Protective teaching mechanisms in case of mild perinatal adversity

Merkelbach, I.

Author: Merkelbach, I.
Title: Protective teaching mechanisms in case of mild perinatal adversities
Issue Date: 2018-11-07

Chapter 5

Differential efficacy to a digital numeracy program in kindergarten

Introduction: Children born late preterm have been shown to be highly susceptible to the quality of the educational environment. Because numeracy and math problems are most firmly established in this group of children, this study tests the effects of a digital early numeracy intervention on delayed kindergartners, especially in late preterm children. *Methods:* In a large random controlled trial, preschool pupils (*N*=375, mean age = 67 months, *SD* = 4.50) were assigned to either a digital, guiding and structured numeracy program or a digital control program. Children worked in a classroom environment for a period of approximately two to three months. *Results:* The group as a whole did not benefit from the intervention, but the program benefited a late preterm subsample (n=40, (*Cohen's d* = .71, *CI* = .07 / 1.36). While these pupils fell behind when working with the control program. *Conclusion*: Digital numeracy interventions can support the early numeracy skills of late preterm children. It might be that highly structured and guiding programs sooth this group of children, whom are generally prone to experience high levels of stress reactivity.

In early childhood mathematical abilities develop long before the formal education of such skills starts. Infants can already discriminate between different small numbers of items and can determine numerical equivalence across perceptual modalities (Wynn, 1992). This sense of numbers keeps developing throughout childhood (Griffin, 2004), and before kindergarten starts most children have learned how to count (Wynn, 1990). Such quantity related skills are predictive for later math performance, as is also the case for other early skills, such as logical thinking (Bryant & Nunes, 2002) and visuo(spatial) skills such as recognizing, comparing, and classifying items by shape and size (Aunio & Niemivirta, 2010). This collection of precursors for later math performance is generally referred to as *early numeracy*. Delays in the development of these early numeracy skills can sort long lasting negative effects on the development of mathematical abilities (Desoete, Ceulemans, De Weerdt, & Pieters, 2010). Research shows however that mathematical performance is particularly susceptible to the effects of interventions (e.g. Gervasoni, 2001), especially when implemented at an early age. If we could identify vulnerable children in time, early interventions could thus possibly prevent long term delays.

One such group that has been found particularly vulnerable for experiencing problems with numeracy and mathematics is the group of children born preterm: although these children generally experience more problems in all cognitive domains (e.g. Chyi, Lee, Hintz, Gould, & Sutcliffe, 2008, Woythaler, McCormick, Mao, & Smith, 2015), problems in the field of numeracy and mathematics are most pronounced (e.g. Poulsen, et al., 2013), resulting in adverse outcomes persisting far into adulthood (Basten, Jaekel, Johnson, Gilmore, & Wolke, 2015). Cognitive problems are not only consistently associated with very preterm birth (born < 32 weeks of pregnancy), but also with late preterm birth (between the 34th and 38th week of pregnancy). Even though late preterm birth is 'only' considered a mild perinatal adversity (Van der Kooy-Hofland, Van der Kooy, Bus, Van IJzendoorn, & Bonsel, 2012), these children thus still consistently show higher levels of cognitive problems (Shah, Kaciroti, Richards, Oh, & Lumeng, 2016, Searle, Smithers, Chittleborough, Gregory, & Lynch, 2017) and for example lower IQ-scores (De Jong, Verhoeven, & Van Baar, 2012) than their peers.

Similar outcomes are found in a another group subject to mild perinatal adversity, those born small for gestational age (below the 10th percentile). In childhood and adolosence this group, too, is at risk for experiencing a range of cognitive problems (e.g. Sommerfelt, et al., 2000, Paz, Gale, Laor, Danon, Stevenson, & Seidman, 1995), among which more frequent as well as higher levels of learning disabilities (O'Keeffe, O'Callaghan, Williams, Najman, & Bor, 2003) and poorer school performance (Larroque, Bertrais, Czernichow, & Léger, 2001). For those born small for gestational age, negative associations between mild perinatal adversity and cognitive outcomes persist into adulthood: people who were small for gestational age at birth tend to show lower levels of academic achievement and professional attainment (e.g. Strauss, 2000; Larroque et al., 2001).

5

Although these mild perinatal adversities are associated with increased chances of negative educational and cognitive outcomes, considering mild perinatal adversities as just a vulnarability factor might be short-sighted. Research shows that people who have experienced mild perinatal adversities might be more susceptible to qualities of the environment, for better and for worse as is described in the *differential susceptibility model* (Belsky & Pluess, 2009). This pattern has been found in rearing settings, in which children with mild perinatal adversities have been found to be more susceptible to the influence of maternal harsh parenting (Windhorst, et al., 2017), but also in educational settings. A small scale experiment shows that kindergartners who have experienced mild perinatal adversiteis are more susceptible to a digital early literacy intervention: *Living Letters*, a program training phonological awareness and alphabetical knowledge (Van der Kooy-Hofland et al., 2012). In this experiment, children with literacy delays were exposed to either *Living Letters* or a control program, with which they worked once a week for ten minutes, over the course of two to three months. For children without mild perinatal adversities *Living Letters* had no effect, neither negative nor positive. While on the other hand, children with mild perinatal adversities fell behind even further when they had worked with the control program, but outperformed their peers when they had worked with *Living Letters.*

