www.jpeds.com • THE JOURNAL OF PEDIATRICS

Restricted, Repetitive, and Reciprocal Social Behavior in Toddlers Born Small for Gestation Duration

Robin Sifre, BS[1], Carolyn Lasch, BA[1], Angela Fenoglio, EdM[1], Michael K. Georgieff, MD[1,2], Jason J. Wolff, PhD[3], and Jed T. Elison, PhD[1,2]

Objective To characterize restricted and repetitive behaviors (RRBs) and reciprocal social behaviors (RSBs) in a large sample of toddlers who represent a range of birth weights and gestation durations.

Study design A battery of questionnaires characterizing demographic information and measuring RRBs and RSBs were completed by parents of toddlers between the ages of 17-26 months (n = 1589 total; n = 98 preterm). The association between birth weight and/or gestation duration and the primary outcome measures (RRBs and RSBs as ascertained through the Repetitive Behavior Scale for Early Childhood and the Video-Referenced Rating of Reciprocal Social Behavior) were tested by using hierarchical multivariate multiple regression.

Results Toddlers born preterm and full term did not differ on RRBs or RSBs. However, there were significant associations between birth weight percentile for gestation duration (BPGD) and RRBs ($\beta = -2.1$, $P = .03$), above and beyond the effects of age, sex, and vocabulary production. Similarly, there was a significant association between BPGD and RSBs ($\beta = -1.8$, $P = .02$), above and beyond the effects of age, sex, and vocabulary production.

Conclusions These findings demonstrate that BPGD better predicted putative antecedents of adverse psychological outcomes—specifically, RRBs and RSBs—than gestation duration alone. These findings provide insight to the link between preterm birth and suboptimal behavioral/psychological outcomes and suggest that high birth weight, which may reflect a more optimal intrauterine environment, may serve as a protective factor irrespective of gestation duration. *(J Pediatr 2018;200:118-24).*

Children born preterm are more likely to exhibit symptoms consistent with autism spectrum disorder (ASD),[1,2] and children born very preterm (<30 weeks of gestation) are at 3 times greater risk for developing psychiatric diagnoses, including anxiety disorders, attention-deficit/hyperactivity disorder, and ASD.[3] Moderate-to-late preterm birth (32-36 weeks of gestation) also is associated with cognitive, language, and motoric delays and deficits in social–emotional competence at 24 months.[4] Although isolating the effects of prematurity from low birth weight on cognitive outcomes is challenging, studies on children born preterm have identified low birth weight as a factor influencing the relationship between prematurity and language delays.[5] A meta-analysis examining cognitive and behavioral outcomes in children born preterm found that mean cognitive scores were proportional to birth weight and gestational age,[6] suggesting that both factors may contribute independently to psychiatric outcomes.

Although there is evidence demonstrating increased risk for adverse psychological outcomes in children born preterm, improved characterization of behavioral risk factors present before maladaptive patterns of behavior consolidate may augment early detection and intervention. This line of inquiry is complicated not only by the observed heterogeneity in clinical manifestation but also the exceptional variability observed among typically developing toddlers and children of preschool age. Measurement tools that better approximate meaningful dimensionality of complex behavior are needed for improved characterization of early emerging risk factors.

The objective of this study was to begin to address this gap by examining whether individual differences in complex behaviors vary as a function of gestation duration and birth weight. Restricted and repetitive behaviors (RRBs) include repetitive motor mannerisms, rituals and routines, circumscribed interests, and insistence on sameness behaviors. Although RRBs can be normative and transient in toddlers,[7,8] increased rates of RRBs can interfere with daily life, are associated with anxiety and phobias in 1- to 7-year-old children,[9] and are a common

ASD	Autism spectrum disorder
BPGD	Birth weight percentile for gestation duration
NICU	Neonatal intensive care unit
RSB	Reciprocal social behavior
RBS-EC	Repetitive Behavior Scale for Early Childhood
RRBs	Restricted and repetitive behaviors
vrRSB	Video-Referenced Rating of Reciprocal Social Behaviors

From the [1]Institute of Child Development; [2]Department of Pediatrics, Division of Neonatology; and [3]Department of Educational Psychology, University of Minnesota Twin Cities, Minneapolis, MN

Supported by a National Institute of Mental Health (NIMH) BRAINS award (R01 MH104324) and McKnight Land-Grant Professorship to J.E. J.W. received grant funding from NIMH K01 MH101653. R.S. is supported by the National Science Foundation Graduate Research Fellowship Program. The authors declare no conflicts of interest.

Portions of this study were presented at the International Meeting for Autism Research, May 10-13, 2017, San Francisco, California.

https://doi.org10.1016/j.jpeds.2018.05.003

feature in children diagnosed with ASD, anxiety disorders,[10] and other neurodevelopmental disorders.[7,11] Reciprocal social behaviors (RSBs) refer to emotionally appropriate and socially contingent communicative behaviors with others.[12] RSBs also emerge in early childhood and are disrupted in toddlers with ASD. Both behaviors span the typical-to-atypical continuum, are normally distributed within the population,[7,12] and when disrupted are associated with adverse behavioral/psychological outcomes. Taken together, these features make RRBs and RSBs good targets for identifying individual differences that may be indicative of early risk for adverse psychological outcomes in children born preterm. A critical first step is to characterize profiles of these complex behaviors early in development in such children.

Toward this goal, we used data from a large community sample of 1589 toddlers that were collected to characterize individual differences in RRBs and RSBs. Of the 1589 toddlers, 98 (6.2%) were born preterm (<37 weeks of gestation). We aimed to examine group differences in RRBs and RSBs and to characterize how individual differences in RRBs and RSBs vary as a function of continuous measures of gestation duration and birth weight.

Methods

Parents of toddlers between 17 and 26 months (n = 4268), recruited from the University of Minnesota Institute of Child Development participant registry, were invited to participate in a study about their child's development between June 2015 to July 2016. The study was approved by the University of Minnesota Human Research Protection Program and institutional review board (#1501S61261), and parents of all participants provided informed consent and permission for their child to participate in this research study. Of the 4268 invited, 2112 (49.5%) chose to participate and completed at least 1 questionnaire. The final sample included 1589 children (833 male), including 98 infants born preterm at 29-36 weeks of gestation, with complete and reliable data (for more information on recruitment and attrition and exclusion, see **Figure 1**, **Table I**, and the **Appendix**; available at www.jpeds.com). A follow-up questionnaire was then sent to parents of toddlers born preterm to characterize additional perinatal risk factors (**Appendix**).

Measures of Birth Weight, Gestation Duration, and Birth Weight Percentile for Gestation Duration

Birth weight was determined through parent report. Gestation duration was determined by computing the difference between the child's expected date of birth and their actual date of birth. Parent report of both birth weight and expected date of birth have been shown to be reliable when compared with medical records 10-15 years postbirth.[13] Birth weight percentile for gestation duration (BPGD) was calculated via the Fenton growth chart for infants born preterm.[14] Toddlers born preterm had significantly lower birth weights (M 2448.5 g, SD 527.5) than toddlers born full term (M 3555.5 g, SD 439.4), (mean difference = 1107.1 g, 95% CI 999.0-1215.1, $t = -20.4$, $P < 1 \times 10^{-15}$, $d = 2.5$) but were statistically equivalent in BPGD (mean preterm = 54.7%, mean full term = 56.4%, $t = -0.68$, $P = .5$, $d = 0.1$).

Outcome Measures

Repetitive Behavior Scale for Early Childhood (RBS-EC). The RBS-EC[7] is a 34-item parent-report questionnaire that is a downward-extension of the Repetitive Behavior Scale–Revised,[10] with good-to-excellent psychometric properties and evidence of validity and reliability[7] (based on a sample of toddlers that partially overlaps with the present sample). The questionnaire is intended to capture normative variation in young children (for distributions of RBS-EC scores in the present sample, see **Figure 2**, A [available at www.jpeds.com], and the **Appendix**). Each item contributes to 2 measures: items endorsed and frequency score. These measures can be summed into an overall composite measure (scored 0-34) or disaggregated into 4 psychometrically validated subscale scores: Repetitive Motor (scored 0-9), Ritual and Routine (scored 0-10), Restricted Behavior (sored 0-8), and Self-directed behavior (scored 0-7). See http://www.cehd.umn.edu/edpsych/research/resources/rbs-ec/ for access to the instrument.

Video-Referenced Rating of Reciprocal Social Behavior (vrRSB).[12] The vrRSB is a 48-item parent questionnaire designed for 18- to 30-month-old subjects and is a downward extension of the Social Responsiveness Scale.[15] The first 13 items refer to a video-displayed exemplar (ie, a typically developing 19-month-old child displaying reciprocal social and communicative behaviors) and ask parents to rate whether their child displays the same behaviors on scale of 0 (not at all) to 4 (more than child in video). The remaining 14-48 items ask parents to check the box that best describes their child's behavior over the past month from 0 (not true) to 4 (almost always true). Greater scores on the vrRSB represent lower developmental capacity for RSBs (for distributions of vrRSB scores in the present sample, see **Figure 2**, B [available at www.jpeds.com] and the **Appendix**).

