Baby's First Years: Summary, Pre-registered Hypotheses, Analysis Strategies

June 19, 2022

Project Summary

In the Baby's First Years (BFY) study, one thousand infants born to mothers with incomes falling at or below the federal poverty threshold in four metropolitan areas in the United States were assigned at random within each of the metropolitan areas to one of two cash gift conditions. The sites are: New York City, the greater New Orleans metropolitan area, the greater Omaha metropolitan area, and the Twin Cities. IRB and challenges with even recruitment across sites led to a distribution of the 1,000 mothers across sites of 121 in one site (the Twin Cities), 295 in two of the other sites (New Orleans and Omaha) and 289 in New York. (We have also randomly sampled 80 of the participating families in the Twin Cities and New Orleans to participate in an in-depth qualitative study, but do not elaborate on those plans in this document.)

Mothers were recruited in maternity wards of the 12 participating hospitals shortly after giving birth and, after consenting, were administered a 30-minute baseline interview. They then were asked to consent to an opportunity to be randomized to receive a monthly cash gift. The "high-cash gift" treatment group mothers (40% of all mothers) are receiving unconditioned cash payments of $333 per month ($4,000 per year) via debit card for 52 months. Mothers in the "low-cash gift" comparator group (60% of all mothers) are receiving a nominal payment – $20 per month, delivered in the same way and also for 52 months. The 40/60 randomization assignment is within site, but not by hospitals within each of the four sites.

BFY was originally designed to study the effects of monthly unconditional cash transfers on child development for the first three years of life, with the cash gifts set to be distributed for 40 months (3 years, 4 months). In response to the COVID-19 pandemic and the need to postpone in-person research activities, the cash transfers were extended for an additional year, through 52 months (4 years, 4 months), enabling us to postpone in-person direct child assessments to age 4. Interviews conducted at child ages 1, 2 and 3 are providing information about family functioning as well as several maternal reports of developmentally-appropriate measures of children's cognitive and behavioral development. The current analysis plan includes university-based assessments at child age 4.

Conditional on participants' consent and our success in securing agreements with state and county agencies, we are also collecting state and local administrative data regarding parental employment, utilization of public benefits such as Medicaid and Supplemental Nutrition Assistance Programs (SNAP), and any involvement in child protective services. (We have worked with state and local officials to ensure to the extent feasible that our cash gifts are not considered countable income for the purposes of determining benefit levels from social assistance programs.)

The compensation difference between families in the high and low cash gift groups will boost family incomes by $3,760 per year, an amount shown in the economics and developmental

psychology literatures to be associated with socially significant and policy relevant improvements in children's school achievement. After accounting for likely attrition, our total sample size of 800 at the child age 4 follow-up, divided 40/60 between high and low cash gift groups, provides sufficient statistical power to detect meaningful differences in cognitive, emotional and brain functioning, and key dimensions of family context (see below).

At the child age 4 follow-up visit we will administer validated, reliable and developmentally sensitive measures of language, executive functioning and socioemotional skills. We will also collect direct EEG- and ERP-based measures of young children's brain development at age 4. Measures and preregistered hypotheses about them as well as family-based measures are shown in the two tables at the end of this document. Child-focused preregistered hypotheses are presented in Appendix Table 7 and maternal and family focused preregistered hypotheses are presented in Appendix Table 8.

The family process measures that we will gather are based on two theories of change on the impact of income supplementation or receipt of a monthly unconditional cash: that increased investment and reduced stress will facilitate children's healthy development. We are obtaining measures of both of these pathways annually. *Investment pathway*: Additional economic resources enable parents to buy goods and services for their families and children that support cognitive development. These include higher quality housing, nutrition and non-parental child care; more cognitively stimulating home environments and learning opportunities outside of the home; and, by reducing or restructuring work hours, more parental time spent with children. *Stress pathway:* A second pathway is that additional economic resources may reduce parents' own stress and improve their mental health. This may allow parents to devote more positive attention to their children, thus providing a more predictable family life, less conflicted relationships, and warmer and more responsive interactions.

Analysis Plan

Pre-registered Hypotheses. We preregistered hypotheses with clinicaltrials.gov within a month after recruitment began (May, 2018) and in September, 2018, preregistered hypotheses with the Registry of Effectiveness Studies and the AEA RCT Registry. Appendix Tables 1 and 2 detail our original hypothesized impacts. Appendix Tables 3 and 4 incorporate minor changes (mostly made to data collection at age 2, with a few changes to age 3 data collection and no changes to Age 1) to the tables that were originally posted in our pre-registrations. Appendix Tables 5 and 6 incorporate minor changes to reflect the COVID-19 disruptions that impacted data collection at age 2, and altered data collection plans at age 3 and ages 48 months. Appendix Tables 7 and 8 reflect updated hypothesized impacts at ages 48 months.

Hypothesis Testing and Power Analysis. Our key aims are to evaluate the impacts of income supplementation on validated, reliable, and developmentally-sensitive measures of cognitive, language, self-regulation, and socio-emotional functioning at child ages 1 (a small subset of these measures), 2 and 3 (a larger subset), and age 4 (almost all) – this is Aim 1 in our original NICHD application; developmentally-sensitive electroencephalographic-based measures

of brain functioning at child ages 1 and 4 (Aim 2); and family expenditures, food insecurity, housing and neighborhood quality, parent stress and parenting practices, and child care arrangements gathered at child ages 1, 2, 3, and 4 (Aim 3).

All of our pre-registered hypotheses focus on full-sample impacts, although we will also estimate in exploratory analyses moderation of impacts by gender, race/ethnicity (African American, Latino, White), family structure at birth and depth of poverty at birth (income to needs $\leq .5$ or not). Before conducting these main analyses, all measures will be examined for psychometric equivalence across race/ethnicity and whether Spanish or English is a primary language spoken at home and we will compare high and low cash gift groups within site on all baseline characteristics to confirm successful implementation of random assignment.

Our basic empirical approach will use the survey and neuroscience data to compare the pooled cross-city $333/month and $20/month groups on a wide range of family process and child outcome measures. Because of random assignment, the low cash gift group average outcomes enable us to identify the average outcomes corresponding to the counterfactual state that would have occurred for individuals in the high cash gift group if they had not been offered the additional $313/month income supplement. Therefore, differences in outcomes for the high compared with the low group (after random assignment) can be interpreted as estimates of causal treatment effects of the $313/month higher income (regardless of whether treatment-group participants actually expend all of the funds.) These are commonly known as intent-to-treat effects.

Estimation strategy. We illustrate our approach to estimation in a simple regression framework. The "Intent-To-Treat effect" (ITT) is captured by the estimate of the coefficient π_1 in a regression of some child or family process outcome (Y) on a dichotomous indicator for assignment (Z) to the high payment group as in (1).

$$(1)\ Y = Z\pi_1 + X\beta_1 + \varepsilon_1$$

We have experienced extremely low rates of "non-compliance" with the offer of cash gifts paid via the debit cards, with less than 10 of the mothers who consented to allow us to see their transactions never having used the study debit card. We will adjust standard errors using robust variance estimation techniques (Cameron et al. 2008). We will estimate (1) without and then with baseline demographic child and family characteristics (X) to improve the precision of our estimates by accounting for residual variation. These baseline measures, all gathered prior to random assignment, have been checked for adequate variation and sufficient independence from other baseline measures. They include: dummy variables for three of the four sites; mother's age, completed schooling, household income, net worth, general health, mental health, race and Hispanic ethnicity, marital status, number of adults in the mother's household, number of other children born to the mother, whether the mother smoked or drank alcohol during pregnancy and whether the father is currently living with the mother; and child's sex, birth weight, gestational age at birth and birth order.

