying “blue” to a red dog and “red” to a blue dog (Beveridge, Jarrold, & Pettit, 2002). The task consisted of 96 trials distributed over four conditions, in which demands on working memory (remembering the name of one or two dogs) and inhibition of the most obvious response (e.g., saying “blue” to a red dog) varied. Incorrect naming and corrections were both scored as errors.

1.4.4.2. Stroop-like task (opposites). Children had to respond with the opposite to contrasting pairs of pictures (e.g., saying “fat” to thin) (based on Berlin & Bohlin, 2002). Incorrect naming and corrections were both scored as errors. This test measured working memory (memorising the names of the pictures) and inhibition.

1.4.4.3. Same tapping. The child copied the experimenter's hammer taps on cubes (Leidse Diagnostische Test; Schroots & Van Alphen de Veer, 1976). Each correct imitation in this working memory task was awarded one point with a maximum score of 12.

1.4.4.4. Peg tapping. The child tapped twice with a pencil after one tap by the experimenter, and vice versa (Diamond & Taylor, 1996). The task measures the ability to inhibit a natural tendency to mimic. The total score was the number of correct responses to 16 items. Intraclass correlation coefficients between two independent coders were high for all four tasks ($r > .97$).

1.4.4.5. Aggregate measure. PCA revealed one component with high loadings (.61–.76) explaining 49% of the variance. Because square root and log transformations failed to normalize the measure, children were classified in three groups using quartiles (1 = first quartile, 2 = second and third quartile, 3 = fourth quartile). The distribution of this new variable was normal for both the treatment and the control group.

1.4.5. Computer behavior

From mouse behavior, registered and stored every tenth of a second, we derived the time between the question and the child’s answer, the total time spent on mouse manipulation, number of mouse clicks, and type and number of support needed to solve the tasks. PCA on these four behaviors resulted in one component that explained 78% of the variance and in component loadings beyond .72. The higher the scores on this component, the more children showed problematic computer behavior, i.e., more mouse clicks, mouse movements and mistakes, and longer response time.

1.5. Data analyses

We conducted an ANCOVA with regression techniques to examine the effect of Living Letters on phoneme skills (Cohen, Cohen, West, & Aiken, 2003). Effect-coded Living Letters was entered in the model, after controlling for age, gender, parental education, PPVT, RCPM, and pretest score on phoneme skills. We hypothesized that treatment effects may be strong among children with average regulatory skills but, due to problems with planning and choosing the right steps, treatment may have a reduced impact in groups with poor regulatory skills and, due to ceiling effects, also in groups with high regulatory skills. Therefore, three categories were created: children scoring
among the lowest 25% on regulatory skills (n = 23), children scoring around average (n = 45), and children scoring among the highest 25% (n = 22). The categories were effect-coded by assigning a value of −1 for the base group (here, the 25% highest-scoring). Each of the other categories was assigned a value of 1 for one code variable and 0 for the other (Cohen et al., 2003). Subjects at each level of regulatory skills were appropriately divided to control and treatment groups. By cross-multiplying the coded level of regulatory skills with the coded variable of the treatment program we tested whether the three levels of regulatory skills responded alike or differently to the treatment.

To examine the effects of regulatory skills on computer behavior in the group of 45 children playing the Living Letters games we used a one-way ANCOVA model. Because relatively poor literacy skills may increase the need for feedback, we adjusted the computer behavior for differences on the pretest for phoneme skills.

1.5.1. Missing values

Incidental computer registrations that were lost due to technical problems were imputed by using mean scores within a set. One child was excluded due to too many missing data.

2. Results

2.1. Impact of Living Letters

Table 2 shows the means and standard deviations for phoneme skills and regulatory skills for treatment and control group. The correlations between predictors were mostly low and moderate at most, as shown in Table 3. The final results of an ANCOVA model to test main and interaction effects of Living Letters and regulatory skills on phoneme skills are presented in Table 4. There were no serious problems of multicollinearity (tolerance values >.10). Gender, age, parental education, PPVT, RCPM, and pretest score on phoneme skills were entered as centered continuous variables or effect-coded category (gender) at step 1. The explained variance equaled 35% (F(6, 76) = 6.91, p < .001). By entering effect-coded treatment and regulatory skills in step 2, the increment to R² was 10%, F(3, 73) = 4.19, p < .01, and by entering the interactions between treatment and regulatory skills in step 3, the increment to R² was 5%, F(2, 71) = 3.82, p < .028.

