COVID-19 first lockdown as a window into language acquisition: associations between caregiver-child activities and vocabulary gains

Natalia Kartushina
MultiLing, Department of Linguistics and Scandinavian Studies, University of Oslo, Norway

Nivedita Mani
Georg-Elias-Müller Institute for Psychology, University of Göttingen, Germany
Leibniz ScienceCampus Primate Cognition, Germany

Aslı Aktan-Erciyes
Department of Psychology, Kadir Has University, Turkey

Khadeejah Alaslani
Department of Linguistics, Purdue University, USA

Naomi J. Aldrich
Department of Psychology, Grand Valley State University, USA

Alaa Almohammadi
Haifa Alroqi
Department of European Languages and Literature, King Abdulaziz University, Saudi Arabia

Lucy M. Anderson
Brigham Young University, Provo, Utah, USA

Elena Andonova
Research Center for Cognitive Science, New Bulgarian University, Bulgaria

Suzanne Aussems
Department of Psychology, University of Warwick, United Kingdom

Mireille Babineau
Laboratoire de Sciences Cognitives et Psycholinguistique, École normale supérieure, PSL University, France
Department of Psychology, University of Toronto, Canada

Mihaela Barokova
Center for Autism Research Excellence, Boston University, USA

Christina Bergmann
Language Development Department, Max Planck Institute for Psycholinguistics, The Netherlands

Cara Cashon
Department of Psychological and Brain Sciences, University of Louisville, USA

Stephanie Custode
University of Miami, USA

Volume 2, Issue 1, 31 December 2022

Cielke Hendriks
Language Development Department, Max Planck Institute for Psycholinguistics, The Netherlands

Tzipi Horowitz-Kraus
Educational Neuroimaging Group, Faculty of Education in Science and Technology, Faculty of Biomedical Engineering, Technion, Israel

Marina Kalashnikova
Basque Center on Cognition, Brain, and Language, Spain
IKERBASQUE, Basque Foundation for Science, Spain

Junco Kanero
Faculty of Arts and Social Sciences, Sabancı University, Turkey

Christina Keller
Centre for Language and Communication Research, Cardiff University, UK

Grzegorz Krajewski
Faculty of Psychology, University of Warsaw, Poland

Catherine Laing
Centre for Language and Communication Research, Cardiff University, UK

Rebecca A. Lundwall
Brigham Young University, Provo, Utah, USA

Magdalena Łuniewska
Karolina Mieszkowska
Faculty of Psychology, University of Warsaw, Poland

Luis Muñoz
Department of Psychology, University of Oslo, Norway

Karli Nave
Faculty of Psychology, University of Warsaw, Poland

Nonah Olesen
Department of Psychological and Brain Sciences, University of Louisville, USA

Lynn Perry
University of Miami, USA

Caroline Rowland
Language Development Department, Max Planck Institute for Psycholinguistics, The Netherlands
Donders Institute for Brain, Cognition & Behaviour, Radboud University, the Netherlands

Daniela Santos Oliveira
University of Tennessee, USA
Jeanne Shinskey
Department of Psychology, Royal Holloway University of London, United Kingdom

Aleksander Veraksa
Lomonosov Moscow State University, Russia

Kolbie Vincent
Department of Psychological and Brain Sciences, University of Louisville, USA

Michal Zivan
Educational Neuroimaging Group, Faculty of Education in Science and Technology, Faculty of Biomedical Engineering, Technion, Israel

Julien Mayor
Department of Psychology, University of Oslo, Norway

Abstract: The COVID-19 pandemic, and the resulting closure of daycare centers worldwide, led to unprecedented changes in children's learning environments. This period of increased time at home with caregivers, with limited access to external sources (e.g., daycares) provides a unique opportunity to examine the associations between the caregiver-child activities and children's language development. The vocabularies of 1742 children aged 8-36 months across 13 countries and 12 languages were evaluated at the beginning and end of the first lockdown period in their respective countries (from March to September 2020). Children who had less passive screen exposure and whose caregivers read more to them showed larger gains in vocabulary development during lockdown, after controlling for SES and other caregiver-child activities. Children also gained more words than expected (based on normative data) during lockdown; either caregivers were more aware of their child's development, or vocabulary development benefited from intense caregiver-child interaction during lockdown, or both. We discuss these results in the context of the extraordinary circumstances of the COVID-19 pandemic and highlight limitations of the study.

Keywords: COVID-19 pandemic; vocabulary development; book reading; passive screen exposure; multi-country

Corresponding author: Natalia Kartushina, MultiLing, Department of Linguistics and Scandinavian Studies, Faculty of Humanities, University of Oslo, Niels Henrik Abels vei 36, 0313, Oslo, Norway. Email: natalia.kartushina@iln.uio.no

ORCID ID: https://orcid.org/0000-0003-4650-5832

Citation: Kartushina, N., Mani, N., Aktan-Erciyes, A., Alaslani, K., Aldrich, N. J., Almohammadi, A., ... & Mayor. J. (2022). COVID-19 first lockdown as a window into language acquisition: associations between caregiver-child activities and vocabulary gains. *Language Development Research*, 2(1), 1—36. https://doi.org/10.34842/abym-xv34

Introduction

Language is acquired knowledge – children need experience with language to learn it. Differences in the quality and quantity of children's language experience may, therefore, influence language learning outcomes. Indeed, the *quantity* of children's language input is positively associated with their vocabulary size and development (in Western, industrialized societies, see Hart & Risley, 1995; Weisleder & Fernald, 2013; but see Casillas et al., 2020 for work on non-Western societies). The *quality* of children's language experience is similarly associated with language development with findings suggesting that the diversity, sophistication and responsiveness of input predict later vocabulary growth (Anderson et al., 2021; Cartmill et al., 2013; Pan et al., 2005).

Much of the work examining factors associated with variability in early language development (Frank et al., 2021) has focused on caregivers' reports of their daily interactions with their children. Such reports do not include input that children routinely receive from other sources (e.g., daycare centers, screen exposure), making it difficult to quantify all of the linguistic input available to children. In early 2020, the COVID-19 pandemic led many countries to implement strict lockdowns such that families had little or no social contact with others outside their household. Schools and daycare centers were shut down in over 160 countries (COVID-19 Educational Disruption and Response. *UNESCO*). Many caregivers worked from home, providing them with a better overview of their child's development and the activities their children were engaged in. Such periods of extended contact between caregivers and children have previously been referred to as "faucet" moments (Entwisle et al., 2001), when shared aspects of the child's environment, e.g., schools and daycare centers, are removed, such that differences in the home environment are particularly weighted in development. The current study capitalized on this "faucet" moment during the first COVID-19 lockdown to examine whether the activities that caregivers and children engaged in correlated with children's vocabulary development during this period.[1]

To achieve these goals, we evaluated, first, the amount of time children spent during lockdown on the following activities[2] (together with a caregiver or alone): shared book

[1] Throughout this manuscript, we will refer to lockdown as the time from March to September 2020 during which daycare centers were closed – and not in the sense of a strict curfew.

[2] To our knowledge, no questionnaire assessing parental activities has been validated across the populations examined in the current study, i.e., 13 countries with children learning 12 different languages. This required us to develop a questionnaire on the activities that caregivers undertook with their children during the COVID19-related lockdown. We acknowledge, however, that this questionnaire has not been validated across the populations tested. It is noteworthy that due to the extra-ordinary time constraints on data collection (the questionnaires needed to be approved by ethics board before launching the study, and sent out as soon as lockdown ended), and, given that children were

reading, structured child-caregiver games (referred to as structured parent-child interaction in the preregistration), free play with their caregiver, singing, speaking, outdoor activities, watching TV, baby shows or cartoons (henceforth, referred to as passive screen exposure), playing digital baby games (henceforth, active screen exposure involving interaction with a device), and playing freely without adults. Then, we assessed whether the time spent on these activities correlated with vocabulary development during lockdown, as indexed by the difference in the child's vocabulary size (in percentile, compared to norms, and in raw scores, where norms were not available) at the beginning and end of the lockdown period. To measure children's vocabulary sizes, we used Communicative Development Inventories (CDI; Fenson et al. 2007) – vocabulary checklists, where parents check words that their child understands or understands and produces. We focused on these activities given prior research finding positive associations between vocabulary development and shared book reading (Shahaeian et al., 2018; Wasik et al., 2016), speaking (Weisleder & Fernald, 2013; Rowe, 2018), singing (Williams et al., 2015), and playing (Hirsh-Pasek et al., 2009); and negative associations between screen exposure (van den Heuvel et al., 2019; Zimmerman et al., 2007) and vocabulary development.

In addition, we also measured caregiver's education (as a proxy for SES) to account for its potential associations with vocabulary development. Previous research suggests that children from higher-SES homes have larger vocabularies than those from lower-SES homes (Pace et al., 2017, Rowe, 2018). SES also moderates the relationship between caregiver-child activities and vocabulary development (Shahaeian et al., 2018, but see Malin et al., 2014). We chose maternal education as a proxy for SES because caregiver education is an important foundation for different developmental outcomes (Davis-Kean et al., 2020). We, therefore, statistically controlled for maternal education attainment in examining the association between caregiver-child activities and vocabulary development during lockdown. In addition, we examined the correlation between maternal education and the activities that caregivers engaged in.