We will apply our regression estimation strategy to the assessment-based measures of cognitive, language, self-regulation, and socio-emotional functioning and EEG measures of brain activity as outlined in Appendix Tables 7. Further information on the EEG hypotheses and analysis plan is described in the section titled *Age-4 Resting EEG Hypotheses and Analysis Plan* below. To investigate family process impacts, we will apply our estimation strategy to maternal and family measures gathered at child ages 1, 2, 3, and 4 as shown in Appendix Table 8.

Attrition. The greatest threat to internal validity is potential bias from sample attrition overall, within site, and differential attrition rates by treatment status overall and within site. We will carefully track response rates by site, by treatment status across sites, and then treatment status within site. Response rates have been very high at ages 1. Of the original 1,000 recruited participants, we secured interviews with 931 at age 1, and 922 at age 2. We expect similarly high response rates at age 3 and aim for 800 completed cases for the child age 4 university-visit.

We will also conduct sensitivity checks to evaluate whether missing data might be biasing estimates. Most sample attrition that is systematically related to our outcomes of interest (Y) would presumably also be related to the distribution of baseline characteristics (X), and so bias due to sample attrition would be evident if our estimates are sensitive to conditioning on baseline characteristics. Some attrition may be due to time-varying (or unobserved) characteristics, and we can approach this problem in two ways. First, we will examine the sensitivity of our results to worst-case bounds, which enable us to bracket the true effects of our treatment without imposing any assumptions about the unobserved outcomes of participants (Manski, 1989; Manski, 1990; Manski, 1995). A second approach to addressing the problem of missing data will be to use multiple imputation strategies with all available data, (including all survey and administrative data on outcomes and predictor variables). Multiple imputation is an appropriate method if, conditional on all observed information, data are missing at random. Finally, because we have permission to collect administrative data from over 75% of mothers so far, we will be able to compare survey respondents and survey non-respondents on formal earnings and receipt of income from social programs.

Interpretation of parameters. The coefficients obtained in our regression models will be used to quantify the causal effects of the $313/month difference in income supplementation on age-1 and 4 child brain circuitry, cognitive development and socioemotional functioning. We will use the same methods to generate causal impact estimates for the family processes in each of the conceptual pathways. Examining the possible explanatory mechanisms in this way uses a series of separate regression equations to estimate program effects on possible treatment mediators, rather than estimating a structural-equation mediation model, and has been effectively used to infer possible mediation in comparable studies. This approach is preferred because it preserves the experimental variation in income receipt generated by random assignment. The underlying insight is that randomization occurred with respect to receipt of the cash gifts and not on the basis of the proposed pathway mediators. With the potential for multiple mediators, a causal interpretation cannot be given to mediational models without very strong, often implausible, assumptions that there are no unobserved confounds of the association between the

mediator and outcome. Still, the pattern of impacts can yield important insight as to which processes are likely to be present and absent and set the stage for future analyses.

Statistical power. The compensation difference between families in the high- and low-cash gift groups amounts to $313 per month and $16,276 over the course of the 52 months. This amount is in the range of income increases associated with child impacts of around .20 sd in studies of welfare experiments and the EITC (Duncan, Morris & Rodrigues, 2011; Morris, Duncan, Clark-Kauffman, 2005; Dahl & Lochner, 2012). After accounting for likely 20% attrition in the child age-4 follow-up visit, and in the absence of adjustments for sample clustering within hospitals or increased precision owing to the inclusion of baseline covariates in our impact estimates, the sample size of 800 at age 4, divided 40%/60% between high and low payment groups, provides 80% statistical power to detect a .219 sd impact at $p < .05$ in a two-tailed test on cognitive functioning and family processes. The use of baseline covariates in estimation models will improve this power, while the use of bootstrap standard errors will decrease it. Based on exploratory analyses of age-3 cognitive outcomes in the Fragile Families study, we expect that these two offsetting factors will have little net impact on the size of our estimated standard errors.

Multiple comparisons. One strength of our study is the collection of survey, neuroscience lab and administrative data on a wide range of outcomes and explanatory pathways. However, the probability of rejecting a true null hypothesis for at least one outcome is greater than the significance level used for each test. We will address the possibility of false positives while minimizing the reduction in statistical power to detect meaningful effects. Best-practice methods differ across disciplines so we will draw from multiple approaches with the goal of ensuring that results from one approach are consistent with results from others (Romano & Wolfe, 2005; Porter, 2018; Benjamini, 2010; Holm, 1979, Westfall & Young, 1993; Schochet, 2008). Where possible we have aggregated measures used to test our pre-registered hypotheses into indexes. In the case of related measures that cannot be aggregated into a single index, we will estimate the statistical significance of the entire family ("familywise error rate") using stepdown resampling methods in Westfall and Young (1993; Westfall, Tobias, Wolfinger, 2011). Pre-registered clusters of measures are identified with grey bars in appendix tables.

Data release. We aim to release data and documentation from our study to the research community 18 months following the end of each data collection wave to enable independent researchers to pursue replication, mediation, moderation as well as other related analytic questions. All data are being deposited at https://www.icpsr.umich.edu/web/DSDR/studies/37871. Baseline and Age 1 data are currently available.

Age-4 Resting EEG Hypotheses and Analysis Plan. For our primary hypothesis, we will test whether the high-cash gift group has more mid- to high-frequency power than the low cash gift group, we will create a single a single composite measure that aggregated across the portion of the spectrum defined by the three mid-to-high-frequency bands (alpha, beta, and gamma power), from 7-45 Hz. Because this approach is focused on estimating intent-to-treat differences

in a single index score, there is no need for multiple-testing adjustments. Covariates will include all preregistered covariates as well as the number of artifact-free epochs contributed by each participant.

As to secondary hypotheses, consistent with the methods used by another prominent RCT examining an early-life intervention on EEG activity (Debnath, Tang, Zeanah, Nelson, & Fox, 2020; Marshall, Fox, & BEIP Core Group, 2004; Vanderwert, Marshall, Nelson, Zeanah, & Fox, 2010; Vanderwert, Zeanah, Fox, Nelson, & III, 2016), we will explore band-specific and regional effects using mixed-design analyses of variance (mixed-ANOVA). Our secondary hypothesis is that there will be an intervention effect on frontal gamma spectral power between the low-cash gift group and high-cash gift group.

In addition, to explore all regional-frequency effects, we will perform separate mixed-ANCOVAs for each frequency band of absolute and relative power with region (frontal, central, parietal, occipital) as a within-subject factor, and group (low-cash, high-cash) as the between-subjects factor. Greenhouse–Geisser correction will be applied for violations of sphericity. Post hoc comparisons will be performed for significant main effects of group. Any main and interaction effects not involving group will not be followed up. Multiple-adjustment corrections will applied for all post hoc comparisons. Covariates will include all preregistered covariates as well as the number of artifact-free epochs contributed by each participant. Models will be examined with and without preregistered baseline covariates as above, and we will conduct sensitivity checks to evaluate whether missing data might be biasing estimates, as described above.

For Appendix Tables 1 through 6, see Statistical Analysis Plan uploaded in July 2020.

