

# Tracking the long-term effects of the Bookstart intervention: Associations with temperament and book-reading habits

Merel G. de Bondt<sup>a</sup>, Adriana G. Bus<sup>b,c,\*</sup>

<sup>a</sup> Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV, Amsterdam, the Netherlands

<sup>b</sup> University of Stavanger, Faculty of Arts and Education, 4036 Stavanger, Norway

<sup>c</sup> ELTE Eötvös Loránd University, Hungary

## ARTICLE INFO

### Keywords:

Bookstart  
Book-reading habits  
Temperament  
Language development  
Differential susceptibility

## ABSTRACT

This study tested whether Bookstart – a program promoting book reading in infancy – continues to have an impact well into Kindergarten. We distinguished between children who were more or less challenging to read to in infancy (more or less temperamentally reactive). Eighty percent ( $n = 471$ ) of a sample participating in a study when the children were one year old – about half involved in Bookstart – agreed to complete a home literacy survey when the children were, on average, 72.1 months. A smaller group ( $n = 318$ ) also consented to collect tests concerning language and math at children's Kindergarten. The findings show that language development when they are about to start learning to read still profits from Bookstart. Especially the temperamentally most reactive 50 % shows benefits ( $d = 0.21$ ). Bookstart also improved children's home literacy environment (longer book reading sessions), but this effect did not explain Bookstart's impact in Kindergarten.

## 1. Introduction

This study focuses on the long-term effects of families' participation in Bookstart. Bookstart aims to promote language and literacy skills — the best predictors of school achievement (e.g., Barnes & Puccioni, 2017; Niklas, Cohrssen, & Tayler, 2016; Pace, Burchinal, Alper, Hirsh-Pasek, & Golinkoff, 2019). The program's primary mechanism is an early start with book reading, preferably in the first year of children's life. Shared book reading is considered one of the most potent stimuli for language and literacy skills (Dickinson & Morse, 2019; Golinkoff, Hoff, Rowe, Tamis-LeMonda, & Hirsh-Pasek, 2019). Compared to other daily activities involving verbal exchanges between parent and child, book reading is a strong incentive for language because parents use richer language than during other activities (e.g., Crain-Thoreson, Dahlin, & Powell, 2001; Sosa, 2016). In addition, picture books are “lexical reservoirs” containing more scarcely used and more sophisticated words and grammar than parent-child dialogue (Debaryshe, 1993; Demir-Lira, Applebaum, Goldin-Meadow, & Levine, 2019; Dickinson, Griffith, Golinkoff, & Hirsh-Pasek, 2012; Majorano & Lavelli, 2014; Montag, 2019; Montag, Jones, & Smith, 2015). Moreover, besides the language input through the narratives, book reading stimulates adults to converse on topics of interest to the children (e.g., Flack, Field, & Horst, 2018).

To promote an early start with book reading, Bookstart parents in the Netherlands receive, apart from valuable tips for sharing books with babies, two baby books for free and (free) access to the local library with a vast collection of books for the very young. The books and folder are packed into a toy case with the Bookstart logo (a Dick Bruna drawing of a parent reading to a child), which presence in the home may remind parents of book reading. It may thus “nudge” parents to initiate and maintain book reading routines (De Bondt, Willenberg, & Bus, 2020). The effect studies in the last three decades in which the program is in use suggest that Bookstart children show a relatively high interest in books and score higher on vocabulary tests when they are around three years old (e.g., High et al., 2014; Jones et al., 2000). A meta-analysis (De Bondt et al., 2020) discusses Bookstart's effects on literacy-promoting aspects of the home environment and children's reading and language skills. The combined set of meta-analyzed studies, mainly focusing on children from zero to three years old and only one involving 5-year-olds (Wade & Moore, 1998), revealed significant effects on the home environment ( $k = 10$ ,  $d = 0.25$ ) and children's reading and language skills ( $k = 5$ ,  $d = 0.23$ ).

Many countries across various continents invest in Bookstart or similar programs. The long-term effects of such programs on children's language and literacy development are of interest to decide whether this

\* Corresponding author at: Houtlaan 7, 2334 CJ Leiden, the Netherlands.

E-mail address: [adriana.g.bus@uis.no](mailto:adriana.g.bus@uis.no) (A.G. Bus).

<https://doi.org/10.1016/j.lindif.2022.102199>

Received 9 July 2021; Received in revised form 20 July 2022; Accepted 26 July 2022

Available online 11 August 2022

1041-6080/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

kind of light-touch parent support in early childhood – lacking, for instance, any personal coaching – is worth the material and financial investment. Aiming at an early start with book reading across the whole population is less worth the investment if there is no beneficial effect in the long run. For example, when parents find it hard to share books with infants (Lin, Reich, Kataoka, & Farkas, 2015), there may be an early vocabulary lag (e.g., High et al., 2014; Jones et al., 2000; van den Berg & Bus, 2014). However, due to regularly hearing stories between three and five years, the delay in language development due to less book reading in the first two years may disappear. In that case, there is no need to stimulate an early start with book reading. Unfortunately, we miss well-designed experiments testing that adverse effects of a late start are still present when children are five to six years (De Bondt et al., 2020). This study addressed this gap.

### 1.1. Differential susceptibility

In testing the effects of Bookstart, we need to consider that program effects may depend on the child's biological and psychological characteristics and the parents' reactions to those (Bronfenbrenner, 1979). We may fail to find any program impact if we do not consider children's behavior that influences parental behavior and thus determines the program's role in the home environment. For instance, we conjecture Bookstart is less valuable when parents are inclined to start early with reading to young children. So, Bookstart might be particularly relevant if parents are not convinced that the youngest children need book reading, as may be the case in low-literate families (e.g., Dickinson & Morse, 2019; Durham, Farkas, Hammer, & Tomblin, 2007). But also, literate parents may be hesitant to read to their children, as is known to occur with so-called temperamentally reactive children (Karrass, Van-Deventer, & Braungart-Rieker, 2003; Machida, Taylor, & Kim, 2002; Vernon-Feagans et al., 2008). Temperamental traits manifest early in life and contribute to individual differences in regulating and modulating emotion, attention, behavior, and motor activity (Gartstein & Rothbart, 2003; Rothbart, 1981). For example, proneness to sadness, anger, and frustration may make temperamentally reactive children's behavior challenging for their parents. As a result, parents may be less inclined to initiate verbal exchanges with the children (e.g., Dixon & Smith, 2000; Karrass, Braungart-Rieker, Mullins, & Lefever, 2002). Unintentionally, they may thus create a less optimal learning environment with no or less book-reading (e.g., Banerjee & Tamis-Lemonda, 2007; Garelo, Viterbori, & Usai, 2012; Spinelli, Fasolo, Shah, Genovese, & Aureli, 2018; Usai, Garelo, & Viterbori, 2009). As a result, their children are more at risk for language delays (e.g., Dixon & Smith, 2000; Karrass et al., 2002; Usai et al., 2009).

In a sample with mainly middle to high educated parents, we expect that, in particular, this group of temperamentally reactive children may benefit from Bookstart stimulating parents to start early with book-reading despite their children's temperament style. The outcome may be that temperamentally reactive infants reach the same score on language skills as less reactive peers (Zuckerman, 1999), which would align with the diathesis-stress model (e.g., Belsky, Bakermans-Kranenburg, & Van IJzendoorn, 2007). It predicts that vulnerable children catch up and perform at their peers' level due to an intervention. However, an alternative differential susceptibility model predicts partly different outcomes (e.g., Belsky et al., 2007; Kegel, Bus, & van IJzendoorn, 2011; Plak, Merkelbach, Kegel, van IJzendoorn, & Bus, 2016). There is evidence in the literature that the temperamentally reactive group, more dependent on an enriched environment, is at the same time more malleable (e.g., Blair, 2002; Poehlmann et al., 2012; Widaman et al., 2012). They not only catch up due to an intervention, but these children may be able to acquire a higher achievement level under more optimal learning conditions and even outperform their temperamentally less reactive peers. In other words, the susceptible group is more sensitive to context, for better and for worse.