We predicted (see https://osf.io/r85fw) that children whose caregivers engaged more in activities known to promote language development would have larger gains in receptive and productive vocabulary by the end of lockdown. In particular, we predicted that the frequency of shared book reading would capture more of the variability in vocabulary development than the frequency of other activities we examined (Montag et al., 2018), and that increased passive screen exposure would be related to smaller gains in vocabulary development (Zimmerman et al., 2007). Furthermore, we predicted that children whose caregivers engaged in more interactive shared book reading (e.g., asking questions, pointing to things) and structured caregiver-child games (Hirsh-Pasek et al., 2009) would show larger gains in vocabulary (Flack et al.,

already in lockdown when the study started, the questionnaires could not be normed - these were not typical circumstances.

2018). We predicted that children of caregivers with lower maternal education would have (a) smaller gains in both receptive and expressive vocabulary size over lockdown than children of caregivers with higher educational attainment, and (b) smaller vocabulary size at the start of lockdown (Pace et al., 2017; Rowe, 2018). However, we also predicted that the relationship between maternal education and vocabulary development would be better explained by the activities that caregivers engaged in with their children: while there may be differences in the activities that caregivers differing in educational attainment engage in with their children (Entwisle et al., 2001; Pace et al., 2017), the duration and the frequency of such activities should be associated with vocabulary gains during lockdown, above and beyond educational attainment. Finally, we also predicted that infants who attended kindergarten before the lock-down period might experience bigger changes in the quantity and quality of parent-child interactions (before vs. during lockdown) as compared to those who did not, which would translate into bigger changes in vocabulary size during lockdown for the former.

Methods

Participants

In total, 5494 caregivers - from 15 countries and 23 labs - filled in the Time 1 (T1) questionnaire at the beginning of lockdown in their country/region (see Supplementary Material 1 for additional sample details) and 2830 caregivers - from 14 countries and 21 labs - filled in the Time 2 (T2) questionnaire at the end of lockdown (see Procedure for details). Among the 2830 caregivers who filled in T2 questionnaires, data regarding 798 children were excluded from the analysis for either not meeting the following inclusion criteria: (a) monolingual children, defined as having a minimum of 90% exposure to their native language, according to caregiver reports, (b) full term babies, defined as born at 37 weeks of gestation or later, (c) no diagnosed developmental disorder, and (d) no hearing/vision impairment; or when we were unable to match participant ID and/or date of birth across both questionnaires. Note that data gathered for bilingual and multilingual children excluded from the present analysis will be analyzed in a separate contribution. In addition, we excluded 16 children who were outside the normative age range of country-specific CDIs (Fenson et al. 2007). Finally, upon careful analysis of the raw data, we excluded 79 children (2.5% of production and 4.4% of comprehension data), whose gains or losses per day in raw CDI comprehension or production scores fell outside of the distribution and were theoretically or practically uninterpretable for a typically developing child (see Analyses.Rmd code on https://osf.io/ty9mn/), likely due to parental inattentiveness or lack of involvement in the project (cf 7-13% exclusion of unreliable caregivers in de Mayo et al. (2021) for suspiciously brief CDI completion times).

Upon application of the inclusion criteria, our final sample comprised 1742 child participants[3] (886 girls and 856 boys; *M* age = 627 days at T1, range = 244-1089) from 18 labs and 13 countries that contributed to both T1 and T2 data. The SES varied between 1 (primary school, 0.57% of the data) and 6 (doctoral degree, 9.7% of the data), with the median education level of 4 (*SD* = 0.9), where 4 is Bachelor degree (27.78% of the data); these data suggest that mothers in this sample had, overall, high education levels, with the largest proportion of mothers having a MA degree (51.5%) and only 2.7% and 6.49% of the mothers having a high school and some college degree, respectively; although there were notable differences across countries (for details, see Analyses_2.html on https://osf.io/ty9mn/). Yet, note that, for the countries for which data on maternal educational attainment were available in wordbank.stanford.edu (Frank et al., 2017), the proportion of mothers with lower education levels (1 and 2 on the maternal education scale) was comparable to that reported in the normative data (see Supplementary Material 3), suggesting that the proportion of mothers with lower educational attainment in our sample was not lower than what can be found in the country-specific normative data, in general. An additional 290 children from Switzerland (for whom the exact age was missing) were included in the analyses of the relationship between SES and activities reported on https://osf.io/ty9mn/ (total *n* = 2033). Information about labs and child participants is included in Table 1.

Materials

T1 Questionnaire

The questionnaire launched at the beginning of lockdown included basic demographic questions about the children (sex, date of birth, estimated proportion of language exposure to each language heard in their daily life, preterm-versus-full-term status, history of ear infections, known hearing or visual impairments, and known developmental concerns), their caregivers (sex, level of education, and native language(s)) and siblings, if any (sex and date of birth). Maternal education (proxy for SES) was measured on a scale from 1 to 6, with 1 – primary school, 2 – high school, 3 – some college/university, 4 – Bachelor degree, 5 – Master degree, and 6 – doctoral degree (see https://osf.io/ty9mn/ for the distribution of maternal education in each country).

We measured children's receptive (for children between 8 and 18 months of age) and expressive (from 8 to 36 months of age) vocabularies at the onset of lockdown using age-appropriate CDIs and their adaptations for the relevant language (or regional variant). Variants included short-CDIs (Mayor & Mani, 2019 – for German) and web-CDIs (de Mayo et al., 2021— for American English, Hebrew, Dutch). CDIs ranged from 303

[3] Note that given that all questions had an option "prefer not to answer", some participants, in the final sample, had no data for some activities or SES.

to 897 words (25 items for the short-CDIs in German). A subset of laboratories collected additional data (not analyzed here) for use in planned follow-up projects.

Table 1. ***Description of the final sample of children (number, mean age in months and sd) included in the analyses of gains in production and comprehension (in percentile and raw CDI score).***

			Production (raw CDI score)		Comprehension (raw CDI score)		Production (percentiles)		Comprehension (percentiles)	
Labid	Country	Language	Age	n	Age	n	Age	n	Age	n
babyling	Norway	Norwegian	21 (6.9)	173	13.1 (2.7)	58	21 (6.9)	173	13 (2.7)	58
bcbl	Spain	Basque	17 (6.5)	18	12.5 (0.9)	10	NA	NA	NA	NA
bcbl	Spain	Spanish	15 (6.6)	19	9.8 (1.7)	10	NA	NA	NA	NA
brc-nijmegen	The Netherlands	Dutch	17 (6.8)	20	12.2 (3.7)	11	NA	NA	NA	NA
brookes	UK	English	19 (7.2)	292	12.6 (2.5)	143	15 (1.1)	83	15 (1.1)	81
clcu	UK	English	20 (7.6)	40	13.1 (3.6)	17	16 (1.6)	10	16 (1.5)	9
cogdevlabbyu	USA	English	12 (3)	39	12.1 (3.0)	38	12 (2.9)	36	12 (2.9)	35
dsc	USA	English	21 (6.6)	5	14.7 (1.3)	2	23 (6.6)	4	14	1
goe	Germany	German	21 (1.6)	37	NA	NA	21 (1.5)	36	NA	NA
HaifaUniv	Israel	Hebrew	21 (5.5)	61	13.5 (2.7)	12	15 (1.4)	11	15 (1.1)	9
ilpll	USA	English	21 (9.0)	49	11.2 (1.9)	16	16 (6.2)	32	11 (1.5)	15
kau-cll	Saudi Arabia	Arabic	22 (6.3)	90	11.3 (1.9)	10	NA	NA	NA	NA
ldl	Canada	English	22 (8.4)	17	12 (3.3)	5	20 (5.8)	12	13 (3.1)	4
mltlab	Turkey	Turkish	24 (6.2)	40	12.8 (2.3)	4	24 (5.5)	36	12 (1.7)	3
msu	Russia	Russian	22 (5.3)	17	15.9 (2.5)	4	23 (5.5)	14	14 (1.8)	2
multilada	Poland	Polish	21 (6.8)	223	13.6 (2.6)	77	21 (6.8)	209	13 (2.4)	69
paris_team	France	French	22 (6.8)	466	12.9 (1.9)	113	NA	NA	NA	NA
rhul_baby_lab	UK	English	15 (1.9)	25	14.4 (1.8)	22	15 (1.1)	23	15 (1.2)	21
technion_il	Israel	Hebrew	22 (7.1)	111	14 (2.5)	33	16 (1.8)	30	15 (1.7)	23
		Total		1742		585		709		330

Note. NA - not available, indicates when CDI norms were not available for a given language and/or CDI instrument. In the Brookes sample, 7 participants in the percentile analysis and 15 in the analysis of raw CDI were exposed to limited daycare during lockdown (means of 1.4 and 1.5 days a week, respectively).

T2 questionnaire

To assess activities that caregivers and their children engaged in during lockdown, a custom-made questionnaire was created and then collaboratively expanded and refined until the launch of the project. Questions evaluated the time spent on the following activities during lockdown: shared book reading, structured child-caregiver games, free play with the child, singing with the child, one-to-one speaking with the child, time spent outdoors, passive screen exposure (watching baby TV, cartoons, shows, with no interaction with a digital device), playing baby games on a digital device, time spent playing without an adult – all on a 10-point scale ranging from "did not do this activity at all" to "more than 4 hours most days." If parents/caregivers indicated that they read to their child at least 15 minutes per day, then they were asked eight yes/no questions (receiving each 1 point for a "yes" answer) on the quality of reading interactions (Han & Neuharth-Pritchett, 2015). The questionnaire also asked about the amount of time caregivers spent working from home and included CDI data to measure vocabulary development over the lockdown period. A subset of laboratories collected additional data (not analyzed here) for use in planned follow-up projects.