References

Benjamini, Y. (2010). Simultaneous and selective inference: Current successes and future challenges. Biometrical Journal, 52(6), 708–721. https://doi.org/10.1002/bimj.200900299

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. The Review of Economics and Statistics, 90(3), 414-427. https://doi.org/10.1162/rest.90.3.414

Dahl, G. B., & Lochner, L. (2012). The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit. American Economic Review, 102(5), 1927–1956. https://doi.org/10.1257/aer.102.5.1927

Debnath, R., Tang, A., Zeanah, C. H., Nelson, C. A., & Fox, N. A. (2020). The long-term effects of institutional rearing, foster care intervention and disruptions in care on brain electrical activity in adolescence. *Developmental science*, *23*(1), e12872. https://doi.org/10.1111/desc.12872

Duncan, G. J., Morris, P. A., & Rodrigues, C. (2011). Does money really matter? Estimating impacts of family income on young children's achievement with data from random assignment experiments. Developmental Psychology, 47(5), 1263-1279. http://dx.doi.org/10.1037/a0023875

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics, 6(2), 65–70. https://www.jstor.org/stable/4615733

Manski, C. F. (1989). Anatomy of the Selection Problem. The Journal of Human Resources, 24(3), 343–360. https://doi.org/10.2307/145818

Manski, C. F. (1990). Nonparametric Bounds on Treatment Effects. The American Economic Review, 80(2), 319–323. https://www.jstor.org/stable/2006592

Manski, C. F. (1995). Learning about social programs from experiments with random assignment of treatments. Institute for Research on Poverty Discussion Papers 1061-95, University of Wisconsin Institute for Research on Poverty.

Marshall, P. J., Fox, N. A., & BEIP Core Group. (2004). A Comparison of the Electroencephalogram between Institutionalized and Community Children in Romania. *Journal of Cognitive Neuroscience, 16*(8), 1327–1338. https://doi.org/10.1162/0898929042304723

Morris, P. A., Duncan, G. J., & Clark-Kauffman, E. (2005). Child well-being in an era of welfare reform: the sensitivity of transitions in development to policy change. Developmental Psychology, 41(6), 919–932. https://doi.org/10.1037/0012-1649.41.6.919

Porter, K. E. (2018). Statistical Power in Evaluations That Investigate Effects on Multiple Outcomes: A Guide for Researchers. Journal of Research on Educational Effectiveness, 11(2), 267–295. https://doi.org/10.1080/19345747.2017.1342887

Romano, J. P. & Wolf, M. (2005). Stepwise multiple testing as formalized data snooping. Econometrica, 73(4), 1237-1282. https://doi.org/10.1111/j.1468-0262.2005.00615.x

Schochet, P. Z. (2008). Guidelines for multiple testing in impact evaluations of educational interventions. Final report. Princeton, NJ: Mathematica Policy Research, Inc. Retrieved from http://www.eric.ed.gov/ERICWebPortal/detail?accno=ED502199

Vanderwert, R. E., Marshall, P. J., Nelson, C. A. III, Zeanah, C. H., & Fox, N. A. (2010). Timing of intervention affects brain electrical activity in children exposed to severe psychosocial neglect. *PLoS ONE, 5*(7), Article e11415. https://doi.org/10.1371/journal.pone.0011415

Vanderwert, R. E., Zeanah, C. H., Fox, N. A., & Nelson, C. A. (2016). Normalization of EEG activity among previously institutionalized children placed into foster care: A 12-year follow-up of the Bucharest Early Intervention Project. *Developmental Cognitive Neuroscience, 17,* 68–75. https://doi.org/10.1016/j.dcn.2015.12.004

Westfall, P. H., Tobias, R. D., & Wolfinger, R. D. (2011). Multiple comparisons and multiple tests using SAS, second edition. Cary, NC: The SAS Institute.

Westfall, P. H. & Young, S. S. (1993). Resampling-based multiple testing: Examples and methods for p-value adjustment. Hoboken, New Jersey: John Wiley & Sons.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Language Development					
Language Milestones	Squires et al., 2009	sensitivity .86 specifity .85		1	Measured using ASQ- Communication Subscale
Vocabulary*	Fenson, 2002; Jackson-Maldonado, 2012	internal consistency .85		2	Measured by short-form versions of the MacArthur Communicative Development Inventories
	Martin & Brownell, 2011		4		Measured by Receptive One Word Picture Vocabulary Test (ROWPVT)
Maternal concern for language delay	Glascoe, 1997		3		Measured by the sum of the two questions included in the PEDS on expressive language and articulation and receptive language: 1. Do you have any concerns about how your child talks and makes speech sounds? (0: No; 1: Yes or a little) 2. Do you have any concerns about how your child understands what you say? (0: No; 1: Yes or a little)
Executive Function and Behavioral Regulation					
Executive Function	Diamond & Taylor, 1996; Weiland& Yoshikawa, 2013; Bierman et al., 2008			4	Measured by the pencil tap test
Executive Function	Carlson, 2017; Carlson, & Zelazo 2014	MEFS: validity .92 test-retest .93	4		Measured by the Minnesota Executive Function Scale
Socio-Emotional Processing					
Social-Emotional Problems	Briggs-Gowan et al., 2004	internal consistency .65-.79		1, 2	Measured by the Brief Infant–Toddler Social and Emotional Assessment (BITSEA)
Behavior/Emotional Problems	Achenbach et al., 2000	parent report reliability .80	3, 4		Measured by a shortened version of the Child Behavior Checklist measuring the following areas: emotionally reactive, anxious/depressed, attention problems, and aggressive behavior

Social-Emotional Behavior	Roggman et al., 2013; Griffen & Friedman, 2007; Belsky, 2007			1	Measured using NICHD SECCYD parent-child-interaction task coding scheme, with child codes Positive Mood, Negative Mood, Activity Level, Sustained Attention, Positive Engagement at age 1 and agency, negativity, persistence, affection at age 4. (Due to funding limitations, this was not feasible to code, and we have no immediate plans to do so).
Maternal concern for behavioral and social-emotional problems	Glascoe, 1997		3		Measured by the sum of the two questions included in the PEDS on behavior and social-emotional: 1. Do you have any concerns about how your child behaves? (0: No; 1: Yes or a little) 2. Do you have any concerns about how your child gets along with others? (0: No; 1: Yes or a little)
IQ					
IQ*	Wechsler, 2012	internal consistency .88 test-retest reliability .77	4		Measured by the Wechsler Nonverbal Scale of Ability
Pre-Literacy					
Pre-Literacy	Hutton et al., 2019; Hutton et al., 2021			4	Measured by The Reading House
Resting Brain Function					
Age-1 Resting Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2004	n/a		1	Measured by low-density mobile electroencephalography at Age 1: we preregistered group differences in theta, alpha, gamma power.

Age-4 Resting Brain Function	Tomalski et al., 2013; Otero et al., 2013; Marshall et al., 2004	n/a	4	4	Measured by high-density in-lab electroencephalography at Age 4: **Primary**: Because of limitations in power expected with multiple testing adjustments, we are preregistering a single composite of mid-to-high-frequency whole-brain power summing across alpha, beta, and gamma bands, from 7 to 45 Hz. **Secondary**: We hypothesizie greater frontal gamma power in the high-cash gift group, and plan to analyze a full model of regions nested within bands, with the plan to report all exploratory outcomes. See attached analysis plan.
Task-Related Brain Function					
Auditory Discrimination Brain Function*	Choeur et al., 2000; Garcia-Sierra et al., 2011; Kuhl et al., 2005	n/a		4	Measured by mismatch negativity (MMN) ERP with larger differences between standard and deviant stimulus in high-cash gift group compared to the low-cash gift group.
Health: BMI					
Body Mass Index (BMI)	Kuczmarski, 2000	n/a		4	Measured by CDC BMI percentile scales
Health: Physiological Stress					
Physiological Stress	Ursache et al., 2017; Meyer et al., 2014; Davenport et al., 2006	n/a		4	Measured by hair cortisol concentration
Health: Sleep					
Sleep problems	Yu et al., 2012	reliability .9	3	1, 2	Measured by PROMIS Sleep Disturbance- Short Form adapted from ECHO; Additive index of the following items: 1. trouble falling asleep (0: never; 1: almost never; 2: sometimes; 3: almost always) 2. sleeping through night (reverse coded) 3. problem with sleep
Health: Other Indicators					