A previous study mainly including middle to high-educated parents

(van den Berg & Bus, 2014) testing the effect of Bookstart on infants' vocabulary at 15 months confirmed the hypothesis that temperamentally reactive infants are more at risk for language delays. Without Bookstart, their parents appeared less inclined to read to the infants who lagged behind their less reactive peers in vocabulary when they were 15 months old. However, in line with the differential susceptibility model, predicting that temperamentally reactive infants are more vulnerable and, at the same time, more malleable, their vocabulary outperformed the less reactive Bookstart peers. This finding confirms that Bookstart had a protective effect when children were at risk of a less stimulating environment and delayed vocabulary development due to their temperament. At the same time, these temperamentally more reactive children were more responsive, and Bookstart enabled them to outperform their peers. However, we could not entirely exclude that other conceivable samples might respond by the diathesis-stress model; temperamentally reactive children catch up due to Bookstart but do not outperform their less reactive peers.

In the present study, we tested Bookstart's effects in the long run in the same sample as in the 2014 study, taking into account the differences in temperament in infancy. The control and Bookstart group gaps found at 15 months might have dissolved due to book-reading and other language experiences since the assessment at 15 months. However, it is also possible that differences remain, perhaps because the early book reading experiences may impact later book reading habits and thus language development. More language proficiency at 15 months may increase children's interest in book-reading, and more print exposure may consolidate or even increase initial differences in language skills. We tested whether, just as at 15 months, scores at 5–6 years align with the differential susceptibility model.

### 1.2. Book-reading habits explaining long-term effects of Bookstart

As children experience more book reading in the first two years, they may, as a result of more language input, know more words and be more efficient in processing familiar words in real-time (e.g., Weisleder & Fernald, 2013). Knowing more words, they may enjoy stories more and be more eager to engage in book-reading after the first two years (e.g., Raikes et al., 2006). An initial advantage in language may bring children into a 'flow,' causing a continuous higher language input through book-reading, which may stimulate their language proficiency (Dawdall et al., 2020). Even at first glance, minor discrepancies in the frequency of book reading between three and five may have enormous consequences for language input. For example, extrapolating from the average length of picture books, Logan, Justice, Yumus, and Chaparro-Moreno (2019) estimated that reading daily and a few times a week could reach an input difference of nearly a million words over four years.

In this line of argumentation, any long-term Bookstart effects might be mediated by book-reading habits between three and five. Furthermore, in line with differential susceptibility, we expect, in particular, the more malleable, temperamentally reactive children to have the best chance to come in a flow, causing a continuous higher language input through book reading. Since these children are more susceptible to the environment, they may profit more from book-reading and, more than their peers, get into an upward spiral. To test this hypothesis, we questioned their parents about the home literacy environment when children were 5–6 years old and tested whether differences in book-reading habits mediate long-term Bookstart effects on language and literacy skills.

### 1.3. The present study

We carried out the present follow-up study to find evidence for the hypothesis that Bookstart continues to impact language and literacy skills well into Kindergarten. We sought to test the following hypotheses:

1. Bookstart has long-term effects but not in the whole group; the temperamentally highly reactive children show impact but not their less reactive peers.
2. In line with the differential susceptibility model, the temperamentally reactive group stays behind in the control condition but outperforms the peers in the Bookstart condition.
3. Besides uniquely impacting language and early literacy skills such as vocabulary, listening comprehension, and phonemic awareness, Bookstart may affect broader executive functions, which would also improve other skills such as math.
4. Bookstart may lead to more frequent and more prolonged reading at 5–6 years, and such differences in book-reading habits may (partially) explain any long-term effects of Bookstart.

## 2. Material and methods

### 2.1. Design

We invited the complete sample involved in a study starting when children were about eight months and with a final test at 15 months (van den Berg & Bus, 2014) to participate in this study aiming at testing long-term Bookstart effects when children were at the edge of beginning primary education. The original experimental group received a Bookstart package – a baby case with two baby books and information about the importance of starting early with book reading – and free library membership for their child. The researchers recruited the control group through 35 child health care centers in municipalities that had not yet introduced the Bookstart program. When their children were around eight months old, parents completed a questionnaire, including a Dutch version of the Infant Behavior Questionnaire – revised (IBQ-r) to assess their temperament (Gartstein & Rothbart, 2003). Approximately five years later, we invited the same participants to complete a questionnaire targeting the frequency of shared book reading and other literacy-promoting aspects of the home environment, like the number of children's books and bedtime routines. Furthermore, we asked parents' permission to obtain language tests, early literacy tests, and math tests from the children's Kindergarten.

### 2.2. Procedure

We contacted parents via an email that included a personalized link to the digital survey, including the request to contact the children's Kindergarten for test results. Parents could receive the survey in print, but no participants wanted a print version. We sent a reminder up to three times if parents had not responded to the previous invitation. Data collection with the online survey took place from November 2017 until January 2018. We combined the data about children's group membership (Bookstart versus control group) and their temperament in infancy with the recent survey about the home literacy environment and standardized tests in Kindergarten to answer the research questions.

All procedures performed were per the institutional and national research committees' ethical standards and were acceptable according to the 1964 Helsinki declaration and later amendments or comparable ethical standards. The Institutional Review Board of the Vrije Universiteit Amsterdam approved the research protocol.

### 2.3. Participants

From the original sample ( $n = 584$ ), 80 % agreed ( $n = 471$ ) to complete the follow-up survey when the children were five to six years. The attrition partly explained by outdated email addresses was comparable with the attrition rate in other longitudinal studies (e.g., Sénéchal & LeFevre, 2002). Three hundred and eighteen parents completing the follow-up survey (67.5 %) gave informed consent to collect tests from children's Kindergarten. Parents who consented to request test results at school did not differ in parental education, age,

and family composition from those who did not consent. However, boys' parents refused consent more often than girls ( $p = .037$ ). The percentage giving consent (66.5 % and 69.1 %, respectively) did not differ between Bookstart and control group ( $p = .714$ ). In some cases, we did not receive useful test data. The Kindergarten used no tests ( $n = 30$ ) or another test than the Cito Language test taken in January of senior Kindergarten year ( $n = 22$ ) or a similar test but taken earlier or later in children's school career ( $n = 28$ ). Sometimes schools did not provide data after repeated requests ( $n = 11$ ) or sent incomplete data insufficient for inclusion ( $n = 18$ ). Four parents withdrew their initial consent realizing that retrieving the test results increased the teacher's burden. Children who completed the Cito Language Test for Kindergarten ( $n = 205$ ) mostly also completed the Cito Math test in the same year ( $n = 198$ ). Fig. 1 shows the flow diagram of the data collection and the reasons for attrition.

Within the sample ( $n = 471$ ), 62 % ( $n = 293$ ) had participated in Bookstart and 38 % ( $n = 178$ ) in the control group. As Table 1 shows, the two groups were equal in sex ( $p = .911$ ), family composition ( $p = .447$ ), and Dutch as primary language ( $p = .471$ ). However, the control group children were slightly older ( $p = .031$ ), and the education level of Bookstart mothers was higher than in the control group ( $p = .001$ ). Concerning parental education, the sample did not reflect the Dutch population. The lowest education level was under-represented in our sample. Only a small group (0.8 %) reported that they only finished primary education, whereas this percentage is much higher (8.8 %) in the total Dutch population (CBS [Statistics in the Netherlands], 2021). However, there were no differences in children's age ( $p = .499$ ) and the mothers' education level ( $p = .501$ ) between the experimental and control groups for which the literacy tests were available. Almost all children completing the Cito Language Test completed the Cito Math Test ( $n = 198$ ). Only seven children were missing. The top half of Table 1 presents scores for the complete sample and the bottom half for this subsample.