In this small-scale study (Van der Kooy-Hofland et al., 2012) children with mild perinatal adversities were treated as a homogenous group,supposedly responding similarly to the qualities of an intervention like *Living Letters.* A large scale replication study showed however that this might not be completely correct (Merkelbach, Plak & Rippe, 2018). In this study, as in the Van der Kooy-Hofland et al. study (2012), children worked with either *Living Letters* or a control program over the course of about two to three months. Because the sample size was large, a distinction could be made between children born late preterm and children who were small for gestational age at birth. As children without mild perinatal adversities, children who were small for gestational age were found not to be susceptible to the qualities of *Living Letters*. However, for late preterm children a differential susceptibility pattern was found, as in the Van der Kooy-Hofland et al. (2012) study, these children fell behind when assigned to the control condition, but outperformed their peers when assigned to *Living Letters.*

Main features of *Living Letters* are structure, repetition of assignments, continous and adaptive feedback, and guidance (Merkelbach et al., 2018), which seem to meet the needs of children born late preterm particularly well. In contrast, they might not fully fit the needs of those born small for gestational age. Because numeracy and mathetical problems are more prevalent in children born late preterm than in their full term peers (Nepomnyaschy, Hegyi, Ostfeld, & Reichman, 2012), and such problems are generally fairly susceptible to the effects of interventions (e.g. Gervasoni, 2001), especially when carried out at an early age, we wondered if late preterm kindergartners would benefit from a program similar in design to *Living Letters*, but targeting early numeracy instead of early literacy skills. In addition, we wondered if children born small for gestational age would respond differently to this intervention, as is suggested in Merkelbach et al. (2018).

We therefore tested the results of the digital early numeracy program *Clever Together* in children born late preterm, children small for gestational age, and children without mild perinatal adversities. Like *Living Letters*, *Clever Together* consists out of short games which train several early numeracy skills and which are repeated several times. In addition, *Clever Together* (just as *Living Letters*) also includes digital tutors that offer the child continuous and adaptive feedback, and high levels of guidance and explanation. *Clever Together* highly resembles *Living Letters* in terms of substantive features, as well as in design (e.g. the same digital tutors) and duration and dosage (ten minutes, once a week for two to three months).

Building on earlier findings, we thus expected children born late preterm to benefit from working with *Clever Together* while no results would be expected for children born small for gestational age and children without mild perinatal adversities. To test this hypothesis, we used standardized numeracy tests, that are administered as a standard part of the student tracking system used in the vast majority of Dutch kindergarten and primary schools (Cito, 2017). Pupils were randomly assigned to *Clever Together* and to the control program consisting of digital animated storybooks, to compare their performance.

We hypothesize that:

1. *Clever Together* stimulates early numeracy skills in kindergarten children lagging behind in such skills, however, effects are only present in vulnerable children;
2. Children who are born late preterm will benefit from working with *Clever Together;*
3. Children without mild perinatal adversities and children small for gestational age will not benefit from working with the program.

Method

Design

In the current study, we tested the benefits of a numeracy program, *Clever Together*, in a large group in which delays in numeracy skills were common(N = 375). In a large-scale experiment consisting of two waves (2012/2013 and 2013/2014), delayed kindergarten children were randomly assigned to the control condition, a book-reading program (*Living Books*), or the experimental condition, *Clever Together.* Because children worked with the programs on an individual basis, children from the same classrooms could be assigned randomly to control and treatment conditions. Given the large sample (*N* = 375) and the stability of the percentage of perinatal adversities, it was reasonable to assume that random assignment by the researcher would result in a comparable number of children born late preterm between conditions (Experimental = 10.1%, Control = 11.2%, Table 1). Teachers could not influence the assignment procedure since they merely logged children in. From that moment on children worked individually with their assigned program. Pre-testing and post-testing of early numeracy skills took place as part of the regular monitoring system applied in Dutch kindergarten classrooms, with a standardized numeracy test administered group wise by the teacher, blind for the hypotheses of the study, in January/February of the second kindergarten year and in May/June, just preceding first grade of primary school. Testing in January/February preceded the intervention while the test in May/June was administered directly after the intervention.

Participants

Kindergarten classrooms of 140 elementary schools, both urban and rural, located across the Netherlands, were included in the trial. Kindergarten teachers selected children from their classroom who were delayed in literacy skills for participation; children had to be five or six years old (Mean age: 67.12 months; *SD* = 4.50). After receiving informed consent from a parent, children were randomly assigned to one of the computer programs (in the overarching research three different programs were used). The target group (*N* = 879) in the current trial was randomly assigned to either *Clever Together* or the control program. Children were excluded due to missing data on the numeracy pretest (*n* =50) or numeracy posttest (*n* = 88). Children were also excluded when there was no consent from the mother to retrieve perinatal information from the national perinatal registry (*n* = 268) or if consent was given but the information provided by parents (home address and date of birth of the mother) was incomplete and we were therefore unable to retrieve perinatal information from the registry (*n* = 96). Lastly, children were excluded when information about the educational level of the father (*n* = 4) was lacking. The final sample therefore consisted of 375 children (Figure 2).