Vocabulary Production. A 400-item checklist measuring expressive vocabulary from the MacArthur-Bates Communicative Development Inventory[16] was used to generate a composite score of expressive language. Parents of toddlers born preterm reported a lower number of words produced (M 86 words, SD 95) than toddlers born at full term (M 108 words, SD 104), (mean difference = 22, 95% CI 3-42, $t = -2.2$, $P = .03$, $d = 0.2$). Although this effect size was small, words produced was included as a control variable in the models as a proxy for overall language ability.

Statistical Analyses

To determine whether birth weight, gestation duration, or BPGD best predicted RRBs and RSBs, we used 3 series of hierarchical multiple multivariable regression analyses using R 3.3.1 (R Core Team; R Foundation for Statistical Computing, Vienna, Austria) via the *stats* package. All 3 predictor variables were continuous measures.[17]

Chronological age, sex, and vocabulary production were entered as step 1 control predictors. Next, the predictive variable of interest (ie, BPGD, birth weight, or gestation duration) was entered as the step 2 predictor to determine the portion of variance accounted for above and beyond the step 1 control variables. For the RBS-EC, each series of regressions was conducted for the composite score and its 4 subscales. Each of these 5 outcomes produces 2 scores (items endorsed and frequency), yielding a total of 10 outcomes tested with 3 predictors. For the vrRSB, each series of regressions was conducted for the total score, tested with 3 predictors. All *P* values were adjusted for multiple comparisons for the RBS-EC (30 total) and vrRSB (3 total) using the Benjamini–Hochberg method[18] with a false discovery rate of 5%, and are reported as *q*-values.

Results

Of the total study sample (n = 1589), 833 were male (52.4%). In total, 92.6% of participating parents reported a household income >$50-74 999, and 89.7% had at least a college degree. In total, 87.6% of the sample was white, 6.9% were more than 1 race, and 5.5% were black or African American, Asian, other, or unknown (**Table II**; available at www.jpeds.com). The average age of participants was 20.1 months (SD 2.0, min 17.2, max 26.5). Toddlers born preterm and full term did not differ in age (preterm = 20.4 months [SD 2.3], full term = 20.0 months [SD 2.0], Welch $t = 1.7$, $P = .1$). The average RBS-EC score for the entire study sample was 11.8 (SD 6.6), and the average vrRSB score for the entire study sample was 20.8 (SD 7.5) (**Table III**; available at www.jpeds.com), the latter of which replicates past findings based on an independent sample of 252 toddler twins.[12]

Restricted and Repetitive Behaviors

There were no significant group differences between toddlers born preterm and full term on the number or frequency of reported behaviors on the RBS-EC or any of its subscales (**Table III**). Significant model results assessing the association between variability in these behaviors and continuous measures of birth weight and gestation duration, are subsequently reported. For full model results, see **Table IV** (available at www.jpeds.com).

Birth Weight Percentile for Gestation Duration

Composite Score. BPGD predicted the composite RBS-EC score, $\beta = -2.1$, $F(4, 1584) = 3.7$, accounting for a significant portion of unique variance above and beyond the effects of age, sex, and vocabulary production, $\Delta R^2 = .01$, $F(1, 1584) = 10.8$, $q = .007$ (**Figure 3**, A). BPGD also predicted the frequency of behaviors on the RBS-EC, $\beta = -5.6$, F(4, 1584) = 5.3, accounting for a significant portion of unique variance, $\Delta R^2 = .01$, F(1, 1584) = 12.35, $q = .007$ (**Figure 3**, B).

Self-Directed. BPGD did not provide explanatory power above and beyond the control variables for predicting the number or frequency of behaviors on the Self-Directed subscale.

Repetitive Motor. BPGD predicted the Repetitive Motor subscale score, $\beta = -0.85$, $F(4, 1584) = 3.5$, accounting for a significant portion of unique variance ΔR^2=.01, $F(1, 1584) = 8.2$, $q = .019$ (**Figure 3**, C). BPGD also predicted the frequency of behaviors on the Repetitive Motor subscale, $\beta = -3.0$, $F(4, 1584) = 5.3$, accounting for a significant portion of unique variance, ΔR^2=.005, $F(1, 1584) = 7.7$, $q = .02$ (**Figure 3**, D).

Restricted Behavior. BPGD predicted the Restricted subscale score $\beta = -0.7$, $F(4, 1584) = 6.1$, accounting for a significant portion of unique variance ($\Delta R^2 = .01$, F(1, 1584) = 10.9, $q = .007$) (**Figure 3**, E). BPGD also predicted the frequency of items reported on the Restricted subscale $\beta = -1.3$, $F(4, 1584) = 7.2$, $q = .03$, accounting for a significant portion of unique variance, $\Delta R^2 = .01$, F(1, 1584) = 10.6, $q = .007$ (**Figure 3**, F).

Ritual and Routine. Although BPGD did not predict the Ritual and Routine subscale score, it did predict the frequency of reported behaviors, $\beta = -0.9$, $F(4, 1584) = 8.5$, accounting for a significant portion of unique variance, ΔR^2=.005, F(1, 1584) = 7.6, $q = .02$ (**Figure 3**, G).

Gestation Duration and Birth Weight

Gestation duration alone did not significantly predict RRBs on the composite RBS-EC or the 4 subscales. Birth weight did not predict RRBs on the composite RBS-EC but did significantly predict the number and frequency of behaviors on the Restricted subscale (**Table IV**).

Reciprocal Social Behaviors

There were no significant group differences between toddlers born preterm and full term on total vrRSB score (**Table III**). BPGD significantly predicted vrRSB scores, $\beta = -1.8$, $F(4, 1584) = 55.6$, $q = .017$, above and beyond control variables, $\Delta R^2 = .004$, $F(1, 1585) = 6.5$, $q = .017$ (**Figure 4**). Gestation duration significantly predicted vrRSB scores $\beta = -.23$, F(4, 1585) = 54.96, $q = .04$, above and beyond control variables, $\Delta R^2 = .002$, $F(1, 1585) = 4.4$, $q = .04$. Birth weight significantly predicted vrRSB scores, $\beta = -.001$, F(4, 1585) = 57.0, $q = .0007$, accounting for a significant portion of unique variance, ΔR^2=.01, F(1, 1584) = 11.4, $q = .0007$.

Further Supplementary Analyses

Perinatal Health Indicators. One possibility is that infants born very small for gestation duration may have been exposed to more perinatal health issues, which could account for the effect of BPGD on RRBs and RSBs. To test for this possibility, mean RBS-EC composite scores and vrRSB total scores were calculated as a function of whether infants born preterm were hospitalized, intubated, or had a cesarean delivery. Of those who responded to the follow-up survey (n = 89, 91%), 51 were delivered vaginally (57.3%), 68 (76.4%) spent at least 1 day in the neonatal intensive care unit (NICU; mean duration = 13 days, SD 14, min 1, max 71), and 12 (13.5%) were intubated for an average duration of 0.7 days (SD 2.8 days, min 5 minutes, max 21 days). Results revealed that