Procedure

On March 12, 2020, the Norwegian government enforced a national lockdown and, among other measures, closed daycare centers. On March 18, the local study on the impact of lockdown on language acquisition among 8- to 36-month-old children in Norway was preregistered and data collection started on March 20. The same day, a call for participation for international partners was issued via various mailing lists, which resulted in the present collaboration, including 23 labs in 15 countries. Each lab was asked to launch the T1 questionnaire as soon as possible upon daycare centers' closure and to launch T2 as close as possible to children starting regular daycare again, or if significant changes took place in local policies that would affect social isolation. Data collection started on March 20, 2020 (Norway) and finished on September 29, 2020 (USA), with a mean time interval between T1 and T2 of 41 days. We welcomed participation from all labs that were able to obtain ethical approval in time to launch the T1 questionnaire close to the daycare centers' closure. No minimum participant numbers were required to join the project.

The whole study was conducted online. We used a variety of means to recruit participants (e.g., social media, lab databases, social platforms, etc.), which allowed us to reach out to larger demographic populations, as compared to those typically tested in the lab (de Mayo et al., 2021). Data collection took part during the first COVID-19 lockdown. The announcement invited parents of 8-36-month-old infants to take part in a research project and included a link to the T1 questionnaire (see Materials), where caregivers were also asked to generate a unique participant identifier and provide a

valid email address, to be used when sending them the T2 questionnaire. Participant compensation varied across labs from no compensation to a small toy, a book or a voucher or a lottery ticket to win gift cards. The research project was approved by the Norwegian Center for Research Data REF536895 and by the ethics committee of the Department of Psychology at the University of Oslo. Collaborating labs obtained ethical approval from their institutions. Central data analyses used exclusively anonymized data.

Transparency statement

Prior to data collection, and prior to the call for an international collaboration, we preregistered our study for the Norwegian sample (https://osf.io/4mhjw). To accommodate for multi-site analyses, and to include modifications made to the questionnaires in the days following the initial preregistration, a multi-site preregistration was made prior to data inspection, visualization and processing (https://osf.io/r85fw). All materials, anonymized data, and analysis codes are available on the project's OSF (https://osf.io/ty9mn/).

Results

Data Processing

Computation of Vocabulary Gains in Percentiles

Our dependent variables were the total number of words that caregivers reported their child understood (between 8 and 18 months of age) and produced (between 8 and 36 months of age). The total number of words on CDIs was transformed into daily percentiles separately for each language using available norming data from wordbank.stanford.edu (Frank et al., 2017), provided that the dataset was dense enough, with a minimum of 50 data points per age (in months), or, for Hebrew, Polish and British English (UK-CDI), via direct contact with the authors who collected the norming data. Monthly percentiles from the norming data were linearly interpolated to establish daily percentiles (i.e., daily norms), then used to compute children's vocabulary size in daily percentiles (cf. https://osf.io/ty9mn/). We were able to derive daily percentiles for 14 labs in 9 countries (cf. Table 1) and computed gains in percentiles (T2-T1) for both comprehension (n = 330) and production (n = 709).

Computation of Normalized Gains in Raw CDI Scores

For 6 CDI instruments from 6 countries, data was either not available on WordBank (Saudi Arabia, the Netherlands, extended OxfordCDI) or the data available on Wordbank was too sparse to ensure reliable computation of percentiles (France, Spain, Is-

rael CDI - WS), despite children meeting the criteria for inclusion in the study. Therefore, these data were only entered into the analyses of raw CDI scores (along with the data from children that entered the percentile analyses).

Given (1) wide variation in the CDI size across languages (from 303 to 897) and (2) that potential gains were constrained by CDI scores at T1 (e.g., a toddler knowing all of the words on the CDI at T1 cannot learn more words on the CDI at T2), we computed a normalized measure of gains for each child that situated her with respect to the average gains from all countries given the same relative number of words known on her respective CDIs at T1 (see Analyses.Rmd on https://osf.io/ty9mn/). To this end, first, we divided the CDI score at T1 by the total number of items on the CDI, thus producing a vocabulary proportion score at T1, that varied between 0 and 1. Second, we fitted a polynomial regression to the T1 proportion score, separately for each tool (CDI Words and Gestures and CDI Words and Sentences) and modality, using the *loess* function and then used *predict* on the model outcomes to compute the average expected gains associated with T1 proportion scores. Then, we subtracted average expected gains associated with the T1 proportion scores from actual gains, resulting in average normalized gains of zero, for all T1 proportion scores (see Supplementary Material 2 for the visualization of non-normalized and normalized gains in vocabulary size). In other words, this procedure allowed us to identify individual deviations from expected gains (controlling for the CDI size and the CDI raw score at T1), and to correlate such deviations from expected gains with activities during social isolation. This normalization procedure for gains in raw CDI scores was conducted separately for each CDI tool and modality, for the entire sample comprising 18 labs from 13 countries in: comprehension (n = 585, 8-18-month-old children) and production (n = 1742, 8-36-month-old children).

Statistical Analyses

Correlations between SES and Activities

Pearson correlations (n = 709, dataset for the analyses of percentile gains in production) between SES and activities are reported in Table 2. Correlation matrix for a larger data set with n = 2033 children (that includes Switzerland and the labs for which norming data for the vocabulary score were not available) is available on the OSF page of the project https://osf.io/ty9mn/. As predicted, maternal education correlated positively with the time spent on shared book reading and negatively with children's passive screen exposure. Moderate correlations (>.30) included: a positive correlation between the time spent on shared book reading and on structured child-caregiver games, and between the time spent on passive screen exposure and playing baby games on a digital device. All other correlations were weaker (<.30). We hypothesized that the relationship between screen exposure and SES might be influenced by parents' availability, indexed by the number of hours they worked from home. A separate

linear model, however, revealed that this interaction was not significant (β = 0.0174, SE = 0.028, t = 0.62, p = 0.534).

Maternal education, activities and gains in production

First, a mixed-effect regression analysis on percentile gains in production, was conducted in R (R Core Team, 2020) for children between 8- and 36-months-of-age (see Table 3) using *lmer* (Bates et al., 2015:4) and *summ* (Long, 2020) to obtain the summary of the model. Fixed factors were time spent on activities that caregivers engaged in with their child during lockdown (mean-centered), maternal education (mean-centered), child's sex, and age (mean-centered in days, at T1), time gap between T1 and T2 in days (mean-centered), and child's daycare attendance before T1 (yes/no). Descriptive statistics for the activities and other variables used in the model can be found in the Supplementary Material 4. Random effects included a maternal education by country slope, hence, country was included as a random factor.[4]

Next, the same analysis was conducted on the second dependent variable, i.e., normalized raw gains in production. The results of the two models are summarized in Table 3. Note that the intercept and the effect of time gap between T1 and T2 need to be interpreted differently across the percentile and raw gains models. The intercept in the percentile model examines whether children (at the reference level of mean-centered age) gained more words than expected during lockdown (given normative data), since we expect children to stay in the same percentile across development. The intercept in the raw gains model is not meaningful as gains were normalized for each instrument. Time gap in the percentile model examines whether children's percentile scores improved linearly with the duration of lockdown, i.e., that they showed greater improvement in their percentile scores, the longer lockdown lasted. Time gap in the raw gains model trivially examines whether children learned more words the longer lockdown lasted.

[4] In order to address a potential issue of cryptic multiple testing raised by one of the reviewers, we performed, as recommended in Forstmeier & Schielzeth (2011), a full-null model comparison for both dependent variables (gains in percentiles and in normalized raw CDI scores), where the full model contained all the factors included in the main model and the null model excluded the activities examined in the paper. The results of the full/null comparison revealed a significant difference between the two models in gains in percentiles (x^2 = 17.6, df = 9, p = .04) and a marginal difference in gains in normalized raw CDI scores (x^2 = 16.2, df = 9, p = .063), suggesting that activities significantly improved the fit of the null model.

Table 2. ***Means, standard deviations, and correlations with confidence intervals between SES and activities.***

Variable	*M*	*SD*	Maternal education	Book reading	Caregiver works @home	Outdoor activities	Free play w.child	Singing	Speaking	Screen exposure	Digital games	Structured games
Maternal education	4.50	0.89										
Book reading	4.06	1.58	**.15**** **[.08, .22]**									
Parent works @home	3.54	3.51	.03 [-.04, .11]	-.03 [-.11, .04]								
Outdoor activities	4.47	2.67	.02 [-.05, .10]	-.00 [-.08, .07]	.03 [-.05, .10]							
Free play w. child	5.83	1.91	.07 [-.01, .14]	**.24**** **[.17, .31]**	.02 [-.06, .09]	**.16**** **[.09, .24]**						
Singing	3.72	1.74	-.03 [-.10, .05]	**.14*** **[.06, .21]**	.03 [-.05, .10]	.11 [.04, .19]	**.21**** **[.14, .28]**					
Speaking	5.94	2.13	-.03 [-.10, .05]	**.20**** **[.13, .27]**	-.04 [-.12, .03]	.04 [-.03, .12]	**.29**** **[.22, .35]**	**.28**** **[.21, .34]**				
Screen exposure	3.24	2.36	**-.16**** **[-.23, -.08]**	**-.12*** **[-.20, -.05]**	.06 [-.02, .13]	**.14**** **[.07, .21]**	-.01 [-.09, .06]	.03 [-.04, .11]	.03 [-.04, .11]			
Digital games	0.52	1.26	-.10 [-.17, -.02]	-.08 [-.16, -.01]	.05 [-.03, .12]	.06 [-.01, .13]	-.03 [-.10, .04]	.04 [-.03, .12]	.01 [-.07, .08]	**.33**** **[.26, .39]**		
Structured games	2.48	1.91	.04 [-.03, .11]	**.41**** **[.35, .47]**	-.07 [-.14, .01]	.04 [-.03, .11]	**.18**** **[.10, .25]**	**.17**** **[.10, .24]**	**.18**** **[.11, .25]**	.11 [.03, .18]	.06 [-.01, .14]	
Free play no adults	5.16	1.90	-.10 [-.17, -.03]	**-.16**** **[-.23, -.08]**	-.00 [-.08, .07]	.09 [.02, .17]	-.00 [-.08, .07]	-.00 [-.08, .07]	.06 [-.01, .14]	**.23**** **[.16, .30]**	**.14**** **[.06, .21]**	.01 [-.07, .08]

Note. * indicates $p < .05$. ** indicates $p < .01$. The Holm method was used to correct for multiple comparisons and adjust p-values.