### 2.4. The Bookstart intervention

The Bookstart intervention started thirty years ago in the UK and is implemented in about 30 countries in Europe, Asia, North America, Central and South America, and Oceania (see for a specification of affiliates: <https://www.booktrust.org.uk/about-us/booktrust-affiliates>). Their number is still growing. Families receive free books through health care centers, librarians, or Bookstart coordinators, and tips for parents on sharing books and finding resources for new books for the very young. The program is open to all families and not just disadvantaged families. For example, in the Netherlands, young parents are contacted through the municipality, sending families a letter to inform them about a present at their local library approximately three months after the youngest child's birth. The gift is an attractive small case containing two baby books- one soft cloth and a cardboard book. The package also includes a flyer promoting an early start with shared book reading and tips for sharing books with babies. When the parents collect the case at the library, the child becomes free of charge a member of the local library, providing access to an extensive collection of infant and children's books in an attractive room, offering parents and children the opportunity to play and share books. The program does not include coaching or other book gifts.

### 2.5. Instruments

#### 2.5.1. Assessments at eight months

**2.5.1.1. Temperament.** When children were eight months old, parents completed 22 items of the Dutch version of the Infant Behavior Questionnaire – revised (IBQ-r), a widely used questionnaire about infant temperament. The items describe child behavior in parent-child interaction (e.g., smiling, fussing, crying) on an 8-point scale (ranging from

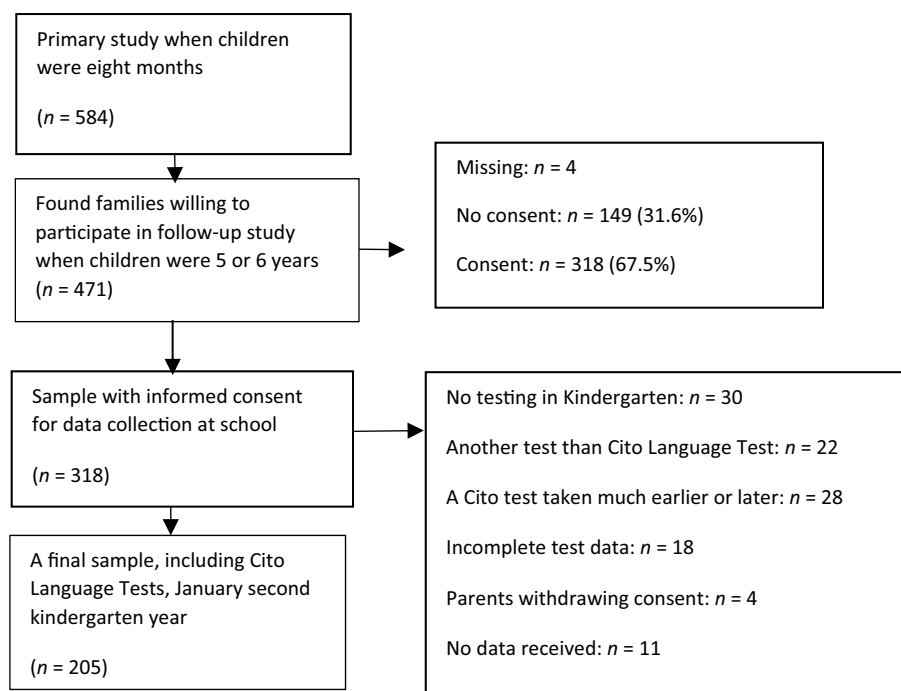


Fig. 1. Flow diagram of data collection.

'always' to 'not applicable'). The questionnaire translated by M. Roest-de Zeeuw and K. van Doesumis was validated in a Dutch study (Klein Velderman, Bakermans-Kranenburg, Juffer, & van IJzendoorn, 2006). A Principal Component Analysis (PCA) was conducted on all items of the IBQ-r and resulted in two factors. The first component (*compliant temperament*) consisted of six items involving children's positive reactions to bathing, washing their faces and hair, laying on their backs, and playing together. Loadings ranged from 0.56 to 0.75. This factor considerably overlaps with Rothbart's smile and laughter scale (Gartstein & Rothbart, 2003; Rothbart, 1981). Cronbach's alpha reliability equaled 0.77. We summarized the scores resulting in a scale ranging from 0 to 42. The second component (*reactive temperament*) consisted of six items reflecting adverse reactions like whining, crying, fussing, squirming, kicking during feeding, washing, carrying the child on the arm, or putting the child to bed. Loadings ranged from 0.58 to 0.76. The Cronbach's alpha reliability equaled 0.67. Higher scores on this scale ranging from 0 to 42 indicated higher temperamental reactivity. Both scales, compliant temperament and reactive temperament, were used in the data analysis.

## 2.5.2. Assessments at 5–6 year

### 2.5.2.1. Survey of the home literacy environment.

We created the digital survey in Qualtrics. The questionnaire included questions about the home literacy environment. We targeted the following features of the target child's home literacy environment: the start of book-sharing, the number of children's books in the household and their location (bedroom or living room), library membership, frequency of library visits (rarely, once per two months, once a month, weekly), participating in book-reading sessions at the library or the bookstore, occasional daytime book reading (one or multiple times a day, a few times per week, incidentally), daily bedtime book-reading sessions (yes/no), other bedtime routines (singing songs, playing games, watching TV or playing a game on tablet or smartphone), duration of reading sessions (1–15 min, 15–30 min, 30–60 min, 60–120 min), and the child's interest in reading independently (one or multiple times a day, one or multiple times a week, incidentally). Parents completed the online survey of the home literacy environment from November 2017 until January 2018,

when their children were five to six years old.

### 2.5.2.2. The Cito Language Test for Kindergarten [Taal voor Kleuters].

The most frequently available test in the current sample's Kindergartens was the Cito Language Test for Kindergarten, the version applied in January in the second Kindergarten year (Lansink & Hemker, 2012). The Committee for Test Quality in the Netherlands [Commissie Testaangelegenheden Nederland] rated the test adequately (Egberink, Leng, & Vermeulen, 2010). In the senior Kindergarten year, teachers can administer this standardized test twice a year (January and June) group-wise (whole class or smaller groups). It includes 60 multiple-choice items with three or four answer alternatives drawn in color, among which one best represents the answer to the question (see Figure 2). The test assesses receptive and expressive vocabulary, story comprehension, rhyming, identifying the first and last word in sentences, sound blending, writing conventions (e.g., reading from left to right), and predicting book content using the book cover. The test comprises two parts assigned on different days or with a long pause, each taking approximately 20 to 30 min. The teachers read aloud the text above the picture in Fig. 2, and the children have ample time to answer the questions. Then, after testing, the teacher marks the wrong answers and calculates the raw and standardized scores. We asked teachers to send us the raw score. Unfortunately, not all schools used this test, and it was, therefore, not available for all children whose parents consented to obtain test results at their school. Our attempts to convert the most common alternative tools' outcomes into a score comparable to the Cito test were unsuccessful.

### 2.5.2.3. The Cito Mathematics Test for Kindergarten [Rekenen voor Kleuters].

To test whether the impact of Bookstart affects broader skills such as executive functions, which would improve other skills such as math, we also collected data about mathematics. The Cito Mathematics Test for Kindergartners (Koerhuis & Keuning, 2011) measures children's emerging numeracy ability. The test targets number comprehension (the number sequence, quantities, numbers), counting (height, volume, weight, and time), and geometry (orienting and locating, constructing, and operating with shapes and figures). The test consists of two parts of respectively 23 and 24 multiple-choice items with three or four answer

**Table 1**

Participant background characteristics and test scores for the total sample (top half) and the subsample for whom standardized test scores were available (bottom half).

