

**Language Through Routines: How Well Do Self-Report Measures Capture Parental
Language Input?**

Radboud Universiteit



MAX PLANCK INSTITUTE
FOR **PSYCHOLINGUISTICS**

Karlijn Beckers (s1119141)

Radboud Universiteit, Nijmegen

Faculty of Arts, Language and Speech Pathology

Master's Thesis (LET-TWM400)

Dr. Prof. Caroline Rowland, Dr. Selim Sametoğlu, and PhD Candidate Magda Matetovici

June 2025

Thesis conducted at the Max Planck Institute for Psycholinguistics, Nijmegen

Abstract

Early language development is a critical predictor of children's cognitive and socio-emotional abilities, influencing long-term academic and career successes. While both genetic and environmental factors contribute to language outcomes, parental language input plays a particularly important role, making reliable and valid assessment of this input essential. This quantitative observational study examined the extent to which a parental input quantity score, derived from the newly developed Family Routines Questionnaire, captures the amount of observed parental language input, as measured by audio recordings. The study 1) evaluated how well the parental quantity input score reflects recorded child-directed speech, 2) assessed the score's psychometric properties, and 3) explored its relationship with children's expressive vocabulary. Correlational analyses showed that the parental input quantity score does not directly correspond to observational audio data. However, the findings show generally good psychometric properties of the score and a moderate relation between the score and children's vocabulary. This suggests that the score may meaningful aspects of the language environment that are not similarly reflected in recordings. Nevertheless, more thorough analyses are needed to be able to make strong claims about the score's validity. These results highlight the value of well-constructed self-report measures as a complementary approach in early language development research.

Keywords: early language development, parental language input, self-report measures, observational measures

Vroege taalontwikkeling is een belangrijke voorspeller van de cognitieve en sociaal-emotionele vaardigheden van kinderen, en heeft invloed op hun latere succes op school en in hun loopbaan. Zowel genetische als omgevingsfactoren dragen bij aan taalvaardigheden, waarbij ouderlijk taalaanbod een cruciale bepalende factor blijkt te zijn. Het is daarom van belang om dit taalaanbod op een betrouwbare en valide manier te kunnen meten. In deze kwantitatieve observationele studie is onderzocht in hoeverre een score voor de hoeveelheid ouderlijk taalaanbod, afgeleid van de nieuw ontwikkelde Family Routines Questionnaire, overeenkomt met daadwerkelijk taalgebruik van ouders, zoals gemeten via audio-opnames. De studie 1) bestudeerde hoe goed de score voor ouderlijk taalaanbod de opgenomen taalgerichte spraak reflecteert, 2) evalueerde de psychometrische eigenschappen van deze score, en 3) onderzocht de relatie tussen deze score en de expressieve woordenschat van kinderen. Correlatieanalyses lieten zien dat de taalaanbodscore niet direct overeenkomt met de geobserveerde

audiogegevens. De bevindingen tonen echter aan dat de score over het algemeen goede psychometrische eigenschappen heeft en een matige relatie vertoont met de woordenschat van kinderen. Dit suggereert dat de score relevante aspecten van de taalomgeving meet die in de geluidsopnames niet op dezelfde manier tot uiting komen. Toch zijn grondige analyses nodig om sterke uitspraken te kunnen doen over de validiteit van de score. Deze resultaten benadrukken het belang van goed ontwikkelde zelfrapportage-instrumenten als waardevolle aanvulling op observatiemethoden binnen onderzoek naar vroege taalontwikkeling.

Trefwoorden: vroege taalontwikkeling, ouderlijk taalaanbod, zelfrapportage-instrumenten, observationele meetinstrumenten

Preface and Acknowledgements

I am pleased to present my master's thesis, which represents the final stage of my studies in Language and Speech Pathology. This research has been carried out with great interest and dedication. Throughout the process, I have gained valuable insights, not only regarding the thesis' subject, but also about the conduct of research itself. Completing and writing this thesis was positively challenging and rewarding. I hope that my work, in some way, is able to meaningfully contribute to the field.

This would not have been possible without the support of several people, whom I would like to sincerely thank. First of all, I want to thank my supervisors from the Max Planck Institute, namely Caroline Rowland, Selim Sametoğlu, and Magda Matetovici, for their guidance and for always offering assistance when needed. I greatly appreciated our collaboration and learned a lot from their expertise. I am also grateful to my fellow students and friends for their support and encouragement, as well as to my family. Finally, I would like to thank the Max Planck Institute as a whole and all the people I've got to collaborate with. They have given me the opportunity to be involved in their research approaches and methodologies, from which I gained valuable knowledge and experience.

Introduction

Language development in early childhood is a fundamental process that not only lays the foundation for general intellectual and cognitive skills, such as executive functions, but also impacts socio-emotional development, including emotion regulation and the acquisition of general social skills. Furthermore, early language skills are significant predictors of long-term outcomes, such as academic achievement and career success (Anderson et al., 2021; Fu & Zheng, 2024; Wang et al., 2020). Given its broad impact, investigating the factors that contribute to variations in children's language development, especially those that can be modified to improve outcomes, is essential. While individual language abilities are partly determined by genetic factors (Anderson et al., 2021; Fu & Zheng, 2024), environmental factors play a critical role as well. In particular, a rich and stimulating language environment significantly influences language development (Anderson et al., 2021; Bennetts et al., 2016; Cates et al., 2023; Coffey & Snedeker, 2025; Fu & Zheng, 2024; Richards et al., 2017; Wang et al., 2020). Parents and caregivers form a key source of such language input (Anderson et al., 2021; Cates et al., 2023; Coffey & Snedeker, 2025; Fu & Zheng, 2024; Wang et al., 2020). This highlights the need to measure the language input provided by parents in an accurate and valid way. Therefore, the present study investigates whether a parental questionnaire captures the same construct, namely the quantity of parental language input, as measured by audio recordings. To comprehend its significance, it is important to understand how strongly parental language input relates to children's language development.

A large body of research has shown that parental speech is significantly associated with children's vocabulary growth, grammatical development, and broader communicative skills from infancy through early childhood (Anderson et al., 2021; Cates et al., 2023; Coffey & Snedeker, 2025; Fu & Zheng, 2024; Wang et al., 2020). Researchers commonly distinguish this input's quantitative (e.g., the number of words or utterances) and qualitative (e.g., vocabulary diversity, morphological complexity) dimensions. However, these aspects are frequently intertwined with each other in naturalistic environments, making it challenging, if not impossible, to separate them analytically (Jones & Rowland, 2017). Some studies have demonstrated that qualitative features tend to have stronger associations with children's language development than quantitative measures (Anderson et al., 2021; Fu & Zheng, 2024), while others reported no significant difference between the predictive strength of language quantity and quality (Coffey & Snedeker, 2025). This suggests that both dimensions of language input are relevant. Given the crucial role of early language input in children's developmental

outcomes, reliable tools that can capture parental language input in an accurate, efficient and scalable way are essential.

A widely used method for assessing parental language input involves the use of observational measurements, such as audio or video recordings of child-parent interactions. Traditionally, these recordings are manually transcribed and analyzed. While this approach remains common, it presents several limitations. First, it is highly time-consuming and resource-intensive, often requiring extensive training of coders, which makes large-scale studies more difficult and expensive to conduct (Cates et al., 2023; Cristia et al., 2024; Wang et al., 2020). Second, these observations are typically based on short recordings, often between 10 and 60 minutes, which may offer only a limited snapshot of a child's language environment and fail to capture day-to-day variability (Scaff et al., 2024; Wang et al., 2020). Moreover, the presence of recording equipment or researchers, especially in artificial laboratory settings, may influence the behavior of parents and children, potentially reducing ecological validity (Scaff et al., 2024). To overcome some of these limitations, automated recording and analysis systems have been developed to capture parental language input more efficiently. LENA (Language ENvironment Analysis) is an example of such systems and is used to quantify children's linguistic environments, using a small, wearable audio recorder and software. It produces measures of Adult Word Count (AWC), Conversational Turns (CTC), Child Vocalization Count (CVC), classifications for overlapping speech, background sounds and silence. From these measures, CTC has shown to be the strongest predictor for children's language development (Wang et al., 2020). This approach reduces the need for manual transcription and allows researchers to capture more naturalistic input across a full day. However, LENA has notable limitations as it is somewhat outdated, including high costs, algorithms primarily optimized for American English, moderate accuracy, and reduced reliability in capturing detailed and linguistically diverse speech features (Räsänen et al., 2021). The ACLEW pipeline (Analyzing Child Language Experiences around the World), in contrast, is inspired by LENA and provides an open-source alternative (Cristia et al., 2024), providing similar metrics with additional measures such as the total duration of vocalizations and the average length of a vocalization. However, ACLEW shares comparable limitations in reliability and demands higher technical expertise (Cristia et al., 2024). Within the ACLEW framework, ALICE (Automatic Linguistic unit Count Estimator) estimates the number of phonemes, syllables, and words using acoustic features rather than lexical recognition, enhancing cross-linguistic generalization. Validation studies indicate that ALICE's output closely aligns with human annotations and performs comparably to, or better than LENA in English corpora, although its accuracy can be affected

by diarization errors (i.e., incorrectly allocating speech to individuals) and has been primarily evaluated in English samples (Räsänen et al., 2021). Although these automated tools are valuable for group-level analyses, their reliability for individual differences is limited (Cristia et al., 2024).

An alternative to observational measures is the use of self-reported measures. These tools are generally cost-effective, time-efficient, and easy to administer on a large scale, making them particularly suitable for research involving diverse or resource-constrained populations (Anderson et al., 2021; Bennetts et al., 2016; Cates et al., 2023). One such instrument is the StimQ2, a valid and reliable parent-report questionnaire designed to assess the cognitive stimulation present in the home environment of young children. It correlates significantly and positively with various child development outcomes, including cognitive, linguistic, and socio-emotional skills (Cates et al., 2023). Additionally, Bennetts et al. (2016) employed several parent-reported measures, including the Sure Start Language Measure (SSLM), the communicative subscale of the Ages and Stages Questionnaire (ASQ), and the MacArthur-Bates Communicative Development Inventory (Short-Form, CDI). Richards et al. (2017) used the Parent Perception Questionnaire and the LENA Developmental Snapshot as parent-report measures, which are overall reliable tools. Despite their advantages, existing self-report instruments are not without limitations. One concern is their validity and reliability, as self-report measures are vulnerable to social desirability bias and may not accurately reflect parent's assessments of their own language input (Bennetts et al., 2016; Cates et al., 2023; Richards et al., 2017). Related to this, parents tend to systematically overestimate how much they talk to their child (Richards et al., 2017), particularly in lower socio-economic environments (Bennetts et al., 2016). Moreover, while the practical advantages of self-report measures are clear, their validity for specifically assessing the quantity of parental language input remains underexplored. Although some studies have compared self-report and observational measures for measuring the same construct, these have mainly focused on whether both methods predict similar child language outcomes. Notably, Bennetts et al. (2016) found moderate agreement between the two approaches. Supporting this, Anderson et al. (2021) reported that parent-reported measures of child language skills, specifically the Communicative Development Inventories (CDIs, Zink & Lejaegere, 2002), show significant alignment with laboratory-based assessments. However, few studies have examined whether the two approaches both measure parental language input, and existing research shows varying findings. For example, Richards et al. (2017) found a weak correlation between the two methods for parents input quantity, while Bleses et al. (2018) reported substantial correlations for parental vocabulary use frequency. In

light of these findings, Bennetts et al. (2016) suggest that self-report tools should not be regarded as direct substitutes for observational methods, but rather as valuable complementary instruments, offering practical benefits such as scalability and accessibility. However, in certain contexts, such as large-scale studies or research involving populations with constrained resources, carefully constructed self-report measures may be the most feasible option for assessing parental language input.

In summary, the discussed studies suggest that the language input children receive from parents plays a crucial role in early language development. Self-reported measurement methods offer a more cost-effective and efficient way to measure parental language input than observational methods. However, there is no real consensus on whether self-reported methods capture the same constructs as observational methods. In particular, research directly comparing these approaches is limited and presents mixed findings. The current study, ‘De Kleine TIJGER’ (DKT), aims to address this gap by developing and evaluating a new self-report instrument, the Family Routines Questionnaire, in terms of how well it captures parental language input quantity. This will be done in three ways: first, we want to analyze how well self-reported parental language input – operationalized as a parental input quantity score derived from a number of questions designed to measure the amount of child-directed speech from parents – aligns with observed input obtained through ALICE processed audio recordings. Second, the study aims to evaluate the psychometric properties of this parental input quantity score. Third, the study investigates the extent to which the parental input score is associated with children’s vocabulary size, as measured by the CDIs (Zink & Lejaegere, 2002). Examining this relationship helps to establish the potential relevance of the questionnaire for child language outcomes, as vocabulary is an important component and predictor of a child’s language development (Capelli et al., 2024), thereby informing future research.

To address these aims, the following overarching research question was formulated: “To what extent does the parental input quantity score, derived from the Family Routines Questionnaire, serve as a valid measure of the amount of child-directed speech, as measured by daylong audio recordings in Dutch families with young children?”. Accordingly, three hypotheses are proposed. First, we expect that the parental input score will show a strong, positive, and statistically significant correlation with the observed amount of parental verbal input. Although literature supporting this is limited, the expectation remains reasonable, as this study uses methods that differ from those in previous research, and both the parental input score and ALICE-derived variables are designed to measure the quantity of child-directed speech. Second, we anticipated that the items of the parental input score will demonstrate satisfactory

psychometric properties, including a high level of internal consistency reliability of the scale, and moderate to strong intercorrelations among the items and item-total correlations. Third, we hypothesize that there will be a small to moderate positive and statistically significant correlation between the parental input score and children's vocabulary, measured through the CDIs. This expectation is based on literature that shows a robust relationship between parental input captured via observational measures and language development in children (Anderson et al., 2021). If our parental input score also captures individual differences in parental input well, we would expect these scores to correlate well with the language development of the parents' children.

Finally, an additional exploratory analysis was conducted to assess the reliability of parental responses, which was not part of the study's pre-registration. In addition to questions on parental language input, the Family Routines Questionnaire included a weekly schedule in which each parent independently reported who typically cared for the child during each part of the day – morning, afternoon, and evening – across a typical week from Monday through Sunday. The level of agreement between parents' responses was then calculated. Given that this schedule reflects relatively objective information (i.e., a parent is either present with the child at a given moment or not), the level of agreement is interpreted as an indicator of the general reliability of parental responses throughout the questionnaire.

This study will provide valuable insights into the validity and reliability of self-reported measures of parental language input and their value for children's language development, thereby contributing to improving assessment tools and providing guidance for future research and policy development in early language development.

Method

The dataset description and the intended analyses were pre-registered prior to data analysis, which are available on the Open Source Foundation (OSF) website (<https://orcid.org/0009-0000-7461-4468>). Following pre-registration, certain aspects related to hypotheses and data analysis were modified. The final analytical approach, reflecting these modifications, has been described below.

Participants

The participants were parents drawn from the ongoing Taal in het BabyBrein (TBB) study, who were originally recruited through the database maintained at the Baby and Child Research Center, located in Nijmegen. These participants were selected as they had already

provided daylong audio recordings and measures of child language development as part of the TBB study.

To ensure developmental typicality and minimize confounding factors, specific criteria were established for the TBB study, which were applied during the TBB recruitment phase and have been maintained for inclusion in the current study. These criteria, as shown in Table 1, include factors such as monolingual Dutch-speaking families, birth and health conditions, developmental disorders, language exposure, and informed consent capacity.

Table 1

Inclusion and Exclusion Criteria for the TBB study

| Inclusion Criteria | Exclusion Criteria |
|-------------------------------------|--|
| Monolingual Dutch-speaking families | Premature birth |
| | Birth weight below 2500 grams |
| | Chronic or recurrent ear infections |
| | Diagnosed developmental disorder |
| | Visual impairment |
| | Exposure to a language other than Dutch for one or more full days per week |
| | Parents unable to provide informed consent due to a learning disability |

Additional inclusion criteria for the DKT study are outlined in Table 2.