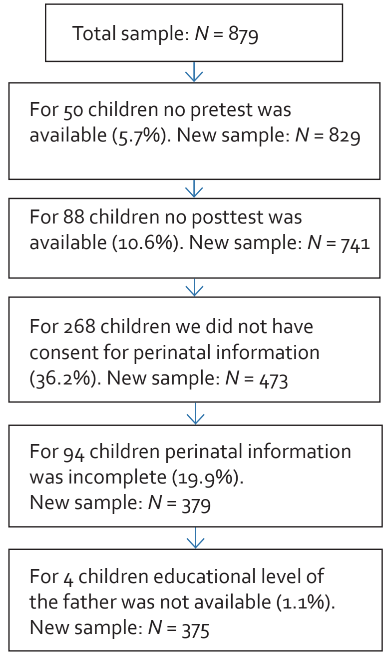


Figure 2. Flowchart of final participant inclusion

Procedure

About 1750 randomly selected schools throughout the Netherlands had received information by mail through brochures and on social media about educational computer programs that might provide extra guidance to kindergarten children who were delayed in academic skills. As a result, 140 schools signed up to participate. Schools were offered three months of free access to educational computer programs that normally require a paid subscription (http://www.bereslim.nl), after completion of the intervention.

Parents provided informed written consent and their email address. Thereafter, parents received a link to a website with frequently asked questions about the project. If they had any further questions, they could also contact the researchers personally (via phone or email). In the first wave, parents' consent for retrieving perinatal information was not a condition for participating in the study. In this wave, parents were asked to provide consent for retrieving perinatal information after the intervention was already completed, because hypotheses regarding differential susceptibility to the educational environment in children with mild perinatal adversities were formulated shortly after collecting general consent. This procedure might explain the high rate of missing perinatal data, which was mainly due to non-consent (67.7%), in that year. In the second year, in reaction to the high rates of missing data, consent for perinatal information was included as a condition for participation in the study and thus for receiving the intervention. This resulted in a much lower rate of missing data (31.7%), largely due to a normal proportion (20%) of matching errors between the registry and our research database.

After assignment to one of the intervention programs, children received access to one of the programs once a week during two to three months, as this was also the dosage used in previous studies (e.g. Van der Kooy-Hofland et al., 2012; Plak, Merkelbach, Kegel, Van IJzendoorn, & Bus, 2016). Children played the games in a classroom setting only receiving adult assistance for logging in. They wore headphones to prevent that the program would attract and distract other children.

Intervention programs

Clever Together. The first 45 tasks (Figure 3) were hide-and-seek games to practice with orienting and locating (e.g. locating objects during hide-and-seek games), as well as with prepositions (e.g. 'in', 'behind'). Sim, one of the characters in the game, asked for help in finding Sanne who is hiding behind one of the objects in the illustration (e.g. "*I am going to hide behind the blue tree*"). In the other 30 games (Figure 3), children had to assemble objects (e.g. a bike) from their parts (e.g. tires, frame, steering wheel), select attributes for an activity (e.g. taking a shower), thereby practicing with spatial prepositions (e.g., 'in', 'behind').

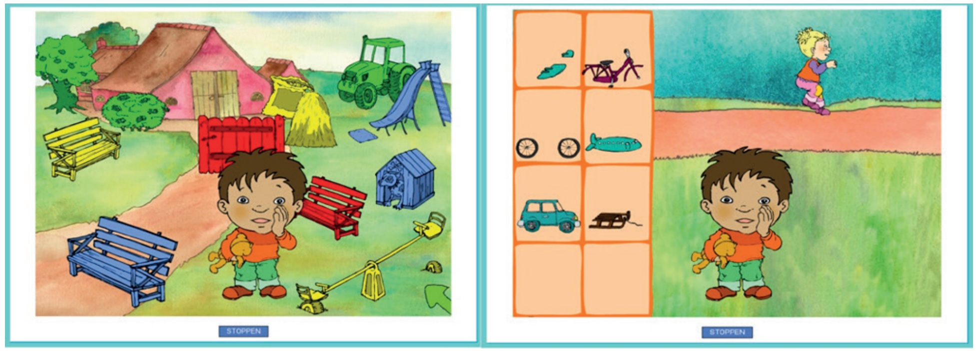


Figure 3. *Clever Together* games: Find Sanne who is hiding behind one of the objects (left) or assemble an object from different parts (right)

In the program, a teddy bear provided responsive replies in a supportive tone. In case of errors a hierarchical set of replies dependent on the child's response was provided (Figure 4). Moreover, assignments that were not answered correctly at a first try were repeated in later sessions followed by similar adaptive feedback to create more opportunities for practicing difficult assignments.