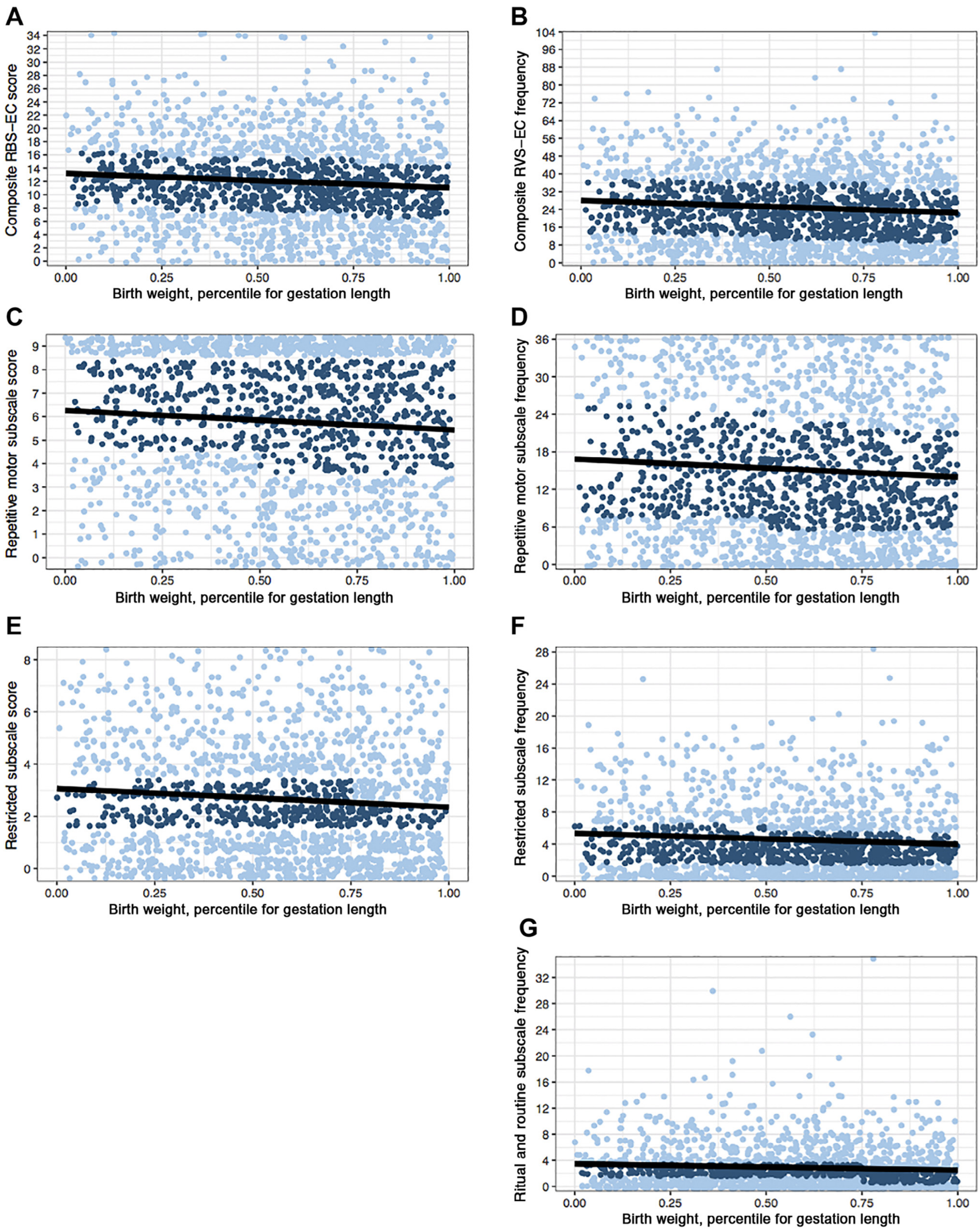

Figure 3. Raw data for **A,** RBS-EC composite score and **B,** frequency score; **C,** repetitive motor score and **D,** frequency; **E,** restricted score and **F,** frequency; and **G,** ritual and routine frequency scores for the entire sample (n = 1589). Darkly shaded data markers indicate the interquartile range (spanning the 25th to 75th percentiles) calculated within each birthweigh quartile. Linear regression lines are plotted over the raw data. Regression models include control variables, and are plotted for the reference group (males with a median vocabulary production score and median age at assessment).

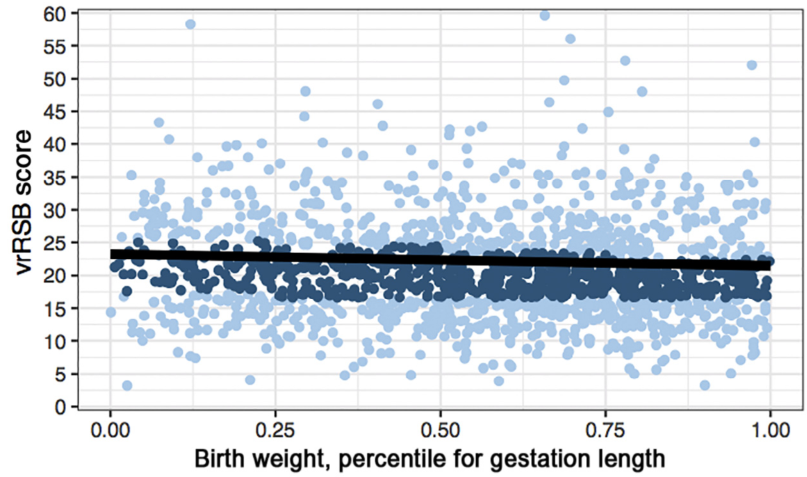

Figure 4. Raw data for the vrRSB for the entire sample (n = 1589). Darkly shaded data markers indicate the interquartile range (spanning the 25th to 75th percentiles) calculated within each birthweight quartile. Linear regression lines are plotted over the raw data. Regression models include control variables, and are plotted for the reference group (males with a median vocabulary production score and median age at assessment). Data show a significant negative association between BPGD and vrRSB scores. Higher scores indicate less sophisticated social behavior.

none of these indicators had a significant effect on RRBs or RSBs (all unadjusted P values ≥ .11), suggesting that BPGD is the primary factor driving later behavioral outcomes, irrespective of adverse perinatal events experienced by infants born preterm. For full statistical results, see **Table V** (available at www.jpeds.com).

Sensitivity Analysis. To examine whether the relationship between BPGD and RRBs and RSBs exists in children born full term, we re-ran these models with a subsample including only children born at full term (n = 1491). All significant model results obtained with the full study sample remained significant in this subsample. For full model results, see **Table VI** (available at www.jpeds.com).

Maximum Likelihood Estimation. Secondary analyses using maximum likelihood fit estimation[19] were conducted on models with significant step 2 predictors to select which of the control variables to include for the most parsimonious model[20] and to provide convergent evidence of BPGD being the strongest predictor of outcomes. Although the most likely models varied in terms of which control variables to include, all the most likely models included BPGD as a parameter, providing convergent evidence that BPGD accounts for variance in the data above and beyond the control variables (**Tables VII-XIV**; available at www.jpeds.com). Furthermore, we used ML estimation to determine which model would be selected when all control variables were held constant (age, sex, vocabulary production), and the candidate set of models included 1 of the 3 predictors of interest (gestation duration, birth weight, or BPGD) (**Tables XV-XXII**; available at www.jpeds.com). The most likely model always included BPGD, with the exception of the Restricted Interest subscale and vrRSB, which selected the model including birth weight as the most likely. Notably, these outcomes also were significantly predicted by birth weight in the regression analyses. For more detailed methods and model results, see the **Appendix**.

Discussion

The present study identified a relationship between variability in BPGD and RRBs and RSBs in 17- to 26-month-old toddlers. Group comparisons between infants born preterm and full term revealed no significant differences for RRBs or RSBs. Analyses that maintained the dimensionality of the primary independent measures (gestation duration and birth weight) revealed the strongest evidence for an index of gestation duration that also accounts for birth weight (BPGD). As BPGD increases, there is a decrease in the number of RRBs and in their frequency. Furthermore, as BPGD increases, the vrRSB total score decreases, suggesting that greater BPGD is associated with more developmentally sophisticated RSBs.

The finding that continuous associations between BPGD and behavior revealed information masked by group-comparisons based on preterm status is significant for 2 reasons. First, it suggests that gestation duration alone may be insufficient for grouping toddlers when examining complex behavioral phenotypes. Alternatively, toddlers may be clinically differentiated based on low and high BPGD. For example, **Figure 5**, A and B demonstrate there is a greater proportion of high RBS-EC scores for toddlers with low BPGD and a greater proportion of low RBS-EC scores for toddlers with high BPGD. A similar pattern is observed for the vrRSB (**Figure 5**, C and D), with a greater proportion of high vrRSB scores (indicating less sophisticated RSBs) for toddlers with