Total sample N = 471	Bookstart n = 293	Control n = 178	Statistics
Boys/girls (n)	153/140	92/86	$\chi^2(1) = 0.013, p = .911$
Age in months (M, SD)	71.7 (3.0)	72.4 (3.0)	$t(469) = 2.161, p = .031$
Percentage two-parent families	93.9	95.5	$\chi^2(1) = 0.577, p = .447$
Percentage speaking Dutch as primary language	95.1	98.3	$\chi^2(1) = 3.547, p = .471$
Education level of mothers <sup>1</sup> (M, SD)	3.7 (0.59)	3.5 (0.63)	$t(460) = -3.344, p = .001$
Positive reactivity (M, SD)	30.0 (6.1)	30.0 (5.5)	$t(427) = -0.006, p = .996$
Negative reactivity (M, SD)	28.2 (5.8)	28.1 (6.4)	$t(434) = -0.053, p = .958$

Cito Language test available n = 205	Bookstart n = 122	Control n = 83	Statistics
Boys/girls (n)	65/57	45/38	$\chi^2(1) = 0.017, p = .859$
Age in months (M, SD)	71.8 (2.9)	72.0 (2.7)	$t(203) = 0.678, p = .499$
Percentage two-parent families	93.4	96.4	$\chi^2(1) = 0.842, p = .359$
Percentage speaking Dutch as primary language	94.9	98.8	$\chi^2(3) = 2.624, p = .453$
Education level of mothers <sup>1</sup> (M, SD)	3.8 (0.61)	3.6 (0.64)	$t(199) = -0.673, p = .501$
Positive reactivity (M, SD)	29.1 (5.9)	29.6 (5.8)	$t(196) = 0.535, p = .593$
Negative reactivity (M, SD)	28.0 (5.7)	29.0 (6.2)	$t(193) = 1.385, p = .168$
Cito Language Test (raw score) (M, SD)	71.70 (12.5)	69.80 (9.4)	$t(203) = -1.25, p = .215$
Cito Mathematics (raw score) <sup>2</sup> (M, SD)	90.8 (13.3)	90.3 (12.7)	$t(196) = -0.296, p = .768$

<sup>1</sup> Scale ranged from 1 (only primary education) to 5 (university degree).

<sup>2</sup> n = 198.

alternatives. The administration of each part takes about 20 to 30 min. Testing and coding are similar to the Cito Language Test for Kindergarten.

2.6. Analyses

To test the long-term effects of Bookstart and which children benefited, in particular (*Hypothesis 1*), we regressed the Cito Language Test for Kindergarten on the predictors: Bookstart (dummy coded), temperament (scores on compliant and reactive temperament), and interactions between temperament and Bookstart. In the first step, we entered the two covariates (sex and the mothers' education level); in the next step, the two temperament scores and Bookstart; in the third step, the interactions between the two temperament scores and Bookstart. The two temperament scores, compliant and reactive, were mean-centered (Aiken, West, & Reno, 1991).

Next, we further inspected, if available, the condition-by-temperament interaction to test whether the data pattern aligns with the differential susceptibility model (*Hypothesis 2*). We expected the temperamentally high reactive group to benefit from Bookstart and the low reactive group not. Projecting the condition (Bookstart versus control) on the x-axis and the language test scores on the y-axis, the two lines for the high and low reactive groups should cross at some point within the range of x-values. Next, we regressed the language test on Bookstart separately for the high and low reactive groups. The slope for the susceptible subgroup (the high reactive group) should be significantly different from zero and, at the same time, considerably steeper than the slope for the non-susceptible subgroup — the low reactive group (Belsky et al., 2007). Finally, we tested whether the crossing point and confidence intervals around the crossing point fell within the 0–1 range on the x-axis. We used Widaman's Extra Material for SPSS (Widaman et al., 2012) to estimate the crossover point and confidence interval.

Next, we repeated the analyses for the Cito Math Test to test Bookstart's effects on math skills (*Hypothesis 3*).

Finally, we tested whether Bookstart and the control group differed in book reading habits and whether habits showing a difference mediated the relationship between Bookstart and outcome measures in Kindergarten (*Hypothesis 4*). We conducted a series of regressions using the SPSS process macros (Baron & Kenny, 1986; MacKinnon, Fairchild, & Fritz, 2007; Preacher & Hayes, 2008). We followed the recommendations from Rucker, Preacher, Tormala, and Petty (2011) and tested whether (a) the independent variable (Bookstart) related to the mediator (book reading habit); and (b) the mediator (book reading habit) related to the dependent variable (outcome measures in Kindergarten). A resampling method was used to yield percentile confidence intervals of the total effect of indirect effects. We thus test whether mediator variables add significantly to the model and mediate the impact of independent variables on outcomes.



Fig. 2. One of the 60 Cito Language Test items. The teacher reads the text above the picture aloud: Where do you see pick up? Underline the picture.

### 3. Results

#### 3.1. Interaction effect between reactive temperament and Bookstart

A Shapiro-Wilk's test and a visual inspection of histograms, normal Q-Q plots, and boxplots showed that the score on the Cito Language test for Kindergarten was approximately normally distributed. Skewness and kurtosis equaled 0.286 ( $SE = 0.219$ ) and  $-0.585$  ( $SE = 0.4$ ) for the Bookstart condition, and 0.165 ( $SE = 0.264$ ) and 0.392 ( $SE = 0.523$ ) for the control condition. For the mean-centered reactive temperament, skewness equaled  $-0.246$  ( $SE = 0.09$ ) and kurtosis  $-0.342$  ( $SE = 0.179$ ), and for the mean-centered compliant temperament, skewness equaled  $-0.359$  ( $SE = 0.09$ ) and kurtosis  $-0.279$  ( $SE = 0.18$ ). Cook's distance, tolerance, and VIF satisfied conditions for multiple regression on the Cito Language Test for Kindergarten. The scatterplot of standardized predicted values versus standardized residuals showed that the data met variance and linearity homogeneity assumptions, and the residuals were approximately normally distributed.

We regressed the Cito Language Test for Kindergarten on the sex of the child and the mothers' education (*first step*), Bookstart (dummy-coded) and two mean-centered temperament scores (*second step*), and the interaction between Bookstart and two indicators of temperament (*third step*); see Table 2 for the results. The first step revealed a significant effect ( $R^2 = 0.067$ ,  $F(2, 192) = 6.86$ ,  $p = .001$ ) to which both covariates contributed. Girls scored higher ( $\beta = 0.18$ ,  $t(192) = 2.57$ ,  $p = .007$ ) and children performed better as mothers were higher educated ( $\beta = 0.19$ ,  $t(192) = 2.67$ ,  $p = .018$ ). After entering compliant temperament and reactive temperament and Bookstart, the model did not significantly improve,  $R^2$  change = 0.024,  $F(3, 189) = 1.67$ ,  $p = .176$ . However, the interactions entered in the third step improved the model,  $R^2$  change = 0.031,  $F(2, 187) = 3.26$ ,  $p = .040$ . The Bookstart X reactive temperament interaction was significant,  $\beta = 0.17$ ,  $t(187) = 2.45$ ,  $p = .015$ . The Bookstart X compliant temperament not,  $\beta = 0.04$ ,  $t(187) = 0.497$ ,  $p = .62$ .

#### 3.2. Examining the condition-by-temperament interaction

We used a median split of reactive temperament to check the condition-by-temperament interaction visually. The raw-score equation in Fig. 3 provided evidence for an ordinal interaction with a crossover point within the dummy coded condition (0 = control, 1 = Bookstart) on the x-axis as expected when the outcomes fit with differential susceptibility. The scores suggest that the high reactive temperament group outperformed the low reactive temperament if children participated in Bookstart. But results were vice versa in the control condition, where the high reactive temperament group remained behind.