Table 2

Additional Inclusion Criteria for the DKT study

| |
|--|
| Monolingual Dutch-speaking families |
| Child aged at least 2 years and 3 months at the time of data collection* |
| Daylong audio recordings available (24 months) |
| N-CDI vocabulary production scores (24 months) |

**Note.* This age was chosen to ensure sufficient time between data collections and to prevent overburdening parents.

A purposive, nested sampling strategy was used, meaning that eligible participants from the existing TBB study cohort were selected for the current study, as they had already provided relevant data. This approach allowed for the integration of existing observational data with newly gathered parental report measures.

For this study, a total of 37 families were initially identified as suitable to participate (see Table 2 for the criteria). All 37 families were invited. However, it was only after the invitation process had been completed that it became clear that one of these families did not, in fact, have available ALICE data. As a result, the number of truly eligible families was reduced to 36. Of these, 28 families completed the Family Routines Questionnaire and were included in the final sample. All of the families were recruited from Nijmegen and surrounding cities.

Tables 3 until 5 provide demographic information for the participating families. Note that since different measures were taken at different times, age information relative to individual measures are illustrated in Table 3. In terms of questionnaire completion, 78.57% of the families had both parents fill it out (N = 22). Among these families, 100% were mother – father pairs (N = 22). For the 6 families where only one parent participated, 83.33% were female (N = 5) and 16.66% were male (N = 1). Regarding the socioeconomic status (SES) of the participants, parental education levels were used as a proxy, with high SES defined as having completed higher education (HBO) or above. Information regarding SES is presented in Table 5.

Table 3*Ages of Participating Parents and Children*

| Group | Age Calculated From* | Mean Age (Years;Months) | SD (Years;Months) | Range (Years;Months) |
|---|--|-------------------------|-------------------|----------------------|
| Parents – TBB study | Date of family questionnaire completion (TBB) | 34;3 | 3;11 | 27;0 – 48;0 |
| Parents – DKT study | Start date data collection current study (DKT) | 36;2 | 3;3 | 29;0 – 47;0 |
| Children – TBB study (N-CDI) | Date of N-CDI completion (TBB) | 1;11 | 0;2 | 1;0 – 2;1 |
| Children – TBB study (audio recordings) | Date of first audio recording (TBB) | 1;11 | 0;2 | 1;0 – 2;0 |
| Children – DKT study | Start date data collection current study (DKT) | 2;6 | 0;2 | 2;3 – 2;11 |

**Note.* Ages were calculated by subtracting the participant's date of birth from the reference date as noted in the second column. For the ages of parents and children in the DKT study, the official start date of data collection was used as the reference point rather than individual questionnaire completion dates. This was done to ensure consistency, as not all families had both parents fill in the questionnaire on the same date.

Table 4*Gender Distribution of Participating Parents and Children*

| Group | Male (%) | Female (%) | Total (N) |
|----------|--------------|--------------|-----------|
| Parents | 46% (N = 23) | 54% (N = 27) | 50 |
| Children | 54% (N = 15) | 46% (N = 13) | 28 |

Table 5*Socioeconomic Status (SES) of Participating Parents*

| Role | High SES (%)* | High SES (N) | Total (N) |
|---------|---------------|--------------|-----------|
| Fathers | 73.91% | 17 | 23 |
| Mothers | 95.59% | 25 | 27 |

**Note.* High SES defined as having completed higher education (HBO) or above.

Ethical Approval and Data Management

Ethical approval for the TBB study was granted by the Ethics Committee of the Faculty of Social Sciences at Radboud University (ECSW-2022-070). The current study was reviewed as an extension of the TBB study, involving minimal additional risk and the re-use of existing data.

Although the DKT study builds on TBB data, participants were clearly informed that DKT is a new and independent study. At the start of the online questionnaire, participants were explicitly asked to provide renewed informed consent, with clear information that participation was entirely voluntary and not linked to their prior involvement. Importantly, they were also informed that their existing TBB data – specifically vocabulary scores and daylong recordings – would be linked to their responses in the current study. It was made clear, however, that researchers did not have access to the actual content of the audio recordings. Only automatically extracted summary information, such as the amount of child-directed speech, was available for analysis. Both caregivers were invited to complete the questionnaire and ideally, both of them would do this. However, participation by only one caregiver was acceptable as well, in cases where the second caregiver was unavailable or declined to participate. Caregivers were instructed to complete the questionnaire independently, without consultation or collaboration. They were reminded that they could withdraw from the study at any time without any consequences, and that requests for deletion of personal data could be made up to one month after participation. Research data collected prior to withdrawal could still be used in anonymized form, as stated in the information and consent forms (see Appendix A).

All data were processed in accordance with the EU General Data Protection Regulation (GDPR). Data were pseudo-anonymized using unique participant ID codes, and personally identifying information was stored separately in a password-protected file accessible only to the authorized research team. All collected data are securely stored on servers at the Max Planck

Institute for Psycholinguistics and will be archived for a minimum of ten years. Fully anonymized summary data may be shared openly upon publication.

Design

This study followed a quantitative observational design with a partially overlapping cohort structure, combining retrospective secondary data from the TBB study with newly collected primary data through the DKT questionnaire.

For the first hypothesis, the parental input quantity score variable and the ALICE variables are included to explore their relationship. The second hypothesis examines the psychometric properties of the parental input score variable. Lastly, the parental input score variable and the child vocabulary score variable are included to investigate their association.

Procedure

Data from the TBB study were collected through full-day in-home recordings and parental completion of an online version of the Dutch version of the MacArthur-Bates Communicative Development Inventories Words & Sentences (N-CDI: WS) (Zink & Lejaegere, 2002) from when the children were approximately 24 months old. For the present study, eligible TBB participants were recontacted via postal invitation and they were asked to fill in a newly created questionnaire. The invitation letter can be found in Appendix B. If no response was received within one week, follow-up contact was made by telephone. A final reminder email was sent one week before the questionnaire closed, specifically targeting participants who had expressed interest during the follow-up call, but had not yet completed the questionnaire. This email was also sent to families in which only one parent had filled it in so far.

Prior to distribution, the questionnaire was reviewed and pre-tested by several staff members of the Max Planck Institute to ensure clarity, accessibility and ease of use. Participants were able to complete the online questionnaire at a time convenient to them. Digital informed consent was obtained at the beginning of the questionnaire, and participants received a modest financial compensation of €5 upon completion.

Measures

Audio Recordings

Parents were invited to participate in optional audio recordings of their children's naturalistic language environment. These were collected and processed with ALICE, an

automated speech analysis algorithm used to extract several aspects of speech, which generates multiple measures of both child speech and the surrounding linguistic input (Räsänen et al., 2021). The recordings were, ideally, collected over two full days (16 hours a day) at the child's age of 24 months. The exact duration of recordings per family is provided in Appendix C.

All the measures derived via ALICE and are presented in Appendix D. Most of these measures are categorized into male (MAL) and female (FEM) speaker variables, with the assumption in this study that MAL reflects mostly paternal speech input and FEM reflects mostly maternal speech input. The following ALICE variables were selected as indicators for the amount of child directed speech produced by the fathers: 'Sum words MAL KCHI', 'Sum time MAL KCHI', 'Total turns MAL', 'Tot Phonemes MAL', 'Tot Syllables MAL', and by the mothers: 'Sum words FEM KCHI', 'Sum time FEM KCHI', 'Total turns FEM', 'Tot Phonemes FEM', 'Tot Syllables FEM'. 'Sum words KCHI' represents the total number of words spoken by a male ('Sum words MAL KCHI') and female ('Sum words FEM KCHI') individual respectively, during their interactions with the child that is wearing the audio recording device. 'Sum time KCHI' reflects the total amount of time the individual interacts with the child. Since both these variables are directly drawn from speech produced in the child's presence and directed at the child, they serve as observational measures of the amount of child-directed speech. 'Total turns', which measures the number of conversational turns between the caregiver and the target child, closely resembles LENA's conversational turns count (CTC), which has been identified as a strong predictor of children's language development (Wang et al., 2020). In addition, 'Tot Phonemes' and 'Tot Syllables' measure the total number of produced phonemes and syllables by either a male or female individual, respectively, and were included to provide more detailed linguistic metrics. Phoneme-level analysis offers a fine-grained view of the speech input and has been shown to slightly outperform word- and syllable-based metrics in predicting early language outcomes (Räsänen et al., 2021). However, syllables remain a useful metric and are thus included alongside phonemes.

Since the duration of the audio recordings varied between families, the ALICE variables needed to be adjusted to allow for comparisons across families. To achieve this, each variable was multiplied by the proportionalized recording durations, which were scaled between 0 and 1.

Children's Expressive Vocabulary Skills

In addition to the audio recordings, the children's expressive vocabulary development at 24 months was assessed using the N-CDI: WS (Zink & Lejaegere, 2002). This is a

standardized 779 item parent-report questionnaire designed to assess early vocabulary development in young children. It includes words that are typically acquired by young children, and parents are instructed to indicate which words their child actively produces. The CDI has been validated as a reliable instrument for measuring expressive vocabulary in Dutch-learning children between 16 and 30 months of age (Zink & Lejaegere, 2002). The questionnaire was administered online to parents when their child was 24 months old via Castor. This is a secure and user-friendly online platform widely used in clinical and behavioral research for designing and distributing digital forms, in accordance with international standards such as GCP and GDPR (Castor, n.d.).

The N-CDI scores were defined as the total number of items marked by parents on the questionnaire. Each item was scored dichotomously (0 = not produced, 1 = produced), and total scores were obtained by summing all produced items per child. Higher scores indicate a larger expressive vocabulary, while lower scores reflect a smaller expressive vocabulary.

Family Routines Questionnaire

For the purposes of this study, we created a new online questionnaire on family routines and implemented it in Qualtrics, requiring approximately 15 minutes per participant to complete. Participants were encouraged to complete the questionnaire independently to minimize the risk of influencing one another's responses. The questionnaire was designed to provide insight into the linguistic environment of young children and covered several areas relevant to the study, including parental responsiveness and language input, the language use of both parents and children, and the extent of media exposure. Additionally, it addressed reading activities, family structure and daily routines, as well as socio-demographic information. The final questionnaire contained approximately 40 items, which were based on previously validated measures and questionnaires. Specifically, they were adapted from the StimQ2 (Toddler and Preschool) (Cates et al., 2023), the Home Literacy Environment Questionnaire (HLEQ) (Marjanovič Umek et al., 2005), the Home Learning Environment Questionnaire (HLEQ) (Krousorati et al., 2022), and the Children's Music-Related Behavior Questionnaire (CMRBQ) (Valerio et al., 2012). Additionally, items were drawn from questionnaires about screen exposure (Martinot et al., 2021), the National Survey of Children's Health (Child and Adolescent Health Measurement Initiative, 2022), the Seven-in-Seven Screen Exposure Questionnaire (Yalçın et al., 2021), and the Digital-Screen Exposure Questionnaire (DSEQ) (Kaur et al., 2021). Furthermore, several questions were adapted from

a more extensive family background questionnaire used in a different study from the Max Planck Institute, which had been created for a new adaptation of the N-CDI.

The development of the questionnaire and the selection of specific items for the parental input score involved extensive deliberation and rigorous review to ensure both relevance and validity. The final composite parental input score consists of 11 carefully selected items designed to capture different dimensions of parental speech quantity, as presented in Table 6. These items were specifically chosen to reflect both how much parents talk with their child and how much language they offer more generally, across different everyday contexts. Together, they provide a meaningful approximation of the amount of parental language input children receive in daily life. As such, the score was intended to serve as a valid and reliable measure of parental input quantity. Both the original items underlying the final questionnaire items and the final questionnaire itself are included in Appendix E.

Table 6

Selected Items for the Parental Input Quantity Score and Response Options

| Items Parental Input Quantity Score |
|---|
| <p>Q21. Frequency of parental language use in the supermarket: “Hoe vaak benoemt u de namen, de kleuren of de grootte van producten in de supermarkt, wanneer u uw kind daarheen meeneemt?”</p> <ul style="list-style-type: none"> <input type="radio"/> Nooit <input type="radio"/> Bijna nooit <input type="radio"/> Soms <input type="radio"/> Vaak <input type="radio"/> Heel vaak |
| <p>Q27. Frequency of parental storytelling and narrative interaction: “Hoe vaak vertelt u uw kind verhalen (zoals sprookjes, verzonden verhalen zonder een boek te gebruiken, of over activiteiten die u in het verleden samen hebt gedaan)?”</p> <ul style="list-style-type: none"> <input type="radio"/> Nooit <input type="radio"/> Bijna nooit <input type="radio"/> Soms <input type="radio"/> Vaak <input type="radio"/> Heel vaak |
| <p>Q30. Frequency of parental talk about the child’s environment: “Sommige verzorgers praten met hun kinderen over hun omgeving en wat er om hen heen gebeurt. Wanneer ouders/verzorgers hiermee beginnen is heel verschillend. Hoe vaak praat u zo met uw kind?”</p> <ul style="list-style-type: none"> <input type="radio"/> Nooit <input type="radio"/> Bijna nooit <input type="radio"/> Soms |

- Vaak
- Heel vaak

Q34. Frequency of parental talk during child's playtime: "Soms praten verzorgers met hun kind als hij/zij met speelgoed speelt, maar soms laten ze hun kind ook alleen spelen. Hoe vaak praat u met uw kind als hij/zij met speelgoed speelt?"

- Nooit
- Bijna nooit
- Soms
- Vaak
- Heel vaak

Q38. Frequency of parental singing and recitation of rhymes during daily activities: "Sommige verzorgers zingen liedjes of zeggen rijmpjes met hun kinderen tijdens dagelijkse activiteiten, zoals in bad gaan, aankleden, opruimen of klaarmaken om ergens naartoe te gaan. Hoe vaak zingt u liedjes of zegt u rijmpjes met uw kind?"

- Nooit
- Bijna nooit
- Soms
- Vaak
- Heel vaak

Q22. Frequency of parental correction of grammatical errors: "Soms maken jonge kinderen grammaticale fouten (zoals 'loopte' in plaats van 'liep', of 'boeks' in plaats van 'boeken'). Als uw kind dit doet, hoe vaak verbetert u uw kind dan?"

- Nooit
- Bijna nooit
- Soms
- Vaak
- Heel vaak

Q18. Frequency of parental verbal language input: "Hoeveel praat u met uw kind in vergelijking tot andere ouders met kinderen van dezelfde leeftijd?"

- Niet veel
- Minder dan de gemiddelde ouder
- Gemiddeld
- Meer dan de gemiddelde ouder
- Veel meer dan de gemiddelde ouder

Q23. Frequency of shared book reading before bedtime: "Denk aan een normale week. Hoeveel dagen per week leest u een boek samen met uw kind vóór het slapengaan?"

- Minder dan één dag per week
- Ongeveer één dag per week
- Twee tot drie dagen per week
- Vier tot zes dagen per week
- Elke dag

Q24. Frequency of shared book reading at other times of the day: "Denk aan een normale week. Hoeveel dagen per week leest u boeken samen met uw kind op andere momenten van de dag dan vóór het slapengaan?"

- Minder dan één dag per week
- Ongeveer één dag per week
- Twee tot drie dagen per week
- Vier tot zes dagen per week
- Elke dag

Q37. Frequency of parental engagement in book-related conversations: “Er zijn verschillende manieren om met kinderen te lezen. Soms hebben ouder en kind gesprekjes over wat er in het boek gebeurt (door vragen te stellen aan het kind, de gebeurtenissen op de pagina te bespreken, etc.). Hoe vaak voert u zulke gesprekken wanneer u met uw kind leest?”

- Nooit
- Bijna nooit
- Soms
- Vaak
- Heel vaak

Q36. Frequency of parental discussion about children’s media exposure: “Hoe vaak praat u met uw kind over tekenfilms of andere (kinder)programma's die uw kind gezien heeft? Denk aan alle (kinder)programma's, niet alleen die op tv, maar ook die op een telefoon, tablet of laptop worden bekeken.”