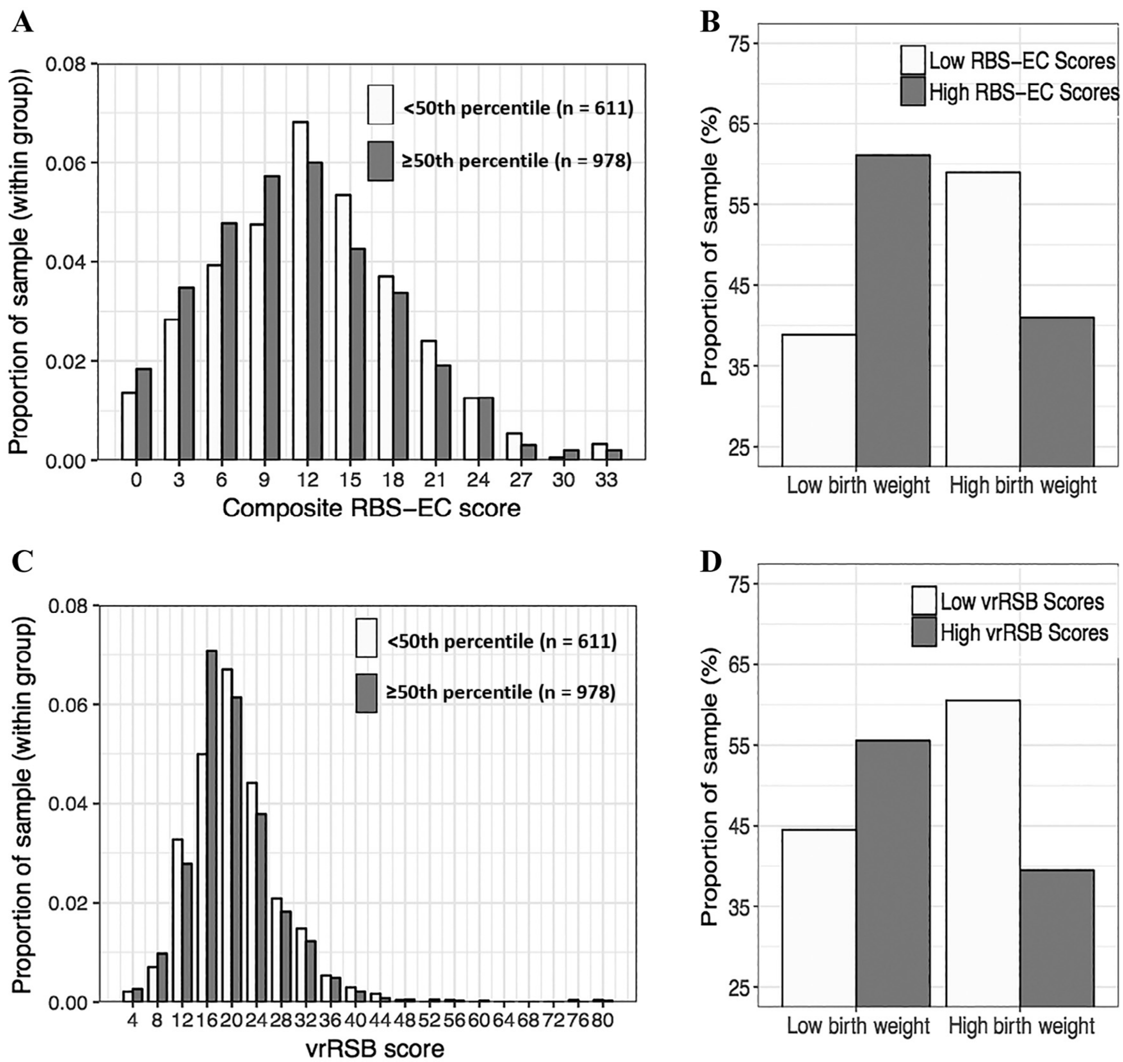

Figure 5. Distribution of low and high **A,** RBS-EC and **C,** vrRSB scores (as defined by below/above the sample mean), plotted as within-sample proportion for toddlers with low- and high-BPGD. Low BPGD was defined as being below the 50th percentile, while high BPGD was defined as >50th percentile. **B,** and **D,** show the relative proportion of toddlers with very low or high scores (as defined by below/above the 10th or 90th percentile) who are also below or above the 10th and 90th percentile for BPGD.

low BPGD, with the opposite pattern for toddlers with high BPGD. The finding that BPGD may be more predictive of RSBs and RRBs than gestation duration alone is supported by convergent evidence provided by maximum likelihood fit estimate analyses. Second, these results suggest that group-based analyses in general may fail to reveal complex associations between these heterogeneous behaviors and measures of birth weight and gestation, and that analyses focusing on continuous dimensionality along the typical-to-atypical continuum may be preferable.[17]

It should be noted that the variable that best predicted these complex behaviors, BPGD, could not be explained by risk factors associated with preterm birth (eg, time spent in NICU), and that the relationship held in the sensitivity analyses excluding toddlers born preterm. This is consistent with the “fetal origins” hypothesis,[21] which posits that uterine conditions have life-long developmental consequences and that fetal growth reflects these latent conditions.[22,23] The demonstrated relationship between BPGD and behavioral outcomes in the present study may be indexing the role that intrauterine experience (maternal age, nutrition, etc)[22] plays in life-long outcomes and may indicate more optimal intrauterine experience as a protective factor against adverse cognitive and behavioral outcomes for infants born at any gestation duration.

A strength of our study is the characterization of complex behaviors by measuring RRBs and RSBs in a large, heterogeneous sample of toddlers born preterm and full term. Through the use of dimensional assessments that capture normative variation in these behaviors (the RBS-EC and vrRSB) and continuous measures of birth weight and gestation, the present study succeeded in identifying a risk factor (BPGD) that predicts behavioral outcomes that may manifest in subtle delays or as subthreshold deficits and are obfuscated when examining categorical predictors.[17]

One limitation of our study is its data collection from a community sample of toddlers, 98 of whom were born preterm. Thus, we did not sample specifically for preterm birth or its associated risk factors. Although the preterm sample spans a continuum of risk factors (low birth weight, days in the NICU, days intubated, etc), they are a relatively healthy preterm sample. Indeed, the current preterm sample's relatively high average BPGD may account for the small effect sizes observed and may temper the observed effect between BPGD and behavioral outcomes. On one hand, given their health status, this allows us to speculate that suboptimal intrauterine conditions (as indexed by low BPGD) rather than illness or other adverse effects associated with preterm birth likely account for later adverse psychological outcomes in this population. This notion is supported by the finding that RRBs and RSBs did not differ as a function of perinatal health issues. On the other hand, a more clinically heterogeneous preterm group, as well as complete clinical characterizations on the entire study sample (preterm and full term), is needed to fully disaggregate the effects of gestation duration, BPGD, and health issues associated with preterm birth.

Follow-up work should enrich for target populations with medical comorbidities including infants born small for gestation duration and supplement parent-reports with direct observation and information from medical records to better understand potential underlying medical conditions leading to growth restriction in infants small for gestation duration. Given the relative racial and socioeconomic homogeneity of the current sample, future work also should collect measures from different racial/ethnic and socioeconomic groups to assess the generalizability of these instruments. Finally, follow-up work should characterize the prevalence of adverse psychological outcomes in these children at older ages, as longitudinal data will be needed to fully establish the relationship between BPGD, RRBs and RSBs, and adverse psychological outcomes. ■

We thank Dustin Gibson for his consultation, Elayne Teska for her efforts on data collection, and we are indebted to the families who contributed their time and effort.

Submitted for publication Jan 12, 2018; last revision received Apr 3, 2018; accepted May 2, 2018

Reprint requests: Jed T. Elison, PhD, University of Minnesota, Institute of Child Development, 51 E. River Rd, Minneapolis, MN 55455. E-mail: jtelison@umn.edu

References

1. Johnson S, Hollis C, Kocchar P, Hennessy E, Wolke D, Marlow N. Autism spectrum disorders in extremely preterm children. J Pediatr 2010;156:525.e2-531.e2.
2. Limperopoulos C, Bassan H, Sullivan NR, Soul JS, Robertson RL, Moore M, et al. Positive screening for autism in ex-preterm infants: prevalence and risk factors. Pediatrics 2008;121:758-65.
3. Treyvaud K, Ure A, Doyle LW, Lee KJ, Rogers CE, Kidokoro H, et al. Psychiatric outcomes at age seven for very preterm children: rates and predictors. J Child Psychol Psychiatry 2013;54:772-9.
4. Cheong JL, Doyle LW, Burnett AC, Lee KJ, Walsh JM, Potter CR, et al. Association between moderate and late preterm birth and neurodevelopment and social-emotional development at age 2 years. JAMA Pediatr 2017;171:e164805.
5. Cusson RM. Factors influencing language development in preterm infants. J Obstet Gynecol Neonatal Nurs 2003;32:402-9.
6. Bhutta A, Casey P, Cradock M, Anand K, Cleves MA. Cognitive and behavioral outcomes. JAMA 2002;288:728-37.
7. Wolff JJ, Boyd BA, Elison JT. A quantitative measure of restricted and repetitive behaviors for early childhood. J Neurodev Disord 2016;8:27.
8. Evans DW, Leckman JF, Carter A, Reznick S, Henshaw D, King RA, et al. Ritual, habit, and perfectionism: the prevalence and development of compulsive-like behavior in normal young children. Child Dev 1997;68:58-68.
9. Evans DW, Gray FL, Leckman JF. The rituals, fears and phobias of young children: insights from development, psychopathology and neurobiology. Child Psychiatry Hum Dev 1998;29:261-76.
10. Bodfish JW, Symons FJ, Parker DE, Lewis MH. Varieties of repetitive behaviour in autism: comparison to mental retardation. J Autism Dev Disord 2000;30:237-43.
11. Moss J, Oliver C, Arron K, Burbidge C, Berg K. The prevalence and phenomenology of repetitive behaviour in genetic syndromes. J Autism Dev Disord 2009;39:572-88.
12. Marrus N, Glowinski AL, Jacob T, Klin A, Jones W, Drain CE, et al. Rapid video-referenced ratings of reciprocal social behavior in toddlers: a twin study. J Child Psychol Psychiatry 2015;56:1338-46.
13. Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. J Clin Epidemiol 1998;51:399-405.
14. Fenton TR, Kim JH, Secker D, De Onis M, Garza C, Victora C, et al. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. BMC Pediatr 2013;13:59.
15. Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, et al. Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. J Autism Dev Disord 2003;33:427-33.
16. Fenson L, Marchman VA, Thal DJ, Dale PS, Reznick S, Bates E. MacArthur-Bates communicative development inventories. User's guide and technical manual. 2nd ed. Baltimore (MD): Brookes; 2007.
17. MacCallum RC, Zhang S, Preacher KJ, Rucker DD. On the practice of dichotomization of quantitative variables. Psychol Methods 2002;7:19-40.
18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing author. J R Stat Soc B 1995;57:289-300.
19. Akaike H, Csaki P. Information theory and an extension of the maximum likelihood principle. In: Petrov BN. and Csaki F, eds. 2nd International Symposium on Information Theory, Armenia, September 2-8, 1971, p. 267-81.
20. Snipes M, Taylor DC. Model selection and Akaike Information Criteria: an example from wine ratings and prices. Wine Econ Policy 2014;3:3-9.
21. Barker DJ. Fetal origins of coronary heart disease. Br Med J 1995;311:171-4.
22. Camerota M, Bollen KA. Birth weight, birth length, and gestational age as indicators of favorable fetal growth conditions in a US sample. PLoS ONE 2016;11:1-15.
23. Bollen K, Noble M, Adair L. Are gestational age, birth weight, and birth length indicators of favorable getal growth conditions? A structural equation analysis of Filipino infants. Stat Med 2013;32:2950-61.

