**\***Data from 6 participants was excluded from the analyses due to indications of non-typical development, as reported by parents in response to items concerning complications experienced during pregnancy, labor, first days, months and years of life in addition to any ongoing concerns at the time of testing. None of the excluded participants are included in the shared dataset.

| **Data Collection** | **Variable** | **Full name** | **Notes** |
| --- | --- | --- | --- |
| **Info** | ID | Participant code | - |
| Group | Participant group | Rural or Urban |
| Age | Participant age in days at time of testing | - |
| Gender | Participant gender | M or F |
| **Testing Battery** | T0Digit\_O | Digit Span Task (Forward) score at Baseline or T0 | Min: 0  Max: 7 |
| T0\_Digit\_I | Digit Span Task (Backward) score at Baseline or T0 | Min: 0  Max: 7 |
| T0\_Prolec | Prolec-R (Same-Different task) score at Baseline or T0 | Min: 0  Max: 20 |
| T0Rav | Raven CPM score at Baseline or T0 | Min: 0  Max: 36 |
| T0TevoCo | Tevopreesc (Concept subtest) score at Baseline or T0 | Two versions were applied depending on participant age group  Version A (standardized for ages 5y to 5y 11m) max: 24  Version B (standardized for ages 6y to 6y 11m) max: 19 |
| T0L1RP | MAIN Retelling Production score in Spanish (L1) at Baseline or T0 | Min: 0  Max: 17 |
| T0L1RC | MAIN Retelling Comprehension score in Spanish (L1) at Baseline or T0 | Min: 0  Max: 10 |
| T0L1RS | MAIN Retelling Structural Complexity score in Spanish (L1) at Baseline or T0 | Min: 0  Max: 3 |
| T0L1TP | MAIN Telling Production score in Spanish (L1) at Baseline or T0 | Min: 0  Max: 17 |
| T0L1TC | MAIN Telling Comprehension score in Spanish (L1) at Baseline or T0 | Min: 0  Max: 10 |
| T0L1TS | MAIN Telling Structural Complexity score in Spanish (L1) at Baseline or T0 | Min: 0  Max: 3 |
| **Respective standardized**  **protocols** | Prolec\_MainIndexes\_st | Prolec-R (Same-Different task) label based on standardized protocol and T0\_Prolec score | - |
| Rav\_Mall\_bracket | Raven CPM age bracket based on participant age | Rows starting with “avg…” represent those participants with NA age. Their age bracket was calculated based on their group’s mean age |
| Rav\_Mall\_pc | Raven CPM percentile based on standardized protocol and T0\_Rav score | - |
| Rav\_Mall\_dx | Raven CPM label based on standardized protocol and Rav\_Mall\_pc | - |
| Tevo\_bracket | Tevopreesc version completed by the participant at Baseline or T0 | - |
| T0TevoCo\_st | Tevopreesc (concepts subtest) label based on standardized protocol and T0\_TevoCo score | - |
| **Parental**  **Questionnaire** | Mage | Mother’s at at Baseline or T0 | - |
| actual.Meducation | Mother’s education level as described by the parent | - |
| Meducation | Mother’s education level category | 0 = Primary school (complete or incomplete)  1 = Secondary school (complete or incomplete)  2 = Incomplete Bachelor’s or Technical degree  3 = Complete Bachelor’s degree  4 = Graduate and Post-graduate school  NA = No info provided |
| actual.Mreading | Mother’s reading habits (categorical levels) | Never  Almost never  Sometimes a month  Sometimes a week  Daily  No info provided |
| Mreading | Mother’s reading habits (numeric levels) | 0 = Never  1 = Almost never  2 = Sometimes a month  3 = Sometimes a week  4 = Daily  NA = No info provided |
| Page | Father’s at at Baseline or T0 | - |
| actual.Peducation | Father’s education level as described by the parent | - |
| Peducation | Father’s education level category | 0 = Primary school (complete or incomplete)  1 = Secondary school (complete or incomplete)  2 = Incomplete Bachelor’s or Technical degree  3 = Complete Bachelor’s degree  4 = Graduate and Post-graduate school  NA = No info provided |
| actual.Preading | Father’s reading habits (categorical levels) | Never  Almost never  Sometimes a month  Sometimes a week  Daily  No info provided |
| Preading | Father’s reading habits (numeric levels) | 0 = Never  1 = Almost never  2 = Sometimes a month  3 = Sometimes a week  4 = Daily  NA = No info provided |
| actual.Bookshome | Number of books available at home (categorical levels) | None  Less than 10  Between 10 and 50  Between 50 and 100  More than 100  NA |
| Bookshome | Number of books available at home (numeric levels) | 0 = None  1 = Less than 10  2 = Between 10 and 50  3 = Between 50 and 100  4 = More than 100  NA = No info provided |
| actual.Booksofparticipant | Number of books available at home meant for the participant to use (categorical levels) | None  Less than 10  Between 10 and 50  Between 50 and 100  More than 100  NA |
| Booksofparticipant | Number of books available at home meant for the participant to use (numeric levels) | 0 = None1 = Less than 10  2 = Between 10 and 50  3 = Between 50 and 100  4 = More than 100  NA = No info provided |
| Income | Monthly household income as described by the parent | - |
| Adj.income | Monthly household income as described by the parent (adjusted) | Rows where Income was stated as a range (e.g. “between 700K to 1M CRC”) were transformed to be equivalent to the mean value of the range provided |
| OnlySpanish | Does the participant listen to a language other than Spanish at home? | Yes/No |
| DetailsOnlySpanish | If yes to the previous question, what language(s) and in what context? | Notes as provided by the participant’s parents |