Overall Health, Medical Care, Diagnosis of Condition or Disability	Child's overall health item source: Idler & Benyamini, 1997 Halim et al., 2013	n/a	3	1, 2	Additive index of the following items*: 1. Child's overall health? (4: excellent, 3: very good, 2: good, 1: fair, or 0: poor) 2.About how many times in the last year did you take child to a doctor because [he/she] was sick? 0-1 times, 2-5 times, 6+ 3. About how many times in the last year did you take child to a doctor because [he/she] was hurt or injured? 4. Did you ever have to take child to the Emergency Room because [he/she] was sick, hurt or injured? (Y/N) 5. How many times ER? 6. Has child been diagnosed with any health condition or disability since birth? (Y/N) *factor analysis of items will be conducted to scale the index
Overall Health, Diagnosis of Health Condition or Disability	Child's overall health item source: Idler & Benyamini, 1997	n/a		4	Additive index of the following items: 1. Child's overall health? (4: excellent, 3: very good, 2: good, 1: fair, or 0: poor) 2.About how many times in the last year was child sick? 0-1 times, 2-3 times, 4-6 times, 7+ 3. Has child been diagnosed with any chronic health condition? (Y/N)
Diagnosis of Developmental Condition					
Diagnosis of Developmental Condition	Study PIs	n/a	4		Has child been diagnosed with any developmental condition, like speech delay, autism, or ADHD? (Y/N)
Child Epigenetic Pace of Aging					
Methylation pace of aging	Belsky et al., 2020; Belsky et al., 2022	n/a		4	Methylation pace of aging was developed from DNA-methylation analysis of Pace of Aging in the Dunedin Study birth cohort. Pace of Aging is a composite phenotype derived from analysis of longitudinal change in 18 biomarkers of organ-system integrity (Belsky et al., 2015). In contrast, so-called epigenetic clocks are trained on chronological age. Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. The second iteration (DunedinPACE) takes into account an additional measurement occasion (collected 20 years after inclusion) and only includes the most reliable DNA methylation probes, i.e. probes with little variation between technical replicates .

Child DNA Methylation					
DNA methylation	McCartney et al, 2022	n/a		4	Salivary DNA-methylation profiles of cognitive functioning, i.e., "Epigenetic-g", can be computed on the basis of weights from a blood-based epigenome wide association study of general cognitive functions (g) in adults (McCartney et al., 2022). General cognitive ability was derived from the first unrotated principal component of logical memory, verbal fluency and digit symbol tests, and vocabulary. Epigenetic-g is conceptually distinct from biological aging.
Child Nutrition					
Consumption of healthy foods	Los Angeles County WIC Survey, 2017			2	Additive index of the number of times per day consumed the following items*: 1. eat fruits 2. eat vegetables
Consumption of unhealthy foods	Los Angeles County WIC Survey, 2017			2	Additive index of the number of times per day consumed the following items*: 1. juice, soda, chocolate milk or other sweet drinks 2. eat sweets
Any Maternal Concern for Developmental Delay					
Parents' Evaluation of Developmental Status (PEDS)	Glascoe, 1997			3	Measured by the total score across categories of components of the PEDS, which includes 10 survey items.
Total "predictive concerns" in the PEDS	Glascoe, 1997			3	Measured by the total number of maternal-reported concerns that are "predictive of developmental delay" in the PEDS
School Achievement & Behavior					
School test scores for target children and siblings	Administrative data	n/a	School age (target child)	School age (siblings)	

Student behavioral data for target children and siblings	Administrative data	n/a	School age (target child and siblings)

Notes. Previous versions of this table specified that "All measures between grey lines will be subject to multiple testing adjustments". This is now changed to be "All measures between grey lines measured *in the same wave* will be subject to multiple testing adjustments".

The previous version of this table refered to "waves" of data collection. For clarity, we have replaced "wave" with "age", with both referring to the age of the baby at planned data Minor, non-substantive changes may be made to the wording of specific items across data collection years.

Due to COVID-19, the age 3 data collection wave is in the form of a phone survey. Thus, sub-domains that were supposed to be measured in-person at ages 2 or age 3 are being postponed to age 4. These domains include: epigenetic age, DNA methylation, BMI, physiological stress, self-regulation, executive function, social-emotional behavior, IQ; resting brain function, auditory discrimination brain function. The sub-domain of child vocalizations was not measured in-person at age 2 (due to COVID-19) and is not being measured at
*Indicates that the sub-domain was called something different in previous versions of this table. The changes are listed below:
-Previously "Communicative Development (Vocabulary)"; presently "Vocabulary".
-Previously "Intelligence; presently "IQ".
-Previously "Language Related Brain Function"; presently "Auditory Discrimination Brain Function".

Domains and sub-domains that were not previously included in this table for pre-registration at age 3 and were added include: Any Maternal Concern for Developmental Delay (domain); Maternal Concern for Behavioral and Social-Emotional Problems (sub-domain); Maternal Concern for Language Delay (sub-domain); Maternal "Predictive Concern" for Language Delay (sub-domain).

Language Development	
Language Milestones	Squires, J., Bricker, D. D., & Twombly, E. (2009). Ages & stages questionnaires. Baltimore, MD: Paul H. Brookes.
Language Processing	Golinkoff, R. M., De Villiers, J. G., Hirsh-Pasek, K., Iglesias, A., Wilson, M. S., Morini, G., & Brezack, N. (2017). *User's Manual for the Quick Interactive Language Screener (QUILS): A Measure of Vocabulary, Syntax, and Language Acquisition Skills in Young Children* . Paul H. Brookes Publishing Company
Vocabulary*	Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Jackson-Maldonado, Donna, Virginia A. Marchman, and Lia C. H. Fernald. 2012. "Short-Form Versions of the Spanish MacArthur-Bates Communicative Development Inventories." Applied Psycholinguistics 34 (4): 837–68. Martin, N. A., & Brownell, R. (2011). ROWPVT-4: Receptive One-Word Picture Vocabulary Test.
Maternal concern for language delay	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.
Executive Function & Self-Regulation	
	Diamond, A., & Taylor, C. (1996). Development of an aspect of executive control: development of the abilities to remember what I said and to "do as I say, not as I do". *Developmental psychobiology* , *29* (4), 315–334. https://doi.org/10.1002/(SICI)1098-2302(199605)29:4<315::AID-DEV2>3.0.CO;2-T
Executive Function - Pencil Tap	Bierman, K. L., Nix, R. L., Greenberg, M. T., Blair, C., & Domitrovich, C. E. (2008). Executive functions and school readiness intervention: impact, moderation, and mediation in the Head Start REDI program. *Development and psychopathology* , *20* (3), 821–843. https://doi.org/10.1017/S0954579408000394 Weiland, C. and Yoshikawa, H. (2013), Impacts of a Prekindergarten Program on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills. Child Dev, 84: 2112-2130. https://doi.org/10.1111/cdev.12099
Executive Function - MEFS	Carlson, S. M., & Zelazo, P. D. (2014). Minnesota Executive Function Scale: Test Manual. St. Paul, MN: Reflection Sciences, Inc. Carlson, S. M. (2017). *Minnesota Executive Function Scale* : *Technical Report, v. 2* . St. Paul, MN: Reflection Sciences, Inc.
Socio-Emotional Processing	