Table 3. ***Fixed effects from the mixed-effect regression on the gains in production (left: percentiles with n = 709, full cases n = 685; right: raw scores with n = 1742). p-values below .05 are marked in bold.***

	Gains in percentiles					Normalized gains in raw CDI scores				
	Est.	SE	t	df	p	Est.	SE	t	df	p
(Intercept)	**3.32**	**1.10**	**3.01**	**685**	**.00**	4.01	3.87	1.04	13.45	.32
SES	-1.08	.68	-1.58	685	.11	-.88	2.22	-.40	8.12	.70
Book reading	.16	.43	0.38	685	.71	**1.71**	**.74**	**2.32**	**1528.0**	**.02**
Structured caregiver-child games	-.06	.37	-0.18	685	.85	.42	.57	.73	1601.0	.47
Passive screen exposure	**-.86**	**.29**	**-2.97**	**685**	**.00**	**-1.14**	**.50**	**-2.27**	**1377.8**	**.02**
Outdoor activities	-.09	.23	-0.40	685	.69	.17	.40	.43	1453.4	.67
Digital games	**1.08**	**.48**	**2.24**	**685**	**.03**	.15	.81	.19	1613.5	.85
Free play w. child	.29	.33	0.89	685	.37	-.42	.55	-.75	1616.7	.45
Singing	-.57	.35	-1.63	685	.10	.32	.63	.50	1603.5	.61
Speaking	.39	.29	1.34	685	.18	.17	.51	.34	1576.0	.74
Free play no adult	-.05	.32	-0.16	685	.88	.10	.51	.19	1618.3	.85
Time gap	-.02	.03	-0.65	685	.52	**.55**	**.07**	**8.14**	**150.4**	**.00**
Daycare before (yes)	1.18	1.22	.97	685	.33	1.05	2.51	.42	1201.8	.68
Gender (m)	.17	1.16	.15	685	.88	-1.37	1.93	-.71	1613.2	.48
Age (T1)	.00	.00	.52	685	.61	-.00	.01	-.33	1587.6	.74

Note. all numeric predictors were mean-centered in the analyses; p-values were calculated using Satterthwaite d.f.

In both analyses, the time spent on passive screen exposure negatively correlated with gains in productive vocabulary. As seen in Figure 1, children with no exposure to screens were reported to have the largest gains relative to the normative (age-matched) data from the CDI measures. Yet, it is noteworthy that regardless of the time spent on screen use, reported gains in production always exceeded or met expectations (a gain of zero is equivalent to what would be expected in the normative data).[5]

We also note that the intercept in the percentile model is significantly above zero, i.e., analyses of caregiver reports suggested that children (at the reference level of mean-centered age) gained more words in their productive vocabularies during lockdown, i.e., daycare closure, when compared to the normative data. A Wilcoxon signed-rank test with continuity correction found no evidence for a difference in children's reported vocabularies relative to normative data at the start of lockdown, at T1 (p = .5, Q1 = 23, median = 50, Q3 = 74), but larger reported vocabularies relative to normative data by the end of lockdown, at T2 (p = .005, Q1 = 28, median = 56, Q3 = 80). As indicated by the significant intercept, a one sample t-test on percentile gains between T1 and T2 revealed that, according to caregiver reports, children gained an average of 4 percentiles by the end of lockdown at T2 (95%CI = [2.7:5.0]; $t(684) = 7.0$, $p < .001$, $d = 0.26$).

The effect of time gap on the normalized gains in production suggests that caregivers reported that the longer the time gap between T1 and T2 was, the more words their children learned. In contrast, we found no evidence that percentile gains in vocabulary size accumulated over lockdown, i.e., that children showed greater vocabulary gains (relative to normative data), the longer lockdown lasted. The effects of digital media games on gains in percentiles, and of shared book reading on normalized raw gains did not replicate across analyses and will not be discussed further. Note also that a positive effect of digital media games on gains in percentiles should be interpreted with caution as 79% of children did not play digital games at all. There were no significant associations between gains in production and children's gender or age.

[5] As preregistered, we re-analyzed the data when >95% and <5% percentiles were excluded to check whether the model outcomes were impacted by these extreme values; the significant intercept and main negative effect of passive screen exposure remained significant (see details on OSF).

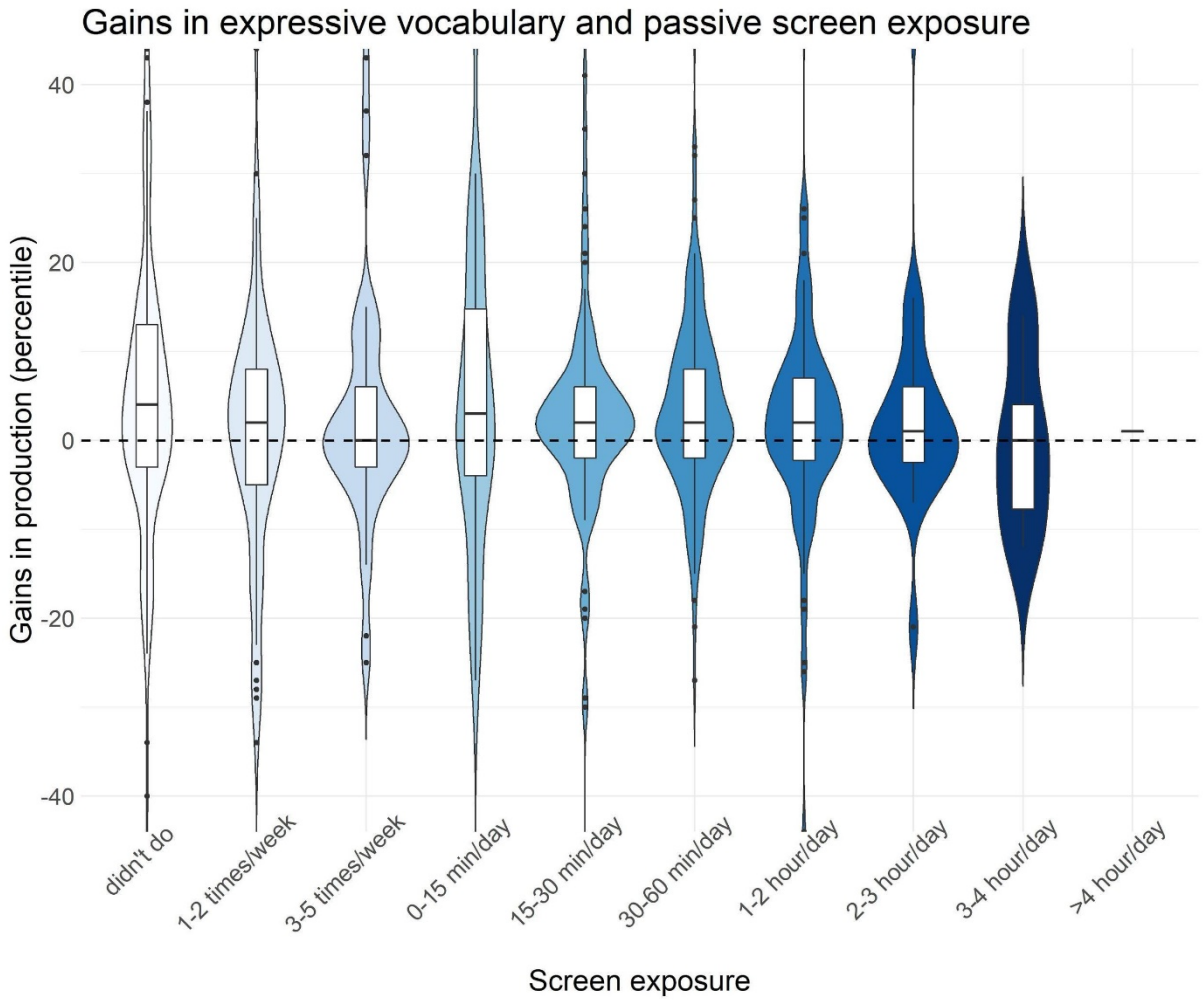

Figure 1. ***Violin plots of the gains in production (percentiles) for different amounts of reported child passive screen exposure. Boxplots display the first quartile and the third quartile, along with the median (the short horizontal bar). Gains of zero (dashed line) correspond to expected gains considering normative data.***

Maternal education, activities and gains in comprehension

A similar mixed-effect regression analysis was run on percentile gains in comprehension for children between 8- and 18-months of age (see Table 4) and on normalized raw gains in comprehension. Similar to the analyses on production, country and variation in maternal education by country were included as random factors.[6] Descriptive statistics for the activities and other variables used in the model can be found in

[6] Similar to the analyses of the production data, in order to address a potential issue of cryptic multiple testing raised by a reviewer, we performed, as recommended in Forstmeier & Schielzeth (2011), a full-null model comparison for both dependent variables (gains in percentiles and in normalized raw

the Supplementary Material 4.