One SD higher on the pretest meant .24 SD higher on the posttest (see Table 4). The group lowest on regulatory skills scored below the grand mean (.33 SD), the intermediate group scored beyond (.24 SD), and the highest group about average (.09 SD higher). The lowest regulatory skill level (RS1) (t = −2.28, p < .05) and the intermediate level (RS2) (t = 2.01, p < .05) differed significantly from the grand mean. The Living Letters group did not score beyond the grand mean (t = 1.32, ns). The significant interaction (RS2 Z) indicates that the Living Letters group outperformed the control group when children scored around average on regulatory skills, as is shown by the predicted values in Fig. 2. The group lowest on regulatory skills scored at a relatively low level at pretest, and their score did not improve as a result of the computer treatment. The group highest on regulatory skills scored just above the grand mean on phoneme skills at pretest, but showed no additional increase as a result of the treatment.

2.2. Computer behavior

Regarding computer behavior we expected especially children scoring lowest on regulatory skills to differ from children scoring in the normal or highest ranges. A planned contrast between the lowest scores and the rest, after controlling for phoneme skills (F(1, 41) = 7.89, p < .01, partial η² = .16), was significant (F(1, 41) = 11.83, p < .01, partial η² = .23 with 95% confidence limits from .11 to .81). Aggregate scores for children low, intermediate, and high on regulatory skills were 1.17, −.14, and −.19, respectively, with high scores indicating more problematic computer behavior.

### Table 3

Correlations for all variables.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age</th>
<th>Edu</th>
<th>PPVT</th>
<th>RCPM</th>
<th>Pre</th>
<th>Post</th>
<th>RS</th>
<th>Comp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>1.00</td>
<td>.19</td>
<td>−.22</td>
<td>−.07</td>
<td>.18</td>
<td>−.22</td>
<td>−.24</td>
<td>−.03</td>
</tr>
<tr>
<td>Edu</td>
<td>1.00</td>
<td>.19</td>
<td>−.09</td>
<td>.10</td>
<td>.10</td>
<td>−.02</td>
<td>.01</td>
<td>−.14</td>
</tr>
<tr>
<td>PPVT</td>
<td>1.00</td>
<td>.09</td>
<td>.39</td>
<td>.25</td>
<td>.25</td>
<td>.22</td>
<td>−.11</td>
<td></td>
</tr>
<tr>
<td>RCPM</td>
<td>1.00</td>
<td>−.04</td>
<td>.10</td>
<td>.03</td>
<td>.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>1.00</td>
<td>.57</td>
<td>−.12</td>
<td>.42</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post</td>
<td>1.00</td>
<td>.21</td>
<td>−.51</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RS</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comp</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For correlations with computer behavior N = 45 and for other correlations N = 90.

*a Correlation is significant at the .05 level (2-tailed).
*b Correlation is significant at the .01 level (2-tailed).

Notes. Gender (−1 = girl, 1 = boy).
Edu = Parental Education.
PPVT = Peabody Picture Vocabulary Test.
RCPM = Raven’s Coloured Progressive Matrices.
Pre = pretest scores of phoneme skills.
Post = posttest scores of phoneme skills.
RS = Regulatory Skills.
Comp = computer behavior.
explains differences in behavioral regulation, as suggested in the
mentioning that the results fail to support the hypothesis that gender
in planning and choosing the right steps.

Our results seem to refute the criticism that regulatory-skills tasks
are not informative about learning problems because they do not relate
to behavior in complex real-world situations (Brown, 1999). Lower
regulatory skills group. In future reports we hope to expand the picture to include
materials that may have outstripped these children’s regulatory skills. The
group scoring lowest on regulatory skills needed more time to
respond, clicked more often, spent more time manipulating the
mouse, and made more mistakes (cf. Bippes et al., 2003). It seemed
that the feedback loops built into Living Letters (e.g., providing cues to
find the correct answer) were insufficient to counterbalance problems
in planning and choosing the right steps.