Social-Emotional Problems	Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant-Toddler Social and Emotional Assessment: screening for social-emotional problems and delays in competence. *Journal of pediatric psychology*, *29*(2), 143-155.
Behavior/Emotional Problems	Achenbach, T. M., & Ruffle, T. M. (2000). The Child Behavior Checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatrics in* review, 21(8), 265-271.
Social-Emotional Behavior^	Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V., & Christiansen, K. (2013). Parenting interactions with children: Checklist of observations linked to outcomes (PICCOLO) in diverse ethnic groups. Infant Mental Health Journal, 34(4), 290-306. Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., Owen, M. T., & NICHD Early Child Care Research Network. (2007). Are there long-term effects of early child care?. *Child development*, *78*(2), 681-701. Griffin, J. A., & Friedman, S. L. (2007). NICHD Study of Early Childcare and Youth Development. National Institute of Health. Adapted script from mother-child-interaction at 15 months.
Maternal concern for behavioral and social-emotional problems	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.
IQ	
IQ*	Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence—fourth edition*. San Antonio, TX: The Psychological Corporation.
Pre-Literacy	
Pre-Literacy	John S. Hutton, Laura Justice, Guixia Huang, Amy Kerr, Thomas DeWitt, Richard F. Ittenbach; The Reading House: A Children's Book for Emergent Literacy Screening During Well-Child Visits. Pediatrics June 2019; 143 (6): e20183843. 10.1542/peds.2018-3843 Hutton, J. S., Dudley, J., Huang, G., Horowitz-Kraus, T., DeWitt, T., Ittenbach, R. F., & Holland, S. K. (2021). Validation of *The Reading House* and Association With Cortical Thickness. *Pediatrics*, *147*(3), e20201641. https://doi.org/10.1542/peds.2020-1641
Resting Brain Function	
Age-1 and Age-4 Resting Brain Function	Tomalski, P., Moore, D. G., Ribeiro, H., Axelsson, E. L., Murphy, E., Karmiloff-Smith, A., ... & Kushnerenko, E. (2013). Socioeconomic status and functional brain development–associations in early infancy. *Developmental Science*, *16*(5), 676-687. Otero, G. A., Pliego-Rivero, F. B., Fernández, T., & Ricardo, J. E. E. G. (2003). EEG development in children with sociocultural disadvantages: a follow-up study. Clinical neurophysiology, 114(10), 1918-1925.

	Marshall, P. J., Fox, N. A., & Group, B. C. (2004). A comparison of the electroencephalogram between institutionalized and community children in Romania. Journal of Cognitive Neuroscience, 16(8), 1327-1338.
Task-Related Brain Function	
Auditory Discrimination Brain Function*	Cheour, M., Leppänen, P. H., & Kraus, N. (2000). Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clinical neurophysiology*, *111* (1), 4-16. Garcia-Sierra, A., Rivera-Gaxiola, M., Percaccio, C. R., Conboy, B. T., Romo, H., Klarman, L., ... & Kuhl, P. K. (2011). Bilingual language learning: An ERP study relating early brain responses to speech, language input, and later word production. Journal of Phonetics, 39(4), 546-557. Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. Developmental science, 8(1), F1-F12.
Health: BMI	
Body Mass Index (BMI)	Kuczmarski, R. J. (2000). CDC growth charts; United States.
Health: Physiological Stress	
Physiological Stress	Ursache, A., Merz, E. C., Melvin, S., Meyer, J., & Noble, K. G. (2017). Socioeconomic status, hair cortisol and internalizing symptoms in parents and children. *Psychoneuroendocrinology*, *78*, 142-150. Meyer, J., Novak, M., Hamel, A., & Rosenberg, K. (2014). Extraction and analysis of cortisol from human and monkey hair. Davenport, M. D., Tiefenbacher, S., Lutz, C. K., Novak, M. A., & Meyer, J. S. (2006). Analysis of endogenous cortisol
Health: Sleep	
Sleep problems	Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., ... & Pilkonis, P. A. (2012). Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. Behavioral sleep medicine, 10(1), 6-24.
Health: Other Indicators	
Overall Health, Medical Care, Diagnosis of Condition or Disability	Halim, M. L., Yoshikawa, H., & Amodio, D. M. (2013). Cross-generational effects of discrimination among immigrant mothers: Perceived discrimination predicts child's healthcare visits for illness. *Health Psychology*, *32* (2), 203.

Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior* , 21-37

Diagnosis of Developmental Condition	
Diagnosis of Developmental Condition	Study PIs
Child Epigenetic Pace of Aging	
Methylation pace of aging	Belsky, W. D. et al. (2020). Quantification of the pace of biological aging in humans through blood test, the DunedinPoAm DNA methylation algorithm. eLife 9:e54870. https://doi.org/10.7554/eLife.54870 Belsky, W. D. et al. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. eLife 11:e73420. https://doi.org/10.7554/eLife.73420
Child DNA Methylation	
DNA methylation	McCartney, D.L., Hillary, R.F., Conole, E.L.S. *et al.* Blood-based epigenome-wide analyses of cognitive abilities. *Genome Biol* **23,** 26 (2022). https://doi.org/10.1186/s13059-021-02596-5
Child Nutrition	
Consumption of healthy foods	Los Angeles County WIC Survey. (2017). Retrievable from: http://lawicdata.org/wp-content/uploads/2014/09/WIC-Parents-Quex-English-FINAL.pdf
Consumption of unhealthy foods	Los Angeles County WIC Survey. (2017). Retrievable from: http://lawicdata.org/wp-content/uploads/2014/09/WIC-Parents-Quex-English-FINAL.pdf
Any Maternal Concern for Developmental Delay	
Parents' Evaluation of Developmental Status (PEDS)	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.
Total "predictive concerns" in the PEDS	Glascoe FP. Parents' Evaluations of Developmental Status: A Method for Detecting and Addressing Developmental and Behavioral Problems in Children. Nashville, TN: Ellsworth & Vandermeer Press, 1997.

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Household Economic Hardship					
Index of economic stress	MTO; Kling, Liebman, Katz, 2007			1, 2, 3, 4	Additive index of dichotomous variables (higher score=more stress): 1. worried about expenses? (0: occasionally or never; 1: frequently or more) 2. whether spent more than income? (0: no; 1: yes) 3. missed rent or mortgage (0 if homeless or not missed; 1 if missed rent or mortgage) 4. Set aside rainy day funds for 1 mo (0: Yes 1: No) 5. Ability to cover expenses for 1 mo with loss of income (0: Yes; 1: No) 6. in past 12 mos, missed payments for water, gas, oil, electricity? (0: no or not applicable; 1: yes) 7. in past 12 mos, gas, water, electricity ever shut off? (0: no; 1: yes) 8. Since child's birth, have you ever been evicted or forced to leave? (0: No; 1: Yes).* 9. needed medical or dental care and did not get it? (0=no; 1=yes) *changes to "in the past 12 months" for surveys at ages 2 through 4
Household Poverty rate	US Census Bureau			1, 2, 3, 4	Measured using the Census Bureau's poverty thresholds by size of family and number of children
Index of food insecurity*	Economic Research Service, USDA, 2012			1, 2, 3, 4	Additive index of 5 dichotomized items (higher score=more food insecurity): 1. Food didn't last, no $ for more (0: Never true, 1: sometimes or often true) 2. Can't afford balanced meals (0: Never true, 1: sometimes or often true) 3. Cut size or skip means (0: No; 1: Yes) 4. Eat less than should (0:No; 1: Yes) 5. Hungry^{+} (0:No; 1: Yes)
Social Services Receipt					
Number of Benefits received by mother	Study PIs			1, 2, 3	Additive index of dichotomized items (higher score=more benefits received): 1. Food stamps SNAP (0: not currently receiving; 1: currently receiving) 2. Free or reduced childcare* 3. Early Head Start or HS* 4. Women, Infants and Children (WIC) 5. State Unemployment 6. Cash assistance/TANF* 7. Medicaid coverage for self 8. Housing assistance 10. LIHEAP / heat/AC assistance* *Indicates benefits that are not being asked about at age 3.
Mother's Labor Market and Education Participation					