Table 4. ***Fixed effects from the mixed-effect regression on the gains in comprehension. (left: percentiles with n = 330, right: raw scores with n = 585). p-values below .05 are marked in bold.***

	Gains in percentiles					Normalized gains in raw CDI scores				
	Est.	SE	t	df	p	Est.	SE	t	df	p
(Intercept)	**6.45**	**2.37**	**2.72**	**15.2**	**.02**	-3.65	7.12	-.51	42.9	.61
Maternal education	-.68	.89	-.76	9.0	.47	-.26	2.38	-.11	2.9	.92
Book reading	**1.48**	**.57**	**2.59**	**316.0**	**.01**	**3.55**	**1.06**	**3.35**	**544.2**	**.00**
Structured caregiver-child games	-.00	.45	-.00	312.9	1.00	1.17	.79	1.48	538.9	.14
Passive screen exposure	.03	.38	.07	268.7	.94	-.04	.78	-.05	538.8	.96
Outdoor activities	-.33	.31	-1.06	296.6	.29	-.38	.56	-.68	541.3	.50
Digital games	.45	.96	.46	311.7	.64	1.37	2.10	.65	526.8	.51
Free play w. child	.03	.42	.06	314.8	.95	-.78	.78	-1.01	534.5	.31
Singing	-.77	.47	-1.63	317.8	.10	-.44	.88	-.50	538.0	.62
Speaking	-.21	.36	.57	283.8	.57	-.34	.64	-.53	529.5	.59
Free play no adult	**-.80**	**.40**	**-2.01**	**311.4**	**.05**	-.67	.71	-.94	532.9	.35
Time gap	-.00	.05	-.07	102.8	.95	**.73**	**.09**	**7.89**	**149.7**	**.00**
Daycare before (yes)	-.93	1.71	-.54	313.3	.59	1.49	3.41	.44	482.7	.66
Gender (m)	-2.20	1.45	-1.51	311.4	.13	-4.95	2.69	-1.84	530.1	.07
Age (T1)	.02	.01	1.53	297.8	.13	-.01	.02	-.39	540.5	.70

Note. all numeric predictors were mean-centered in the analyses; p-values were calculated using Satterthwaite df.

CDI scores), where the full model contained all the factors included in the main model and the null model excluded the activities examined in the paper. The results of the full/null comparison revealed a significant difference between the two models in both gains in percentiles ($x^2 = 17.3$, $df = 9$, $p = .044$) and in normalized raw CDI scores ($x^2 = 19.8$, $df = 9$, $p = .019$), suggesting that the activities caregivers engaged their children with significantly improved the fit of the null model.

In both analyses, the time spent on shared book reading significantly correlated with gains in receptive vocabulary. As seen in Figure 2, children whose caregivers read 2-3 hours a day to them were reported to have the largest gains in receptive vocabulary size relative to the normative (age-matched) data. Yet, it is noteworthy that even participants with moderate exposure to books (more than 15 minutes per day) were reported to have gained more words than expected considering the (age-matched) norms.[7]

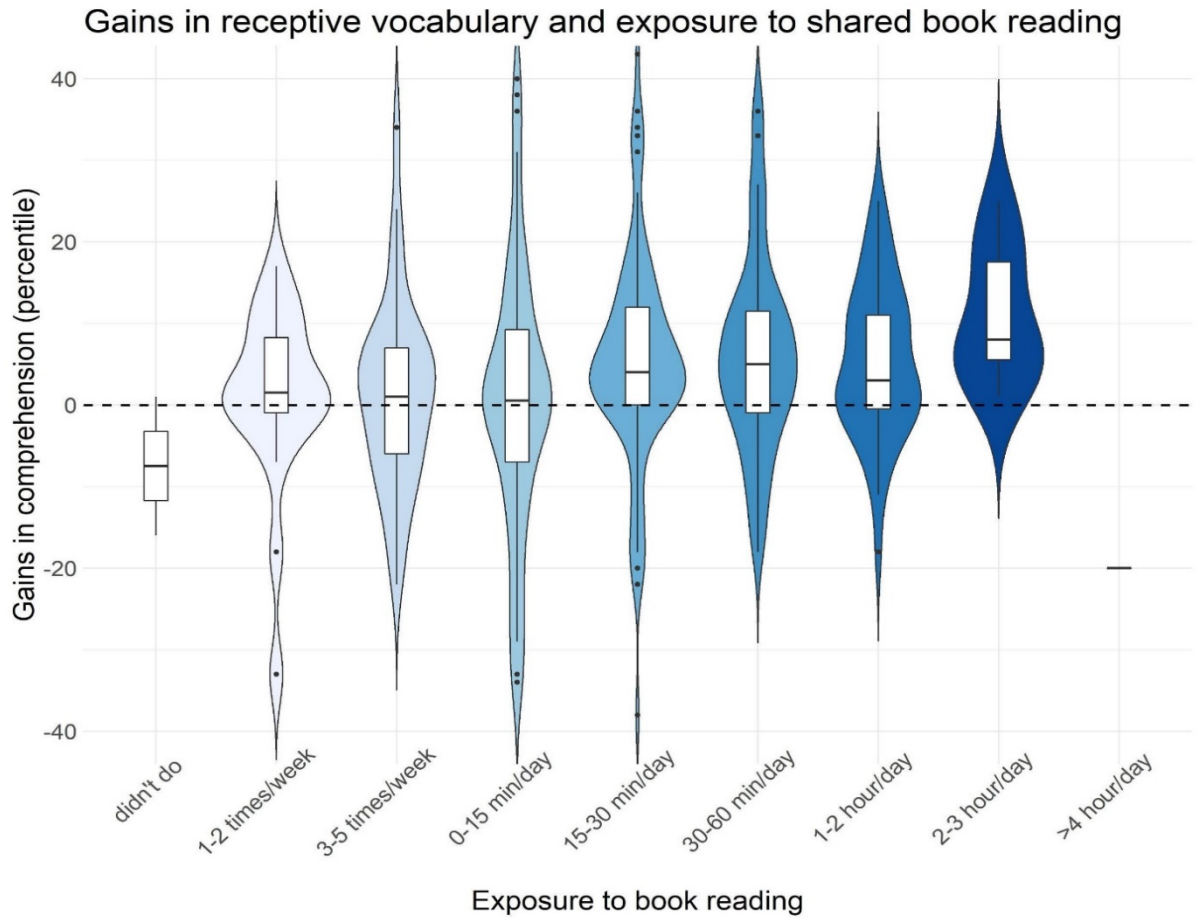

Figure 2. ***Gains in receptive vocabulary (in percentiles) for different amounts of reported shared book reading time. Gains of zero (dashed line) correspond to expected gains considering normative data.***

[7] As preregistered, similar to the analyses on production, we re-analyzed the data when >95% and <5% percentiles were excluded to check whether the model outcomes were impacted by these extreme values; the significant intercept and main positive effect of book reading remained significant (see details on OSF).

The quality of book reading, however, did not robustly correlate with gains in vocabularies, i.e., not across both measures of gains (see Analyses.Rmd on https://osf.io/ty9mn/ for the full analysis).

Similar to our analysis of production scores, analysis of caregiver reports suggested that young children (at the reference level of mean-centered age) gained more words in their receptive vocabularies during lockdown, i.e., daycare closure, when compared to the (age-matched) normative data (see Table 4 – the intercept is significantly above zero in the analysis on percentiles). A Wilcoxon signed-rank test with continuity correction found no evidence for a difference in children's reported vocabularies relative to normative data as children entered lockdown, at T1 (p = .9, Q1 = 23, median = 50, Q3 = 76), but found larger vocabularies relative to normative data at the end of lockdown (p = .01, Q1 = 29, median = 56, Q3 = 79). As indicated by the significant intercept, a one sample t-test in percentile gains between T1 and T2 revealed that, according to caregiver reports, children gained an average of 3.8 percentiles by T2 (95% CI [2.3, 5.2]; $t(317) = 5.0$, $p < .001$, $d = 0.28$)

A strong effect of time gap was also reported for the normalized gains in raw CDI scores, i.e., caregivers' vocabulary reports suggested that their children gained words throughout the lockdown. The additional effect of time spent playing without an adult in the percentile analysis did not replicate across analyses and will not be discussed further. There were no significant associations between children's gender or age and vocabulary development.

Maternal education and Vocabulary at T1

To estimate the extent to which maternal education was associated with expressive and receptive vocabulary at T1, in percentiles[8], we fitted two generalized linear mixed models with beta error structure and logit link function (McCullagh & Nelder, 1989; Bolker, 2008) using *glmmTMB* (Brooks, 2017). We fitted models with beta error structure due to issues with the homogeneity and normality of the residuals in the preregistered Gaussian model. The model revealed no effect of maternal education in either production ($\beta = 0.073$, $SE = 0.046$, $x^2 = 2.28$, $df = 1$, $p = .131$), or comprehension ($\beta = 0.041$, $SE = 0.058$, $x^2 = 0.501$, $df = 1$, $p = .479$, see Supplementary Material 5 for the full analysis). There were no significant associations between children's gender and receptive vocabulary at T1.