Appendix

Supporting Information Document

Recruitment and Data Collection. Based on state birth records, parents voluntarily select into the participant registry pool. The registry largely reflects the racial/ethnic proportions of the broader Minneapolis–St. Paul metropolitan area but under-represents the socioeconomic diversity of this region. Data collection occurred online[1]; following extant literature,[2] parents received an e-mail introducing the study 2-3 days before receiving an e-mail that included a consent form and link to the battery of questionnaires. Follow-up e-mails were sent 1- and then 2-weeks following the initial invitation, inviting parents to participate or to finish partially completed questionnaires. Parents were reimbursed with a $10 gift card and their name was entered into a drawing for a $50 gift card (1 of 150 for every complete data set) for completing all of the questionnaires.

Attrition and Exclusionary Criteria. Participants who did not complete the entire battery (n = 274), which took ~40 minutes, were excluded from analyses. Furthermore, 227 children were excluded after quality control efforts identified for unrealistic questionnaire completion times (n = 216), ages outside of the target age recruitment range (n = 8), or unrealistic or missing gestation duration or birth weight (n = 3) (**Figure 1**). An additional 20 toddlers born post-term (≥42 weeks of gestation) were excluded from analysis, as postterm birth has unique risk factors that differ clinically from preterm birth,[3] the main focus of the present study. Two toddlers born extremely preterm (<28 weeks gestation) were excluded, as the sample size was not large enough to adequately represent this clinical group. Complete information on attrition and exclusion can be found in **Figure 1**, and comparisons between included and excluded participants can be found in the online supplement and in **Table I**.

Clinical Characterization of Preterm Sample. All parents of toddlers with a gestation duration of <37 weeks were sent a follow-up questionnaire to characterize perinatal risk factors and to verify their child's gestation duration and birth weight. Parents were asked questions about whether their child was delivered vaginally or via cesarean, and whether their child spent time in the NICU, was intubated, experienced postnatal infection or postnatal brain injury. Responses were coded for yes/no or unknown responses, as well as the duration of hospitalization and intubation. Gestation duration and birth weight were verified by asking parents to provide this information a second time. If there was a gestation length discrepancy of more than one day between their initial and follow-up responses (n = 20, average discrepancy = 2 days, range = 1-5 days), or if birth weight differed by more than 100g (n = 4, average discrepancy = 383 g, range = 311-454 g), we used the more recently reported measures that were ascertained through a shorter questionnaire or phone call, and were less likely to contain errors.

Eighty-nine (91% of the preterm sample) responded the follow-up survey. Acknowledging the clinical heterogeneity of preterm birth,[4] we chose to include the 9 toddlers born preterm whom we were unable to clinically characterize in analyses.

Information on Excluded Participants. Although 2112 participants chose to participate in the study, 490 were excluded for failing to complete all questionnaires (n = 274) or for providing unreliable data on these questionnaires (n = 216). We were interested in whether these parents who were unable to complete the survey battery may qualitatively differ from parents who completed the entire survey battery. For example, we hypothesized that these participants may have been more likely to have a child born preterm with health issues and therefore have less time to participate in research. To assess this, we compared participants who provided demographic information but failed to complete the RBS-EC, vrRSB, or MacArthur-Bates Communicative Development Inventories, to participants who were included in our final study sample. Specifically, we tested for group differences in gestation duration, birth weight, BPGD, sex ratio, parental age at time of birth, household income, and parent education. Of the 490 excluded participants who submitted a demographics form, 213 had incomplete or unreliable demographics forms (eg, missing fields, unrealistic gestation duration, etc) yielding 277 excluded participants who could be compared with the study sample (n = 1589). Due to these asymmetrical sample sizes, in addition to performing 2-sample *t* tests, we calculated Cohen *d* to estimate the effect size of group differences.

Results from these comparisons can found in **Table I**. There were no significant group differences in birth weight, sex ratio, household income, or parent education. Although there were significant group differences in gestation duration, BPGD, and parental age at time of birth, these effect sizes were small to negligible, and these differences were not meaningful (eg, included participants had an average gestation duration of 39.4 weeks while excluded participants had an average gestation duration of 39.21 weeks). These results suggest that excluded and included participants were comparable on these measures.

Model Selection

Although regression analyses are optimal for determining whether there is a significant relationship between two variables of interest while controlling for other population variables, such models may under-fit and not adequately capture the true nature of the data, or over-fit and increase variability in the estimation ratio.[4] To select the most parsimonious model with parameters that maximize the likelihood given the current data, and to provide convergent evidence with the results from the multiple multivariate linear regression analyses, planned post-hoc maximum likelihood fit estimation[5] was carried out on outcome measures that demonstrated a statistically significant association with BPGD based on linear regression results (composite RBS-EC score and frequency, Repetitive Motor subscale score and frequency, Restricted subscale score and frequency, Ritual and Routine subscale frequency, and vrRSB score). Two series of analyses were performed using Maximum Likelihood Fit Estimation. First, to select which control variables to include for the most parsimonious model, several models with different combinations of the predictor variable (BPGD) and control variables (age, sex, and vocabulary production) were entered. Two interaction terms (sex × vocabulary production and birth weight percentile × vocabulary production) were also examined based on a priori hypotheses generated through the literature demonstrating sex[6] and birth weight[7] effects on early vocabulary production. Second, maximum likelihood fit estimation was carried out with models including all control variables (age, sex, and vocabulary production) and (1) BPGD, (2) birth weight, or (3) gestation duration. The purpose of these analyses was to provide convergent evidence for the multiple multivariate linear regression results of the predictive power of BPGD.

Analyses were done using R 3.3.1 via the *AICcmodavg* package, and the Akaike information criteria was used to select the model closest to capturing the veridical relationship between the birth weight/gestation duration variables and RRBs/RSBs without losing generality from over-fitting. The log evidence ratio (LER) was used to determine the relative likelihood of a pair of models, enabling comparison across models. The terms "minimal," "substantial," "strong," and "decisive" correspond to LERs between model probabilities greater than 0, 0.5, 1, and 2, respectively.[4] Each model is compared to the best fitting model.

The first series of analyses, which was intended to select the most parsimonious model among the models including BPGD, revealed that for each behavioral outcome, the best-fitting model included a term for BPGD, providing convergent evidence that BPGD accounted for a unique portion of variance above and beyond the control variables (**Tables VII-XIV**). Furthermore, based on the LERs, there was "decisive" (RBS-EC score and frequency, Restricted score, Ritual and Routine Score, and vrRSB Score) or "strong" (Restricted frequency, Repetitive motor score and frequency) evidence that the best fitting model including BPGD was more predictive than any of the models with only control variables. Which control variables were included in the most parsimonious model differed across outcome variables. For example, no control variables were included in the best fitting model for RBS-EC Score (**Table VII**), whereas Sex and Age were included in the best fitting model for the Restricted Score (**Table X**).

The second series of analyses (**Tables X-XXII**), which was done to provide convergent evidence that BPGD was more predictive than birth weight or gestation duration alone, confirmed that the model with BPGD was the best fitting model for the RBS-EC Score and Frequency, Repetitive Motor Score and Frequency, Restricted frequency, and Ritual Frequency. Based on the LERs, there was "strong" to "decisive" evidence that model including BPGD better fit the data than the models including birth weight or gestation duration for RBS-EC score and frequency, and the Repetitive Motor subscale score and frequency. There was "minimal" evidence that the model including BPGD better fit the data than the model including birth weight for the Restricted subscale frequency. However, there was "decisive" evidence that this model was more predictive than the model including gestation duration.

Model likelihood estimation results revealed that the model including birth weight was the most likely model for the Restricted subscale score, and for the vrRSB total score. For the Restricted subscale score, there was "minimal" evidence in favor of the model including birth weight over the model including BPGD; however, there was "decisive" evidence in favor of the birth weight model over the gestation duration model. For the vrRSB, there was "strong" evidence in favor of the model including birth weight, over both the model including BPGD and gestation duration.

Post-Hoc Analyses of Perinatal Health Indicators. Toddlers with clinical characterization data (n = 89) were grouped into whether they were hospitalized (n = 68) or not (n = 21), whether they were intubated (n = 12) or not (n = 77), and whether they were delivered vaginally (n = 51) or via cesarean (n = 38). Two-sample *t* tests were then carried out based on these grouping to calculate mean differences with 95% CIs for RBS-EC composite scores and vrRSB total scores, as well as *P* values to test for statistical significance.

When comparing hospitalized and nonhospitalized toddlers, we found that there were no significant group differences in mean RBS-EC composite scores (mean difference = 1.82, 95% CI −1.54 to 5.17, *P* = .29) or mean vrRSB scores (mean difference = 3.36, 95% CI −0.8 to 7.52, *P* = .11). When comparing intubated and nonintubated toddlers, there were no significant group differences in mean RBS-EC composite scores (mean difference = 0.9, 95% CI −4.80 to 4.62, *P* = .97) or mean vrRSB scores (mean difference = 2.27, 95% CI −3.36 to 7.9, *P* = .40). When comparing toddlers delivered vaginally and those delivered via cesarean, there were no significant group differences in mean RBS-EC composite scores (mean difference = 0.11, 95% −3.14 to 2.92, *P* = .94) or vrRSB scores (mean difference = 1.29, 95% CI −4.93 to 2.35, *P* = .48). Complete summary statistics can be found in **Table IV**.