- Nooit
- Bijna nooit
- Soms
- Vaak
- Heel vaak

To construct the parental input quantity score, responses to all 11 selected items were first coded to ensure a uniform response scale across items. While each item originally included five response options, the content of these options varied slightly between items. Specifically, eight items used frequency-based options (Q21, Q27, Q30, Q34, Q38, Q22, Q37 and Q36) ranging from ‘never’ to ‘very often’ (see Table 6), two items referred to the number of days per week (Q23 and Q24), and one item compared the respondent to an average parent (Q18). Despite these differences, all higher options consistently indicated greater or more frequent parental input. This allowed all responses to be recoded onto a scale from 1 to 5, to make comparability possible. Subsequently, these values were proportionalized on a scale ranging from 0 to 1. Specifically, scores of 1 were recoded as 0.00, 2 as 0.25, 3 as 0.50, 4 as 0.75 and 5 as 1.00, where 0 represents the lowest and 1 the highest level of reported input. For each of the 50 participating parents, an individual average proportion score was then calculated. This was done by summing the proportional scores across all 11 items and dividing the total by the number of items. As a result, the final parental input scores reflect the average level of speech quantity per parent.

Data Analysis

The data were exported from Qualtrics and cleaned and analyzed using Excel (Microsoft Corporation, 2019) and R version 4.5.0 (R Core Team, 2025) within the RStudio® Integrated Development Environment (IDE) version 2023.12.0+369 (RStudio Team, 2023). The full and detailed data analysis can be found in the R scripts provided in Appendix F.

The significance level was set at $p < .05$ for all statistical tests, with 95% confidence intervals to indicate estimate precision. The statistical hypotheses are presented in Table 7. Regarding the first research question, a strong, positive ($r \geq 0.70$) (Gignac & Szodorai, 2016) and statistically significant correlation ($p < .05$, one-tailed) between the derived parental input quantity score and observational ALICE measures of the amount of the child-directed speech was expected. This is based on methodological improvements over previous studies and the fact that both the score and ALICE measures intend to measure the same construct. These results would imply that a meaningful relationship exists between the parental input and the observed speech data. To examine this, a Pearson's correlation test was conducted between the parental input score, calculated separately for fathers and mothers as the average of proportioned individual item scores, and child-directed speech, as measured by the selected ALICE variables. This analysis was performed separately for fathers and mothers.

The second research question asked whether the psychometric properties of the items comprising the parental input score would be at least satisfactory. Specifically, we predicted that the following would be true: a) the overall level of internal consistency reliability of the scale, as measured by Cronbach's alpha, would be high ($\alpha \geq 0.70$) (Adamson & Prion, 2013; Jensen, 2003), b) the inter-item correlations, reflecting the degree of consistency among individual items, would be moderate ($r \geq 0.30$) to strong ($r \geq 0.50$) (Hajjar, 2018), and c) the item-total correlations, interpreted as the contribution of each item to the composite score, would be moderate ($r \geq 0.30$) to strong ($r \geq 0.50$) (Gliem & Gliem, 2003). These expectations were assessed through psychometric analyses including the calculation of internal consistency reliability, inter-item correlations, and item-total correlations.

For the third and last question, we predicted a small to moderate ($0.10 \leq r \leq 0.30$) (Gignac & Szodorai, 2016) positive and statistically significant correlation ($p < .05$, one-tailed) between the parental input quantity score and the N-CDI production vocabulary scores of the children. This would show that our parental input score has a good association with the vocabulary size in children. To test this, a Pearson's correlation test was used separately for fathers and mothers, to investigate the relationship between the parental input score and the children's CDI scores.

Table 7*Overview Research Questions and Statistical Hypotheses*

| Research Question | Statistical Hypothesis (H_0 and H_1) |
|--|---|
| 1. Does the parental input quantity score correlate with the ALICE derived variables? | <p>H_0: The correlation between the parental input quantity score and child-directed speech is no different from 0.</p> <p>H_1: The correlation between the parental input quantity score and child-directed speech is positive and significantly different from 0.</p> |
| 2. Do the items that make up the parental input quantity score have good psychometric item properties? | <p>H_0: The items do not meet satisfactory psychometric properties.</p> <p>H_1: The items meet satisfactory psychometric properties, as defined in the text.</p> |
| 3. Does the parental input quantity score correlate with children's expressive vocabulary size? | <p>H_0: The correlation between the parental input quantity score and children's vocabulary scores is no different than 0.</p> <p>H_1: The correlation between the parental input quantity score and children's vocabulary scores is positive and significantly different from 0.</p> |

Exploratory Analysis

Ultimately, 22 families participated in which both parents completed the study. This allowed for an additional exploratory analysis to assess the extent to which parents agreed on the distribution of time spent with the child throughout the week, as reported independently in the weekly schedule. Since this schedule captures relatively objective data, the level of agreement is considered as an estimate of the general reliability of the parents' responses across the questionnaire. This analysis was not included in the original pre-registration since it was

not expected to have complete data from a large number of families. Thus, it was added solely as an exploratory analysis.

To calculate the mean agreement score, which was done manually in Excel, the response options to the questions regarding the weekly schedule were coded. Parents were informed to complete the schedule independently. For each part of the day (morning, afternoon, and evening) and each day of the week, they selected one of six response options: ‘uzelf’, ‘de andere ouder’, ‘beide ouders’, ‘dagopvang’, ‘andere persoon dan ouders’, or ‘zelfstandig spelen/slappen’. A specific coding adjustment was made to address situations where parents selected different options that referred to the same caregiver. For example, if parent 1 selected ‘uzelf’ and parent 2 selected ‘de andere ouder’, they were referring to the same person, despite choosing different response options. In this case, responses were coded so that ‘uzelf’ selected by parent 1 and ‘the other parent’ selected by parent 2 were both classified as ‘parent A’. In contrast, if parent 1 selected ‘de andere ouder’ and parent 2 selected ‘uzelf’, these were coded as ‘parent B’. The other response options remained unchanged in the coding process. Agreement was then assessed daily and weekly, expressed as percentages. For each day, agreement was coded as 1 if both parents provided the same (coded) answer at a given time point, and 0 if their answers differed. Daily agreement per family was calculated by dividing the number of matching responses between the two parents by the number of time blocks that day. Weekly agreement per family was then determined by summing all instances of agreement across the seven days and dividing by the total number of time blocks in the week. Finally, overall mean agreement scores across all families were calculated as percentages by aggregating the daily and weekly instances of agreement from each family.

Results

Table 8 presents the descriptive statistics for the language input provided by fathers (MAL) and mothers (FEM), based on the ALICE derived variables. To control for variation in recording lengths across families (see Appendix C), the raw scores were not used. Instead, all variables were proportionalized by multiplying each variable by the scaled recording duration. Therefore, the values reported in the table reflect duration-adjusted language input, meaning that they should be interpreted as relative measures. Higher values indicate greater language input relative to the recording length. On average, fathers produced 1618 words during their interactions with the target child (‘Sum words MAL KCHI’) (SD = 1059.49, median = 1615.03, 95% CI [1321.81 – 1915.12]), whereas mothers spoke significantly more (‘Sum words FEM KCHI’), with a mean word count of 5251.01 (SD = 2735.83, median = 5466.48, 95% CI

[4484.98 – 6017.05]. A similar pattern was found in terms of the total duration (in seconds) of interaction time between the parent and child: fathers ('Sum time MAL KCHI') spoke for an average of 518.94 seconds (SD = 379.75, median = 506.75, 95% CI [412.61 – 625.27], compared to 1875.45 seconds for mothers ('Sum time FEM KCHI') (SD = 1029.19, median = 1931.65, 95% CI [1587.28 – 2163.62]). This difference also emerged in the number of conversational turns between parent and child. Fathers (Total turns MAL') contributed an average of 1035.64 turns (SD = 685.24, median = 959.65, 95% CI [843.77 – 1227.51]), while the mean turns for mothers ('Total turns FEM') was 3013.02 (SD = 1534.41, median = 3372.18, 95% CI [2583.39 – 3442.66]). The wide ranges observed in these measures highlight substantial individual variation in language input quantity.

Table 8

Descriptives ALICE Variables

| Variables | Mean | Median | Range | SD | 95% CI |
|-----------------------|----------|----------|-----------------|----------|--------------------------|
| Sum words MAL KCHI | 1618.47 | 1615.03 | 0.00 – 4730.22 | 1059.49 | [1321.81 – 1915.12] |
| Sum time MAL KCHI | 518.94 | 506.75 | 0.00 – 1783.73 | 379.75 | [412.61 – 625.27] |
| Total turns MAL | 1035.64 | 959.65 | 0.00 – 3044.32 | 685.24 | [843.77 – 1227.51] |
| Tot Phonemes MAL | 43626.05 | 45941.09 | 0.00 – 90306.05 | 26597.81 | [36178.66 – 51073.43] |
| Tot Syllables MAL | 20931.36 | 21953.91 | 0.00 – 43743.93 | 12844.62 | [17334.87 – 24527.86] |
| Sum words FEM KCHI | 5251.01 | 5466.48 | 0.00 – 10034.99 | 2735.83 | [4484.98, 6017.05] |

| | | | | | |
|---------------|----------|----------|-----------------|------------|----------------------|
| Sum time | 1875.45 | 1931.65 | 0.00 – 3970.75 | 1029.19 | [1587.28 – 2163.62] |
| FEM KCHI | | | | | |
| Total turns | 3013.02 | 3372.18 | 0.00 – 5505.16 | 1534.41 | [2583.39 – 3442.66] |
| FEM | | | | | |
| Tot Phonemes | 97025.17 | 97106.58 | 0.00 – | 48063.82 | [83567.30– |
| FEM | | | 160496.39 | 110483.03] | |
| Tot Syllables | 46236.49 | 45531.15 | 0.00 – 76733.18 | 23045.39 | [39783.78– 52689.20] |
| FEM | | | | | |

The children's total CDI scores are displayed in Table 9, based on an analysis of 641 items. The mean score was 184.53 (SD = 115.46, median = 176.64, 95% CI [152.20 – 216.86]), with scores ranging from 0 to 478. This indicates considerable variability in vocabulary development.

Table 9

Descriptives CDI Scores Children

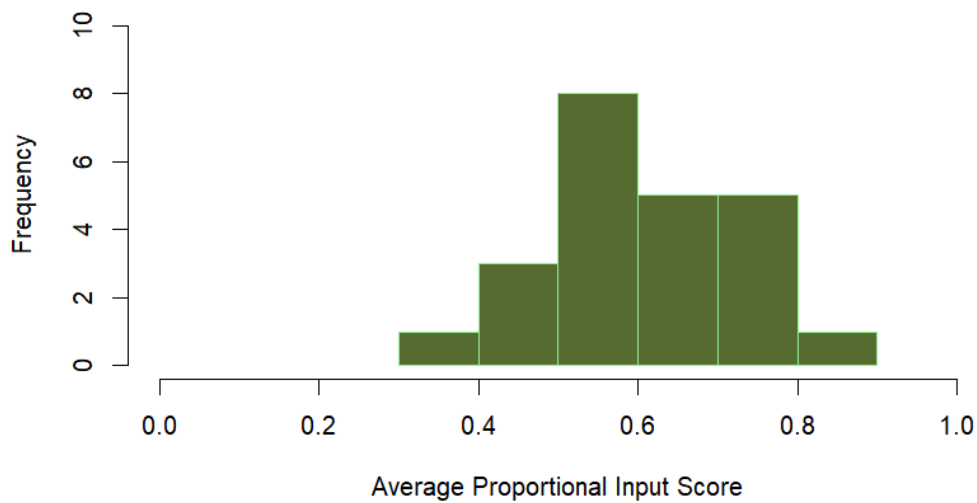
| Measure | Mean | Median | Range | SD | 95% CI |
|-----------------|--------|--------|---------------|--------|-------------------|
| Total CDI Score | 184.53 | 176.64 | 0.00 – 478.00 | 115.46 | [152.20 – 216.86] |

Descriptive statistics were also computed for the parental input quantity score for both fathers and mothers, as presented in Table 10. For fathers, the mean score is 0.61 (SD = 0.12, median = 0.59, 95% CI [0.55 – 0.66]), while mothers show a slightly higher mean at 0.66 (SD = 0.13, median = 0.66, 95% CI [0.60 – 0.71]), suggesting slightly higher language input in mothers than in fathers. Ranges are 0.36 – 0.82 for fathers and 0.36 – 0.86 for mothers, indicating adequate spread for detecting individual differences.

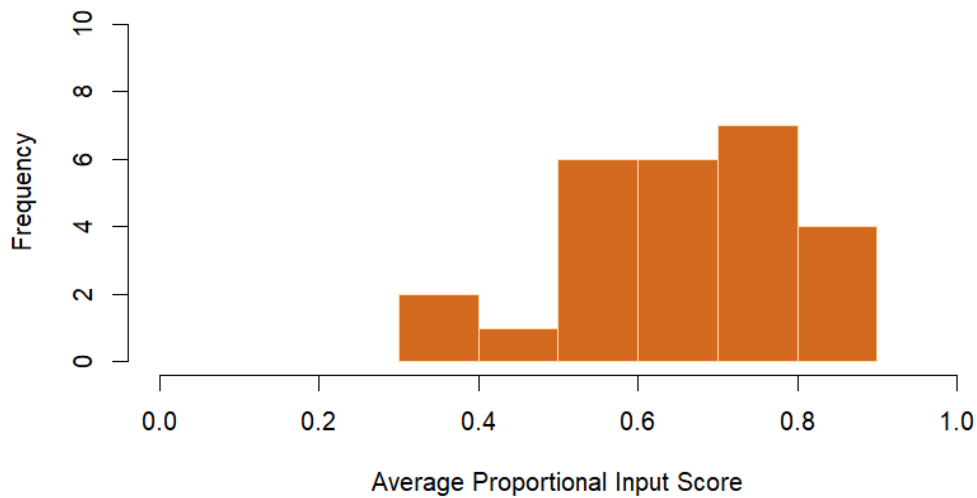
Table 10*Descriptives Parental Input Quantity Score for Fathers and Mothers*

| | Mean | Median | Range | SD | 95% CI |
|---------|------|--------|-------------|------|---------------|
| Fathers | 0.61 | 0.59 | 0.36 – 0.82 | 0.12 | [0.55 – 0.66] |
| Mothers | 0.66 | 0.66 | 0.36 – 0.86 | 0.13 | [0.60 – 0.71] |

The distributions of the parental input score, as illustrated in Figure 1 and Figure 2, appear approximately symmetrical, with medians equal or closely aligned with the means, and no substantial skew or outliers. The standard deviations reflect low variability within both groups, suggesting minimal individual differences among participants. Overall, the parental input quantity scores for both fathers and mothers demonstrate a balanced distribution.

Figure 1*Distribution Parental Input Quantity Scores Fathers*

*Note. N = 23 fathers

Figure 2*Distribution Parental Input Quantity Scores Mothers*

*Note. N = 27 mothers

The first aim of this study was to assess whether our parental input quantity score correlated with ALICE derived observational variables of parental language input. A strong, positive, and statistically significant correlation ($r \geq 0.70$) was expected.

However, the results did not align with this prediction. For both fathers and mothers, all correlations were small and close to zero, with some positive and some negative, and none reaching statistical significance. Among fathers, correlation coefficients are positive but small, indicating weak and non-significant correlations (Table 11). As shown in Table 12, the slightly negative correlations for mothers in ‘Sum words FEM KCHI’ and ‘Total turns FEM’ suggest that higher input scores are marginally linked to a smaller amount of produced words and less conversational turns, respectively. Taken together, these findings indicate that the parental input quantity score shows only limited correlation with the observational quantity input as measured by the ALICE system, particularly for mothers.

Table 11*Correlation Parental Input Quantity Score Fathers and ALICE Variables Fathers*

| ALICE Variables | <i>r</i> | <i>p</i> |
|--------------------|----------|----------|
| Sum words MAL KCHI | 0.15 | .495 |
| Sum time MAL KCHI | 0.13 | .556 |
| Total turns MAL | 0.16 | .479 |
| Tot Phonemes MAL | 0.14 | .533 |
| Tot Syllables MAL | 0.14 | .534 |

Table 12*Correlation Parental Input Quantity Score Mothers and ALICE Variables Mothers*

| ALICE Variables | <i>r</i> | <i>p</i> |
|--------------------|----------|----------|
| Sum words FEM KCHI | - 0.01 | .979 |
| Sum time FEM KCHI | 0.00 | .998 |
| Total turns FEM | - 0.09 | .650 |
| Tot Phonemes FEM | 0.04 | .844 |
| Tot Syllables FEM | 0.05 | .809 |

The second research question focused on evaluating whether the items included in the parental input quantity score showed satisfactory psychometric properties, including a high level of internal consistency ($\alpha \geq 0.70$), moderate ($r \geq 0.30$) to strong ($r \geq 0.50$) inter-item correlations and moderate ($r \geq 0.30$) to strong ($r \geq 0.50$) item-total correlations.