[8] Given that raw CDI sizes varied considerably across languages/tools (as number of items varied considerably across tools), correlated with age and we had wide variations in participants' ages across instruments, it was not possible to perform those analyses on raw CDI scores.

Discussion

Three findings stand out from the reported analyses. First, children who had less passive screen exposure during lockdown showed larger gains in their expressive, but not receptive, vocabulary size. Second, children whose caregivers read more to them during lockdown showed larger gains in their receptive, but not expressive, vocabulary size. Third, overall, based on caregivers' reports, children's receptive and expressive vocabularies showed larger increases during lockdown relative to their pre-lockdown, age-matched peers, i.e., using normative data collected pre-lockdown. We discuss these and other reported findings as well as provide potential explanations for these effects.

First, children who had more passive screen exposure during lockdown were reported to have lower gains in expressive vocabulary size (see Figure 1). Children who had no passive exposure to screens showed modest gains in expressive vocabulary relative to their pre-lockdown peers and smaller gains with increasing exposure to screens. There was no influence of passive screen exposure on children's receptive vocabulary across analyses. This differential association between screen exposure on receptive and expressive vocabulary size aligns with recent results in toddlers (Dynia et al., 2021). We suggest that the negative association between expressive vocabulary size and screen consumption may be explained by the fact that there is no requirement to respond to asynchronous digital content. This, in turn, may lead to longer stretches where children are not actively engaged in interacting with others, thereby providing them with little opportunity to expand their productive repertoire. In other words, digital media exposure may have an "opportunity cost" in that it takes time away from other interactions where children may have more opportunities to expand their expressive vocabulary. We did not collect information on the context of screen exposure, yet, recent research suggests that the context in which children are exposed to TV (e.g., during family meals, free day time, etc.) can have differential effects on language development (Martinot et al., 2021). A spin-off project on digital exposure provides more detail on digital practices in children and parents during the first covid lockdown (Bergmann et al., in press).

Second, we found that shared book reading explained more of the variance in gains in receptive vocabulary than any of the other examined activities (c.f. Montag et al., 2018). As shown in Figure 2, children whose caregivers did not engage in shared book reading at all were reported to have lower receptive vocabulary gains relative to pre-lockdown age-matched peers, whereas children whose caregivers engaged in more than 15-30 minutes of shared book reading per day were reported to have an increase in receptive vocabulary relative to pre-lockdown age-matched peers. There was no

similarly consistent association between shared book reading and children's expressive vocabulary size[9], nor between the quality of shared book reading and children's expressive or receptive vocabulary size. Our results highlight the association between book reading and some aspects of children's language development. Indeed, shared book reading includes more referential language than other routines and activities (Tamis-LeMonda et al., 2019); presents the child with higher frequencies of rare words than in everyday conversation (Montag et al., 2018) and allows children to explore words and worlds beyond the here and now.

It is noteworthy that reported receptive and expressive vocabulary growth during lockdown outpaced vocabulary growth in normative age-matched peers. There were no differences in the vocabulary increase between those infants who attended a daycare before the lockdown and who did not. While we did not predict such a lockdown boost, we suggest, post-hoc, alternative explanations for this finding. First, we may, perchance, be tapping into a demographic which differs from the sample used to calculate vocabulary norms. We suggest this to be unlikely given that we found no evidence that vocabulary sizes at T1 in our sample differed from normative data, nor did we find substantial differences in the distribution of maternal education in our sample and the one used to derive the vocabulary norms for the countries for which these data were available (see Supplementary Material 3). Second, many caregivers were working from home during lockdown and were with their child for longer stretches during the day relative to pre-lockdown. Thus, they had more opportunity to assess their child's development and might have been more aware of the words their child understood and produced, leading to more complete responding on the parent report forms we used and, hence, higher CDI scores. Third, social contact restrictions and closing of child-care facilities may have led to increased family and quality time between caregivers and children, providing them with more opportunities for activities that boost vocabulary knowledge, e.g., shared book reading. We are currently unable to disentangle the latter two interpretations of our findings and advocate caution in interpreting this lockdown boost in receptive vocabulary growth. Yet from a broader perspective these two interpretations need not be mutually exclusive: greater knowledge of children's vocabulary may allow caregivers to fine-tune the type and amount of input they provide to their child, in turn potentially leading to better outcomes (Fusaroli et al., 2019). Equally, children who showed greater improvements verbally may also have elicited particular interactions with their parents, e.g., increased amounts of time spent on shared book reading and less screen exposure. Other factors that might have modulated the role of activities are the household structure, the presence (and, if so, the number) of siblings, which is examined in a separate spin-off project, and the circumstances of data collection. Given that the data were collected during the first COVID-19 lockdown, it is possible that parents' engagement

[9] The relationship between book reading and gains in expressive vocabulary was only revealed for the normalized gains in vocabulary.

in the study was affected by the ongoing pandemic and differed from the non-COVID-19 times, when parents have other demands on their time and attention and feel less stressed. Recent studies reported that the pandemic affected mothers in particular (Langin, 2021), as mothers spent more time to take care about the child and the household than fathers, and mothers' experience of pandemic (not measured in the current study) might have influenced their behavior and responsiveness (Evans et al., 2021).

Importantly, children entered the lockdown with a range of vocabulary sizes and had been exposed to learning environments differing in quality prior to daycare closure. The associations between shared book reading, screen time and receptive and expressive vocabulary development, respectively, reported above are considerable, as they capture associations between momentary modulations in the child's learning environment (over an average of just 41 days) and vocabulary development. This is especially so, given recent findings suggesting that parental input shapes children's language skills even after controlling for potential genetic confounds (Coffey et al., 2021). Other activities (outdoor activities, caregiver-child interactions/games), that did not predict gains in receptive and expressive vocabulary size, contributed to other aspects of the child's development, such as the child's well-being during the lockdown (currently being investigated in a separate spin-off project). In contrast to book reading and screen exposure – the two activities that have been systematically analyzed in recent child development research - there are no standardized questionnaires that cover the wide spectrum of languages used in the present study, to examine, retrospectively, child-parent engagement across the wider set of activities used in the current study, e.g., singing, outdoor activities. Therefore, the lack of a significant effect of other activities on vocabulary gains might be attributed, to the lack of salience of other activities to parents, to unknown psychometric properties of reports associated with some activities (e.g., most infants did not use digital games in our study), or to limited reliability when parents are asked to recall past activities (Nivison et al., 2021). However, the analysis, over the same cohort, of the impact of activities on a child's well-being - the focus of a separate contribution (see https://osf.io/ns6gh/?view_only=bee2c0f1686542e9b006ea04e36f0c88)- suggests that parental reports can be used across a range of activities, and that varying activities might have differential effects on child's language development and well-being.

Contrary to our hypothesis, maternal education did not correlate with receptive or expressive vocabulary growth during lockdown or vocabulary size at the onset of lockdown. Note that the absence of an effect of maternal education on gains in receptive or expressive vocabulary size should be taken with caution, as there were relatively few participants with the maternal education lower than a Bachelor degree, which was level 4 on a scale from 1 to 6 in our study (14% of the comprehension data and 10% of the production data) and few participants with the high-school education level only, which was level 2 on our scale (5% of the comprehension data and 3% of the production data). Although the proportion of mothers with low education level in the

current sample was comparable to that reported in the normative data for some of the countries in wordbank.stanford.edu (see Supplementary Material 3), research on a sample with a more homogeneous distribution of maternal education is required to further address this question. Therefore, the extent to which these findings generalize to families from lower socioeconomic backgrounds (as indexed by lower education level in the current study) and less industrialized countries, who were hit hardest by the pandemic, remains uncertain. Nonetheless, the absence of the effect of maternal education is consistent with the modest effects of maternal education on vocabulary reported in data from Wordbank (excluding the USA; Frank et al., 2021) particularly in children under 24 months, especially since a large percentage of the current sample involved children below this age (68%). However, maternal education did correlate positively with time spent on shared book reading, and negatively with time the child spent with digital media. Thus, while there were differences in the activities that caregivers with differing levels of educational attainment engaged in with their child (Entwisle et al., 2001; Pace et al., 2017), our results suggest that the activities that caregivers engaged in with their children, rather than caregivers' educational attainment, correlated with children's receptive and expressive vocabulary development during lockdown. The conjunction of these results highlights some of the pathways through which maternal education (as a proxy for SES) may explain variability in vocabulary development in other studies (Fernald et al., 2013; Pace et al., 2017, Rowe, 2018).

Conclusion

This large-scale multinational study (1742 participants, 13 countries) offers a unique window into associations between features of the home environment and children's longitudinal receptive and expressive vocabulary development. Taken together, the results suggest, that in our sample, caregiver education, children's age or sex were not associated with children's receptive and expressive vocabulary development as much as some of the activities that caregivers reported undertaking with their children.

In particular, the frequency and duration of shared book reading and screen exposure were related to respective receptive and expressive vocabulary gains in lockdown – children whose caregivers read more to them and who had less passive screen exposure showed larger receptive and expressive vocabulary gains, respectively, – and that children's reported receptive and expressive vocabulary development was boosted compared to pre-pandemic CDI norms.