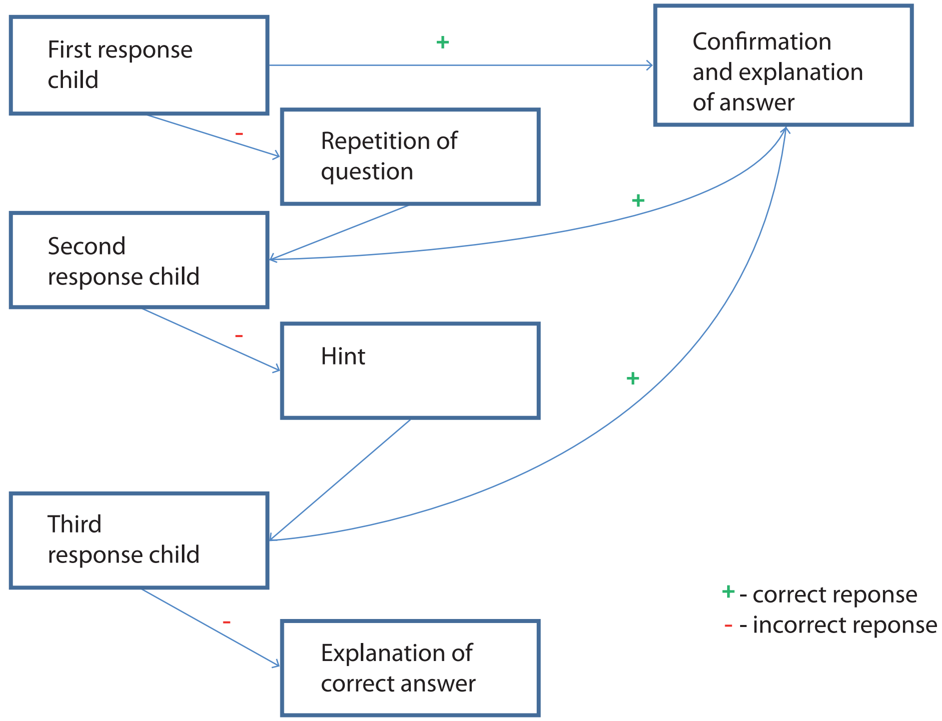


Figure 4. Feedback circle *Clever Together*

Control condition. The control program consisted out of eight digital, age-appropriate, multimedia storybooks with oral text, each read twice. The story text matched the nonverbal, film-like information including animated pictures, music, and sounds. Each story was interrupted four times by digital tutors for a question about difficult words that appeared in the text or about story events, followed by a similar set of hierarchical replies as is offered in *Clever Together*. However, in this program the questions and answers only occupied a small part of the session, about 10% of the total duration.

Measures

Background variables

Children's age (in months) and sex, and the educational level of the father were assessed. Following the rationale of Van der Kooy et al. (2012) that educational level of the father is more strongly associated with mild perinatal adversities than educational level of the mother (as was the case in the current study), we made use of father's educational level instead of that of the mother. The sex and the date of birth of the child

5

were reported by the teacher of the child. The educational level of the father was reported by the parent(s) on a 7-point scale (ranging from no education to university or higher).

Cito Numeracy Skills

The Cito Numeracy Test for Kindergarten Pupils (CNT) is a group-administered standardized numeracy test for kindergarten children administered by the teachers in January/February and May/June in the senior year of kindergarten[1] when children were five to six years of age (Koerhuis & Keuning, 2011). The test consists of 48 questions that focus on number sense (e.g. '*Where do you see three rabbits?*'), classification (e.g. '*Which dog is the biggest?*'), and geometry (e.g. '*Which shadow matches this picture?*'). Based on normative scores, the pretest score of the CNT January/February was dichotomized and coded into scoring at the 40th percentile or lower (0, score lower than 78) or scoring average and above (1, score of 78 or higher). As the posttest, the full range of scores on the CNT May/June was used. Versions of the CNT administered in January/February and the CNT administered in May/June were similar in content and design but included different items.

Perinatal Data

The Netherlands Perinatal Registry (Centraal Bureau voor de Statistiek, 2013) combines data about duration of pregnancy and weight of the child at birth from three registries: the national obstetric database by midwives, the national obstetric database by gynecologists, and the national neonatal/pediatric database (Méray, Reitsma, Ravelli, & Bonsel, 2007) and covers about 96% of all pregnancies in the Netherlands.

Duration of pregnancy was dichotomized into being born full term (0) or being born late preterm (1) which was defined as a gestational age at birth of 34 weeks – 37 weeks + 6 days. Small for gestational age was dichotomized into 'not small for gestational age at birth' (0) and 'small for gestational age at birth' (1), which was defined as lower than the 10th percentile of birth weight for gestational age, taking into account gender and parity.

Statistical analyses

Dependent measures were regressed on the intervention, late preterm birth and small for gestational age, and the interactions between late preterm birth and intervention, and small for gestational age and intervention. For both susceptibility markers a dummy variable was created. Children could thus be in both groups, as was the case for two children. Using the likelihood-ratio test, we tested whether the model fit would improve when we allowed intercepts and slopes to vary across schools (multilevel approach).