Internal consistency. The 11 items of the questionnaire that together constitute the input score had a Cronbach's alpha value of 0.73 ($\alpha = 0.73$), which is in line with the hypothesis and indicates a good internal consistency. This implies that the scale offers reasonably consistent measurements of the amount of parental language input, yet there remains potential for improvement. Table 13 presents the alpha values when each item is excluded. An increase of the value above 0.73 indicated that the reliability of the score would improve by removing that item. This analysis was conducted to identify items that may have an effect on the overall reliability.

Table 13*Cronbach's Alpha If Item Is Dropped*

| Item Number Questionnaire | α |
|---------------------------|----------|
| Q18 | 0.73 |
| Q21 | 0.73 |
| Q22 | 0.70 |
| Q23 | 0.73 |
| Q24 | 0.70 |
| Q27 | 0.73 |
| Q30 | 0.70 |
| Q34 | 0.70 |
| Q36 | 0.68 |
| Q37 | 0.72 |
| Q38 | 0.71 |

Items Q18 (verbal language input), Q21 (language use in supermarket), Q23 (shared book reading before bedtime), and Q27 (storytelling and narrative interaction) maintain the alpha at 0.73 when removed, staying identical to the overall scale's reliability. This suggests that these items neither improve nor reduce the internal consistency of the scale. In contrast, items Q22 (correction of grammatical errors), Q24 (shared book reading), Q30 (conversation about child's environment), Q34 (input during playtime) and Q38 (signing and rhyming) cause a slight decrease in the alpha-value. This indicates that these items contribute positively to the scale's consistency. Notably, item Q36 (media conversations) reduces the alpha the most, meaning that it has the strongest positive contribution to the main construct and plays an important role in the reliability of the scale. Despite these minor variations, all alpha values remain close to the expected threshold ($\alpha \geq 0.70$), meaning that the items fit well in the measured construct. Thus, we do not recommend removing any items from the scale.

Inter-item correlation. Correlations between individual items ranged from $r = -0.07$ to 0.50, as illustrated in Table 14. Item pair Q30 (conversations about child's environment) and Q38 (singing and rhyming) show a strong correlation ($r = 0.50$), whereas others were weaker, and some even negative. Remarkably, item Q37 correlates negative with several other items. This suggests that this particular item might be measuring a different or opposing aspect of the construct of verbal language input compared to the other items. The presence of negative

correlations could imply that higher scores on Q37 tend to correspond with lower scores on the items it negatively correlates with. A more detailed examination of item Q37 shows a mean value 0.61, a median of 0.50, and a standard deviation of 0.25 (mean = 0.61, median = 0.50, SD = 0.25), suggesting a moderate frequency and variability in responses. To conclude, the strength of inter-item correlations was variable, with most inter-item correlations being lower than expected. The items with lower correlations might be less related or possibly measuring other dimensions within parental quantity input.

Table 14

Inter-item Correlation of Items Parental Input Quantity Score

| | Q18 | Q21 | Q22 | Q23 | Q24 | Q27 | Q30 | Q34 | Q36 | Q37 | Q38 |
|-----|--------|--------|------|------|------|--------|------|--------|------|--------|------|
| Q18 | – | 0.16 | 0.05 | 0.05 | 0.19 | 0.29 | 0.29 | 0.15 | 0.06 | - 0.03 | 0.38 |
| Q21 | 0.16 | – | 0.10 | 0.23 | 0.12 | 0.02 | 0.06 | 0.36 | 0.26 | - 0.07 | 0.04 |
| Q22 | 0.05 | 0.10 | – | 0.08 | 0.22 | 0.22 | 0.25 | 0.31 | 0.41 | 0.45 | 0.24 |
| Q23 | 0.05 | 0.23 | 0.08 | – | 0.28 | 0.07 | 0.01 | 0.20 | 0.45 | 0.02 | 0.04 |
| Q24 | 0.19 | 0.12 | 0.22 | 0.28 | – | 0.11 | 0.27 | 0.46 | 0.40 | 0.27 | 0.23 |
| Q27 | 0.29 | 0.02 | 0.22 | 0.07 | 0.11 | – | 0.24 | - 0.01 | 0.09 | - 0.02 | 0.17 |
| Q30 | 0.29 | 0.06 | 0.25 | 0.01 | 0.27 | 0.24 | – | 0.33 | 0.33 | 0.23 | 0.50 |
| Q34 | 0.15 | 0.36 | 0.31 | 0.20 | 0.46 | - 0.01 | 0.33 | – | 0.40 | 0.09 | 0.14 |
| Q36 | 0.06 | 0.26 | 0.41 | 0.45 | 0.40 | 0.09 | 0.33 | 0.40 | – | 0.31 | 0.29 |
| Q37 | - 0.03 | - 0.07 | 0.45 | 0.02 | 0.27 | - 0.02 | 0.23 | 0.09 | 0.31 | – | 0.23 |
| Q38 | 0.38 | 0.04 | 0.24 | 0.04 | 0.23 | 0.17 | 0.50 | 0.14 | 0.26 | 0.23 | – |

Item-total correlation. As shown in Table 15, the item-total correlations varied between $r = 0.35$ and $r = 0.71$. These correlations reflect the extent to which each individual item aligns with the overall scale score, excluding the item itself. In this analysis, item Q27 (storytelling and narrative interaction) shows the lowest item-total correlation ($r = 0.35$), suggesting that, of all the items, this item aligns the least with the rest of the scale, although still moderate. In contrast, item Q36 (media conversations) demonstrates a high item-total correlation ($r = 0.71$), indicating that the item contributes well to the internal consistency of the scale and likely measures the same underlying construct. Thus, item Q36 behaves consistently with the other items and contributes meaningfully to the total score. Overall, all items demonstrated at least a moderate correlation, with six of the 11 items showing a strong

correlation ($r \geq 0.50$). This suggests that the majority of the items align very well with the scale and likely reflect the construct of parental input.

Table 15

Item-total Correlation of Items Parental Input Quantity Score

| Item Number Questionnaire | <i>r</i> |
|---------------------------|----------|
| Q18 | 0.39 |
| Q21 | 0.37 |
| Q22 | 0.60 |
| Q23 | 0.43 |
| Q24 | 0.66 |
| Q27 | 0.35 |
| Q30 | 0.59 |
| Q34 | 0.58 |
| Q36 | 0.71 |
| Q37 | 0.46 |
| Q38 | 0.57 |

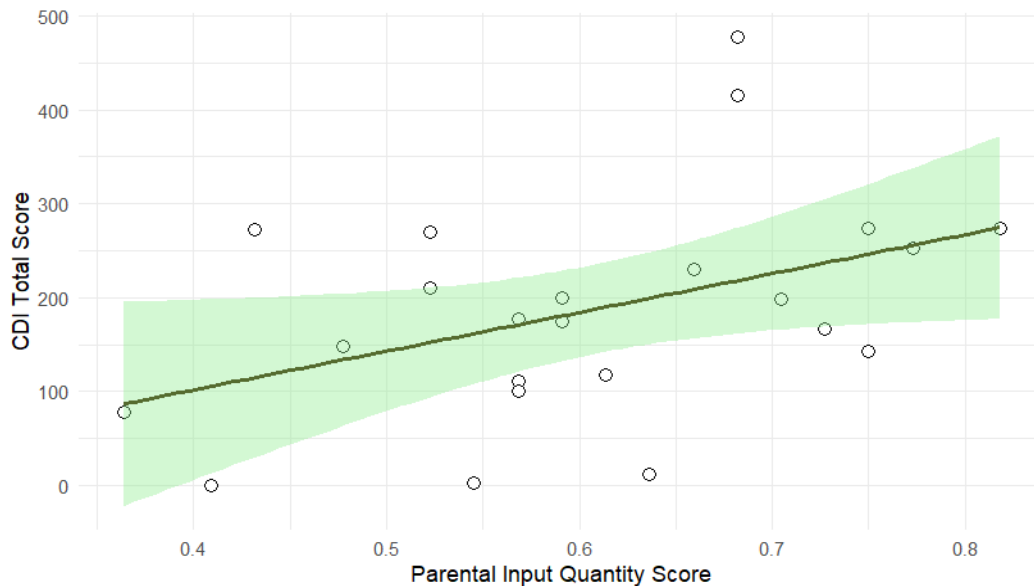
The purpose of the third research question was to assess whether there exists a positive and statistically significant correlation between the parental input quantity score and children's productive vocabulary score, as measured by the N-CDI. A small to moderate correlation ($0.10 \leq r \leq 0.30$) was expected.

The findings support this hypothesis for both parental groups. As illustrated in Table 16, for both fathers and mothers, the correlation with child vocabulary was higher than predicted ($r = 0.42$; $r = 0.39$ respectively), thus indicating moderate statistically significant correlations ($p < .05$). This suggests that increased parental language input is associated with a larger expressive vocabulary in children.

Table 16*Correlation Parental Input Quantity Score and N-CDI Scores*

| | <i>r</i> | <i>p</i> |
|------------|----------|-------------|
| PR Fathers | 0.42 | .045 |
| PR Mothers | 0.39 | .047 |

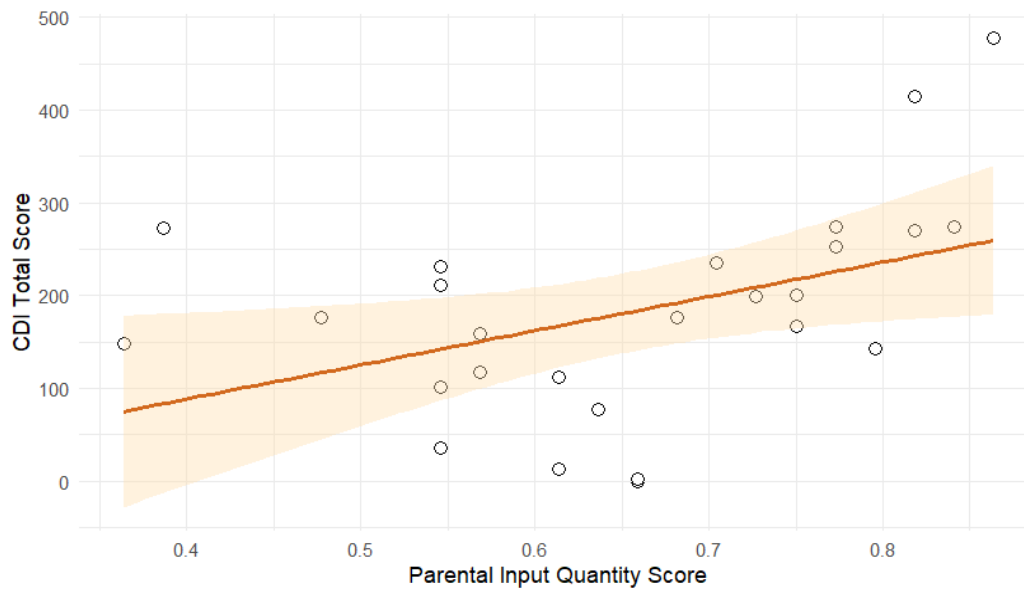
Figure 3 and Figure 4 provide a visual representation of these correlations. The scatterplot illustrating the fathers' data (Figure 3) displays a clear upward trend, indicating that higher paternal responsiveness is associated with stronger expressive vocabulary skills in children. The data points are relatively scattered around the regression line, suggesting an association that, while positive, is modest and varying across individuals. Similarly, the scatterplot for mothers (Figure 4) also presents a positive relationship. Together, these findings support the hypothesis and imply that parental verbal input is related to the development of children's expressive vocabulary, although the strength of this relationship appears to be moderate.

Figure 3*Correlation Parental Input Quantity Score Fathers and N-CDI Scores Children*

*Note. N = 23 fathers

Figure 4

Correlation Parental Input Quantity Score Mothers and N-CDI Scores Children



*Note. N = 27 mothers

Exploratory analysis

The additional exploratory analysis aimed to assess the degree of agreement between parents within families regarding their reported distribution of time spent with the child during the week. This analysis was conditional upon having a sufficient number of families in which both parents completed the questionnaire, which was the case in the present sample size. Table 17 shows the mean percentage agreement rates across parent dyads. Across the parent-pairs, agreement rates were consistently high. The average agreement across all seven days was 89.83%, indicating a generally strong alignment in reported time distribution and reliable responses. Since this information is relatively objective, the high agreement level indicates consistent and reliable parental reporting regarding the questionnaire overall.

Table 17

Mean Agreement Rates Across Parent Dyads on the Weekly Schedule (Daily & Weekly %)

| Monday | Tuesday | Wednesday | Thursday | Friday | Saturday | Sunday | Entire week |
|--------|---------|-----------|----------|--------|----------|--------|-------------|
| 93.94 | 90.90 | 89.39 | 84.85 | 90.90 | 89.39 | 89.39 | 89.83 |

Discussion

The goal of this study was to assess the extent to which a new parent-report measure could capture parental input quantity. We did this in three ways: 1) we assessed the alignment between parents' self-reported language input, expressed as a proxy score for the amount of parental speech input, and observed language input obtained from ALICE processed recordings, 2) we evaluated the psychometric properties of this parental input score, and 3) we investigated the relationship between this score with children's vocabulary size. The results regarding the first hypothesis indicate that, based on this sample, we could not detect an association between the self-reported and observed parental input, in contrast to our expectation. Despite this, the findings related to the second hypothesis show that the items comprising the parental input score exhibited acceptable psychometric qualities, including good internal consistency and adequate item-total correlations, demonstrating that it reliably measured a coherent construct. Moreover, the score was moderately and significantly associated with children's expressive vocabulary, which is in line with the third hypothesis.

The first research question explored the extent to which the parental input quantity score aligned with ALICE-derived variables. Contrary to expectations, weak and statistically non-significant correlations were found between the parental input quantity score and the ALICE variables, both for fathers and mothers. These findings stand in contrast to earlier research suggesting that parent-reports can moderately predict observational speech metrics (Anderson et al., 2021; Bleses et al., 2018; Cates et al., 2023). However, there are various ways in which the findings can be interpreted, which could possibly give explanations for these results. First, it is plausible that the parental input score and the ALICE-derived measures are capturing different constructs altogether. In other words, self-reported and observational measures may tap into distinct dimensions of parental language input. The input score is designed to reflect parental perceptions of their typical language engagement, averaged across multiple settings and timeframes. It includes behaviors such as talking with their child in different contexts, storytelling, and shared book reading, potentially influenced by parents' values, beliefs, and habitual routines. This may reflect more qualitative aspects of language input. By contrast, ALICE provides an observational snapshot of the actual quantity of parental speech over a constrained time period within the home environment, making it more sensitive to situational variation. These differences further highlight the challenge of disentangling quantity and quality in parental language input (Jones & Rowland, 2017).

Second, social desirability and self-perception biases may also contribute to the findings (Bennetts et al., 2016). While such biases can affect both self-report and observational measures

– since parents may modify their behavior when being recorded (Scaff et al., 2024) – it is generally much harder to maintain artificial behavior in natural interactions with children, especially over extended periods such as a full day. Therefore, self-report measures may be influenced by social desirability to a greater extent than observational measures. In line with this, some studies have shown that parents overestimate how much language they provide (Richards et al., 2017). This is particularly the case when responding to questionnaires that implicitly reflect ideals of ‘good parenting’. Even when questionnaire items are framed behaviorally, e.g., ‘How often do you talk with your child when...’, the responses may reflect what parents believe is expected of them, rather than their actual behavior. Due to this inherent bias, self-report measures may capture a slightly different construct than observational measures. This bias may also explain why the input score for mothers was higher than that of fathers, and the correlations between that score and the mothers’ ALICE variables were lower. Mommersteeg et al. (2024) suggest that women are, more often than men, confronted with societal expectations regarding, among other things, caregiving and household responsibilities, and experience more societal pressure because of this. Thus, it is possible that mothers overestimate the quantity of their language input, as they want to meet this expectation. As a result, their input score was higher while their correlations were lower, possibly so because their self-reported behavior deviated more from the observational data. However, the ALICE variables also showed that mothers spoke considerably more than fathers, suggesting that their higher input scores reflect both actual differences in language use and potential overestimation in self-reports.