References

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking Quality and Quantity of Parental Linguistic Input to Child Language Skills: A Meta-Analysis. *Child Development*, 92(1). https://doi.org/10.1111/cdev.13508.
Baayen, R.H. (2008). Analyzing Linguistic Data: A Practical Introduction to Statistics. Cambridge: Cambridge University Press.

Barr, D.J., Levy, R., Scheepers, C. & Tily, H.J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014. http://dx.doi.org/10.18637/jss.v067.i01

Bergmann, C., Dimitrova, N., Alaslani, K., Almohammadi, A., Alroqi, H., Aussems, S., Barokova, M., Davies, C., Gonzalez-Gomez, N., Gibson, S. P., Havron, N., Horowitz-Kraus, T., Kanero, J., Kartushina, N., Keller, C., Mayor, J., Mundry, R., Shinskey, J. L., & Mani, N. (in press). Young children's screen time during the first COVID-19 lockdown in 12 countries. *Scientific reports* https://doi.org/10.31219/osf.io/p5gm4

Bolker, B. M. (2008). *Ecological models and data in R*. Princeton University Press. ISBN: 9780691125220

Brooks, M. E., Kristensen, K., Van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., ... & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, 9(2), 378-400. https://doi.org/10.32614/RJ-2017-066

Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences*, 110(28), 11278-11283. https://doi.org/10.1073/pnas.1309518110

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Tseltal Mayan village. *Child Development*, 91(5), 1819-1835. https://doi.org/10.1111/cdev.13349

Coffey, J. R., Shafto, C. L., Geren, J. C., & Snedeker, J. (2021). The effects of maternal input on language in the absence of genetic confounds: Vocabulary development in internationally adopted children. *Child Development*, 93(1).

https://doi.org/10.1111/cdev.13688

Davis-Kean, P. E., Tighe, L. A., & Waters, N. E. (2021). The Role of Parent Educational Attainment in Parenting and Children's Development. *Current Directions in Psychological Science,* 30(2), 186–192. https://doi.org/10.1177/0963721421993116

deMayo, B., Kellier, D., Braginsky, M., Bergmann, C., Hendriks, C., Rowland, C. F., Frank, M. C., & Marchman, V. A. (2021). Web-CDI: A system for online administration of the MacArthurBates Communicative Development Inventories. *Language Development Research,* 1(1), p 55-98. https://doi.org/10.34758/kr8e-w591

Dynia, J. M., Dore, R. A., Bates, R. A., & Justice, L. M. (2021). Media exposure and language for toddlers from low-income homes. *Infant Behavior and Development,* 63, 101542. https://doi.org/10.1016/j.infbeh.2021.101542

Entwisle, D. R., Alexander, K. L., & Olson, L. S. (2001). Keep the faucet flowing summer learning and home environment. *American Educator,* 25(3), 10-15. ISSN: ISSN-0148-432X.

Evans, D. K., Jakiela, P., & Knauer, H. A. (2021). The impact of early childhood interventions on mothers. *Science,* 372(6544), 794-796. https://doi.org/10.1126/science.abg0132

Fenson, L. (2007). *MacArthur-Bates communicative development inventories.* Baltimore, MD: Paul H. Brookes Publishing Company. ISBN 13: 978-1557668882

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science,* 16(2), 234-248. https://doi.org/10.1111/desc.12019

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. https://psycnet.apa.org/doi/10.1037/dev0000512

Forstmeier, W., & Schielzeth, H. (2011). Cryptic multiple hypotheses testing in linear models: Overestimated effect sizes and the winner's curse. *Behavioral Ecology and Sociobiology,* 65(1), 47–55. https://doi.org/10.1007/s00265-010-1038-5

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language,* 44(3), 677. https://doi.org/10.1017/S0305000916000209

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project*. MIT Press. https://doi.org/10.1162/opmi_a_00026

Fusaroli, R., Weed, E., Fein, D., & Naigles, L. (2019). Hearing me hearing you: Reciprocal effects between child and parent language in autism and typical development. *Cognition*, 183, 1-18. https://doi.org/10.1016/j.cognition.2018.10.022

Hirsh-Pasek, K., Golinkoff, R. M., Berk, L. E., & Singer, D. G. (2009). *A mandate for playful learning in preschool: Presenting the evidence*. Oxford University Press, USA. Print ISBN-13: 9780195382716

Han, J., & Neuharth-Pritchett, S. (2015). Meaning-related and print-related interactions between preschoolers and parents during shared book reading and their associations with emergent literacy skills. *Journal of Research in Childhood Education*, 29(4), 528-550. https://doi.org/10.1080/02568543.2015.1073819

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing. ISBN: 978-1-55766-197-5

Langin, K. (2021). Pandemic hit academic mothers hard, data show. *Science*, 371(6530), p 660. https://doi.org/10.1126/science.371.6530.660

Long, J. A. (2020). Package 'jtools'.

Malin, J. L., Cabrera, N. J., & Rowe, M. L. (2014). Low-income minority mothers' and fathers' reading and children's interest: Longitudinal contributions to children's receptive vocabulary skills. *Early Childhood Research Quarterly*, 29(4), 425-432. https://doi.org/10.1016/j.ecresq.2014.04.010

Mayor, J., & Mani, N. (2019). A short version of the MacArthur–Bates Communicative Development Inventories with high validity. *Behavior Research Methods*, 51(5), 2248-2255. https://doi.org/10.3758/s13428-018-1146-0

Martinot, P., Bernard, J. Y., Peyre, H., De Agostini, M., Forhan, A., Charles, M.-A., Plancoulaine, S., & Heude, B. (2021). Exposure to screens and children's language development in the EDEN mother–child cohort. *Scientific Reports*, 11(1), 11863. https://doi.org/10.1038/s41598-021-90867-3

McCullagh, P., & Nelder, J. A. (1989). Monographs on statistics and applied probability. *Generalized linear models*, 37. ISBN-13: 978-0412317606

Montag, J. L., Jones, M. N., & Smith, L. B. (2015). The words children hear: Picture books and the statistics for language learning. *Psychological Science*, 26(9), 1489-1496. https://doi.org/10.1177%2F0956797615594361

Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and diversity: Simulating early word learning environments. *Cognitive Science*, 42, 375-412. https://doi.org/10.1111/cogs.12592

Nivison, M. D., Vandell, D. L., Booth-LaForce, C., & Roisman, G. I. (2021). Convergent and Discriminant Validity of Retrospective Assessments of the Quality of Childhood Parenting: Prospective Evidence From Infancy to Age 26 Years. Psychological Science, 32(5), 721-734.

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, 3, 285-308. https://doi.org/10.1146/annurev-linguistics-011516-034226

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development*, 76(4), 763-782. https://doi.org/10.1111/1467-8624.00498-i1

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org.

Rowe, M. L. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives*, 12(2), 122-127. https://doi.org/10.1111/cdep.12271

Schielzeth, H. & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. *Behavoral Ecology*, 20, 416-420.

Shahaeian, A., Wang, C., Tucker-Drob, E., Geiger, V., Bus, A. G., & Harrison, L. J. (2018). Early shared reading, socioeconomic status, and children's cognitive and school competencies: Six years of longitudinal evidence. *Scientific Studies of Reading*, 22(6), 485-502. https://doi.org/10.1080/10888438.2018.1482901

Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2019). Routine language: Speech directed to infants during home activities. *Child Developmen*t, 90(6), 2135-2152 https://doi.org/10.1111/cdev.13089

UNESCO (2020). COVID-19 educational disruption and response. UNESCO.
van den Heuvel, M., Ma, J., Borkhoff, C. M., Koroshegyi, C., Dai, D. W., Parkin, P.

C., … & Birken, C. S. (2019). Mobile media device use is associated with expressive language delay in 18-month-old children. *Journal of Developmental and Behavioral Pediatrics*, 40(2), 99. https://doi.org/10.1097/DBP.0000000000000630

van den Heuvel, M., Ma, J., Borkhoff, C. M., Koroshegyi, C., Dai, D. W. H., Parkin, P. C., Maguire, J. L., & Birken, C. S. (2019). Mobile Media Device Use is Associated with Expressive Language Delay in 18-Month-Old Children. *Journal of Developmental and Behavioral Pediatrics*, 40(2), 99–104. https://doi.org/10.1097/DBP.0000000000000630

Wasik, B. A., Hindman, A. H., & Snell, E. K. (2016). Book reading and vocabulary development: A systematic review. *Early Childhood Research Quarterly*, 37, 39-57. https://doi.org/10.1016/j.ecresq.2016.04.003

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143-2152. https://doi.org/10.1177/0956797613488145

Williams, K. E., Barrett, M. S., Welch, G. F., Abad, V., & Broughton, M. (2015). Associations between early shared music activities in the home and later child outcomes: Findings from the Longitudinal Study of Australian Children. *Early Childhood Research Quarterly*, 31, 113-124. https://doi.org/10.1016/j.ecresq.2015.01.004

Zimmerman, F. J., Christakis, D. A., & Meltzoff, A. N. (2007). Associations between media viewing and language development in children under age 2 years. *The Journal of Pediatrics*, 151(4), 364-368. https://doi.org/10.1016/j.jpeds.2007.04.071

Data, code and materials availability statement

Data, code and materials that support the findings of this study are openly available on the OSF repository at https://osf.io/ty9mn/

Ethics statement

The research project was approved by the Norwegian Center for Research Data REF536895 and by the ethics committee of the Department of Psychology at the University of Oslo. Collaborating labs obtained ethical approval from their institutions.