Likewise, among children with high regulatory skills the computer
treatment was no incentive for phoneme skills. They had relatively
high scores on the pretest (approximately .53 SD above the grand
mean), which may have made Living Letters less challenging to this
group. However, given the fact that these children scored among the
lowest 30% despite relatively strong regulatory skills, we can also
suppose that relatively more subjects in this group are most at risk for
reading problems due to a phonological deficit (Snowling, 2000). This
comes even more plausible when we take into account that this
group showed the same computer behavior as the middle-regulatory
skill group. In future reports we hope to expand the picture to include
follow-up tests in grades 1 and 2.

Our results seem to refute the criticism that regulatory-skills tasks
are not informative about learning problems because they do not relate
to behavior in complex real-world situations (Brown, 1999). Lower
scores on regulatory-skills tests typically coincide with problematic
computer behavior. The alternative to this explanation is that poor
literacy skills cause a need for more feedback, and more time and mouse
clicks to solve the games. However, the present findings refute this
hypothesis because after controlling for pre-tested phoneme skills the
relation between computer behavior and regulatory skills still exists.

Another interesting result is that regulatory skills are predictors of
school success beyond verbal or nonverbal intelligence. It is also worth
mentioning that the results fail to support the hypothesis that gender
explains differences in behavioral regulation, as suggested in the
literature (e.g., Ponitz et al., 2008).

3.1. Future directions and limitations

As the program uses the child’s own name, treatments are somewhat different for most children. This may have an impact on
generalizability. In so far as we could test differential effects by
contrast- ing effects of the proper name (n = 36) with effects of mama
(n = 9), there was no evidence for different outcomes.

Further experiments should consolidate the differential effects of
the program by testing regulatory skills prior to the experiment and by
assigning equal numbers of children with low, intermediate, and high
levels of regulatory skills to treatment and control groups.

The feedback and help options in the present program anticipated
problematic regulatory skills, but were evidently insufficient to scaffold
learning behavior and to correct for uncontrolled mouse behavior and
distraction from the task. Given that the children with poor regulatory
skills did not benefit from the intervention, there clearly is a need to
individualize games by adapting content (e.g., more games practicing the
same) and providing appropriate feedback (e.g., after one or more
ersors starting each new task with a reminder of relevant steps).

3.2. Implications

Computer programs can be used to stimulate early phoneme skills
of poorly performing kindergarten children, although our current
results also point to the weaknesses of computer programs. Children
with poor regulatory skills did not benefit from the computer inter-
vention, probably due to their failure to ignore distracters and to
choose an adequate problem-solving strategy.

The program can also be used as a diagnostic tool to detect poor
regulatory skills as a barrier to learning, thus also making it a valuable
~aching aid.

Acknowledgements

Preparation of this manuscript was supported by an award from the
Dutch organization Kennisnet/ICT op School to Adriana G. Bus.
Development of the computer program Living Letters was made possible by Education through Information and Communication Tech-
ology (EDITC), The Netherlands.

References


Pilot Pientere Peuter. Deel II. Inhoudelijke spelenontwerp [Pilot study Clever Toddler, Part II.
Intrinsic Game Design]. Groningen: Stichting Edict, Molendrift B.V. & Vertis B.V.

belief understanding to emerging math and literacy ability in kindergarten. Child
Development, 78, 647–663.

phoneme awareness training with children of dyslexic parents. In C. Hulme, &
M. Snowling (Eds.), Dyslexia: Biology, cognition and intervention (pp. 235–253).

London: Whurr.

Does the name have a special role in understanding the symbolic function of
writing? Literacy Teaching and Learning, 12, 37–55.

Both-de Vries, A. C., & Bus, A. G. (2009). The proper name as starting point for basic
reading skills. Reading and Writing, 22, 1007:0071114508009858-2.


Brown, T. E. (1999). Does ADHD diagnosis require impulsivity hyperactivity? A respon-
se to Gordon & Barkley. ADHD Report, vol. 7 (pp. 1–7).

A meta-analysis of experimental training studies. Journal of Educational Psychology,
91, 403–414.

Byrne, B. (1998). The foundation of literacy: The child’s acquisition of the alphabetic

correlation analysis for the behavioral sciences. Mahwah, NJ: Lawrence Erlbaum
Associates.