[1] In the Netherlands children attend two years of kindergarten. In both years the emphasis is on play instead of formal education.

Because the level of missing data was high (57.3%) we followed Little's (1986) MCAR procedure to see if complete cases analysis would be allowed for. This would be so only when data were missing completely at random (MCAR). In addition, other procedures were applied to account for missing data and models were compared. Models were estimated with a full information maximum likelihood (ML) approach and a multiple imputation (MI) approach. Using ML, models were estimated on the basis of all information available, both from complete and incomplete cases (n=879), yielding the most likely association parameters. Using MI, missing values were imputed (m=100) via chained equations. Applying predictive mean matching, linear and logistic regression prediction where appropriate. The imputation scheme includes all model variables, interactions as well as exogenous variables. Estimates of parameters and standard errors were pooled over imputed datasets, yielding very precise parameter estimates, however potentially showing slightly increased standard errors to account for multiple estimation of missing information. To assess robustness of results, estimates and standard errors were compared between the three approaches (complete case, ML, and MI). Similarity of estimates indicates robustness, while considerable differences signal that results derived from complete case analysis might be strongly biased.

Results

Missing data

Based on Little's MCAR test (1986), we could reject the null hypothesis that data were not missing completely at random (χ^2 = .08, p = .777); therefore it also made therefore sense to apply complete case analysis including only individuals with complete data.

Sample characteristics

Sample characteristics are presented in Table 1. A small majority of children was male (54.9%), in accordance with the general finding that more boys than girls are delayed in the early years of schooling (Gurian, 2010). In this sample of 375 children, 40 children (10.7%) were born late preterm and 94 children were born small for gestational age (25.1%). Sample characteristics did not differ between groups.

Table 1: *Percentages or means (standard deviations) for the complete group of children with complete cases and for the experimental (Clever Together) and control conditions (Living Books) separately (N = 375); p-values for Chi-square or t-test.*

	Complete group (*n* =375)	**Experimental group (*n* =179)**	**Control group (*n* =196)**	***p***
Male	54.9%	55.3%	54.6%	.889
Age (in months)	67.12 (*4.50*)	67.58 (*4.64*)	66.70 (*4.33*)	.060
Father's education (max = 6)	3.74 (*1.50*)	3.72 (*1.50*)	3.77 (*1.51*)	.774
Mild perinatal adversities	32.5%	35.2%	30.1%	.293
Late preterm	10.7%	10.1%	11.2%	.714
Small for gestational age	25.1%	27.4%	23.0%	.324
CNT pretest (max = 137)	78.67 (*10.30*)	79.54 (*11.44*)	77.88 (*9.09*)	.119
Percentage delayed children[1]	45.3%	40.8%	49.5%	.091

1.Below 40^{th} percentile on CNT pretest math

Effects of Clever Together

Table 2 shows the mean standardized outcome scores on the standardized numeracy test for late versus full term children and for the group as a whole.

Table 2. *Means and Standard Deviations for standardized numeracy post-tests by condition and mild perinatal adversities*

	CNT posttest (standardized)			
	Clever Together	***n***	***Living Books***	***n***
Full term	.04 (*1.05*)	161	-.04 (*.93*)	174
Late Preterm	.43 (*1.26*)	18	-.33 (*.88*)	22
Not SGA[1]	.05 (*.97*)	130	-.11 (*.87*)	151
SGA	.16 (*1.30*)	49	.04 (*1.11*)	45
Total	.08 (*1.07*)	179	-.07 (*.93*)	196

1. SGA = small for gestational age

Effects of Clever Together

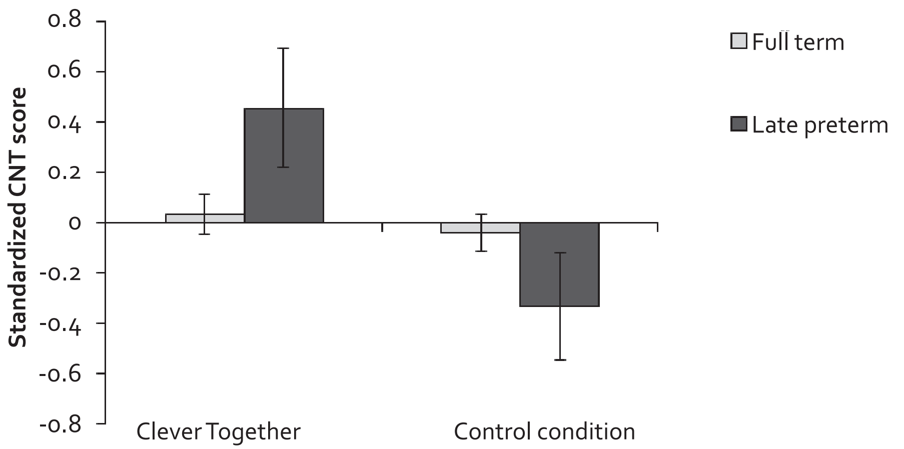
The CNT (June) was regressed on dichotomized numeracy pretest score, late preterm versus full term, small for gestational age vs. normal for gestational age, and the two two-way interaction: late preterm X condition, and small for gestational age X condition. First, we tested whether or not it was necessary to allow the intercept and slope to differ between schools in the regression model (Bickel, 2007). 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 .94. Following a chi-square distribution with one degree of freedom, this difference was not significant (*p* > .10). Likewise, 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 (.00) was not significant (*p* > .10). This indicates that variability in scores on the numeracy test administered after the intervention was similar across schools, therefore a non-hierarchical ordinary least squares (OLS) was applied.