References

1. Weigold A, Weigold IK, Russel EJ. Examination of the equivalence of self-report survey-based paper-and-pencil and internet data collection methods. Psychol Methods 2013;18:53-70.
2. Singer E, Ye C. The use and effects of incentives in surveys. Ann Am Acad Pol Soc Sci 2013;645:121-41.
3. Olesen AW, Westergaard JG, Olsen J. Perinatal and maternal complications related to postterm delivery: a national register-based study, 1978-1993. Am J. Obstet Gynecol 2003;189:222-7.
4. Snipes M, Taylor DC. Model selection and Akaike information criteria: an example from wine ratings and prices. Wine Econ Policy 2014;3:3-9.
5. Akaike H, Csaki P. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, eds. 2nd International Symposium on Information Theory, Armenia, September 2-8; 1971. p. 267-81.
6. Marrus N, Glowinski AL, Jacob T, Klin A, Jones W, Drain CE, et al. Rapid video-referenced ratings of reciprocal social behavior in toddlers: a twin study. J Child Psychol Psychiatry 2015;56:1338-46.
7. Barre N, Morgan A, Doyle LW, Anderson PJ. Language abilities in children who were very preterm and/or very low birth weight: a meta-analysis. J Pediatr 2011;158:766.e1-774.e1.

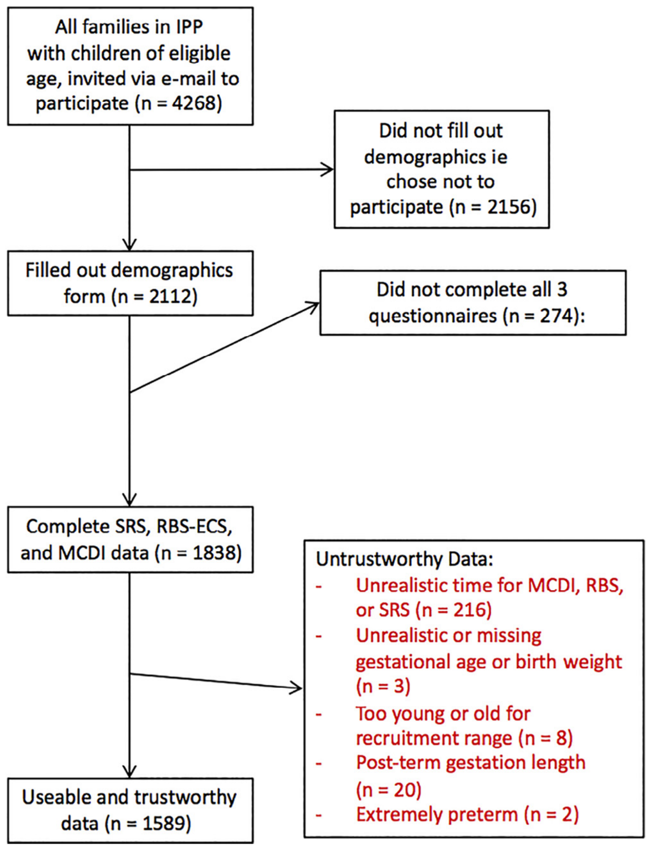

Figure 1. Flowchart describing survey participation and compliance. *IPP*, Institute of Child Development Participant Pool; *MCDI*, MacArthur-Bates Communicative Development Inventories; *SRS/vr-RSB*, Video-Referenced Rating of Reciprocal Social Behavior.

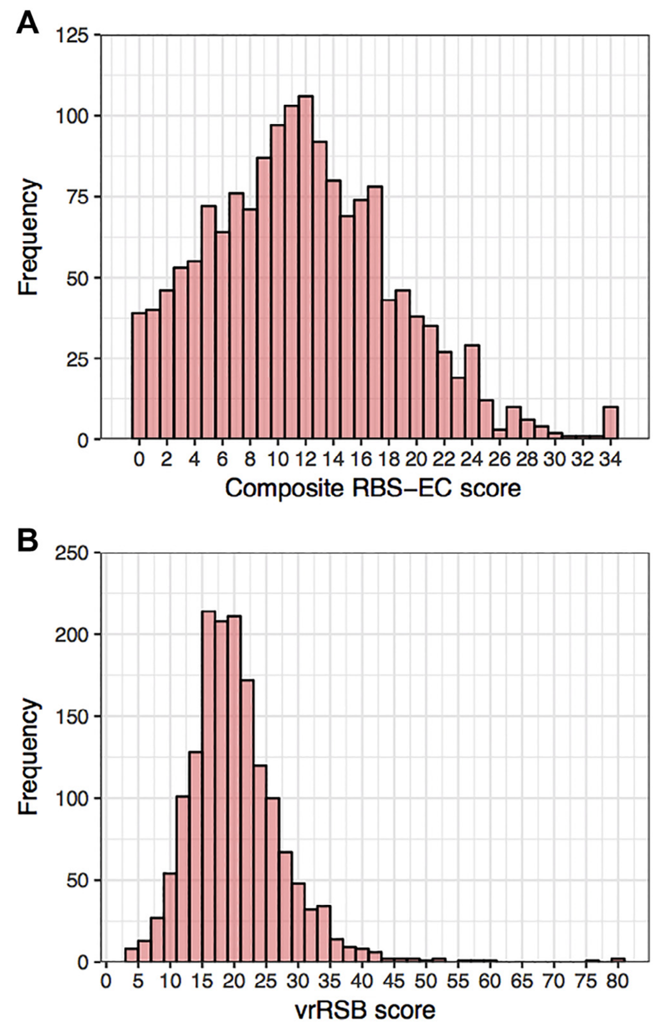

Figure 2. Distribution of **A,** composite RBS-EC scores and **B,** vrRSB scores for the entire sample (n = 1589).

Table I. Demographic comparisons between excluded participants with demographic data (n = 277) and study sample (n = 1589)

Outcome variables	Mean included (n = 1589)	Mean excluded (n = 277)	t	P	Cohen d
Gestation duration	39.40	39.21	4.57	.000005*	.08 (negligible)
Birth weight	3488.21g	3485.69	0.19	.85	.004 (negligible)
BPGD	56.57%	58.71%	–3.28	.001*	.08 (negligible)
Sex ratio	.52	.54	–.53	.60	–.03 (negligible)
Parent age	31.80	30.89	2.69	.008*	.22 (small)
Household income*	5.56	5.49	0.70	.49	.05 (negligible)
Parent education level†	5.78	5.70	.95	.34	.06 (negligible)

*Household incomes were coded as factors. These results indicate that the average participant fell between brackets 5 ($75-99 999) and 6 ($100-149 999) for both groups.
†Parent education level was coded as a factor. These results indicate that the average participant fell between bracket 5 (some graduate school) and 6 (graduate degree).

Table II. Demographic information for study sample (n = 1589), broken down by infants born preterm (n = 98) and full term (n = 1491)

Demographics	Preterm, n (%)	Full term, n (%)
Household income		
<$24 999	0 (0.00)	17 (1.14)
$25-34 999	2 (2.04)	30 (2.01)
$35-49 999	5 (5.10)	66 (4.43)
$50-74 999	16 (16.33)	219 (14.69)
$75-99 9999	21 (21.43)	312 (20.93)
$100-149 999	31 (31.63)	485 (32.53)
$150-199 999	14 (14.29)	224 (15.02)
>$200 000	9 (9.18)	138 (9.26)
Parent education		
Junior high	0 (0.00)	0 (0)
High school degree	1 (1.02)	1 (0.07)
Some college/2-year degree	1 (1.02)	15 (1.01)
College degree	13 (13.27)	133 (8.92)
Some graduate school	41 (41.84)	623 (41.78)
Graduate degree	6 (6.12)	97 (6.51)
Race		
White	92 (93.88)	1300 (87.19)
Black or African American	0 (0.00)	10 (0.67)
Asian	1 (1.02)	29 (1.95)
Unknown	2 (2.04)	23 (1.54)
Other	0 (0.00)	22 (1.48)
More than one race	3 (3.06)	107 (7.18)
Age of primary caregiver on child's date of birth, y		
≤25	2 (2.04)	52 (3.49)
25-29.9	23 (23.47)	466 (31.28)
30-34.9	43 (42.88)	689 (46.24)
35-39.9	26 (26.53)	249 (16.71)
40-44.9	3 (3.06)	30 (2.01)
≥45	1 (1.02)	4 (0.27)

Percentages calculated as within-group proportions.

Table III. RBS-EC and vrRSB scores information for study sample (n = 1589), broken down by infants born preterm (n = 98) and full term (n = 1491)

		Total sample (n = 1589)		Preterm (n = 98)		Full term (n = 1491)			
Scales	Scoring scale, min/max	Mean	SD	Mean	SD	Mean	SD	*t* stat*	*P*†
RBS-EC scale									
Composite topographies	0/34	11.75	6.62	12.94	7.06	11.67	6.58	1.73	.2
Self-directed topographies	0/7	1.52	1.68	1.71	1.66	1.5	1.68	0.57	.68
Repetitive motor topographies	0/9	5.77	3.04	5.78	3.02	5.77	3.04	1.22	.35
Ritual and routine topographies	0/10	1.94	1.95	2.32	2.25	1.92	1.93	0.5	.68
Restricted topographies	0/8	2.52	2.15	3.13	2.23	2.48	2.14	.02	.98
Composite frequency	0/136	23.99	16.2	24.96	17.64	23.92	16.11	–0.61	.68
Self-directed frequency	0/28	2.1	2.68	2.22	2.65	2.09	2.68	1.7	.2
Repetitive motor frequency	0/36	14.93	11.06	14.31	10.38	14.97	11.1	1.38	.31
Ritual and routine frequency	0/40	2.78	3.31	3.39	4.58	2.74	3.21	2.82	.06
Restricted frequency	0/32	4.18	4.24	5.04	4.84	4.12	4.2	1.83	.2
vrRSB	0/81	20.8	7.52	22.8	8.8	20.7	7.4	2.33	.12

There were no significant between-group differences for the RBS-EC, or any of its subscales, or for the vrRSB.
*T-statistics generated using the Welch *t* test.
†All *P* values adjusted for multiple comparisons using the Benjamini–Hochberg method.