Next, we performed simple regressions per temperament group to determine the steepness of the slopes for Bookstart, controlling for the two covariates (sex and mothers' education). We tested whether the slope for the susceptible subgroup was significantly different from zero and steeper than the non-susceptible subgroup's (Cohen, Cohen, West, & Aiken, 2013). There was no significant effect of Bookstart in the group

scoring relatively low on reactive temperament,  $\beta = -0.15$ ,  $p = .439$ ,  $d = -0.08$  (blue line in Fig. 3). However, Bookstart positively influenced the Cito Language Test for Kindergarten when children scored relatively high on reactive temperament (red line in Fig. 3),  $\beta = 0.43$ ,  $p = .029$ ,  $d = 0.21$ , explaining an additional 4 % of the variance.

We used Widaman et al.'s (2012) Extra Material to test whether the crossover point may fall outside the boundaries of the dummy coded condition and whether alternative samples may not show the same pattern. Table 3 summarizing the results with Widaman et al.'s (2012) Extra Material indicates the crossover point fell within the 0–1 range of Bookstart,  $\hat{C} = 0.28$  ( $SE = 0.30$ ), but the Confidence Interval fell partly outside the 0–1 range, 95 % CI [ $-0.32, 0.87$ ]. The upper limit (0.87) fell within Bookstart's range, but the lower limit ( $-0.32$ ) did not. So, we can expect the same pattern for the Bookstart condition in the population, but the control condition may differ.

#### 3.3. Effects of Bookstart on the mathematics test

We tested whether the impact of Bookstart was broader than language and early literacy skills and ran the same regression on the Cito Mathematics test for Kindergarten. This sample included 198 children, overlapping with the group that had a score on the Cito Language Test for Kindergarten. Bookstart had no main effect ( $p = .694$ ), and the interaction between Bookstart and reactive temperament was nonsignificant ( $p = .231$ ).

#### 3.4. Book Reading habits as a mediator between Bookstart and long-term effects

Concerning monthly library visits, the number of books in the home, and book reading routines, the home literacy environment was the same in the Bookstart and control condition (see the upper half of Table 4). However, five variables showed differences (an early start with book reading, library membership, taking part in reading sessions at the library and the bookstore, and length of reading sessions), all favoring the Bookstart group. The first three variables were corollaries of the Bookstart program and, therefore, no mediators. Possible mediators are attending reading sessions at the book store and the reading session duration. The difference in session duration – 15–30 min versus <15 min – seems slight. However, a difference of 5 min means one extra picture book each session, which will grow into a substantial amount of language over a longer period.

The only group eligible for mediation analysis was the temperamentally highly reactive group, as they showed long-term effects. Therefore, in this sub-sample showing similar differences in the home literacy environment as the whole group except for taking part in reading sessions at the bookstore (see the lower half of Table 4), we tested whether session duration mediated the relationship between Bookstart and the Cito Language Test for Kindergarten. As we used bootstrapping and robust standard errors, there was no need to worry about normality and homoscedasticity. The scatterplots between each predictor and the dependent variables showed that the data met the

**Table 2**  
Results of regressing the Cito Language Test for Kindergarten on predictor variables.

	B	SE B	$\beta$	95 % CI		p
				LL	UL	
Constant	57.15	9.69		38.04	76.26	<.001
Mothers' education	3.11	1.32	0.19	0.50	5.71	.018
Sex (0 = F, 1 = M)	4.01	1.58	0.18	0.91	7.12	.007
Bookstart (0 = control, 1 = Bookstart)	-16.90	10.82	-0.74	-38.24	4.44	.168
Reactive temperament	-0.10	0.20	-0.05	-0.50	0.29	.071
Compliant temperament	0.07	0.22	0.04	-0.37	0.51	.279
BookstartXReactive	0.60	0.27	0.17	0.07	1.13	.015
BookstartXCompliant	0.08	0.28	0.04	-0.47	0.63	.620

Note. LL = lower level. UL = upper level.

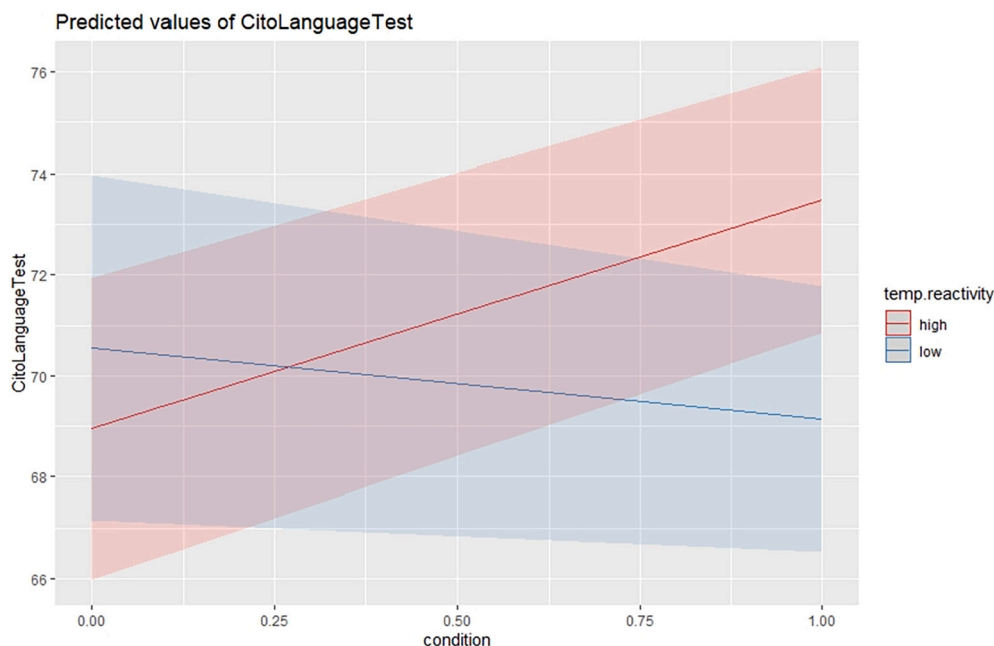


Fig. 3. Plots of the regression lines and their 95 % confidence intervals when we regressed the Cito Language Test on condition (0 = control versus 1 = Bookstart) with reactive temperament as the grouping variable (high versus low); we controlled sex and education (covariates).

**Table 3**  
Results for standard and re-parameterized regression models for language skills.

Standard parameterization		Re-parameterized model		
Parameter	Raw scores	Parameter	Crossover centered	95 % CI
$B_0$	60.93(6.60)	$A_0$	70.18(1.04)	68.13, 72.23
$B_1$	-0.11(0.18)	$B_1$	4.41(2.02)	0.43, 8.40
$B_2$	-14.62(7.11)	$C$	0.28(0.30)	-0.32, 0.87
$B_3$	0.59(0.24)	$B_2$	-1.37(2.20)	-5.72, 2.97

assumption of linearity.

Bookstart was a significant predictor of the Cito Language Test for Kindergarten (c-path; point estimate = 5.11 ( $SE = 2.308$ ),  $t(98) = 2.21$ ,  $p = .029$ ). The covariates sex of the child and the mother's education were nonsignificant ( $p$ 's > .05) (sex of the child; point estimate = 3.84;  $t(95) = 1.80$ ,  $p = .074$ ; mother's education level; point estimate = 4.01,  $t(95) = 1.81$ ,  $p = .073$ ). The c-path remained significant if we entered Bookstart simultaneously with duration of reading sessions (c'-path; point estimate = 4.43 ( $SE = 2.19$ );  $t(96) = 2.03$ ,  $p = .046$ ), indicating that the latter variable was not a complete mediator. The a-path from participation in Bookstart to the duration of reading sessions was positive but only approached significance (a-path; point estimate = 0.20 ( $SE = 0.11$ );  $t(97) = 1.83$ ,  $p = .070$ ). The b-path from duration of reading to outcomes on the Cito Language Test for Kindergarten was also positive but not significant (b-path; point estimate = 2.67 ( $SE = 1.96$ );  $t(96) = 1.36$ ,  $p = .176$ ). In other words, there was no indirect effect ( $IE = 0.54$ ) from Bookstart on the Cito Language Test for Kindergarten via duration of reading sessions (point estimate = 0.54,  $SE = 0.59$ , 95 % CI [-0.23, 2.02]).