Results of the OLS are presented in Table 3. The CNT pretest (*F* (1, 373) = 171.78, *p* <.001) showed a main effect, children with an average or above score on the pretest scored higher (*Mean* = 93.22 (*11.03*)) than children with a below average score on the pretest (*Mean* = 79.69 (*8.54*)). No main effects were found for late preterm birth (*F* (1, 373) = 2.55, *p* = .111) and small for gestational age (*F* (1, 373) = .31, *p* = .577). There was no significant interaction between small for gestational age X condition (*F* (1, 373) = .38, *p* = .537), however the interaction, born late preterm X condition was significant (*F* (1, 373) = 5.63, *p*=.018). Children born late preterm scored higher on the posttest than their peers when working with *Clever Together* but lagged further behind with *Living Books*, the control condition (see Figure 5). Four CLT scores were outliers (more than three *SD*s above the sample mean).

Repetition of the analysis using MI and ML yielded highly similar results and thus similar substantive conclusions indicating that results derived from complete case analysis were not biased. Estimates and standard errors were highly comparable across all parameters (Supplementary Table 1), including the interaction between late preterm and condition. Estimates for complete cases were: 7.88 (*3.32*); for MI: 7.84 (*3.32*), and for ML: 6.90 (*2.94*).

Table 3. *Numeracy skills regressed on CNT pretest, Clever Together, born late preterm, small for gestational age, and interactions between conditions and mild perinatal adversities (N = 375)*

Measure	**Beta (*SE*)**	***F (df=373)***	**Mean Square**	***p*-value**
Intercept	82.41 (*2.13*)	1499.96	147995.72	<.001
Main effects				
Cohort	-1.67 (*1.12*)	2.55	251.63	.138
CNT pretest	13.55 (*1.03*)	171.78	16949.26	<.001
Clever Together (vs. *Living Books*)	.20 (*1.23*)	.03	2.58	.872
Late preterm	-3.55 (*2.23*)	2.55	251.63	.111
Small for gestational age	.94 (*1.68*)	.31	30.80	.577
Two-way interaction				
Late preterm * *Clever Together*	7.88 (*3.32*)	5.63	555.01	.018
Small for gestational age * *Clever Together*	-1.46 (*2.36*)	.38	37.64	.537



$N_{full\ term/Clever\ Together}$= 161; $N_{preterm/Clever\ Together}$= 18; $N_{full\ term/Control}$ = 174; $N_{preterm/Control}$= 22.

Figure 5. Adjusted, standardized mean scores on the standardized CNT posttest for children born late preterm versus full term assigned to *Clever Together* or the control condition

Effect sizes of the intervention were calculated for the group as a whole and separately for children born late preterm and children born full term (Table 4). For the group as a whole, a small, non-significant, positive effect of *Clever Together* on numeracy skills at the end of senior kindergarten year was found (*Cohen's d* = .15, *CI* = -.05 / .35). In the group born full term, the effect size was close to zero (*Cohen's d* = .08, *CI* = -.13 /.30). However, *Clever Together* produced a large effect in the late preterm group (*Cohen's d* = .71, *CI* = .07 / 1.36).

Table 4. *Effect sizes of Clever Together for the complete group, children born late preterm and children born full term separately.*

		CNT posttest		
Dataset	**group**	***n***	***Cohen's d***	***95% CI***
Complete sample [*](*N* = 375)	Full term	335	.08	-.13 / .30
	Late preterm	40	.71	.07 /1.36
	Total group	375	.15	-.05 /.35

Discussion

Previous studies showed that late preterm children attending kindergarten, although generally at risk for developing academic delays (Chyi et al., 2008), are highly susceptible to a digital early literacy intervention (*Living Letters*); after working with this program they even tend to outperform their peers. This result was typical for late preterm children, while both children with another mild perinatal adversity (i.e. small for gestational age) and children without mild perinatal adversities did not benefit from working with the program (Merkelbach et al., 2018), suggesting that when it comes to susceptibility to the (educational) environment, children with mild perinatal adversities should not be treated as a homogenous group. In line with these results we expected late preterm children also to benefit from *Clever Together*, a digital intervention highly similar to *Living Letters* in approach and design, but targeting a different academic skill: early numeracy. Children born small for gestational age and children without mild perinatal adversities were however not expected to benefit.