Table IV. Multiple linear regression models for all 3 potential outcomes

Key parameters	Step 1 R^2*	Step 1 *P*	Step 2 β	Step 2 ΔR^2	Step 2 *F*	Step 2 *P*
BPGD						
RBS-EC composite topographies	**.003**	**.26**	**–2.1**	**.01**	**10.8**	**.007**
RBS-EC self-directed topographies	.004	.09	–.21	.001	1.691	0.301
RBS-EC rep motor topographies	**.004**	**.12**	**–.85**	**.01**	**8.2**	**.019**
RBS-EC ritual & routine topographies	.012	.0002	–.37	.002	3.9	0.121
RBS-EC restricted topographies	**.01**	**.004**	**–.7**	**.01**	**10.9**	**.007**
RBS-EC composite frequency	**.006**	**.03**	**–5.6**	**.01**	**12.35**	**.007**
RBS-EC self-directed frequency	.006	.02	–.34	.001	1.64	0.301
RBS-EC rep motor frequency	**.008**	**.01**	**–3.0**	**.005**	**7.7**	**.020**
RBS-EC ritual & routine frequency	**.016**	**8.5×10^{-6}**	**–.90**	**.005**	**7.6**	**.020**
RBS-EC restricted frequency	**.01**	**.0005**	**–1.3**	**.01**	**10.58**	**.007**
vrRSB total score	**.12**	**2.2×10^{-16}**	**–1.8**	**.004**	**55.6**	**.016**
Gestation duration						
RBS-EC composite topographies			.05	.0001	.2335	0.674
RBS-EC self-directed topographies			–.007	0	.06	0.828
RBS-EC rep motor topographies			.07	.001	2.0	0.301
RBS-EC rituual & routine topographies			.03	.0005	.87	0.390
RBS-EC restricted topographies			–.04	.0009	1.4	0.346
RBS-EC composite frequency			.24	.0006	.95	0.390
RBS-EC self-directed frequency			.01	0	.03	0.869
RBS-EC rep motor frequency			.25	.001	2.1	0.301
RBS-EC ritual & routine frequency			.05	.0006	1.04	0.390
RBS-EC restricted frequency			–.06	.0005	.88	0.390
vrRSB total score			**–.23**	**.003**	**4.4**	**.04**
Birth weight						
RBS-EC composite topographies			–.001	.0036	5.67	.052
RBS-EC self-directed topographies			–.0001	.001	1.758	0.301
RBS-EC rep motor topographies			–.0002	.001	1.8	0.301
RBS-EC ritual & routine topographies			–.0001	.0007	1.2	0.376
RBS-EC restricted topographies			**–.0004**	**.01**	**11**	**.007**
RBS-EC composite frequency			–.002	.003	4.9	.075
RBS-EC self-directed frequency			–.0001	.001	1.5	0.390
RBS-EC rep motor frequency			–.0007	.001	1.7	0.301
RBS-EC ritual & routine frequency			.0003	.002	2.6	0.238
RBS-EC restricted frequency			**–.0007**	**.01**	**10.3**	**.008**
vrRSB total score			**–.001**	**.006**	**11.44**	**.002**

R^2 and *P* values are reported for Step 1 models, which included the control variables (age at assessment, vocabulary production, and sex). Step 2 regression statistics are reported for the β value of the predictor of interest (birth weight percentile for gestation length, birth weight, or gestation length), as well as the ΔR^2 from the Step 1 model, and corresponding *F* and *P* values from comparison with Step 1 model. All *P* values adjusted for multiple comparisons using the Benjamini-Hochberg method, with rows in bold indicating $P < .05$.

Table V. Mean RBS-EC Composite Score and vrRSB Total Scores for preterm sample with available clinical characterization information (n = 89), broken down by perinatal health indicators (whether they were hospitalized, spent intubated, or were delivered vaginally)

Clinical variable	Mean (SD)	Mean(SD)	Mean difference (95% CI)	*P* value*
Hospitalized	No (n = 21)	Yes (n = 68)		
RBS-EC	11.52 (6.40)	13.34 (7.34)	1.82 (–1.54, 5.17)	.29
vrRSB	20.14 (0.38)	23.5 (0.13)	3.36 (–0.8, 7.52)	.11
Intubated	No (n = 77)	Yes (n = 12)		
RBS-EC	12.92 (7.19)	12.83 (7.09)	.09 (–4.80, 4.62)	.97
vrRSB	22.4 (0.11)	24.67 (0.71)	2.27 (–3.36, 7.9)	.40
Delivery method	Cesarean (n = 38)	Vaginal (n = 51)		
RBS-EC	12.97 (6.92)	12.86 (7.36)	0.11 (–3.14, 2.92)	.94
vrRSB	23.45 (0.21)	22.16 (0.18)	1.29 (–4.93, 2.35)	.48

*Unadjusted *P* values generated by 2-sample *t* test.

Table VI. Sensitivity analyses, with multiple linear regression results for toddlers born full term

	BPGD					
Outcome variable	Step 1 R^2*	Step 1 *p*	Step 2 β	Step 2 ΔR^2	Step 2 *F*	Step 2 *q*
RBS-EC composite topographies	**.002**	**.4**	**−2.15**	**.01**	**10.5**	**.003**
RBS-EC self-directed topographies	.004	.10	−.21	.001	1.81	0.200
RBS-EC rep motor topographies	**.003**	**.16**	**−.86**	**.01**	**7.8**	**.008**
RBS-EC ritual & routine topographies	.014	.0001	−.38	.002	3.8	.067
RBS-EC restricted topographies	**.007**	**.005**	**−.7**	**.01**	**10.3**	**.003**
RBS-EC composite frequency	**.004**	**.002**	**−5.9**	**.01**	**13.37**	**.003**
RBS-EC self-directed frequency	.006	.03	−.34	.001	1.61	0.204
RBS-EC rep motor frequency	**.008**	**.006**	**−3.3**	**.006**	**8.62**	**.005**
RBS-EC ritual & routine frequency	**.018**	**3.2×10^{-6}**	**−.94**	**.006**	**8.65**	**.005**
RBS-EC restricted frequency	**.01**	**.003**	**−1.4**	**.01**	**10.45**	**.003**
vrRSB total score	**.12**	**2.2×10^{-16}**	**−1.9**	**.004**	**7.1**	**.008**

R^2 and *P* values are reported for Step 1 models, which included the control variables (age at assessment, vocabulary production, and sex). Step 2 regression statistics are reported for the β-value of the predictor of interest (birth weight percentile for gestation length, birth weight, or gestation length), as well as the ΔR^2 from the Step 1 model, and corresponding *F* and *P* values from comparison with Step 1 model. All *P* values adjusted for multiple comparisons using the Benjamini–Hochberg method, with rows in bold indicating $P < .05$.

Table VII. Model selection and AIC for composite RBS-EC frequency

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	13348.71	0	0.27	0
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	13348.77	.06	0.26	.01
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	13350.72	2.01	0.1	0.44
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	13350.74	2.03	0.1	0.44
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	13351.58	2.86	.07	0.62
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	13352.23	3.52	.05	0.76
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + ϵ	4	13352.26	3.55	.05	0.77
$\beta_0 + \beta_1$Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	13352.29	3.58	.05	0.78
$\beta_0 + \beta_1$Birth Weight Percentile + ϵ	3	13353.16	4.45	.03	0.97
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age + ϵ	4	13353.16	4.45	.03	0.97
$\beta_0 + \beta_1$Vocabulary Production + ϵ	3	13358.87	10.16	0	2.21
$\beta_0 + \beta_1$Sex + ϵ	3	13363.1	14.39	0	3.12
$\beta_0 + \beta_1$Age + ϵ	3	13364.18	15.47	0	3.36

AIC, Akaike information criteria.