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Time to labor market reentry from birth	Current Population Survey			1	Continuous outcome: # of months until mom's reentry into labor market from birth of child derived from the following items: 1. did you ever work for pay since child's birth? 2. in what months did you work for pay?
Time to full-time labor market reentry from birth	Current Population Survey			1	Continuous outcome: # of months until mom's full-time reentry into labor market from birth of child derived from the following items: 1. did you ever work full time since child's birth? 2. in what months did you work full time?
Mother's education and training attainment	Current Population Survey			1, 2, 3	Dichotomous variable indicating that mother participated in education and/or job training activities since birth* *changes to "in the past 12 months" for surveys at ages 2 and 3
Mother's Labor Market Participation				4	Dichotomous variable indicating whether mother is participating in the labor market using the item "do you currently work for pay?"
Maternal Earnings	PSID			4	Mother's Earnings in the previous calendar year
Child-Focused Expenditures					
Index of child-focused expenditures (since birth)	Lugo-Gil, Yoshikowa, 2006			1	Additive index of the following dichotomous items (higher score=more purchased): Since child's birth, purchased… 1. Crib? 2. Car seat? 3. High chair? 4. Safety covers for outlets? 5. Latches for cabinets? 6. Gate? 7. Smoke detector? 8. books (yes/no)?
Index of child-focused expenditures (in past 30 days)	Lugo-Gil, Yoshikowa, 2006			1, 2, 3, 4	Continuous dollar amount of age-relevant items*: Past 30 days, total $ amount spent on... 1. books 2. toys 3. clothes 4. diapers 5. videos for age 1; 1. books 2. toys 3. clothes 4. activities 5. videos for ages 2, 3, and 4
Cost of paid child care	National Study of Early Care and Education			1, 2, 3, 4	Out of pocket spending on child care last week. 1. altogether, about how much money did you spend out-of-pocket on all of [CHILDNAMEF]'s child care arrangements last week?
Use of center-based care	National Study of Early Care			1	1. Has child spent any time in childcare or day care? (Y/N)

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
[illegible]	of Early Care and Education			2, 3, 4	1. Has child spent 5 or more hours in a child care or day care center last week? (Y/N)
Housing and Neighborhoods					
Index of perceptions of neighborhood safety	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	Additive index of two items (higher score=feels more safe). 1. how safe during day? (3: very safe, 2: safe, 1: unsafe, 0: very unsafe) 2. how safe during night? (3: very safe, 2: safe, 1: unsafe, 0: very unsafe)
Index of housing quality	MTO; Kling, Liebman, Katz, 2007			1	Additive index of 7 items (higher score=higher quality): 1. Bad walls (0: big problem; 1: small problem; 2: not problem) 2. bad plumbing 3. rodents 4. cockroaches 5. bad windows 6. bad heat 7. overall condition (3: excellent, 2: good 1: fair, 0: poor)
				2	Additive index of 8 items (higher score=higher quality): 1. Bad walls (0: big problem; 1: small problem; 2: not problem) 2. bad plumbing 3. rodents 4. cockroaches 5. bad windows 6. bad heat 7. bad air condition 8. bad locks~ 9. overall condition (3: excellent, 2: good 1: fair, 0: poor)
Homelessness	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	Additive index of two dichotomized items (higher score=more homelessness): 1. Since child's birth, have you been homeless?* (0: Yes; 1: No) 2. Since birth, have you been in a group shelter?* (0: Yes; 1: No) *changes to "in the past 12 months" for surveys at ages 2 and 3
				4	Dichotomous indicator of whether mom experienced "homelessness, eviction, or sudden loss of housing in the past 12 months".
Excessive Residential mobility	MTO; Kling, Liebman, Katz, 2007			1, 2, 3	Moved three or more times since birth of baby* (Y/N) *changes to "in the last 12 months" for surveys at ages 2 and 3
Neighborhood poverty	Kling, Liebman, Katz, 2007			1, 2, 3, 4	# of residents below poverty line in census tract divided by total number of residents in census tract
Family and Maternal Perceived Stress					

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Perceived stress	Cohen et al., 1994, 1983	alpha: .86		1, 2	Perceived Stress Scale (PSS): additive index of 9 items (0: never; 1: almost never; 2: sometimes; 3: fairly often; 4: very often) 1. upset because of something unexpected 2. felt unable to control important life things 3. felt nervous and stressed 4. confident in ability to handle personal probs (reverse coded - rc) 5. couldn't cope with all things to do 6. control of irritations in life (rc) 7. "on top of things" (rc) 8. angered bc of things outside control 9. could not overcome difficulties
				3	Perceived Stress Scale (PSS): additive index of 10 items (0: never; 1: almost never; 2: sometimes; 3: fairly often; 4: very often) 1. upset because of something unexpected 2. felt unable to control important life things 3. felt nervous and stressed 4. confident in ability to handle personal probs (reverse coded - rc) 5. couldn't cope with all things to do 6. control of irritations in life (rc) 7. "on top of things" (rc) 8. angered bc of things outside control 9. could not overcome difficulties 10. felt things were going "your way" (rc)~
Parenting stress	Items 1-4: Project GAIN Items 5-7: PSID-Child Development Supplement			1, 2, 4	Aggravation in Parenting Scale: additive index of 7 items (0: Strongly agree-5: Strongly disagree): 1. confidence in parenting abilities 2. feels good about parenting abilities 3. thinks good parent 4. kids will say she was wonderful 5. giving up more for kids than ever expected 6. feels trapped (rc) 7. unable to do different things bc of kids (rc)
Maternal Happiness and Optimism					
Global happiness	The General Social Survey from NORC			1, 2, 3	One-item with 3-point response scale"Taken altogether, how happy are you these adys?" (0: not happy; 1: pretty happy; 2: very happy)

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Maternal Agency	Snyder et al., 1991	alpha: .86 test-retest: .81		1, 2, 3	HOPE Scale: additive index of 8 items with 5-point response scale (0: definitely false; 5: definitely true) 1.think of ways to get out of a jam 2.. energetic pursuit of goals 3. lot of ways around any problem 4. ways to get what's important 5. solves problems 6. past has prepared me for future 7. pretty successful in life 8. meets goals set for oneself
Maternal Physiological Stress					
Maternal hair cortisol	Ursache et al., 2017			1, 4	Measured using a sample of hair that is >=15mg in weight and ~3cm long; analyzed with sensitive and specific enzyme-linked immunosorbent assay; assay readout converted to pg cortisol per mg dry hair weight
Maternal Mental Resources					
Maternal cognitive resources	Carlson, 2017; Carlson, & Zelazo 2014			4	Minnesota Executive Function Scale
Maternal Mental Health					
Index of maternal depression	Kroenke & Spitzer, 2002			1, 2, 3, 4	PHQ-8: additive index of 8 items (0: not at all; 1: several days; 2: more than half of days; 3: every day) 1. little interest or pleasure doing things 2. feeling down, depressed, hopeless 3. trouble sleeping or sleep too much 4. feel tired and no energy 5. poor appetite or overeating 6. feel like a failure 7. trouble concentrating 8. moving slowly or fidgety
Index of maternal anxiety	Steer & Beck, 1997	alpha: .92 test-retest: .75		1, 3	Beck Anxiety Inventory: additive index of 21 common anxiety symptom items (0: not at all; 1: mildly; 2: moderately; 3: severely bothersome)
	Spitzer et al., 2006	alpha: .92 test-retest: .83		2, 3, 4	GAD-7: additive index of 7 items (0: not at all; 1:several days; 2: more than half the days; 3: nealy every day)
Maternal Substance abuse[×]					