Results offer support for the hypothesis that *Clever Together* can boost early numeracy skills, but only in subgroups susceptible to its particular features. No interaction between small for gestational age and condition was found, and when considering the group as a whole, not making a distinction between those with or without perinatal adversities, also no positive effects of *Clever Together* could be distinguished (*Cohen's d* = .15, *CI* = -.05/.35), while late preterm children clearly benefitted from working with the program (*Cohen's d* = .71, *CI* = .07 / 1.36). Consistent with the differential susceptibility model (Belsky & Pluess, 2009), when assigned to the control condition, late preterm children lagged behind as compared to their peers, while they outperformed their peers after having worked with *Clever Together*.

Features central to both *Clever Together* and *Living Letters* thus seem to meet the needs of late preterm children particularly well. Although both programs also show strong resemblance in design, substantive features are most likely to have led to the learning gains experienced by late preterm children. We hypothesize that the high levels of repetition, structure, guidance, and feedback, central to both *Clever Together* and *Living Letters*, might facilitate learning in late preterm children. A positive effect of these features on especially late preterm children is plausible, because preterm birth is associated with increased levels of maternal stress during pregnancy (Dole, Savitz, Hertz-Picciotto, Siega-Riz, McMahon, & Buekens, 2003; Mulder, Robles de Medina, Huizink, Van den Bergh, Buitelaar , & Visser, 2002), which in turn is predictive for increased levels of fearfulness (Pike, 2004) and stress reactivity (Meaney, 2001) in offspring. In an educational environment these features are likely to be expressed as performance- and test anxiety, which are known to have detrimental effects on performance (McDonald, 2001). In normal school settings increased levels of stress reactivity might thus cause children to shut themselves from learning experiences (Van der Kooy-Hofland et al., 2012). In *Clever Together* however the repetition, structure, feedback, and guidance central to the program help clarify the task and hand. Because task clarity is associated with lowering of levels of cardiovascular reactivity to stress (Richter & Gendolla, 2006), we could assume that the features of *Clever Together* (and *Living Letters*) thus result in lower levels of stress through providing high levels of clarity and predictability, thereby facilitating learning in children normally possibly too stressed to effectively process all information presented. This proposed biological mechanism underlying the increased susceptibility to the digital learning environment found in late preterm children, is however still highly speculative. More research is needed to identify which exact features support the learning of late preterm children and why.

Conclusion

The digital early numeracy intervention *Clever Together* can boost the early numeracy performance of kindergartners born late preterm, while other children do not benefit from this intervention. On the other hand, late preterm children fall behind when assigned to a control condition, following the pattern as described by the differential susceptibility model. This pattern does not hold for children born small for gestational age. As a possible explanation for the effectivity of *Clever Together* in preterm children we expect that structure, guidance, and feedback provided by this program might have a soothing effect on children born late preterm, a group expected to experience higher levels of stress reactivity. However, more research is needed to found this speculative hypothesis.

Unavoidably this study has some limitations. We tested the effect of late preterm birth on the results of a digital program in the field of numeracy, expecting a similar result as was found for a literacy program. However, it should be noted that the studies looking into effects of *Living Letters* (Merkelbach et al., 2018) and *Clever Together* (current study) are not completely independent. In both studies the same control condition was used, thus including largely the same sample of children. Additionally, the current overarching study (in which thus three digital programs were included: *Clever Together*, *Living Letters*, and the control condition) has been designed with the prime purpose of exploring and stimulating early literacy skills. Therefore, teachers selected children on the basis of possible problems in the field of early literacy, instead of in the field of early numeracy. Children in this study who experienced problems in the field of numeracy, are thus those children who experience problems in both domains. These children might, even though

literacy- and numeracy skills are highly correlated (in total sample: $r = .589$, $p < .001$), differ from children who only experience problems in the field of numeracy. Another limitation is that we can only speculate about effective functionalities in *Clever Together* and the biological mechanisms explaining this effectivity.

Although details of the *Clever Together* numeracy intervention need further study, we can conclude that those children having experienced mild perinatal adversity, but only when born preterm, benefit from numeracy training via *Clever Together*.

References

Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, 427-435.

Basten, M., Jaekel, J., Johnson, S., Gilmore, C., & Wolke, D. (2015). Preterm birth and adult health. *Psychological Science*, 1608-1619.

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

Bickel, R. (2007). *Multilevel Analysis for Applied Research It's just regression!* New York: The Guilford Press.

Centraal Bureau voor de Statistiek. (2013). *Documentatierapport Perinatale Registratie (PRN).*

Chyi, L., Lee, H., Hintz, S., Gould, J., & Sutcliffe, T. (2008). School outcomes of late preterm infants: Special needs and challenges for infants born at 32 to 36 weeks gestation. *The Journal of Pediatrics*, 25-31.

De Jong, M., Verhoeven, M., & Van Baar, A. (2012). School outcome, cognitive functioning, and behaviour problems in moderate and late preterm children and adults: a review. *Seminars in Fetal and Neonatal*, 163-169.

Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2010). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, 64-81.

Dole, N., Savitz, D., Hertz-Picciotto, I., Siega-Riz, A., McMahon, M., & Buekens, P. (2003). Maternal stress and preterm birth. *American Journal of Epidemiology*, 14-24.