Table VIII. Model selection and AIC for composite RBS-EC score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Birth Weight Percentile + ϵ	3	10509.27	0	1	0
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + ϵ	4	10509.6	0.33	0.85	.07
$\beta_0 + \beta_1$Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	10509.88	0.61	0.74	0.13
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	10510.56	1.29	0.53	0.28
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	10510.95	1.68	0.43	0.36
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age + ϵ	4	10511.24	1.97	0.37	0.43
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	10511.57	2.31	0.32	0.5
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	10511.8	2.53	0.28	0.55
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	10512.53	3.26	0.2	0.71
$\beta_0 + \beta_1$Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	10513.31	4.04	0.13	0.88
$\beta_0 + \beta_1$Vocabulary Production + ϵ	3	10518.97	9.7	.01	2.11
$\beta_0 + \beta_1$Sex + ϵ	3	10519.03	9.76	.01	2.12
$\beta_0 + \beta_1$Age + ϵ	3	10520.54	11.27	0	2.45

Table IX. Model selection and AIC for restricted subscale frequency

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + ϵ	4	9081.9	0	0.35	0
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	9082.3	0.4	0.29	0.09
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	9083.79	1.89	0.14	0.41
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	9084.16	2.26	0.11	0.49
$\beta_0 + \beta_1$Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	9085.3	3.4	.06	0.74
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	9086.12	4.22	.04	0.92
$\beta_0 + \beta_1$Sex + ϵ	3	9090.91	9.01	0	1.96
$\beta_0 + \beta_1$Birth Weight Percentile + ϵ	3	9095.9	14	0	3.04
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age + ϵ	4	9096.26	14.35	0	3.12
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	9097.26	15.36	0	3.34
$\beta_0 + \beta_1$Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	9097.89	15.98	0	3.47
$\beta_0 + \beta_1$Age + ϵ	3	9104.7	22.79	0	4.95
$\beta_0 + \beta_1$Vocabulary Production + ϵ	3	9106.53	24.63	0	5.35

Table X. Model selection and AIC for restricted subscale score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	6932.58	0	0.37	0
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	6934.1	1.52	0.17	0.33
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + ϵ	4	6934.22	1.64	0.16	0.36
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	6935.1	2.52	0.11	0.55
$\beta_0 + \beta_1$Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	6935.49	2.9	.09	0.63
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	6936.04	3.46	.07	0.75
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age + ϵ	4	6939.2	6.62	.01	1.44
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	6939.91	7.33	.01	1.59
$\beta_0 + \beta_1$Birth Weight Percentile + ϵ	3	6940.89	8.31	.01	1.80
$\beta_0 + \beta_1$Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	6942.9	10.32	0	2.24
$\beta_0 + \beta_1$Sex + ϵ	3	6943.8	11.22	0	2.44
$\beta_0 + \beta_1$Age + ϵ	3	6948.19	15.6	0	3.39
$\beta_0 + \beta_1$Vocabulary Production + ϵ	3	6952.25	19.67	0	4.27

Table XI. Model selection and AIC for repetitive motor subscale frequency

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	12133.73	0	0.27	0.00
$\beta_0 + \beta_1$Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	12134.26	0.53	0.21	0.11
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	12135.38	1.65	0.12	0.36
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Age + ϵ	4	12135.44	1.71	0.12	0.37
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	12135.73	2	0.1	0.43
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	12136.2	2.47	.08	0.54
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	12137.43	3.7	.04	0.80
$\beta_0 + \beta_1$Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	12137.51	3.78	.04	0.82
$\beta_0 + \beta_1$Vocabulary Production + ϵ	3	12139.48	5.75	.02	1.25
$\beta_0 + \beta_1$Age + ϵ	3	12141.78	8.05	0	1.75
$\beta_0 + \beta_1$Birth Weight Percentile + ϵ	3	12143.01	9.28	0	2.02
$\beta_0 + \beta_1$Birth Weight Percentile + β_2Sex + ϵ	4	12145	11.27	0	2.45
$\beta_0 + \beta_1$Sex + ϵ	3	12150.78	17.05	0	3.70

Table XII. Model selection and AIC for Repetitive motor subscale score

Model names	df	AIC	dAICc	Model LL	Evidence ratio	LER
β_0+ β_1Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	8032.13	0	0.22	1.00	0.00
β_0+ β_1Birth Weight Percentile + β_2Age + ϵ	4	8032.27	0.13	0.21	1.07	0.03
β_0+ β_1Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	8032.92	0.78	0.15	1.48	0.17
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	8033.88	1.75	.09	2.4	0.38
β_0+ β_1Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	8034.23	2.09	.08	2.85	0.45
β_0 + β_1Birth Weight Percentile + ϵ	3	8034.44	2.3	.07	3.16	0.50
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	8034.76	2.63	.06	3.72	0.57
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	8034.98	2.84	.05	4.15	0.62
β_0+ β_1Birth Weight Percentile + β_2Sex + ϵ	4	8036.39	4.25	.03	8.39	0.92
β_0+ β_1Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	8036.65	4.52	.02	9.58	0.98
β_0+ β_1Vocabulary Production + ϵ	3	8037.99	5.86	.01	18.70	1.27
β_0+ β_1Age + ϵ	3	8038.87	6.74	.01	29.05	1.46
β_0+ β_1Sex + ϵ	3	8042.61	10.48	0	188.51	2.28

Table XIII. Model selection and AIC for ritual and routine subscale frequency

Model names	df	AIC	dAICc	Model LL	LER
β_0+ β_1Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	8288.29	0	0.28	0.00
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	8288.43	0.15	0.26	0.03
β_0+ β_1Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	8288.64	0.35	0.23	0.08
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	8290.45	2.16	.09	0.47
β_0+ β_1Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	8290.47	2.19	.09	0.48
β_0+ β_1Birth Weight Percentile + β_2Age + ϵ	4	8291.76	3.47	.05	0.75
β_0+ β_1Age + ϵ	3	8297.75	9.46	0	2.05
β_0+ β_1Birth Weight Percentile + β_2Sex + ϵ	4	8304.47	16.19	0	3.51
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	8305.44	17.16	0	3.73
β_0 + β_1Birth Weight Percentile + ϵ	3	8307.66	19.37	0	4.21
β_0+ β_1Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	8309.15	20.86	0	4.53
β_0+ β_1Sex + ϵ	3	8311.24	22.95	0	4.98
β_0+ β_1Vocabulary Production + ϵ	3	8315.91	27.62	0	6.00

Table XIV. Model selection and AIC for the vrRSB total score

Model names	df	AIC	dAICc	Model LL	LER
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Vocabulary Production + ϵ	5	10721.26	0	1	0.00
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + ϵ	6	10722.27	1.01	0.6	0.22
β_0+ β_1Birth Weight Percentile + β_2Sex+ β_3Age+ β_4VocabularyProduction + β_5Sex x Vocabulary Production + ϵ	7	10722.8	1.54	0.46	0.33
β_0+ β_1Birth WeightPercentile + β_2Sex + β_3Age + β_4VocabularyProduction + β_5Birth Weight Percentile x Vocabulary Production + ϵ	7	10724.04	2.78	0.25	0.60
β_0+ β_1Birth Weight Percentile + β_2VocabularyProduction+ϵ	4	10727.67	6.41	.04	1.39
β_0+ β_1Birth Weight Percentile + β_2Age+ β_3VocabularyProduction + ϵ	5	10728.12	6.86	.03	1.49
β_0+ β_1Vocabulary Production + ϵ	3	10732.27	11.01	0	2.39
β_0+ β_1Birth Weight Percentile + β_2Sex + β_3Age + ϵ	5	10861.94	140.68	0	30.55
β_0+ β_1Birth Weight Percentile + β_2Age + ϵ	4	10880.24	158.99	0	34.52
β_0+ β_1Age + ϵ	3	10887.77	166.51	0	36.16
β_0+ β_1Birth Weight Percentile + β_2Sex + ϵ	4	10899.11	177.86	0	38.62
β_0+ β_1Sex + ϵ	3	10905.64	184.39	0	40.04
β_0 + β_1Birth Weight Percentile + ϵ	3	10916.6	195.34	0	42.42

Table XV. Model Selection and AIC for the RBS-EC composite frequency score

Model names	df	AIC	dAICc	Model LL	LER
β_0+ β_1Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	13350.74	0	1	0
β_0+ β_1Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	13358.22	7.473577	.024	1.623
β_0+ β_1Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	13362.13	11.38752	.003	2.473

Table XVI. Model selection and AIC for the RBS-EC composite Score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	10511.8	0	1	0
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	10516.95	5.151223	.076	1.119
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	10522.4	10.594768	.005	2.301

Table XVII. Model selection and AIC for the repetitive motor frequency score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	12135.73	0	1	0
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	12141.32	5.592782	.061	1.214
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	12141.71	5.985363	.05	1.3

Table XVIII. Model selection and AIC for the repetitive motor score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	8034.758	0	1	0
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	8040.874	6.115255	.047	1.328
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	8041.097	6.338803	.042	1.376

Table XIX. Model selection and AIC for the restricted subscale frequency score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	9084.16	0	1	0
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	9084.642	0.4823039	0.786	0.105
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	9093.854	9.6939405	.008	2.105

Table XX. Model selection and AIC for the restricted subscale score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	6933.278	0	1	0
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	6934.104	0.8257144	0.662	0.179
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	6943.63	10.352515	.006	2.248

Table XXI. Model selection and AIC for the ritual and routines frequency score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	8288.434	0	1	0
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	8293.376	4.941893	.085	1.073
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	8294.999	6.56499	.038	1.426

Table XXII. Model selection and AIC for the vrRSB total score

Model names	df	AIC	dAICc	Model LL	LER
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight	6	10717.35	0	1	0
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4Birth Weight Percentile	6	10722.27	4.922013	.085	1.069
$\beta_0 + \beta_1$Sex + β_2Age + β_3VocabularyProduction + β_4GestationDuration	6	10724.36	7.017884	.03	1.524