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Alcohol and cigarette use	MTO; Kling, Liebman, Katz, 2007			1, 3	Additive index of the following items (0: never in last year; 1: less than 1x per month; 2: several times per month; 3: several times per week; 4: everyday): 1. How often do you smoke cigarettes? 2. How often drink alcohol?
Opioid use	MTO; Kling, Liebman, Katz, 2007			1, 3	Number of times of opioid use in the past year (0: never in last year; 1: less than 1x per month; 2: several times per month; 3: several times per week; 4: everyday):
Chaos in Home					
Index of chaos in the home	Evans et al., 2005	alpha: .77 test-retest: .93		1, 2	Home Environment Chaos Scale: additive index of 20 items (higher score=more chaos): (0: not true; 1: true) 1. can find things (reverse coded - rc) 2. little commotion in home (rc) 3. always rushed 4. can "stay on top of things" (rc) 5. always late 6. "zoo" in home 7. can talk wo interruption (rc) 8. always a fuss 9. family plans don't work out 10.can't hear oneself think at home 11. drawn into others' arguments 12. can relax at home (rc) 13. phone takes up a lot of time 14. atmosphere is calm at home (rc) 15. regular morning routine (rc) 16. eat together during daily (rc) 17. evening routine with child (rc) 18. regular late afternoon routine with child (rc) 19. child goes to bed at regular time (rc) 20. set aside for talking with child daily (rc)
Maternal Relationships[×]					
Physical Abuse	Fragile Families and Child			1,2	1. Ever abused? (1: yes; 0: no)
Frequency of Arguing				1,2	1. How often argue about things that are important to you? (1: never; 2: almost never; 2: sometimes; 3: fairly often; 4: very often)

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Relationship quality	Wellbeing Study			1	Additive index of the following items (higher score=higher qual rel) 1. Partner fair and willing to compromise? (3: Often; 2: sometimes; 1: never) 2. partner expressed affection or love? (3: Often; 2: sometimes; 1: never) 3. partner insulted or criticized you or your ideas (0: Often; 1: sometimes; 2: never) 4. partner made you feel down or bad about yourself during an argument? (0: Often; 1: sometimes; 2: never) 5. partner encouraged or helped you to do things that were important to you? (2: Often; 1: sometimes; 0: never) 6. partner isolated you? (0: Often; 1: sometimes; 2: never) 7. partner hurt you physically (0: Often; 1: sometimes; 2: never) 8. partner sexually abused you? (0: Often; 1: sometimes; 2: never) 9. partner listened to you? (3: Often; 2: sometimes; 1: never) 10. partner made you feel afraid? (0: Often; 1: sometimes; 2: never) 11. partner threatened or hurt your child/children?+ (0: Often; 1: sometimes; 2: never)
				2, 3, 4	Dichotomous indicator of current or recent relationship quality, where poor quality is defined as 1 if the mother is in a relationship and has a score of 26 or below on the relationship quality scale (approximately the bottom tercile of the low cash gift group distribution of scores) and a 0 either if the mother is not in a relationship or is in a relationship and has a relationship quality index score of 27 or above (approximately in the top two terciles of the distribution).
Maternal Physical Health					
Global health	Idler & Benyamini, 1997			1, 2	One item with 5-point response scale "overall, how would you describe your health…" (0: excellent-5: poor)
Sleep	Yu et al., 2012			1, 3	Additive index of the following items (higher score=higher qual sleep): 1. Quality of sleep (0: very poor-5: very good) 2. Difficulty falling asleep (0: not atll; 5: very much) (rc) 3. Felt tired (0: not at all-5: very much) (rc)
Mother's BMI	CDC scales			4	Measured by CDC BMI percentile scales
Parent-Child Interaction Quality					
Adult word count	Xu et al (2009), LENA foundation			1	Measured using LENA processing software
Conversational turns	Xu et al (2009), LENA foundation			1	Measured using LENA processing software

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
Index of mother's positive parenting behaviors	Roggman, et al., 2013; Griffen & Friedman, 2007; Belsky, et al., 2007	inter-rater reliability varies by domain: .69-.80; alpha: .78		1, 4	Measured using PICCOLO coding of parenting behaviors from three sub-scales (affection, responsiveness, encouragement and teaching) with responses ranging from 0: absent, 1: barely, 2: clearly. Parent child interaction task and script adapted from the NICHD Study of Early Child Care and Youth Development.
Epigenetic Pace of Aging					
Methylation pace of aging	Belsky et al., 2020; Belsky et al., 2022			4	Methylation pace of aging was developed from DNA-methylation analysis of Pace of Aging in the Dunedin Study birth cohort. Pace of Aging is a composite phenotype derived from analysis of longitudinal change in 18 biomarkers of organ-system integrity (Belsky et al., 2015). In contrast, so-called epigenetic clocks are trained on chronological age. Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. The second iteration (DunedinPACE) takes into account an additional measurement occasion (collected 20 years after inclusion) and only includes the most reliable DNA methylation probes, i.e. probes with little variation between technical replicates.
Maternal DNA Methylation					
DNA methylation	McCartney et al, 2022			4	Salivary DNA-methylation profiles of cognitive functioning, i.e., "Epigenetic-g", can be computed on the basis of weights from a blood-based epigenome wide association study of general cognitive functions (g) in adults (McCartney et al., 2022). General cognitive ability was derived from the first unrotated principal component of logical memory, verbal fluency and digit symbol tests, and vocabulary. Epigenetic-g is conceptually distinct from biological aging.
Frequency of Parent Child Activity					
Self-Report of Parent-child activities	Rodriguez & Tamis-LeMonda, 2011			1	Additive index of 4 items with response scale (higher score=higher frequency of activities): 1. read books (0: rarely or never; 1: a few times/month; 2: a few times/week ; 4:everyday) 2. tell stories 3. play together 4. play groups

Appendix Table 8: Maternal and Family Focused Preregistered Hypotheses

Domains (in gray) and sub-domains	Measure/Item source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures (All measures between grey lines measured during the same wave will be subject to multiple testing adjustments)
				2, 3	Additive index of 5 items with response scale (higher score=higher frequency of activities): 1. read books (0: rarely or never; 1: a few times/month; 2: a few times/week ; 4:everyday) 2. tell stories 3. play together 4. play groups (not asked at age 3 due to COVID) 5. play pretend games
Child meal and sleep routine index	Study PIs			4	Additive index of 2 survey items (higher score=more routines): 1. eat meals together (0: 0-2 days; 1: 3+ days) 2. had regular bedtime (0: no; 1: yes)
Time on mother-focal child activities	Rodriguez & Tamis-LeMond a, 2011			4	Additive index of activities where the number of days reported doing the activity are multiplied by the number of minutes on a given day. Example below: 1. How many days did you participate in [activity]? (0: no days; 1.5: 0-1 days; 4: 3-5 days; 6.5: 6-7 days) 1a. On those days, how many minutes do you do [activity]? (2: 4 minutes or less; 7.5: 5-10 minutes; 15.5: 11-20 minutes; 25.5: 21-30 minutes; 35: more than 30 minutes).
Maternal Discipline[×]					
Spanking discipline strategy	Reichman et al., 2001			1, 2, 3	Dichotomous indicator using the following item: 1. In past month, have you spanked child due to misbehavior (1: yes; 2:no)

Notes. The previous version of this table refered to "waves" of data collection. For clarity, we have replaced "wave" with "age", with both referring to the age of the baby at planned data collection. Minor, non-substantive changes may be made to the wording of specific items across data collection years.

+ indicates that items were omitted or programmed incorrectly in the age 1 survey administered to mothers and cannot be used to calculate outcomes. These include item 5 from the index of food insufficiency ("hungry"), and item 11 from the relationship quality index ("partner threatened or hurt your child/children? ").

[×]indicates outcomes that were not administered at age 1 once in-person interviews switched to phone interviews due to COVID-19 .

˜Indicates that item was omitted from previous pre-registrations but was administered to mothers and is being included in the outcome analyses.

*Indicates that the sub-domain was called something different in previous versions of this table. The sub-domain "Food Insecurity" was previously referred to as "Food Insufficiency" .