Gervasoni, A. (2001). Special programs for children who are low attaining in mathematics: do they help? *24th Annual MERGA conference*.

Griffin, S. (2004). Building number sense with Number Worlds: a mathematics program for young children. *Early Childhood Research Quarterly*, 173-180.

Gurian, M. (2010). *Boys and girls learn differently! A guide for teachers and parents.* San Fransisco: Jossey-Bass.

Koerhuis, I., & Keuning, J. (2011). *Wetenschappelijke verantwoording van de toetsen rekenen voor kleuters.* Arnhem: Cito.

Larroque, B., Bertrais, S., Czernichow, P., & Léger, J. (2001). School difficulties in 20-year-olds who were born small for gestational age at term in a regional cohort study. *Pediatrics*, 111-115.

Little, R. (1986). Survey nonresponse adjustments for estimates of means. *International Statistique Review*, 139-157.

McDonald, A. (2001). The prevalence and effects of test anxiety in school children. *Educational Psychology*, 89-101.

Meaney, M. (2001). Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual Review of Neuroscience*, 1161-1192.

Méray, N., Reitsma, J., Ravelli, A., & Bonsel, G. (2007). Probabilistic record linkage is a valid and transparent tool to combine databases without patient identification number. *Journal of Clinical Epidemiology*, 883e1-883e11.

Merkelbach, I., Plak, R., & Rippe, R. (2018). Reproducibility of young learners' susceptibility to the learning context. *Learning and Individual Differences*, 167-175.

Mulder, E., Robles de Medina, P., Huizink, A., Van den Bergh, B., Buitelaar , J., & Visser, G. (2002). Prenatal maternal stress: effects on pregnancy and the (unborn) child. *Early Human Development*, 3-14.

Nepomnyaschy, L., Hegyi, T., Ostfeld, B., & Reichman, N. (2012). Developmental outcomes of late-preterm infants at 2 and 4 years. *Maternal and Childen Health Journal*, 1612-1624.

O'Keeffe, M., O'Callaghan, M., Williams, G., Najman, J., & Bor, W. (2003). Learning, cognitive and attentional problems in adolescents born small for gestational age. *Pediatrics*, 301-307.

Paz, I., Gale, R., Laor, A., Danon, Y., Stevenson, D., & Seidman, D. (1995). The cognitive outcome of full-term small for gestational age infants at late adolosence. *Obstetrics & Gynecology*, 452-456.

Pike, I. (2004). Maternal stress and fetal responses: Evolutionary perspectives on preterm delivery. *American Journal of Human Biology*, 55-65.

Plak, R., Merkelbach, I., Kegel, C., Van IJzendoorn, M., & Bus, A. (2016). Brief computer interventions enhance emergent academic skills in susceptible children: a gene-by-environment experiment. *Learning and Instruction*, 1-8.

Poulsen, G., Wolke, D., Kurinczuk, J. J., Boyle, E. M., Field, D., Alfirevic, Z., & Quigley, M. A. (2013). Gestational age and cognitive ability in early childhood: A population based cohort study. *Paediatric and Perinatal Epidemiology*, 371-379.

Richter, M., & Gendolla, G. (2006). Incentive effects on cardiovascular reactivity in active coping with unclear task difficulty. *International Journal of Psychophysiology*, 215-225.

Searle, A., Smithers, L., Chittleborough, C., Gregory, T., & Lynch, J. (2017). Gestational age and school achievement: a population study. *Archives of Disease in Childhood: Fetal & Neonatal Edition*, 409-415.

Shah, P., Kaciroti, N., Richards, B., Oh, W., & Lumeng, J. (2016). Developmental outcomes of preterm infants from infancy to kindergarten. *Pediatrics*, 1-10.

Sommerfelt, K., Andersson, H., Sonnander, K., Ahlsten, G., Ellersten, B., Markestad, T., . . . Bakketeig, L. (2000). Cognitive development of term small for gestational age children at five years of age. *Archives of Disease in Childhood*, 25-30.

Strauss, R. (2000). Adult functional outcomes of those born small for gestational age: twenty-six-year follow-up of the 1970 British Birth Cohort . *Journal of the American Medical Association*, 625-632.

Van der Kooy-Hofland, V., Van der Kooy, J., Bus, A., Van IJzendoorn, M., & Bonsel, G. (2012). Differential susceptibility to early literacy intervention in children with mild perinatal adversities: short- and long-term effects of a randomized controlled trial. *Journal of Educational Psychology*, 337-349.

Windhorst, D., Rippe, R., Mileve-Seitz, V., Verhulst, F., Jaddoe, V., Noppe, G., . . . Bakermans-Kranenburg, M. (2017). Mild perinatal adversities moderate the association between maternal harsh parenting and hair cortisol: Evidence for differential susceptibility. *Developmental Psychobiology*, 324-337.

Woythaler, M., McCormick, M., Mao, W., & Smith, V. (2015). Late preterm infants and neurodevelopmental outcomes at kindergarten. *Pediatrics*, 424-431.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, 155-193.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 749-750.