#### 4. Discussion

##### 4.1. Bookstart effects

Bookstart has no long-term impact on language and literacy skills in the complete sample. However, in the 50 % with a relatively reactive temperament in infancy, we find long-term Bookstart effects, explaining 4 % of language and literacy skills variation at 5–6 years (*Hypothesis 1*).

Despite Bookstart being a light-touch intervention in the first two years of life, the 50 % showing relatively challenging behavior in infancy in interactions with their caregivers benefits from Bookstart well into Kindergarten.

The data pattern at 15 months aligning with differential susceptibility (van den Berg & Bus, 2014) is replicated five years later — at the edge of learning to read (*Hypothesis 2*). We find the temperamentally highly reactive group to benefit from Bookstart ( $d = 0.21$ ) but not the low reactive group ( $d = -0.08$ ). The pattern corroborates differential susceptibility (Belsky et al., 2007; Kegel et al., 2011; Plak et al., 2016): the temperamentally more reactive children exhibit poorer outcomes in the control condition without incentives to start reading early. However, when parents receive incentives to begin early with book reading (Bookstart), the more reactive children show outcomes that are superior to the less reactive.

Even though the pattern reported at 15 months aligns with the findings five years later, there are also some shifts in the outcomes. In the control group, the gap between high and low reactive children decreased, maybe because learning conditions for the two groups became more similar between two and five. During that period, most children growing up in the middle to high-educated families are involved in verbal interactions and hear stories if not at home in school. The more malleable group – the temperamentally reactive children – thus have a chance to catch up. In the Bookstart group, by contrast, the gap widens. At 15 months, the temperamentally reactive Bookstart children scored higher but not significantly. Five years later, they score considerably higher than their less reactive peers on language and literacy skills. As the early start with book-reading and early language skills raise reactive children's interest in book-reading, this more pliable group may come in a 'flow,' causing a continuous higher language input through book reading.

We found a long-term impact of Bookstart but only for language and literacy skills, not math, even though we assessed both at about the same age for about the same group of children. This finding corroborates the theory that Bookstart uniquely influences language-related skills (*Hypothesis 3*). There is no evidence for more general effects of Bookstart due to the parent being more involved in children's activities and providing more guidance leading to more knowledge or better executive functions.

**Table 4**  
 Characteristics of the home literacy environment in the Bookstart and control group in the whole group (upper half) and the sub-sample of highly reactive children (lower half).

Activity N = 471	Bookstart n = 293	Control n = 178	
An early start of book reading	55.0 %	40.1 %	$\chi^2(1) = 9.247, p = .002$
Child membership library	95.2 %	79.2 %	$\chi^2(1) = 29.389, p = .001$
Reading sessions at the library	54.6 %	38.8 %	$\chi^2(1) = 11.126, p = .001$
Reading sessions at the bookstore	15.7 %	7.3 %	$\chi^2(1) = 7.124, p = .008$
Monthly library visits	31.1 %	33.7 %	$\chi^2(4) = 9.244, p = .055$
26–50 children's books at home	47.4 %	50.0 %	$\chi^2(3) = 1.780, p = .619$
Children's books in the bedroom	96.2 %	92.7 %	$\chi^2(1) = 2.884, p = .089$
Bedtime book reading routine	83.3 %	83.1 %	$\chi^2(1) = 0.001, p = .971$
Occasional daytime book reading	79.9 %	76.4 %	$\chi^2(2) = 3.435, p = .180$
Usually lengthy reading sessions (15–30 min)	51.5 %	32.0 %	$\chi^2(4) = 21.314, p = .001$

n = 116	Bookstart n = 68	Control n = 48	
An early start of book reading	58.2 %	39.5 %	$\chi^2(1) = 7.31, p = .007$
Child membership library	95.1 %	80.2 %	$\chi^2(1) = 12.77, p = .001$
Reading sessions at the library	57.7 %	42.9 %	$\chi^2(1) = 4.93, p = .026$
Reading sessions at the bookstore	16.9 %	12.1 %	$\chi^2(1) = 1.01, p = .316$
Monthly library visits	35.2 %	35.2 %	$\chi^2(4) = 6.12, p = .190$
26–50 children's books at home	47.2 %	48.4 %	$\chi^2(3) = 4.35, p = .226$
Children's books in the bedroom	97.9 %	93.4 %	$\chi^2(1) = 3.00, p = .083$
Bedtime book reading routine	85.2 %	90.1 %	$\chi^2(1) = 1.19, p = .276$
Occasional daytime book reading	84.5 %	75.8 %	$\chi^2(2) = 2.91, p = .234$
Usually lengthy reading sessions (15–30 min)	51.4 %	33.0 %	$\chi^2(4) = 9.73, p = .021$

4.2. Mediating role of book reading habits

We also tested whether Bookstart improves book reading habits at 5–6 years and whether those explain the long-term effect of Bookstart in the temperamentally highly reactive 50 % (Hypothesis 4). Control and Bookstart groups report the availability of books and ample access to children's books in the homes. Furthermore, it seems typical for that age to regularly read to children — >85 % report reading daily bedtime stories to their children, and >70 % occasional daytime book reading. However, in the Bookstart group, the duration of reading sessions is longer. It may look slightly different — Bookstart parents mostly read 15 to 30 min/session and control families <15 min. However, it is, in fact, a difference that may considerably impact children's experiences. A daily five-minute difference would mean one extra picture book per day, adding more than a thousand books over three years, which can have enormous consequences for the language input (Logan et al., 2019).

The hypothesis that the highly reactive group's long-term effects on language and literacy result from longer book-reading sessions (Hypothesis 4) is not confirmed. We failed to prove that longer book-reading sessions “snowball” temperamentally reactive children's language development. Instead, we argue that early emerging language skills

increase children's interest in linguistic activities such as book reading between three and five and may result in more reading and a faster increase in language skills. A survey completed by the parents when the children were 5–6 years old showed differences in session duration, indicating that children display more interest in book-reading and may read more books. However, we could not prove that the longer duration of book-reading sessions has a mediating role and partly or wholly explains the long-term effects of Bookstart in this group.

We probably need more specific information about book-reading and children's responses to it to learn which qualities mediate between Bookstart and language and literacy skills at the edge of learning to read. Apart from the book-reading sessions' duration, other variations may contribute to the steeper growth curve in the temperamentally reactive group. For instance, children may be more enthusiastic about book reading and, because of that, more attentive. Alternative measures for assessing the home literacy environment, such as young children reflecting on literacy practices in their homes (Evans & Hulak, 2019), may give insights into the book-reading qualities that are more telling than surveys completed by parents.

4.3. Limitations

As always, this study has limitations that complicate interpreting the results. First, participation in Bookstart implies self-selection. Bookstart parents indeed lived in municipalities where the program was operational, but they had a choice to participate or not (Sharif, Rieber, Ozuah, & Reiber, 2002). So, assuming that only parents with some interest in book reading agreed to join Bookstart, self-selection may have reduced the natural variation in book reading.

Furthermore, the mother's educational level was higher in the Bookstart than in the control group. Despite a statistical correction, this difference may have favored the Bookstart group, thus interfering with a critical test of Bookstart.

Bookstart is free, but the lowest-educated families seem less inclined to join. Consequently, families in need of interventions like Bookstart do not participate, thus reducing the variance that Bookstart explained in the reported study. On the other hand, the program might have shown a more substantial effect when less educated families joined.