Authorship and Contribution Statement

NK and JM conceptualized and designed the study. NK and JM preregistered the original analyses for the Norwegian data. NK, JM, and NM created material for the study. NK, NM, AEA, KA, NA, AA, HA, LA, EA, SA, MB, MDB, CB, CC, SC, ADC, ND,

AD, RF, CF, AF, MG, HGS, SG, KG, NGG, EH, EEH, NH, JH, CH, THK, MK, JK, CK, GK, CL, RL, MŁ, KM, KN, NO, LP, CR, DSO, JS, AV, KV, MZ and JM contributed to data collection. NK, JM, CL, CB, SA, MB, GK, JK, NA, NH, NM and MK preregistered planned analyses for the full sample. NK, LM and JM processed the data, and NK led the data analysis with JM. NK, NM and JM interpreted the data with input from SA, CB, NH, JK and MF. NK, NM and JM wrote the manuscript. NK, NM, JM, CL, CB, SA, MB, ND, THK, JK, MB, NA, RL, HA, AA, KA, NGG, SG, LP, SC, MK, CR, MF, EH reviewed and revised the manuscript. All authors proof-read and approved the final version of the manuscript for submission.

Acknowledgments

We thank Roger Mundry for his analysis of the role of SES on vocabularies at T1. We are grateful to families who took part in the study in these challenging times. We would like to thank reviewers for their insightful comments and suggestions. Funding for Polish sample: grant from National Science Center NCN (Poland), no 2018/31/B/HS6/03916. NK was partly supported by the Research Council of Norway through its Centres of Excellence funding scheme [project number 223265].

Appendices

Supplementary Material 1

Table 1. ***Sample sizes for participating labs, for production data.***

Labid	Language	Country	T1 sample	T2 sample	Final T1-T2 sample	Average T1-T2 gap (days)
kau-cll	Arabic	Saudi Arabia	336	171	90	73
nbu	Bulgarian	Bulgaria	69	18	0	
brc-nijmegen	Dutch	The Nether-lands	26	25	20	39
brookes	English	UK	565	341	292	74
cogdevlabbyu	English	USA	93	89	39	23
clcu	English	UK	123	56	40	35
dsc	English	USA	32	14	5	86
ilpll	English	USA	263	115	49	73
ldl	English	Canada	63	29	17	33
Louisville	English	USA	62	nc	na	
owll	English	USA	10	nc	na	
rhul_baby_lab	English	UK	55	34	25	25
unlv	English	USA	56	27	0	
paris_team	French	France	654	535	466	28
goe	German	Germany	84	69	37	63
HaifaUniv	Hebrew	Israel	343	103	61	26
technion_il	Hebrew	Israel	335	164	111	37
babyling	Norwegian	Norway	786	182	173	20
multilada	Polish	Poland	670	246	223	27
hetsl	French	Switzerland	nc	400	ca	
msu	Russian	Russia	255	24	17	41
bcbl	Spanish	Spain	157	131	37	65
mltlab	Turkish	Turkey	57	57	40	31
		Total	5094	2830	1742	41

nc - data not collected
na - does not apply, giving that data in one sample was missing
ca - due to the lack of child's exact age, Swiss final data (n = 290) was used in the analyses of the relationship between maternal education and activities (cf project's OSF)
Note. Final T1-T2 sample contains data points that have passed the inclusion criteria after the merge of the matching T1 and T2 questionnaires.

Supplementary Material 2

Figure 1. ***Non-normalized (top) and normalized (bottom) gains in comprehension vocabulary as a function of the adjusted CDI score at T1 for the CDI tools Words and Gestures (wg) and Words and Sentences (ws). See Analyses_2.Rmd code on https://osf.io/ty9mn/ for data on production.***

Supplementary Material 3

We report below (Table 2) the fraction of participants having completed primary (level 1) and "some secondary" (level 2) education in the current sample, as well as in the sample used to derive vocabulary norms (from WordBank). We restricted our comparison to the handful of instruments for which we have maternal education information in both samples, as well as having commensurate measures of maternal education.

Table 2. Percentage of participants in the first two levels of maternal education scale (primary, and some secondary), for the norming sample (WordBank) and our sample. Differences in the maternal education between the Wordbank sample and the German and Spanish samples in the current study are likely attributed to smaller sample sizes in these two countries in our study.

Instrument (CDI)	Percentage of participants on WordBank	Percentage of participants in our sample (and sample size)
American English CDI	5.3%	5.5% (110)
Norwegian CDI	5.0%	4.0% (173)
French CDI	0%	1.1% (466)
German CDI	37.1%	13.5% (37)
Spanish CDI	5.4%	8.1% (37)

Supplementary Material 4

Table 3. ***Descriptive statistics of the variables used in the analyses of the production data (in percentiles).***

	Maternal education	Age at T1 (in days)	Book reading	Structured caregiver-child	Passive screen exposure	Outdoor activities	Digital games	Free play w. child	Singing	Speaking	Free play no adult	Time gap
Mean	4.5	588.9	4.1	2.5	3.3	4.4	0.5	5.8	3.7	5.9	5.2	35.8
SD	0.9	195.8	1.6	1.9	2.4	2.7	1.3	1.9	1.7	2.1	1.9	21.5
Min	1	245	0	0	0	0	0	0	0	0	0	4
Max	6	1075	9	8	9	9	7	9	9	9	9	111

Table 4. ***Descriptive statistics of the variables used in the analyses of the comprehension data (in percentiles).***

	Maternal education	Age at T1 (in days)	Book reading	Structured caregiver-child	Passive screen exposure	Outdoor activities	Digital games	Free play w. child	Singing	Speaking	Free play no adult	Time gap
Mean	4.4	418.6	3.9	2.0	2.4	3.8	0.2	5.9	3.9	6.0	4.9	40.3
SD	0.9	73.9	1.5	1.9	2.3	2.6	0.8	1.8	1.7	2.3	1.9	22.9
Min	1	245	0	0	0	0	0	0	0	0	0	5
Max	6	566	9	7	8	9	5	9	9	9	9	111

Supplementary Material 5

Impact of SES on Vocabulary at T1

To estimate the extent to which language comprehension and production depended on maternal education we fitted two Generalized Linear Mixed Models (GLMM; Baayen 2008) with beta error structure and logit link function (McCullagh & Nelder 1989; Bolker 2008). We used a beta rather than a Gaussian error function since the residuals of the Gaussian model were neither normally distributed nor homogeneous. Both models were identical in their fixed and random effects: As fixed effects, we included maternal education while controlling for sex, i.e., two fixed factors. We included random intercepts of country and random slopes of both predictors within country (cf Schielzeth & Forstmeier 2009; Barr et al. 2013). We excluded parameters for the correlations among the random intercept and slopes due to model convergence issues.

Maternal education was z-transformed (M=0, SD=1) to ease model convergence and the random effect of sex was manually dummy coded and centered. We fitted the model in R (version 4.0.3; R Core Team 2020) using the function glmmTMB (version 1.0.2.1; Brooks 2017). We determined the significance of individual fixed effects by comparing the respective full model with reduced models lacking them one at a time, utilizing likelihood ratio tests (Dobson 2002). We determined confidence intervals of model estimates by means of a parametric bootstrap (function simulate of the package glmmTMB) and estimated model stability by dropping countries one at a time and comparing estimates of models fitted to the respective subsets of the data to those obtained for the full data set. This revealed both models to be of moderate to good stability (see results). Neither of the two models was overdispersed (dispersion parameters; comprehension model: 1.00; production model: 1.048). The samples analysed for the two models comprised a total of 352 children from eight countries (comprehension model) and a total of 729 children from nine countries (production model).

As can be seen in Tables 1 and 2, neither maternal education nor sex were significant in either of the two models. However, sex was only marginally non-significant in the production model and all model estimates had the hypothesized sign (Tables 1 and 2).

Table 5. ***Results of the comprehension model (estimates together with standard errors, confidence limits, significance tests, and range of estimates (min, max) when dropping countries one at a time).***

term	Estimate	SE	lower Cl	upper Cl	x^2	df	P	min	max
Intercept	0.092	0.164	-0.227	0.429			[1]	-0.096	0.185
mat. educ[2]	0.041	0.058	-0.075	0.162	0.501	1	0.479	0.031	0.083
sex[3]	-0.146	0.151	-0.454	0.136	0.840	1	0.360	-0.254	-0.040

[1] not indicated because of having a very limited interpretation
[2] z-transformed to a mean of zero and a standard deviation (sd) of one; mean and sd of the original variable were 4.375 and 0.944, respectively
[3] dummy coded with female being the reference category

Table 6. ***Results of the production model (estimates together with standard errors, confidence limits, significance tests, and range of estimates (min, max) when dropping countries one at a time).***

term	Estimate	SE	lower Cl	upper Cl	x^2	df	P	min	max
Intercept	0.053	0.115	-0.169	0.296			-1.000	-0.018	0.121
mat. educ[2]	0.073	0.046	-0.017	0.162	2.285	1	0.131	0.040	0.105
gender[3]	-0.268	0.130	-0.520	-0.013	3.061	1	0.080	-0.349	-0.178

[1] not indicated because of having a very limited interpretation
[2] z-transformed to a mean of zero and a standard deviation (sd) of one; mean and sd of the original variable were 4.505 and 0.882, respectively
[3] dummy coded with female being the reference category

License