Third, it is important to acknowledge the limitations of the ALICE data itself. While the ALICE system is suitable for large-scale analysis, its automated measures are not as precise as human annotations and introduce noise, which can weaken the strength of correlations. Therefore, caution is warranted when using them to assess variation at the individual level (Cristia et al., 2024). Besides this, it is important to note that the weak and statistically non-significant correlations should not be interpreted as evidence that the input score is invalid or that parental self-reports are intrinsically unreliable. However, they do suggest limited convergence between the two methods in this specific study. Previous studies have shown varying results regarding comparisons between self-reports and observational measures (Anderson et al., 2021; Bennetts et al., 2016; Bleses et al., 2018; Cates et al., 2023; Richards et al., 2017), possibly reflecting broader challenges in aligning these measurement approaches. At the same time, the absence of a correlation underscores the principle that while correlation is required to infer a possible causal relationship, it is not sufficient on its own. When no

correlation is found, it can be concluded that there is no linear relationship between the two variables in the context examined. However, this does not imply that the parental input score is entirely unrelated to the underlying construct.

Another ALICE related limitation is the length of the recordings, which were limited to a maximum of two consecutive days. In addition, all families completed the full duration of 32 hours. Although longer sampling periods are preferable for robust estimates of language input (Räsänen et al., 2021), a maximum of 32 hours might not be representative of a child's language environment. Especially in this case, as the recordings were conducted in the home environment, excluding different settings where verbal input are also frequent. Another ALICE related issue concerns ALICE's technical limitations. Although ALICE performs reliably in English-language datasets, its application to Dutch has not been validated yet (Räsänen et al., 2021). Errors in speaker diarization, particularly in distinguishing between male and female voices, may have compromised the accuracy of parental speech identification. Moreover, the assumption that the 'MAL' and 'FEM' variables correspond directly to fathers and mothers respectively, may not always be correct, potentially influencing the findings if some of the measured input actually came from other individuals. To conclude, the research question can be answered as follows: the findings question the assumption that self-report measures correlate with observational measures of parental language input. Instead, they support the view that parents' perceptions and their actual behavior present related, but distinct aspects.

The second research question addressed the psychometric properties of the parental input quantity score items, and the findings offered partial support. The internal consistency of the input score, as indicated by Cronbach's alpha, was good. This may suggest that the eleven items collectively constitute a reasonably cohesive and stable instrument for measuring the quantity of parental language input. Still, there remains room for improvement. The alpha value if individual items were removed was investigated, and all items generally fit well within the scale. However, more detailed examination of inter-item correlations indicate several limitations. While the overall internal reliability was robust, inter-item correlations were generally low and exhibited considerable variability. Notably, item 37, which addresses parent-child discussions about books, demonstrated negative correlations with several other items. Specifically, with items 18 (general parental language input), 21 (input in supermarket), and 27 (storytelling and narrative interactions). Further analysis of item 37 revealed moderate variance. This suggests that book-related interactions may reflect a distinct communicative style or context, potentially emphasizing structured, literacy-oriented input as opposed to more everyday language exchanges captured by the other items. It is also possible that parents

interpreted this item in a different way than intended. For example, as referring to passive reading rather than interactive discussion. However, without qualitative data being available regarding this item, this interpretation remains speculative. Finally, item-total correlations were examined. All eleven items demonstrated at least a moderate correlation, with six of the items showing high correlations. This suggests that the scale is generally coherent and capture parental language input very well. To conclude, the items of the parental quantity input score are reliable. Nevertheless, the possibility of underlying subdimensions of parental input or differences in interpretation needs further research to be able to validate the scale's structure and construct validity more thoroughly.

The third and last research question related to the association between the parental input score and children's vocabulary size. The hypothesis for this question was confirmed, indicating that the input scores for both fathers and mother showed moderate, statistically significant correlations with children's expressive vocabulary scores. These correlations suggest that the input score may capture important aspects of the home language environment that contributes to language development. This interpretation is consistent with previous research emphasizing the role of response caregiving, shared verbal routines, and interaction quality in early vocabulary growth (Anderson et al., 2021; Donnelly & Kidd, 2021; Rowe, 2012). Meta-analytic studies have further shown that interactional quality and turn-taking are stronger predictors of language outcomes than the amount of words spoken (Anderson et al., 2021; Fu & Zheng, 2024). Additionally, Donnelly & Kidd (2021) propose a bidirectional dynamic in which responsive conversational turns facilitate vocabulary growth, which in turn supports interactions. Importantly, these correlations are not the same as causality, meaning that it's impossible to conclude that parental language input has a direct effect or predictive value on the vocabulary size of children, solely based on these findings. The observed correlation may be influenced by other factors that affect both parental language behavior, as reflected by the parental input score, and the child's language development. These factors may include genetic traits, other aspects of the home environment, or cognitive abilities (Coffey & Snedeker, 2025; Jones & Rowland, 2017). Taken together, the findings indicate that there is indeed a relationship between parental language input and children's vocabulary.

Besides the research questions, an exploratory analysis was conducted in families with two participating parents to assess the level of agreement between the parents regarding their reported distribution of time spent with their child during a typical week. The responses to this weekly schedule showed a high average agreement rate, suggesting several important implications. First, this strong alignment indicates high inter-rater reliability. Given the

objective measures of this schedule, the high agreement suggests that parents were reliable and consistent in how they approached the questionnaire. When two parents report similar time-use patterns independently, it suggests that the items were clearly understood and interpreted accurately. High agreement implies that the items of the parental input quantity score are based on consistent and observable behaviors. Second, the consistency may reflect structured family routines. In dual-parent, middle- to high-SES households like those in this study, caregiving roles are often stable and well-defined (Don et al., 2013). This likely helps both parents form accurate perceptions of each other's involvement, increasing reporting accuracy. Nevertheless, it is important to acknowledge that the observed level of agreement may not extend to more diverse populations, e.g., households with weaker communication between parents. In conclusion, this finding suggests that self-report instruments like the one used in this study may accurately capture consistent and mutually recognized patterns of parental behavior, rather than relying solely on subjective perceptions. Nevertheless, it does not provide direct validation against observational measures of parental language input.

Collectively, the results offer insight into validity of the parental input quantity score, derived from a family routines questionnaire, as an indicator of children's exposure to parental speech, as measured through observational daylong audio recordings. This study suggests that, based on the specific methodological and contextual conditions of the study, the parental input quantity score does not operate as a valid proxy for actual parental language input. This is mainly concluded by the absence of statistically significant correlations between the parental input quantity score and ALICE-derived measures of child-directed speech, therefore indicating limited convergent validity. Nonetheless, the parental input score exhibited strong internal consistency, although more thorough analyses like CFA are needed to make strong claims about the score's validity. Moreover, the score showed moderate associations with children's expressive vocabulary, possibly supporting its construct and predictive validity. These results indicate that while the parental input score do not align directly with the automated observational measures produced by the ALICE algorithm, it may still cover relevant dimensions of the home language environment. This discrepancy likely stems from inherent differences in what each method captures: self-reported questionnaires tend to reflect generalized parental routines and behaviors in various contexts, whereas daylong audio recordings provide a more narrowly defined, temporally specific snapshot of verbal input. Accordingly, the two approaches could be viewed as providing complementary insights into the language environments of children, rather than serving as convertible measures.

Limitations and Future Research

While the present study offers valuable insights into self-reported measures for parental language input, several methodological limitations warrant careful consideration, as they affect the generalizability, interpretability, and precision of the findings.

First, the study was based on a relatively small and socioeconomically homogeneous sample (N = 50 parents from 28 families), comprising monolingual Dutch families, mostly highly educated parents, and all residing in or near Nijmegen. Participants had previously engaged in the TBB study, suggesting a sample that was likely highly motivated. These characteristics limit the applicability of the findings to more diverse populations, particularly those from lower socioeconomic backgrounds or bilingual families, where language input, parenting roles, and reporting accuracy may differ considerably. Additionally, the sample size reduces the statistical power.

In addition, the evaluation of the psychometric item properties of the parental input quantity score consisted of simple statistics such as internal consistency reliability (Cronbach's alpha), inter-item correlations, and item-total correlations. While these analyses offer an initial indication of reliability and coherence, they do not provide a thorough understanding of the scale's structure and validity, which would have been possible with more in depth and complex analyses. Confirmatory factor analysis (CFA) would provide a more thorough evaluation. Specifically, CFA offers fit indices that test how well the hypothesized model fits the data, resulting in stronger claims.

Finally, there was a temporal misalignment between the study's three data sources. While the ALICE and CDI data were collected around the child's second birthday, the family routines questionnaire was conducted at a later stage. This temporal gap may have introduced noise due to changes in development of routines. Nevertheless, previous research suggests that parental speech habits tend to remain relatively stable throughout the first years of a child (Hart & Risley, 1992). This suggests that the questionnaire, even when conducted at a later moment, still offers a representative image of the parental speech input during the period in which the other data was collected. However, speech habit can change over time due to changes in family structures, child development, or daily routines (Richards et al., 2017). This may impact the accuracy of retrospective reports, although the time gap between the different data collection moments in this was relatively short which minimizes major changes in speech habits.

Taken together, these limitations highlight the importance of future research that applies larger and more demographically diverse samples, including a wider range of cultural and educational backgrounds to increase generalizability. Moreover, it is crucial that future studies

use longer audio recordings in different contexts to capture more of the child's language environment and gain more complete and comprehensive insights into this. Another suggestion is related to the initially planned CFA. Therefore, it is recommended that CFA or other advanced methods are applied in future studies, in order to assess dimensionality, model fit, and validity more thoroughly.

Conclusion

The current 'De Kleine TIJGER' study aimed to examine the extent to which a parental input quantity score corresponds with observationally measured language input derived from audio recordings, to evaluate the psychometric properties of this score, and to investigate its relationship with children's expressive vocabulary. The findings showed no significant correlation between the parental input score and observational language input, indicating that these methods may reflect different aspects of parental language input. However, the parental input score itself proved to be a reliable and coherent measure, with good internal consistency. Moreover, a moderate, significant association with children's vocabulary was found, supporting the possible predictive use of the input score. These findings emphasize that self-reports and observational recordings provide complementary insights: the input score possibly captures broader perceptions of language input, while ALICE processed recordings offer a snapshot of this input. In order to be able to draw strong conclusions about whether self-reports of parental language input are a valid measure of observed parental language input, more in-depth analyses are required. This study highlights the complexity of measuring parental language input and the value of using multiple measures for a more comprehensive understanding. In conclusion, despite the absence of direct validation, well-designed self-report tools can yield important information about the home language environment of young children, which is crucial for understanding early language development.

References

- Adamson, K. A., & Prion, S. (2013). Reliability: measuring internal consistency using Cronbach's α . *Clinical simulation in Nursing*, 9(5), e179-e180.
<https://doi.org/10.1016/j.ecns.2012.12.001>
- 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(2), 484-501. <https://doi.org/10.1111/cdev.13508>
- Bennetts, S. K., Mensah, F. K., Westrupp, E. M., Hackworth, N. J., & Reilly, S. (2016). The agreement between parent-reported and directly measured child language and parenting behaviors. *Frontiers in Psychology*, 7, 1710. <https://doi.org/10.3389/fpsyg.2016.01710>
- Bleses, D., Vach, W., & Dale, P. S. (2018). Self-reported parental vocabulary input frequency for young children. *Journal of child language*, 45(5), 1073-1090.
 doi:10.1017/S0305000918000089
- Capelli, E., Dondena, C., Lorusso, M. L., Mascheretti, S., Pozzoli, R., Salandi, A., ... & Cantiani, C. (2024). Predictive Measures in Child Language Development: The Role of Familial History and Early Expressive Vocabulary. *Journal of Speech, Language, and Hearing Research*, 67(10), 3714-3732. https://doi.org/10.1044/2024_JSLHR-23-00815
- Castor (n.d.). *About us*. Castor EDC. Retrieved from <https://www.castoredc.com/about-us/#we-want-to>
- Cates, C. B., Roby, E., Canfield, C. F., Johnson, M., Raak, C., Weisleder, A., ... & Mendelsohn, A. L. (2023). Validation of the StimQ2: A parent-report measure of cognitive stimulation in the home. *PloS one*, 18(7), e0286708.
<https://doi.org/10.1371/journal.pone.0286708>
- Child and Adolescent Health Measurement Initiative. (2022). *2022 National Survey of Children's Health (NSCH) data query* [Data set]. Data Resource Center for Child and Adolescent Health. Retrieved from <https://www.childhealthdata.org>
- Coffey, J. R., & Snedeker, J. (2025). How strong is the relationship between caregiver speech and language development? A meta-analysis. *Journal of Child Language*, 1-36.
<https://doi.org/10.1111/cdev.13508>
- Cristia, A., Gautheron, L., Zhang, Z., Schuller, B., Scaff, C., Rowland, C., ... & Soderstrom, M. (2024). Establishing the reliability of metrics extracted from long-form recordings using LENA and the ACLEW pipeline. *Behavior Research Methods*, 56(8), 8588-8607.
<https://doi.org/10.3758/s13428-024-02493-2>

- Don, B. P., Biehle, S. N., & Mickelson, K. D. (2013). Feeling like part of a team: Perceived parenting agreement among first-time parents. *Journal of Social and Personal Relationships, 30*(8), 1121-1137. <https://doi.org/10.1177/0265407513483105>
- Donnelly, S., & Kidd, E. (2021). The longitudinal relationship between conversational turn-taking and vocabulary growth in early language development. *Child Development, 92*(2), 609-625. <https://doi.org/10.1111/cdev.13511>
- Fu, B., & Zheng, X. (2024). A Three-Level Meta-Analysis of the Relationship Between Parental Language Input and Child Language Outcomes. *International Journal of Early Childhood, 1-22*. <https://doi.org/10.1007/s13158-024-00406-0>
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. <https://doi.org/10.1016/j.paid.2016.06.069>
- Gliem, J. A., & Gliem, R. R. (2003). *Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales*. Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education. Retrieved from <https://scholarworks.indianapolis.iu.edu/server/api/core/bitstreams/976cec6a-914f-4e49-84b2-f658d5b26ff9/content>
- Hajjar, S. T. (2018). Statistical analysis: Internal-consistency reliability and construct validity. *International Journal of Quantitative and Qualitative Research Methods, 6*(1), 27-38. Retrieved from <https://www.eajournals.org/wp-content/uploads/Statistical-Analysis-Internal-Consistency-Reliability-and-Construct-Validity.pdf>
- Hart, B., & Risley, T. R. (1992). American parenting of language-learning children: Persisting differences in family-child interactions observed in natural home environments. *Developmental Psychology, 28*(6), 1096–1105. <https://doi.org/10.1037/0012-1649.28.6.1096>
- Jensen, M. P. (2003). Questionnaire validation: a brief guide for readers of the research literature. *The Clinical journal of pain, 19*(6), 345-352. Retrieved from https://journals.lww.com/clinicalpain/fulltext/2003/11000/questionnaire_validation__a_brief_guide_for.2.aspx
- Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive psychology, 98*, 1-21. <https://doi.org/10.1016/j.cogpsych.2017.07.002>