The outcome pattern for more and less reactive children seems robust in the Bookstart group but not in the control group. So, for replication studies, we can expect that in the Bookstart group, the temperamentally highly reactive children outperform less reactive children, but in the control group, children may not differ (Widaman et al., 2012). In that case, findings indicate that Bookstart does not prevent a language proficiency lag but does help temperamentally reactive children to achieve optimally.

We relied on the parental report about the home literacy environment. Parents might have given social-desirable responses about book-reading habits, possibly reducing variation and thus correlations. Other ways of assessing the home literacy environment, for instance, interviewing the children (Evans & Hulak, 2019), might shed a different light on whether Bookstart enhances the book reading routines in the long run and if those routines are mediators.

4.4. Practical implications

The findings align with the hypothesis that an early start with book reading is essential for language and literacy skills (cf. Dickinson & Morse, 2019; Raikes et al., 2006) and will improve children's academic success and success in life (Golinkoff et al., 2019). We find evidence for the relevance of a light-touch intervention like Bookstart, stimulating caregivers to promote language development by actively sharing stories starting early. So, it is a necessary policy to remind young parents of the relevance of book reading and the need to keep trying to begin early, even if their child often responds negatively to their attempts.

Another interesting finding is that an intervention targeting



language development is valuable in a sample that mainly includes middle- to high-educated families, thus supporting the Bookstart policy to invite all young parents to participate. The downside of this all-inclusive policy is that the typical Bookstart approach may not allude to the needs of all parents. The method is somewhat distal, not involving personal contact with caregivers, which may not work well for less literate parents. As they may feel discomfort with reading to very young children or not know how to read with an infant (Justice, Logan, & Damschroder, 2015; Lin et al., 2015), personal contact and direct help might be indispensable. A recent meta-analysis (De Bondt et al., 2020) indicates the possible benefits of multiple personal contacts of a particular kind with a health care professional.

The current findings highlight how essential the differential susceptibility concept is to evaluate intervention programs. We would not have included children's temperaments in the final model if we had not been aware of this model, and we would not have found any long-term effects of Bookstart. Targeting the temperamentally more reactive children enabled us to make visible that Bookstart is indispensable to prevent unnecessary lags in many young children's language and literacy development.

### Declaration of competing interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was supported by the Spinoza grant to Marinus H. van IJzendoorn and a grant of the Reading Association (Stichting Lezen) to Adriana G. Bus.

### References

- Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. New York, NY: Sage.
- Banerjee, P., & Tamis-LeMonda, C. (2007). Infants' persistence and mothers' teaching as predictors of toddlers' cognitive development. *Infant Behavior and Development, 30*, 479–491. <https://doi.org/10.1016/j.infbeh.2006.12.001>
- Barnes, E., & Puccioni, J. (2017). Shared book reading and preschool children's academic achievement: Evidence from the Early Childhood Longitudinal Study—Birth Cohort. *Infant and Child Development, 26*(6), Article e2035. <https://doi.org/10.1002/icd.2035>
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Belsky, J., Bakermans-Kranenburg, M. J., & Van IJzendoorn, M. H. (2007). For better and for worse: Differential susceptibility to environmental influences. *Current Directions in Psychological Science, 16*(6), 300–304. <https://doi.org/10.1111/j.1467-8721.2007.00525.x>
- Blair, C. (2002). Early interventions for low birth weight, preterm infants: The role of negative emotionality in the specification of effects. *Development and Psychopathology, 14*, 311–332. <https://doi.org/10.1017/S0954579402002079>
- Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Centraal Bureau voor de Statistiek [Statistics Netherlands] (CBS). (2021). Bevolking; onderwijsniveau en migratieachtergrond 2003-2021 [Working population; education level, and migration background 2003-2021]. <https://opendata.cbs.nl/statline/#/CBS/nl/dataset>. Retrieved from.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York, NY: Routledge.
- Crain-Thoreson, C., Dahlin, M. P., & Powell, T. A. (2001). Parent-child interaction in three conversational contexts: Variations in style and strategy. *New Directions for Child and Adolescent Development, 2001*(92), 23–38. <https://doi.org/10.1002/cd.13>
- Dawdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P. J. (2020). Shared picture book reading interventions for child language development: A systematic review and meta-analysis. *Child Development, e383*–e399. <https://doi.org/10.1111/cdev.13225>
- De Bondt, M., Willenberg, I. A., & Bus, A. G. (2020). Do book giveaway programs promote the home literacy environment and children's literacy-related behavior and skills? *Review of Educational Research, 90*(3), 349–375. <https://doi.org/10.3102/0034654320922140>
- DeBaryshe, B. D. (1993). Joint picture-book reading correlates of early oral language skills. *Journal of Child Language, 20*(2), 455–461. <https://doi.org/10.1017/S0305000900008370>
- Demir-Lira, Ö. E., Applebaum, L. R., Goldin-Meadow, S., & Levine, S. C. (2019). Parents' early book reading to children: Relation to children's later language and literacy outcomes controlling for other parent language input. *Developmental Science, 22*(3), Article e12764. <https://doi.org/10.1111/desc.12764>
- Dickinson, D., Griffith, J., Golinkoff, R. M., & Hirsh-Pasek, K. (2012). How reading books fosters language development around the world. *Child Development Research, 2012*, 1–15. <https://doi.org/10.1155/2012/602807>
- Dickinson, D. K., & Morse, A. B. (2019). *Connecting through talk. In Nurturing children's development with language*. Baltimore, MD: Brookes.
- Dixon, W. E., & Smith, P. H. (2000). Links between early temperament and language acquisition. *Merrill-Palmer Quarterly, 46*(3), 417–440.
- Durham, R., Farkas, G., Hammer, C. S., & Tomblin, J. B. (2007). Kindergarten oral language skills: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility, 25*, 294–305. <https://doi.org/10.1016/j.rssm.2007.03.001>
- Egberink, I.J.L., Leng, W.E. de, & Vermeulen, C.S.M. (2010). COTAN beoordeling 2010, Taal voor Kleuters [COTAN review 2010, Language Test for Kindergarten]. Retrieved from [www.cotandocumentatie.nl](http://www.cotandocumentatie.nl).
- Evans, M. A., & Hulak, L. (2019). Learning to read at home: Kindergarten children's report in relation to parent observed behavior. *Early Childhood Research Quarterly, 50*(2), 38–45. <https://doi.org/10.1016/j.ecresq.2018.09.010>
- Flack, Z. M., Field, A. P., & Horst, J. S. (2018). The effects of shared storybook reading on word learning: A meta-analysis. *Developmental Psychology, 54*(7), 1334–1346. <https://doi.org/10.1037/dev0000512>
- Garello, V., Viterbori, P., & Usai, M. C. (2012). Temperamental profiles and language development: A replication and an extension. *Infant Behavior and Development, 35*(1), 71–82. <https://doi.org/10.1016/j.infbeh.2011.09.003>
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development, 26*(1), 64–86. [https://doi.org/10.1016/S0163-6383\(02\)00169-8](https://doi.org/10.1016/S0163-6383(02)00169-8)
- Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development, 90*(3), 985–992. <https://doi.org/10.1111/cdev.13128>
- High, P. C., Klass, P., Donoghue, E., Glassy, D., DelConte, B., Earls, M., Lieser, D., McFadden, T., Mendelsohn, A., Scholer, S., Schulte, E. E., Takagishi, J., Vanderbilt, D., & Williams, P. G. (2014). Literacy promotion: An essential component of literacy care pediatric practice. *Pediatrics, 134*, 404–409. <https://doi.org/10.1542/peds.2014-1384>
- Jones, V. F., Franco, S. M., Metcalf, S. C., Popp, R., Staggs, S., & Thomas, A. E. (2000). The value of book distribution in a clinic-based literacy intervention program. *Clinical Pediatrics, 39*(9), 535–541. <https://doi.org/10.1177/000992280003900905>
- Justice, L. M., Logan, J. R., & Damschroder, L. (2015). Designing caregiver-implemented shared-reading interventions to overcome implementation barriers. *Journal of Speech, Language, and Hearing Research, 58*(6), S1851–S1863. [https://doi.org/10.1044/2015\\_JSLHR-L-14-0344](https://doi.org/10.1044/2015_JSLHR-L-14-0344)
- Karrass, J., Braungart-Rieker, J. M., Mullins, J., & Lefever, J. B. (2002). Processes in language acquisition: The roles of gender, attention, and maternal encouragement of attention over time. *Journal of Child Language, 29*, 519–543. <https://doi.org/10.1017/S0305000902005196>
- Karrass, J., VanDeventer, M. C., & Braungart-Rieker, J. M. (2003). Predicting parent-child book reading in infancy. *Journal of Family Psychology, 17*, 134–146. <https://doi.org/10.1037/0893-3200.17.1.134>
- Kegel, C. A. T., Bus, A. G., & van IJzendoorn, M. H. (2011). Differential susceptibility in early literacy instruction through computer games: The role of the dopamine D4 Receptor Gene (DRD4). *Mind Brain and Education, 5*(2), 71–78. <https://doi.org/10.1111/j.1751-228X.2011.01112.x>
- Klein Velderman, M., Bakermans-Kranenburg, M. J., Juffer, F., & van IJzendoorn, M. H. (2006). Effects of attachment-based interventions on maternal sensitivity and infant attachment: Differential susceptibility of highly reactive infants. *Journal of Family Psychology, 20*(2), 266–274. <https://doi.org/10.1037/0893-3200.20.2.266>
- Koerhuis, I., & Keuning, J. (2011). *Rekenen voor kleuters Groep 1 en 2 [Mathematics for Kindergarten]. Handleiding [Manual]*. Arnhem, The Netherlands: Cito.
- Lansink, N., & Hemker, B. (2012). *Wetenschappelijke verantwoording van de toets taal voor kleuters voor groep 1 en 2 uit het cito volgsysteem primair onderwijs [Scientific justification of Language for Kindergarten]*. Arnhem, The Netherlands: Cito.
- Lin, J., Reich, S. M., Kataoka, S., & Farkas, G. (2015). Maternal reading self-efficacy associated with perceived barriers to reading. *Child Development Research, 2015*, Article 218984. <https://doi.org/10.1155/2015/218984>
- Logan, J. A., Justice, L. M., Yumus, M., & Chaparro-Moreno, L. J. (2019). When children are not read to at home: The million-word gap. *Journal of Developmental & Behavioral Pediatrics, 40*(5), 383–386. <https://doi.org/10.1097/DBP.0000000000000657>
- Machida, S., Taylor, A. R., & Kim, J. (2002). The role of maternal beliefs in predicting home learning activities in head start families. *Family Relations, 51*, 176–184. <https://doi.org/10.1111/j.1741-3729.2002.00176.x>
- MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology, 58*, 593–614. <https://doi.org/10.1146/annurev.psych.58.110405.085542>
- Majorano, M., & Lavelli, M. (2014). Maternal input to children with specific language impairment during shared book reading: Is mothers' language in tune with their children's production? *International Journal of Language & Communication Disorders, 49*, 204–214. <https://doi.org/10.1111/1460-6984.12062>
- Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear in picture books and the statistics for language learning. *Psychological Science, 26*, 1489–1496. <https://doi.org/10.1177/0956797615594361>