Due to COVID-19, the age 2 and age 3 data collection wave is in the form of a phone survey. Thus, sub-domains that were supposed to be measured in-person at ages 2 or age 3 are being postponed to ages 45-48 months. These domains include: index of mother's positive parenting behaviors, epigenetic age, DNA methylation, BMI, physiological stress, cognitive resources. Additionally, sub-domains that we had not intended to include in pre-registeration at age 3 have been been added to the phone survey at age 3 and to the pre-registration table. These include: self-report of parent-child activities,

Certain sub-domains were pre-registered at age 3 and are no longer preregistered because they are not being included in the age 3 data collection (due to time constraints). These include: global health, physical abuse, index of chaos in the home, parenting stress, index of housing quality.

Measure description	Bibliography	
Preregistered measures	**Source 1**	**Source 2**
Household Economic Hardship		
Index of economic stress **Index of food insecurity** **Household poverty rate**	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119. https://www.ers.usda.gov/media/8282/short2012.pdf Fontenot, Kayla, Jessica Semega, and Melissa Kollar, U.S. Census Bureau, Current Population Reports, P60-263, Income and Poverty in the United States: 2017, U.S. Government Printing Office, Washington, DC, 2018.	http://www2.nber.org/mtopublic/
Social Services Receipt		
Number of Benefits received by mother	Study PIs	
Mother's Labor Market and Education Participation		
Time to labor market reentry from birth **Time to full-time labor market reentry from birth** **Mother's education and training attainment**	Current Population Survey, retrieved from: https://www.census.gov/programs-surveys/cps/technical-documentation/questionnaires.html	
Mother's Earnings	Panel Study of Income Dynamics https://psidonline.isr.umich.edu/	
Child-Focused Expenditures		
Index of child-focused expenditures	Lugo-Gil, J., Yoshikawa, H. (2006). Assessing expenditures on children in low-income, ethnically diverse, and immigrant families. National Poverty Center Working Paper Series, 06-36.	
Child-focused expenditures **Cost of paid child care** **Use of center-based care**	National Study of Early Care and Education	
Housing and Neighborhoods		
Index of perceptions of neighborhood safety **Index of housing quality** **Residential mobility** **Homelessness** **Neighborhood poverty**	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.	
Family and Maternal Perceived Stress		
Perceived stress	Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. Measuring stress: A guide for health and social scientists.	Cohen, S., Kamarck, T., Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior,* 24(4), 385-396.

Parenting stress	PSID-CDS Aggravation in Parenting Scale https://psidonline.isr.umich.edu/cds/cdsi_usergd.pdf for items 5-7 5. giving up more for kids than ever expected 6. feels trapped (rc) 7. unable to do different things bc of kids (rc)	Project GAIN (Gaining Access to Income Now) https://preventionboard.wi.gov/Pages/OurWork/ProjectGAIN.aspx for items 1-4 1. confidence in parenting abilities 2. feels good about parenting abilities 3. thinks good parent 4. kids will say she was wonderful
Maternal Happiness and Optimism		
Global happiness	The General Social Survey from NORC at the University of Chicago, retrieved from: http://gss.norc.org/Get-Documentation/questionnaires	
Maternal Agency	Snyder, C.R., Harris, C., Anderson, J.R., Holleran, S.A., Irving, L.M., Sigmon, S.T., Yoshinobu, L., Gibb, J., Langelle, C., Harney, P. (1991). The will and the ways: development and vaildation of an individual-differences measure of hope. *Journal of Personality and Social Psychology, 60*(4), 570-585.	
Maternal Epigenetic Pace of Aging		
Epigenetic age	Belsky, W. D. et al. (2020). Quantification of the pace of biological aging in humans through blood test, the DunedinPoAm DNA methylation algorithm. eLife 9:e54870. https://doi.org/10.7554/eLife.54870	Belsky, W. D. et al. (2022). DunedinPACE, a DNA methylation biomarker of the pace of aging. eLife 11:e73420. https://doi.org/10.7554/eLife.73420
Maternal DNA Methylation		
DNA methylation	McCartney, D.L., Hillary, R.F., Conole, E.L.S. *et al.* Blood-based epigenome-wide analyses of cognitive abilities. *Genome Biol* **23,** 26 (2022). https://doi.org/10.1186/s13059-021-02596-5	
Maternal Physiological Stress		
Maternal hair cortisol	Ursache, A., Merz, E.C., Melvin, S., Meyer, J., Noble, K.G. (2017). Socioeconomic status, hair cortisol and internalizing symptoms in parents and children. *Psychoneuroendocrinology*, 78, 142-150.	
Maternal Mental Resources		

Maternal cognitive resources	Carlson, S. M., & Zelazo, P. D. (2014). Minnesota Executive Function Scale: Test Manual. St. Paul, MN: Reflection Sciences, Inc.	Carlson, S. M. (2017). *Minnesota Executive Function Scale*: *Technical Report, v. 2*. St. Paul, MN: Reflection Sciences, Inc.
Maternal Mental Health		
Index of maternal depression	Kroenke, K. & Spitzer, R.L. (2002). The PHQ-9: a new depression diagnostic and severity measure. Psychiatric annals, 32(9), 509-515.	
Index of maternal anxiety	Steer, R.A. & Beck, A.T., (1997). Beck Anxiety Inventory. In C.P. Zalaquett & R.J. Wood (Eds), Evaluating stress: A book of resources (pp. 23-40). Lanham, MD, US: Scarecrow Education	
Index of maternal anxiety	Spitzer RL, Kroenke K, Williams JBW, Löwe B. A Brief Measure for Assessing Generalized Anxiety Disorder: The GAD-7. *Arch Intern Med.* 2006;166(10):1092–1097. doi:10.1001/archinte.166.10.1092	
Maternal Physical Health		
Global health	Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: a review of twenty-seven community studies. *Journal of health and social behavior*, 21-37.	
Sleep	Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., ... & Pilkonis, P. A. (2012). Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. Behavioral sleep medicine, 10(1), 6-24.	
Mother's BMI	Kuczmarski, R. J. (2000). CDC growth charts; United States.	
Maternal Substance abuse		
Alcohol and cigarette use	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.	
Opioid use		
Chaos in Home		
Index of chaos in the home	Evans, G.W., Gonnella, C., Marcynyszyn, L.A., Gentile, L, & Salpekar, N. (2005). The role of chaos in poverty and children's socioemotional adjustment. *Psychological Science*, 16(7), 560-565.	
Maternal Relationships		
Physical Abuse	User's Guide for the Fragile Families and Child Wellbeing Study Public Data, Year 3. (2018). Retrieved from: https://fragilefamilies.princeton.edu/sites/fragilefamilies/files/year_3_guide.pdf#page=84	
Frequency of Arguing		
Relationship quality		
Parent-Child Interaction Quality		
Adult word count	Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA Language Environment Analysis System in young children's natural home environment. *LENA Foundation*	

Conversational turns	environment. *LENA Foundation*.	
Index of mother's positive parenting behaviors	Roggman, L.A., Cook, G.A., Innoccenti, M.S., Norman, V.J., Christiansen, K. (2013). Observations Linked to Outcomes (PICCOLO) Of Diverse Ethnic Groups. *Infant Mental Health Journal, 34(4), 290-306.*	Griffin, J. A., & Friedman, S. L. (2007). NICHD Study of Early Childcare and Youth Development. *National Institute of Health*
Frequency of Parent Child Activity		
Self-Report of Parent-child activities	Rodriguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5 years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child development*, *82*(4), 1058-1075.	
Child meal and sleep routine index	Study PIs	
Time on mother-focal child activities	Rodriguez, E. T., & Tamis-LeMonda, C. S. (2011). Trajectories of the home learning environment across the first 5 years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child development*, *82*(4), 1058-1075.	
Maternal Discipline		
Spanking discipline strategy	Reichman, N.E., Teitler, J.O., Garfinkel, I., MclAnahan, S.S. (2001). Fragile Families: Sample and design. *Children and Youth Services Review, 23*(4-5), 303-326.	