- Kaur N, Gupta M, Kiran T, Malhi P, Grover S (2021) Development and evaluation of the digital-screen exposure questionnaire (DSEQ) for young children. *PLoS ONE* 16(6): e0253313. <https://doi.org/10.1371/journal.pone.0253313>
- Krousorati, K., Gregoriadis, A., Tsigilis, N., Grammatikopoulos, V., & Evangelou, M. (2022). Introducing the Home Learning Environment Questionnaire and examining the profiles of home learning environments in Greece. *Frontiers in Education*, 7, 987131. <https://doi.org/10.3389/educ.2022.987131>
- Martinot, P., Bernard, J. Y., Peyre, H., De Agostini, M., Forhan, A., Charles, M. A., ... & 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>
- Marjanovič Umek, L., Podlesek, A., & Fekonja, U. (2005). Assessing the home literacy environment. *European Journal of Psychological Assessment*, 21(4), 271-281. <https://doi.org/10.1027/1015-5759.21.4.271>
- Max Planck Institute for Psycholinguistics. (n.d.). *MPI logo* [Image]. Retrieved from <https://www.mpi.nl/idlas-nl/faq>
- Microsoft Corporation. (2019). *Microsoft Excel* (Version 1808, Build 10417.20012) [Spreadsheet software]. Retrieved from <https://www.microsoft.com>
- Mommersteeg, P. M., van Valkengoed, I., Lodder, P., Juster, R. P., & Kupper, N. (2024). Gender roles and gender norms associated with psychological distress in women and men among the Dutch general population. *Journal of health psychology*, 29(8), 797-810. <https://doi.org/10.1177/1359105323120729>
- R Core Team. (2025). *R: A language and environment for statistical computing* [Software, Version 4.5.0]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Radboud University. (n.d.). *Radboud University logo* [Image]. Retrieved from <https://www.ru.nl/medewerkers/services/campusfaciliteiten-gebouwen/communicatie-en-promotie/vormgeving-en-opmaak/huisstijl/logo>
- Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, 53, 818-835. <https://doi.org/10.3758/s13428-020-01460-x>
- Richards, J. A., Gilkerson, J., Xu, D., & Topping, K. (2017). How much do parents think they talk to their child? *Journal of Early Intervention*, 39(3), 163-179. <https://doi.org/10.1177/1053815117714567>

- Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762–1774. <https://doi.org/10.1111/j.1467-8624.2012.01805.x>
- RStudio Team. (2023). *RStudio: Integrated Development Environment for R* [Software, Version 2023.12.0+369]. RStudio, PBC. Retrieved from <https://posit.co/download/rstudio/>
- Scaff, C., Casillas, M., Stieglitz, J., & Cristia, A. (2024). Characterization of children's verbal input in a forager-farmer population using long-form audio recordings and diverse input definitions. *Infancy, 29*(2), 196-215. Retrieved from <https://doi.org/10.1111/infa.12568>
- Valerio, W. H., Reynolds, A. M., Morgan, G. B., & McNair, A. A. (2012). Construct validity of the children's music-related behavior questionnaire. *Journal of Research in Music Education, 60*(2), 186-200. <https://doi.org/10.1177/0022429412444450>
- Wang, Y., Williams, R., Dilley, L., & Houston, D. M. (2020). A meta-analysis of the predictability of LENA™ automated measures for child language development. *Developmental Review, 57*, 100921. <https://doi.org/10.1016/j.dr.2020.100921>
- Yalçın, S. S., Tezol, Ö., Çaylan, N., Erat Nergiz, M., Yildiz, D., Çiçek, Ş., & Oflu, A. (2021). Evaluation of problematic screen exposure in pre-schoolers using a unique tool called “seven-in-seven screen exposure questionnaire”: cross-sectional study. *BMC pediatrics, 21*, 1-11. <https://doi.org/10.1186/s12887-021-02939-y>.
- Zink I., & Lejaegere M. (2002). *N-CDIs Lijsten voor Communicatieve Ontwikkeling: aanpassing en hernormering van de MacArthur CDIs van Fenson et al.* Leuven: Uitgeverij Acco

Appendix A

Information and Consent Form



INFORMATIEBLAD VOOR PROEFPERSONEN: de Kleine TIJGER-studie

In de "Kleine TIJGER-studie" ontwikkelen we een korte vragenlijst om meer te weten te komen over de taalomgeving van jonge kinderen. We willen deze vragenlijst later ook gaan gebruiken in een groot online onderzoek met veel gezinnen.

Om te controleren of de korte vragenlijst goed werkt, vragen we eerdere deelnemers van de studie "Taal in het BabyBrein" (de TBB-studie) om deze in te vullen. U heeft voor die studie al uitgebreidere gegevens gegeven. Dit helpt ons om belangrijke vragen te beantwoorden, zoals:

- Geeft de korte vragenlijst vergelijkbare informatie als de informatie die we hebben geanalyseerd uit de geluidsopnames in huis van de TBB-studie?
- Kunnen de antwoorden iets zeggen over de taalvaardigheid van kinderen en ouders?

We vragen beide ouders om de vragenlijst **apart van elkaar** in te vullen. Zo kunnen we zien of ouders de vragen op dezelfde manier begrijpen. Bijvoorbeeld:

- Vinden beide ouders de vragenlijst even makkelijk om in te vullen?
- Geven ouders uit hetzelfde gezin vergelijkbare antwoorden over de taalomgeving van hun kind, of zien zij die anders?

Uw deelname helpt ons om de vragenlijst te verbeteren, zodat we deze in de toekomst kunnen gebruiken bij de grote TIJGER-studie. Met de grote TIJGER-studie willen we kijken hoe taalvaardigheden van ouders worden doorgegeven aan hun kinderen.

De metingen

Vragenlijst

Het invullen van de vragenlijst duurt ongeveer 15 minuten per persoon. De vragenlijst bestaat uit vragen die gaan over de taalomgeving van uw kind. We zijn hierbij vooral benieuwd naar:

- gebruik van media (tv, tablet, smartphone);
- leesactiviteiten (voorlezen, boekjes bekijken, etc.);
- interacties tussen ouder en kind;
- het weekschema van het kind (bijvoorbeeld hoeveel tijd uw kind thuis of op de opvang doorbrengt).

De vragen zijn vergelijkbaar met die uit de langere vragenlijst over de taalomgeving die u eerder al heeft ingevuld in de TBB-studie.

Eerdere geluidsopnames in huis en taalmetingen

We zullen de antwoorden van de nieuwe "Kleine TIJGER-studie" koppelen aan de gegevens die al zijn verzameld in de TBB-studie. Voor deze studie heeft u een tijd geleden meegedaan met de geluidsopnames in huis. Daarnaast ontving u van ons een vragenlijst over uw eigen woordenschat en (indien dit van toepassing was) de woordenschat van de andere ouder van uw kind. Zowel de uitkomsten van de woordenschattaak voor volwassenen als de uitkomsten van het computermodel (dus bijvoorbeeld het aantal woorden dat er tegen uw kind gezegd werd op een dag) willen we nu meenemen in onze "Kleine TIJGER-studie". Het is belangrijk om te weten dat wij geen toegang hebben tot de inhoud van de eerder verkregen geluidsopnames, maar alleen tot de getallen uit het computermodel, zoals het aantal minuten dat er tegen uw kind werd gepraat.

Toestemming

We hopen dat beide ouders willen deelnemen aan dit onderzoek. Deelname is echter volledig vrijwillig, en het is niet verplicht dat beide ouders meedoen. Het is ook mogelijk dat slechts één ouder deelneemt.

Daarnaast is het zo dat uw deelname aan de TBB-studie u op geen enkele manier verplicht om aan dit onderzoek mee te doen. De "Kleine TIJGER-studie" is een nieuw onderzoek waarover u dus ook opnieuw een beslissing mag maken om hier wel of niet aan deel te nemen.

Aangezien deelname aan onderzoek in het Baby & Child Research Center geheel vrijwillig is, kunt u op elk moment, zonder opgave van reden, besluiten af te zien van uw deelname. Ook kunt u tijdens het invullen van de vragenlijsten op ieder moment stoppen. U kunt uw toestemming voor gebruik van persoonlijke gegevens (zoals bijv. geboortedatum) altijd intrekken. De anonieme onderzoeksgegevens die zijn verzameld als u uw toestemming intrekt, kunnen nog wel worden gebruikt in het onderzoek.

Krijg ik betaald voor deelname en zijn er mogelijk risico's?

Als u besluit deel te nemen, bedanken we u na afloop met een financiële vergoeding van €5 per ouder. Er zijn geen risico's aan deze studie verbonden. Graag beantwoorden we al uw vragen aan de telefoon of via e-mail. Dit onderzoek is getoetst door de Ethiek Commissie Faculteit Sociale Wetenschappen (ECSW) van de Radboud Universiteit Nijmegen.

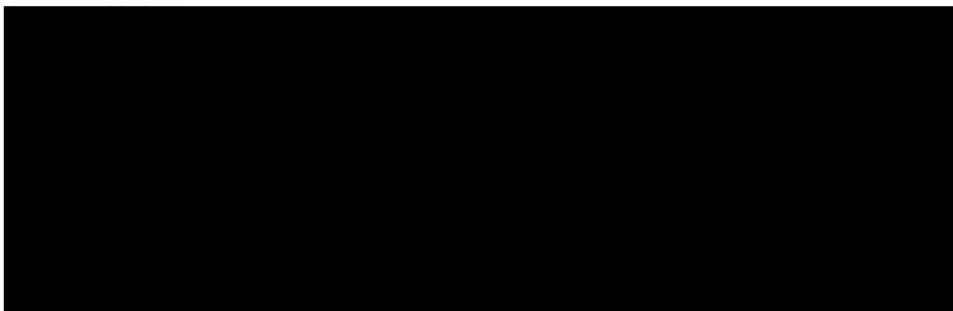
Wat gebeurt er met mijn gegevens?

Mocht u besluiten deel te nemen, dan worden de volgende gegevens verzameld: uw e-mailadres; geboortedatum van ouder(s) en kind; geslacht van ouder(s) en kind; informatie over de gezondheid van ouder(s) en kind; uw antwoorden op de vragenlijst.

Wij maken gebruik van de online vragenlijstsoftware Qualtrics (meer informatie vindt u hier: <https://www.qualtrics.com/privacy-statement/>).

Mocht u een vergoeding willen ontvangen, dan hebben wij ook uw contactgegevens, geboortedatum, IBAN en BSN nodig. Deze zullen door onze financiële afdeling verwerkt worden en staan los van uw onderzoeksgegevens. De verwerking van uw betaalgegevens is verplicht door de Belastingdienst. Deze gegevens worden 10 jaar bewaard.

Als u na deelname besluit dat u de data vernietigd wilt hebben, neemt u dan alstublieft zo spoedig mogelijk contact met ons op, uiterlijk tot 1 maand na afronding van het onderzoek. Alle verzamelde data worden bewaard op een server die met een wachtwoord beveiligd is, in een beveiligde omgeving op het Max Planck Instituut voor Psycholinguïstiek. De data zijn alleen toegankelijk voor werknemers met specifieke autorisatie. Onderzoeksdata (verzameld tijdens het onderzoek) worden apart van persoonlijke data bewaard. Alle onderzoeksdata worden zo snel mogelijk geanonimiseerd maar uiterlijk op het moment dat de dataverzameling gereed is, indien de data niet al anoniem verzameld zijn. De verzamelde data zullen alleen gebruikt worden voor wetenschappelijke doeleinden, en alleen beschikbaar zijn voor de onderzoeksgroep en samenwerkende onderzoekers aan dit project op het Max Planck Instituut voor Psycholinguïstiek. Zij zullen overeenkomstig het EU-regelement voor de gegevensbescherming van persoonlijke data (GDPR) handelen, ook als ze buiten de EU gesitueerd zijn. We hopen de onderzoeksdata te publiceren in december 2027. Graag willen wij u vragen er rekening mee te houden dat wij onderzoeksdata niet kunnen vernietigen na publicatie van het onderzoek. Alle informatie over gegevensbescherming binnen het BRC en uw rechten kunt u nalezen op onze website www.babyandchild.nl/privacybeleid.



Toestemming voor deelname aan onderzoek: de Kleine TIJGER-studie (online studie)

Ik stem in met deelname van mij en mijn kind aan een onderzoek van het Baby & Child Research Center. Hierbij verklaar ik dat ik naar tevredenheid ben geïnformeerd en dat ik de informatiebrief heb gelezen. Ik weet dat deelname aan dit onderzoek geheel vrijwillig is en ik op ieder moment zonder consequenties kan besluiten om zelf niet meer deel te nemen en om mijn kind niet meer deel te laten nemen. Ik heb begrepen dat de gegevens die tijdens het onderzoek worden verzameld alleen zullen worden gebruikt voor wetenschappelijk onderzoek. Gebruik voor niet-wetenschappelijke (bijv. commerciële) doeleinden is uitgesloten. Ik weet dat de verzamelde gegevens minimaal 10 jaar beschikbaar zullen zijn voor controle, hergebruik en replicatie. Ik weet dat de verzamelde gegevens beschikbaar kunnen worden gesteld voor gebruik van externe onderzoekers, maar alleen zonder alle gegevens die tot identificatie van personen zouden kunnen leiden. Ik weet dat tot de persoon herleidbare data alleen toegankelijk zijn voor het onderzoeksteam van het Baby & Child Research Center o.l.v. Dr. Sabine Hunnius. Dit onderzoek is goetst door de Ethiek Commissie Faculteit der Sociale Wetenschappen (ECSW) van de Radboud Universiteit Nijmegen.

*Ik geef toestemming voor het verzamelen, bewaren en gebruiken van de volgende (deels persoonlijke): mijn e-mailadres; mijn geboortedatum en van mijn kind; geslacht van mij en mijn kind; informatie over de gezondheid van mij en mijn kind; mijn antwoorden op de vragenlijsten.

*Ik geef toestemming om de gegevens uit dit onderzoek te koppelen aan de eerder verzamelde gegevens over mij en mijn kind in de Taal en het Babybrein-studie.

Naam & achternaam van het kind

Nijmegen, datum

Handtekening ouder

Naam & achternaam van de ouder

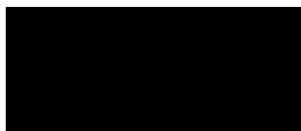
**Ik begrijp dat ik op alle bovengenoemde punten met een *sterretje* 'ja' moet antwoorden om aan het onderzoek deel te kunnen nemen.*

Appendix B

Invitation Letter



Max Planck Instituut voor Psycholinguïstiek
Radboud Universiteit Nijmegen
Radboudumc



Onderzoek van het Baby & Child Research Center: de Kleine TIJGER-studie (online studie)

Beste ouder(s)/verzorger(s),

Een tijd geleden heeft u samen met uw kind deelgenomen aan de studie Taal in het Babybrein (TBB). Op dit moment loopt er een nieuw onderzoek (de Kleine TIJGER-studie). Hiervoor worden deelnemers van de TBB-studie uitgenodigd, omdat we van u al veel data hebben mogen verzamelen. Zo heeft u voor ons uitgebreide vragenlijsten ingevuld over de woordenschat van uw kind (bij 12, 18 en 24 maanden) en heeft u thuis geluidsopnames gemaakt (bij 12 en 24 maanden). Nogmaals hartelijk dank hiervoor!

Over het onderzoek

De Kleine TIJGER-studie is onderdeel van het grotere TIJGER-project (Taal In Je Genen en je ERvaringen). In dat project willen we kijken hoe taalvaardigheden van ouders worden doorgegeven aan kinderen.

Ter voorbereiding starten we nu met de Kleine TIJGER-studie, waarin we een vragenlijst ontwikkelen die later in de grote TIJGER-studie gebruikt zal worden. Via deze brief willen we u vragen of u interesse heeft om deel te nemen aan de Kleine TIJGER-studie. Het gaat om een korte vragenlijst over de dagelijkse routines die verband houden met de taalomgeving en de taalontwikkeling van uw kind. De antwoorden uit de Kleine TIJGER-studie willen we graag koppelen aan de gegevens die eerder al zijn verzameld in de TBB-studie. Op deze manier komen we erachter of we met behulp van de korte vragenlijst de belangrijkste informatie over de taalomgeving van een kind kunnen meten.

In dit onderzoek nodigen we, indien mogelijk, beide ouders uit om thuis onze online vragenlijst in te vullen. Dit duurt ongeveer 15 minuten per persoon.

Aanvullende informatie als u besluit deel te nemen

Toestemming

We hopen dat beide ouders willen deelnemen aan dit onderzoek. Deelname is echter volledig vrijwillig, en het is niet verplicht dat beide ouders meedoen. Het is ook mogelijk dat slechts één ouder deelneemt. Daarnaast is het zo dat uw eerdere deelname aan de TBB-studie u op geen enkele manier verplicht om ook aan dit onderzoek mee te doen. De Kleine TIJGER-studie is een nieuw onderzoek waarover u dus ook opnieuw een beslissing mag maken om hier wel of niet aan deel te nemen.