- Montag, J. L. (2019). Differences in sentence complexity in the text of children's picture books and child-directed speech. *First Language*, 39(5), 527–546. <https://doi.org/10.1177/0142723719849996>
- Niklas, F., Cohrsen, C., & Tayler, C. (2016). The sooner, the better: Early reading to children. *SAGE Open*, 6(4). <https://doi.org/10.1177/2158244016672715>
- Pace, A., Burchinal, P., Alper, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Early Childhood Research Quarterly*, 46, 112–125.
- Plak, R. D., Merkelbach, I., Kegel, C. A. T., Van IJzendoorn, M. H., & Bus, A. G. (2016). Brief computer interventions enhance emergent academic skills in susceptible children: A gene-by-environment experiment. *Learning and Instruction*, 45, 1–8. <https://doi.org/10.1016/j.learninstruc.2016.06.002>
- Poehlmann, J., Hane, A., Burnson, C., Maleck, S., Hamburger, E., & Shah, P. E. (2012). Preterm infants who are prone to distress: Differential effects of parenting on 36-month behavioral and cognitive outcomes. *Journal of Child Psychology and Psychiatry*, 53, 1018–1025. <https://doi.org/10.1111/j.1469-7610.2012.02564.x>
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40(3), 879–891. <https://doi.org/10.3758/BRM.40.3.879>
- Raikes, H., Alexander Pan, B., Luze, G., Tamis-LeMond, C. S., Brooks-Gunn, J., Constantine, J., Rodriguez, E. T., ... (2006). Mother-child book reading in low-income families: Correlates and outcomes during the first three years of life. *Child Development*, 77(4), 924–953. <https://doi.org/10.1111/j.1467-8624.2006.00911.x>
- Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, 52, 569–578. <https://doi.org/10.2307/1129176>
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5(6), 359–371. <https://doi.org/10.1111/j.1751-9004.2011.00355.x>
- Sénéchal, M., & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skills: A five-year longitudinal study. *Child Development*, 73(2), 445–460. <https://doi.org/10.1111/1467-8624.00417>
- Sharif, I., Rieber, S., Ozuah, P. O., & Reiber, S. (2002). Exposure to reach out and read and vocabulary outcomes in inner-city preschoolers. *Journal of the National Medical Association*, 94(3), 171–177..
- Sosa, A. V. (2016). Association of the type of toy used during play with the quantity and quality of parent-infant communication. *JAMA Pediatrics*, 170(2), 132–137. <https://doi.org/10.1001/jamapediatrics.2015.3753>
- Spinelli, M., Fasolo, M., Shah, P. E., Genovese, G., & Aureli, T. (2018). The influence of early temperament on language development: The moderating role of maternal input. *Frontiers in Psychology*, 9, 1527. <https://doi.org/10.3389/fpsyg.2018.01527>
- Usai, M. C., Garello, V., & Viterbori, P. (2009). Temperamental profiles and linguistic development: Differences in the quality of linguistic production in relation to temperament in children of 28 months. *Infant Behavior and Development*, 32, 322–330. <https://doi.org/10.1016/j.infbeh.2011.09.003>
- Van den Berg, H., & Bus, A. G. (2014). Beneficial effects of BookStart in temperamentally highly reactive infants. *Learning and Individual Differences*, 36, 69–75. <https://doi.org/10.1016/j.lindif.2014.10.008>
- Vernon-Feagans, L., Pancsofar, N., Willoughby, M., Odom, E., Quade, A., & Cox, M. (2008). Predictors of maternal language to infants during a picture book task in the home: Family SES, child characteristics, and the parenting environment. *Journal of Applied Developmental Psychology*, 29, 213–226. <https://doi.org/10.1016/j.appdev.2008.02.007>
- Wade, B., & Moore, M. (1998). An early start with books: Literacy and mathematical evidence from a longitudinal study. *Educational Review*, 50(2), 135–145. <https://doi.org/10.1080/0013191980500205>
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24, 2143–2152. <https://doi.org/10.1177/0956797613488145>
- Widaman, K. F., Helm, J. L., Castro-Schilo, L., Pluess, M., Stallings, M. C., & Belsky, J. (2012). Distinguishing ordinal and disordinal interactions. *Psychological Methods*, 17(4), 615–622. <https://doi.org/10.1037/a0030003>
- Zuckerman, M. (1999). *Vulnerability to psychopathology: A biosocial model*. Washington, DC: American Psychological Association.