Aangezien deelname aan onderzoek van het Baby & Child Research Center geheel vrijwillig is, kunt u op elk moment, zonder opgave van reden, besluiten af te zien van uw deelname. U kunt uw toestemming voor gebruik van persoonlijke gegevens (geboortedatum) altijd intrekken. De anonieme onderzoeksgegevens die zijn verzameld als u uw toestemming intrekt, kunnen nog wel worden gebruikt in het onderzoek.



Krijg ik betaald voor deelname en zijn er mogelijke risico's?

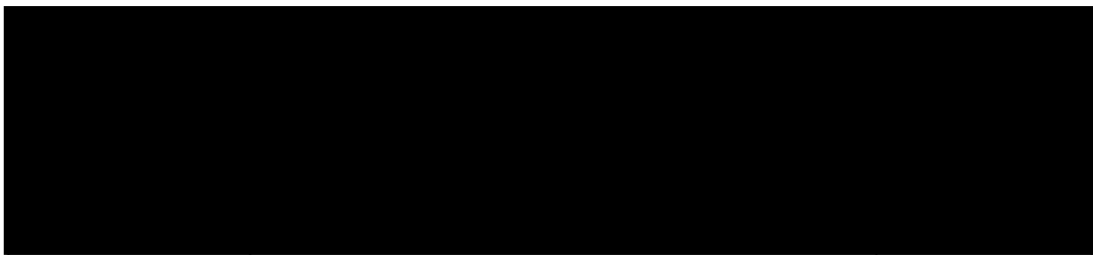
Als u besluit deel te nemen, bedanken we u met een financiële vergoeding van €5 per ouder. Er zijn geen risico's aan deze studie verbonden. Graag beantwoorden we al uw vragen aan de telefoon of via e-mail. Ons onderzoek is getoetst door de ethische commissie van de Faculteit Sociale Wetenschappen van de Radboud Universiteit.

Wat gebeurt er met mijn gegevens?

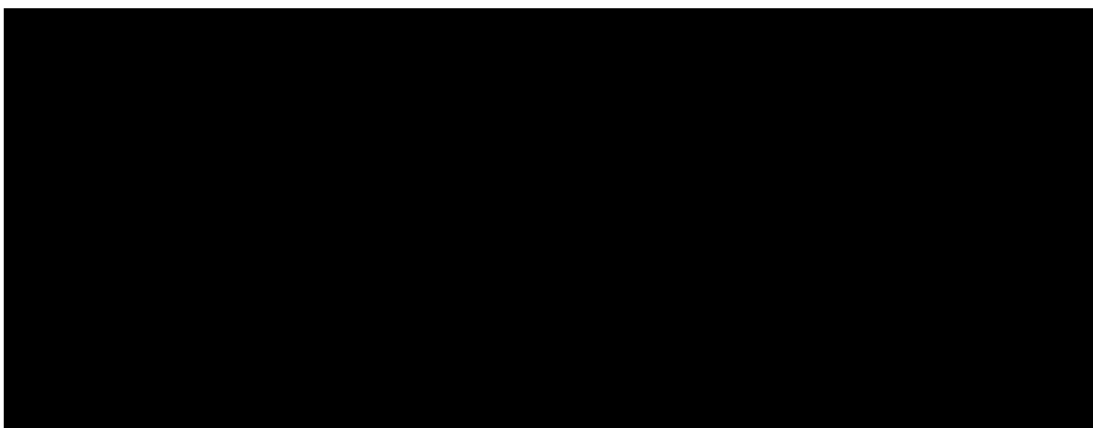
Mocht u besluiten deel te nemen, dan worden er de volgende (bijzondere) persoonlijke gegevens verzameld: uw e-mailadres; geboortedatum van ouders en kind; geslacht van ouders en kind; informatie over de gezondheid van ouders en kind; uw antwoorden op de vragenlijsten. De gegevens zullen alleen gebruikt worden voor onderzoeksdoeleinden en zullen alleen beschikbaar worden gesteld aan externe onderzoekers zonder alle gegevens die tot identificatie van personen kunnen leiden. Al onze onderzoekers zullen overeenkomstig het EU-reglement voor de gegevensbescherming van persoonlijke data (GDPR) handelen. Meer informatie over hoe we met gegevens omgaan vindt u op: www.babyandchild.nl/privacybeleid.

Wat moet u doen als u wilt deelnemen?

Om de vragenlijst te starten kunt u de QR-code scannen of de link gebruiken. Aan het begin van de vragenlijst wordt u gevraagd om een participantnummer in te vullen. De participantnummers staan hieronder vermeld, één voor elke ouder. Bij deze brief vindt u ook twee kaartjes met daarop dezelfde QR-code, link en participantnummers als hieronder. Deze kaartjes kunt u en/of de andere ouder gebruiken als u de vragenlijst op een ander moment wil invullen. Als slechts één ouder de vragenlijst invult, gebruik dan de eerste van de twee participantnummers.



U kunt de vragenlijst invullen tot en met 6 mei 2025.



Appendix C

Duration Audio Recordings per Participating Family

| TBB Study Family ID | DKT Study Family ID | Duration Audio Recordings (Hours:Minutes:Seconds) |
|---------------------|---------------------|--|
| TBB002 | DKT002 | 13:41:46 |
| TBB006 | DKT006 | 20:20:06 |
| TBB008 | DKT008 | 15:19:33 |
| TBB009 | DKT009 | 21:15:10 |
| TBB010 | DKT010 | 23:10:51 |
| TBB011 | DKT011 | 27:22:44 |
| TBB012 | DKT012 | 25:38:29 |
| TBB013 | DKT013 | 17:16:44 |
| TBB015 | DKT015 | 06:23:38 |
| TBB016 | DKT016 | 22:50:08 |
| TBB017 | DKT017 | 03:55:43 |
| TBB018 | DKT018 | 06:47:12 |
| TBB019 | DKT019 | 22:04:34 |
| TBB022 | DKT022 | 28:43:58 |
| TBB024 | DKT024 | 05:02:03 |
| TBB025 | DKT025 | 23:17:06 |
| TBB029 | DKT029 | 20:31:04 |
| TBB035 | DKT035 | 22:53:44 |
| TBB036 | DKT036 | 28:19:56 |
| TBB038 | DKT038 | 21:53:31 |
| TBB041 | DKT041 | 24:52:34 |
| TBB042 | DKT042 | 23:57:25 |
| TBB043 | DKT043 | 25:03:52 |
| TBB044 | DKT044 | 23:49:07 |
| TBB045 | DKT045 | 22:05:25 |
| TBB046 | DKT046 | 19:58:20 |
| TBB047 | DKT047 | 22:31:59 |
| TBB050 | DKT050 | 23:26:03 |

Appendix D

All ALICE-derived Variables and Definitions

| Variable | Definition |
|--------------------|--|
| Sum words MAL KCHI | Total number of words spoken by the male individual during interactions with the target child using the device. |
| Sum words FEM KCHI | Total number of words spoken by the female individual during interactions with the target child using the device. |
| Sum time MAL KCHI | Total duration (in seconds) of interaction time between the male individual and the target child using the device. |
| Sum time FEM KCHI | Total duration (in seconds) of interaction time between the female individual and the target child using the device. |
| Total turns MAL | Total number of conversational turns between the male individual and the target child using the device. |
| Total turns FEM | Total number of conversational turns between the female individual and the target child using the device. |
| Total turns | Combined total number of conversational turns (max. 5 secs) between both male and female individuals and the target child. |
| Count MAL | Total number of utterances produced by the male speaker. |
| Count FEM | Total number of utterances produced by the female speaker. |
| Tot Duration MAL | Total speaking duration of the male speaker (in seconds). |
| Tot Duration FEM | Total speaking duration of the female speaker (in seconds). |
| Tot Phonemes MAL | Total number of phonemes produced by the male speaker. |
| Tot Syllables MAL | Total number of syllables produced by the male speaker. |
| Tot Words MAL | Total number of words produced by the male speaker. |
| Tot Phonemes FEM | Total number of phonemes produced by the female speaker. |
| Tot Syllables FEM | Total number of syllables produced by the female speaker. |
| Tot Words FEM | Total number of words produced by the female speaker. |

| | |
|---------------------|---|
| Tot Duration KCHI | Total speaking duration of the target child using the device (in seconds). |
| Tot Duration SPEECH | Total duration of unclassified speech (in seconds). This refers to speech-like audio that could not be confidently attributed to any specific speaker category by the pretrained model (ALICE). |
| Tot Duration CHI | Total speaking duration of other children (not the target child) (in seconds). |

Appendix E

Original Items Underlying Final Items of the Family Routines Questionnaire

- Do you have the chance to point out the names, the colors or the sizes of items in the grocery store when taking your child there, or are you too busy getting your shopping done? (StimQ2 Toddler, Cates et al., 2023)
- Do you tell your child stories (such as folktales, made up stories without using a book, or stories about activities you have done together in the past)? (StimQ2 Toddler, Cates et al., 2023)
- Some caregivers talk to their children about their surroundings and what is happening around them. Have you started to talk to your child in this way or is she too young for that? (StimQ2 Toddler, Cates et al., 2023)
- When your child is playing with his/her toys, do you usually talk to your child or does your child usually play on his/her own? (StimQ2 Preschool, Cates et al., 2023)
- I sing songs or perform rhythms/rhymes for my child. (Children’s Music-Related Behavior Questionnaire (CMRBQ), Valerio et al., 2012)
- I sing songs or chant rhythms/rhymes with my child during daily routines such as bathing, dressing, cleaning up toys, getting ready to go somewhere. (Children’s Music-Related Behavior Questionnaire (CMRBQ), Valerio et al., 2012)
- I correct my child's use of the plural and encourage them to use them correctly. (Home Learning Environment Questionnaire (HLEQ), Krousorati et al., 2022)
- I correct my child's use of the past and future tenses and encourage them to use them correctly. (Home Learning Environment Questionnaire (HLEQ), Krousorati et al., 2022)
- How much do you talk to your child? (Parent Perception Questionnaire, Richards et al., 2017)
- How much does your child talk (or vocalize)? (Parent Perception Questionnaire, Richards et al., 2017)
- How often is the television on in the dining room while the child is eating at home? (Martinot et al., 2021)
- On most weekdays, about how much time does your child usually spend alone in front of a TV, computer, cellphone or other electronic device watching programs, playing games, accessing the internet or using social media, not including schoolwork? (National Survey of Children’s Health, Child and Adolescent Health Measurement Initiative, 2022)

- Do you share screen media experiences with your child? (Seven-in-Seven Screen Exposure Questionnaire, Yalçin et al., 2021)
- The child uses digital media gadgets to learn letters, words, vocabulary or language online. (Digital-Screen Exposure Questionnaire (DSEQ), Kaur et al., 2021)
- I talk to my child about what he/she has seen on TV. (Home Learning Environment Questionnaire (HLEQ), Krousorati et al., 2022)
- How many children's books do you have in your house? (More extensive family background questionnaire used at the Max Planck Institute)
- How many days per week do you read books together with your child at times of day other than bedtime? (StimQ2 Toddler, Cates et al., 2023)
- How many days per week do you read a book together with your child at bedtime? (StimQ2 Toddler, Cates et al., 2023)
- For how long do you usually read to your child each time you read to them? (More extensive family background questionnaire used at the Max Planck Institute)
- I encourage my child to narrate while looking at pictures. (Home Literacy Environment Questionnaire (HLEQ), Marjanovič Umek et al., 2005)
- I allow my child to interrupt me and ask questions while I am reading to her. (Home Literacy Environment Questionnaire (HLEQ), Marjanovič Umek et al., 2005)
- I allow my child to create her own stories while I am reading to her. (Home Literacy Environment Questionnaire (HLEQ), Marjanovič Umek et al., 2005)
- Has your child ever been identified as having a problem that might affect their development by a professional (e.g. very premature, developmental delay, hearing problem etc.)? (More extensive family background questionnaire used at the Max Planck Institute)
- Would you like to specify what this problem is? (More extensive family background questionnaire used at the Max Planck Institute)
- How many siblings younger than 5 years old does your child have at home? (More extensive family background questionnaire used at the Max Planck Institute)
- How many siblings with an age of 5 years old and older does your child have at home? (More extensive family background questionnaire used at the Max Planck Institute)
- What is the highest level of education completed by parent 1? (More extensive family background questionnaire used at the Max Planck Institute)

- What is the highest level of education completed by parent 2? (More extensive family background questionnaire used at the Max Planck Institute)
- Have you, your partner or any of the grandparents of your child been diagnosed by a professional with one of the following? (More extensive family background questionnaire used at the Max Planck Institute)

Final Items and Response Options in the Family Routines Questionnaire

- Deze vragenlijst wordt ingevuld door:
 - o De biologische moeder van het kind
 - o De biologische vader van het kind
 - o De niet-biologische moeder van het kind
 - o De niet-biologische vader van het kind
 - o Anders. Wat is uw relatie tot het kind? Ik ben de...
- Wat is uw geboortedatum? (dd-mm-jjjj)?
- Wat is de geboortedatum van uw kind (dd-mm-jjjj)?
- Wat is de hoogste opleiding die u heeft afgemaakt?
 - o Basisschool
 - o VMBO/MAVO
 - o HAVO
 - o VWO
 - o MBO
 - o HBO
 - o Universiteit Bachelor
 - o Universiteit Master
 - o Post-master degree (bijv. arts-specialisatie of PhD)
 - o Anders, namelijk: ...
 - o Zeg ik liever niet
- Is uw kind een jongen, een meisje of een ander geslacht?
 - o Jongen
 - o Meisje
 - o Anders, namelijk: ...
 - o Zeg ik liever niet

- Is er bij uw kind door een deskundige ooit een ontwikkelingsprobleem of -stoornis vastgesteld die de taalontwikkeling zou kunnen beïnvloeden (bijv. spraak- of gehoorproblemen, hersenverlamming, autisme, fragiele X-syndroom, spierdystrofie, Di George-syndroom, Down-syndroom, Williams-syndroom)?
 - o Ja
 - o Nee
 - o Zeg ik liever niet
- Vult u hier alstublieft in wat dit probleem is. Als u deze vraag liever niet wil beantwoorden, kies dan de optie 'Zeg ik liever niet'.
 - o Licht hier toe...
 - o Zeg ik liever niet
- Heeft u, uw partner of een van de biologische grootouders van uw kind ooit een diagnose gekregen voor een van de volgende aandoeningen? Wij vragen dit omdat bepaalde aandoeningen in de familie de taalontwikkeling van kinderen kunnen beïnvloeden. Let op! U kunt "Geen van bovenstaande" en "Zeg ik liever niet/Weet ik niet" niet combineren met andere antwoordopties.
 - o Autisme
 - o Dyslexie
 - o Taalstoornis
 - o Geen van bovenstaande
 - o Zeg ik liever niet/Weet ik niet
- Wordt er thuis alleen Nederlands gesproken?
 - o Ja
 - o Nee
- Hoeveel broers en/of zussen jonger dan 5 jaar wonen er bij uw kind thuis? Als uw kind in meer dan één huis woont, vertel ons dan over het huis met de meeste kinderen jonger dan 5 jaar.
 - o Geen
 - o Eén
 - o Twee
 - o Drie tot vier
 - o Vijf of meer

- Hoeveel broers en/of zussen van 5 jaar en ouder wonen er bij uw kind thuis? Als uw kind in meer dan één huis woont, vertel ons dan over het huis met de meeste kinderen ouder dan 5 jaar.
 - Geen
 - Eén
 - Twee
 - Drie tot vier
 - Vijf of meer
- Hoe vaak vertelt u uw kind verhalen (zoals sprookjes, verzonden verhalen zonder een boek te gebruiken, of over activiteiten die u in het verleden samen hebt gedaan)?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Hoe vaak benoemt u de namen, de kleuren of de grootte van producten in de supermarkt, wanneer u uw kind daarheen meeneemt?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Denk aan een normale week. Hoeveel dagen per week leest u een boek samen met uw kind vóór het slapengaan?
 - Minder dan één dag per week
 - Ongeveer één dag per week
 - Twee tot drie dagen per week
 - Vier tot zes dagen per week
 - Elke dag
- Denk aan een normale week. Hoeveel dagen per week leest u boeken samen met uw kind op andere momenten van de dag dan vóór het slapengaan?
 - Minder dan één dag per week
 - Ongeveer één dag per week

- Twee tot drie dagen per week
 - Vier tot zes dagen per week
 - Elke dag
- Hoeveel praat uw kind in vergelijking tot leeftijdsgenootjes?
 - Niet veel
 - Minder dan het gemiddelde kind
 - Gemiddeld
 - Meer dan het gemiddelde kind
 - Veel meer dan het gemiddelde kind
- Hoeveel praat u met uw kind in vergelijking tot andere ouders met kinderen van dezelfde leeftijd?
 - Niet veel
 - Minder dan de gemiddelde ouder
 - Gemiddeld
 - Meer dan de gemiddelde ouder
 - Veel meer dan de gemiddelde ouder
- Hoeveel kinderboeken heeft u (ongeveer) in huis? Geef een schatting. Tel zowel papieren boeken als elektronische boeken mee.
 - 0 - 5
 - 6 - 15
 - 16 - 40
 - 41 - 75
 - 76 - 150
 - Meer dan 150
- Wanneer uw kind eet, hoe vaak staat de tv dan aan in dezelfde ruimte als waar uw kind eet?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Soms maken jonge kinderen grammaticale fouten (zoals 'loopte' in plaats van 'liep', of 'boeks' in plaats van 'boeken'). Als uw kind dit doet, hoe vaak verbetert u uw kind dan?

- Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Deze vraag gaat over een normale week. Hoeveel tijd brengt uw kind meestal alleen door voor de televisie, computer, mobiele telefoon of een ander elektronisch apparaat om programma's te kijken, spelletjes te spelen of het internet te gebruiken (bijvoorbeeld YouTube Kids)?
 - Minder dan 1 uur per dag (1)
 - Tussen de 1 en 2 uur per dag (2)
 - Tussen de 2 uur en 3 uur per dag (3)
 - Tussen de 3 en 4 uur per dag (4)
 - Meer dan 4 uur per dag
- Tijdens het voorlezen, hoe vaak moedigt u uw kind aan om te vertellen over wat uw kind ziet op de plaatjes?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Hoe vaak deelt u samen met uw kind schermtijd, zoals het samen kijken van films, spelen van spelletjes of het browsen op sociale media?
 - Minder dan 1 uur per dag
 - Tussen de 1 en 2 uur per dag
 - Tussen de 2 uur en 3 uur per dag
 - Tussen de 3 en 4 uur per dag
 - Meer dan 4 uur per dag
- Tijdens het voorlezen, hoe vaak moedigt u uw kind aan om te vertellen over wat uw kind ziet op de plaatjes?
 - Nooit
 - Bijna nooit
 - Soms

- Vaak
 - Heel vaak
- Hoe vaak deelt u samen met uw kind schermtijd, zoals het samen kijken van films, spelen van spelletjes of het browsen op sociale media?
 - Nooit
 - Minder dan één keer per week
 - Ongeveer één keer per week
 - Meerdere keren per week
 - Dagelijks
- Hoe oud was uw kind toen u begon met regelmatig voorlezen?
 - 0 maanden – 6 maanden
 - 7 maanden – 12 maanden
 - 13 maanden – 24 maanden
 - 25 – 36 maanden
 - Ik ben nog niet begonnen
- Als u weleens boekjes voorleest aan uw kind, hoe lang leest u dan voor, in het algemeen?
 - We lezen nog geen boekjes voor
 - Minder dan 5 minuten per keer
 - 5 minuten per keer
 - 10 minuten per keer
 - 15 minuten per keer
 - 30 minuten per keer
 - Meer dan 30 minuten per keer
- Hoe vaak laat u het toe dat uw kind u onderbreekt en vragen stelt, terwijl u voorleest?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Sommige verzorgers praten met hun kinderen over hun omgeving en wat er om hen heen gebeurt. Wanneer ouders/verzorgers hiermee beginnen is heel verschillend. Hoe vaak praat u zo met uw kind?
 - Nooit

- Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Hoe vaak gebruikt uw kind digitale apparaten om online letters, woorden of taal te leren?
 - Nooit
 - Minder dan één keer per week
 - Ongeveer één keer per week
 - Meerdere keren per week
 - Dagelijks
- Hoe vaak laat u uw kind zijn/haar eigen verhaal verzinnen, terwijl u voorleest?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Hoe vaak is uw kind, in het algemeen, geïnteresseerd in een boek als u hem/haar voorleest?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Soms praten verzorgers met hun kind als hij/zij met speelgoed speelt, maar soms laten ze hun kind ook alleen spelen. Hoe vaak praat u met uw kind als hij/zij met speelgoed speelt?
 - Nooit – Mijn kind speelt altijd zelfstandig
 - Bijna nooit – Ik praat af en toe met mijn kind terwijl hij/zij speelt
 - Soms – Ik praat ongeveer de helft van de tijd met mijn kind tijdens het spelen
 - Vaak – Ik praat regelmatig met mijn kind tijdens het spelen
 - Heel vaak – Ik praat bijna altijd met mijn kind tijdens het spelen

- Hoe vaak praat u met uw kind over tekenfilms of andere (kinder)programma's die uw kind gezien heeft? Denk aan alle (kinder)programma's, niet alleen die op tv, maar ook die op een telefoon, tablet of laptop worden bekeken.
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Er zijn verschillende manieren om met kinderen te lezen. Soms hebben ouder en kind gesprekjes over wat er in het boek gebeurt (door vragen te stellen aan het kind, de gebeurtenissen op de pagina te bespreken, etc.). Hoe vaak voert u zulke gesprekken wanneer u met uw kind leest?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Sommige verzorgers zingen liedjes of zeggen rijmpjes met hun kinderen tijdens dagelijkse activiteiten, zoals in bad gaan, aankleden, opruimen of klaarmaken om ergens naartoe te gaan. Hoe vaak zingt u liedjes of zegt u rijmpjes met uw kind?
 - Nooit
 - Bijna nooit
 - Soms
 - Vaak
 - Heel vaak
- Wie is er met uw kind op een normale maandag/dinsdag/woensdag/donderdag/vrijdag/zaterdag/zondag? Geef voor elk tijdstip aan met wie uw kind het meeste contact heeft.
 - Uzelf
 - De andere ouder
 - Beide ouders
 - Dagopvang
 - Andere persoon dan ouders

- Zelfstandig spelen/slapen
- Zou u een aanvullende beschrijving willen geven van het weekschema van uw kind om ons een duidelijker beeld te geven? Zo ja, vermeld dan de dag, het tijdstip en met wie het kind deze tijd doorbrengt. Bijvoorbeeld, als uw kind 's nachts wakker wordt, met wie is hij dan en op welke dagen? Vermeld bijvoorbeeld ook als één van de ouders langere tijd per jaar niet thuis is. Zo nee, laat dit veld dan leeg.
- Hoeveel uur per dag besteedt u op een normale doordeweekse dag aan face-to-face gesprekjes met uw kind? Denk hierbij aan alle interacties gedurende de dag. Geef een schatting.
 - Ongeveer 15 minuten of minder
 - Ongeveer 30 minuten
 - Ongeveer 1 uur
 - Ongeveer 2 uur
 - Ongeveer 3 uur
 - Ongeveer 4 uur of meer
- Hoeveel uur per dag besteedt u op een normale weekenddag aan face-to-face gesprekjes met uw kind? Denk hierbij aan alle interacties gedurende de dag. Geef een schatting.
 - Ongeveer 15 minuten of minder
 - Ongeveer 30 minuten
 - Ongeveer 1 uur
 - Ongeveer 2 uur
 - Ongeveer 3 uur
 - Ongeveer 4 uur of meer
- Dit was het einde van de vragenlijst. Hartelijk bedankt voor het invullen van alle vragen! Wij horen graag hoe het was om de vragenlijst in te vullen. Denkt u bijvoorbeeld aan:
 - Vragen die onduidelijk waren
 - Antwoordopties. Miste u antwoordopties bij vragen? Zo ja, bij welke vragen? (U mag hier een algemeen antwoord geven)
 - Leesbaarheid
 - Waren alle vragen duidelijk leesbaar? Denk hierbij bijvoorbeeld aan de opmaak van de vragenlijst.

Alle andere feedback en suggesties om deze vragenlijst te verbeteren zijn ook van harte welkom.

Appendix F

R Script Data Analysis

1. Cleaning R & Loading Packages

```
# Load necessary packages
library(readxl)
library(dplyr)
library(psych)
library(ggplot2)

df <- read_excel("dkt_merged.xlsx")
df_father <- read_excel("dkt_merged_father.xlsx")
df_mother <- read_excel("dkt_merged_mother.xlsx")

# Turning the tibbles into dataframes
df <- as.data.frame(df)
df_father <- as.data.frame(df_father)
df_mother <- as.data.frame(df_mother)
```

2. Descriptives

2.1 Descriptive Statistics Parental Responsiveness (PR)

```
# Descriptive statistics fathers
mean_father <- round(mean(df_father$Average_prop, na.rm = TRUE), 2)
sd_father <- round(sd(df_father$Average_prop, na.rm = TRUE), 2)
median_father <- round(median(df_father$Average_prop, na.rm = TRUE), 2)
range_father <- round(range(df_father$Average_prop, na.rm = TRUE), 3)
ci_father <- round(t.test(df_father$Average_prop)$conf.int, 2)

cat("Proxy score fathers: Mean =", mean_father,
    ", SD =", sd_father,
    ", Median =", median_father,
    ", Range =", paste0(range_father[1], "-", range_father[2]),
    ", 95% CI =", paste0(ci_father[1], "-", ci_father[2]), "\n")
```

```
# Descriptive statistics mothers
mean_mother <- round(mean(df_mother$Average_prop, na.rm = TRUE), 2)
sd_mother <- round(sd(df_mother$Average_prop, na.rm = TRUE), 2)
median_mother <- round(median(df_mother$Average_prop, na.rm = TRUE), 2)
range_mother <- round(range(df_mother$Average_prop, na.rm = TRUE), 3)
ci_mother <- round(t.test(df_mother$Average_prop)$conf.int, 2)

cat("Proxy score mothers: Mean =", mean_mother,
    ", SD =", sd_mother,
    ", Median =", median_mother,
    ", Range =", paste0(range_mother[1], "-", range_mother[2]),
    ", 95% CI =", paste0(ci_mother[1], "-", ci_mother[2]), "\n")
```

2.2 Distribution Parental Input Quantity Scores Fathers and Mothers

```
# Distribution Input Scores
```

```
hist(df_father$Average_prop,
     main = "Parental Input Quantity Score Fathers",
     xlab = "Average Proportional Input Score",
     ylab = "Frequency",
     col = "darkolivegreen",
     border = "lightgreen",
     xlim = c(0.0, 1.0),
     ylim = c(0, 10))
```

```
hist(df_mother$Average_prop,
     main = "Parental Input Quantity Score Mothers",
     xlab = "Average Proportional Input Score",
     ylab = "Frequency",
     col = "chocolate",
     border = "navajowhite",
     xlim = c(0.0, 1.0),
     ylim = c(0, 10))
```

3. Hypothesis 1

```

# For the fathers
cor.test(df_father$Average_prop, df_father$Sum_words_MAL_KCHI)
cor.test(df_father$Average_prop, df_father$Sum_time_MAL_KCHI)
cor.test(df_father$Average_prop, df_father$Total_turns_MAL)
cor.test(df_father$Average_prop, df_father$Tot_Phonemes_MAL)
cor.test(df_father$Average_prop, df_father$Tot_Syllables_MAL)

# Correlation table all ALICE variables fathers
vars <- c("Sum_words_MAL_KCHI", "Sum_time_MAL_KCHI", "Total_turns_MAL",
         "Tot_Phonemes_MAL", "Tot_Syllables_MAL")

get_cor_results <- function(varname) {
  x <- df_father$Average_prop
  y <- df_father[[varname]]

  valid <- complete.cases(x, y)
  test <- cor.test(x[valid], y[valid])

  data.frame(
    Variable = varname,
    Correlation = round(test$estimate, 3),
    P_value = round(test$p.value, 4),
    Significance = ifelse(test$p.value < 0.05, "Yes", "No"))}

results_father <- do.call(rbind, lapply(vars, get_cor_results))
print(results_father)

# For the mothers
cor.test(df_mother$Average_prop, df_mother$Sum_words_FEM_KCHI)
cor.test(df_mother$Average_prop, df_mother$Sum_time_FEM_KCHI)
cor.test(df_mother$Average_prop, df_mother$Total_turns_FEM)
cor.test(df_mother$Average_prop, df_mother$Tot_Phonemes_FEM)
cor.test(df_mother$Average_prop, df_mother$Tot_Syllables_FEM)

```

```

# Corelation table all ALICE variables mothers
vars_mother <- c("Sum_words_FEM_KCHI", "Sum_time_FEM_KCHI",
"Total_turns_FEM", "Tot_Phonemes_FEM", "Tot_Syllables_FEM")

get_cor_results_mother <- function(varname) {
  x <- df_mother$Average_prop
  y <- df_mother[[varname]]
  valid <- complete.cases(x, y)
  test <- cor.test(x[valid], y[valid])

  data.frame(
    Variable = varname,
    Correlation = round(test$estimate, 3),
    P_value = round(test$p.value, 4),
    Significance = ifelse(test$p.value < 0.05, "Yes", "No"))}

results_mother <- do.call(rbind, lapply(vars_mother, get_cor_results_mother))
print(results_mother)

```

4. Hypothesis 2

```

# Hypothesis 2
names <- colnames(df[3:13]) # Extracting the names of the PR items
items <- df[, names] # Subset of the bigger dataframe, where only the PR items are present
results_reliability <- alpha(items)

# All Cronbach's items
results_reliability

# Cronbach's alpha
results_reliability$total$raw_alpha

# Alpha value if item dropped
alpha_if_dropped <- data.frame(

```

```

item = rownames(results_reliability$alpha.drop),
alpha_if_item_dropped = results_reliability$alpha.drop$raw_alpha) #raw_alpha
print(alpha_if_dropped)

```

```
# Inter-item correlations
```

```

inter_item_cor <- round(cor(items, use = "pairwise.complete.obs"), digits = 3)
inter_item_cor

```

```
# Q37
```

```

c(mean = mean(items[, 10]),
  median = median(items[, 10]),
  sd = sd(items[, 10]))

```

```
# Item-total correlations
```

```

item_total_correlations <- data.frame(
  item = rownames(results_reliability$item.stats),
  item_total_correlation = results_reliability$item.stats$r.drop) #r.drop
print(item_total_correlations)

```

5. Hypothesis 3

```
# Correlation between PR and CDIs should be positively correlated, with small to moderate correlations
```

```
# Correlation fathers
```

```

cor.test(df_father$Average_prop, df_father$CDI_total_score)
cor.test()

```

```
# Correlation mothers
```

```
cor.test(df_mother$Average_prop, df_mother$CDI_total_score)
```

```
# Scatterplot for the fathers
```

```

ggplot(df_father, aes(x = Average_prop, y = CDI_total_score)) +
  geom_point(color = "black", fill = "white", shape = 21, size = 2.5, stroke = 0.8) + #
  zwart/wit punten

```

```
geom_smooth(method = "lm", se = TRUE, color = "darkolivegreen", fill = "lightgreen") + #
lijn + CI
labs(title = "Relation between Parental Input Quantity Score and CDI Score (Fathers)",
x = "Parental Input Quantity Score", y = "CDI Total Score") + theme_minimal()

# Scatterplot for the mothers
ggplot(df_mother, aes(x = Average_prop, y = CDI_total_score)) +
  geom_point(color = "black", fill = "white", shape = 21, size = 2.5, stroke = 0.8) + #
  zwart/wit punten
  geom_smooth(method = "lm", se = TRUE, color = "chocolate", fill = "navajowhite") + #
  lijn + CI
labs( title = "Relation between Parental Input Quantity Score and CDI Score (Mothers)",
x = "Parental Input Quantity Score",
y = "CDI Total Score") + theme_minimal()
```