NCT03593356 July 1, 2026

Household Income and Child Development in the First Years of Life (Baby's First Years)
Statistical Analysis Plans for Phase 1 and Phase 2, including a Phase 2 Addendum

Our posted Analysis Plan for Phase 1 (dated December 13, 2022) covers hypotheses for data proposed to be collected in our age 4 follow-up. (Analytic plans for hypotheses for data collected in our ages 1, 2 and 3 follow-ups are available through the "Record History" link on the study's clinicaltrials.gov page.)

Our most recent posted Phase 2 Analysis Plan (submitted and dated July 3, 2025; owing to other complications not published until June 2, 2026) covered hypotheses for data proposed to be collected in our ages 6 and 8 follow-ups. This version of the Analysis Plan updates the posted Phase 2 Plan.

This document begins with the December 13, 2022 Phase 1 Plan, followed by the July 3, 2025 Phase 2 Plan. It then provides an Addendum to the Phase 2 plan that includes an updated plan for analyzing age-8 EEG data. Analysis plans for epigenetic data gathered at ages 4 and 6 have been posted on the Open Science Foundation's website:

Age 4 OSF links:

Child & mother clinical trial outcomes: https://osf.io/ahv2p/overview
Child & mother epigenome wide analyses: https://osf.io/4djtw/files/xftsz (osf.io/jvqx9).

Age 6 OSF links:

Child clinical trial outcomes: https://osf.io/75f2s/overview
Mother clinical trial outcomes: https://osf.io/ju7ra/overview
Child & mother longitudinal epigenome wide analyses: https://osf.io/jvqx9/files/dbgrp

Analysis plans for Age 8 data have not yet been posted.

As with prior versions of our Analysis Plans, an appendix to this document includes summary charts of all current hypotheses and their associated measures for data collected at ages 1, 2, 3, 4, 6, and 8. Child-related hypotheses and measures are listed in Table 11; maternal- and family-related hypotheses and measures are listed in Table 12.

Statistical Analysis Plan Phase 1
Baby's First Years Phase 1: Summary, Pre-registered Hypotheses, Analysis Strategies
Posted December 13, 2022

Project Summary

In the Baby's First Years (BFY) study, one thousand infants born to mothers with incomes falling 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 recruiting issues 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 the cash gifts. 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 care 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 stratified by site, but not by hospitals, within each of the four sites.

BFY was originally formulated 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 lab-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 age 4 years, divided 40/60 between high and low payment 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 age 4 lab 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 surrounding the income supplements: 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 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 ≤ .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 1,000 participants never having charged anything on their debit cards. 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.

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, 922 at age 2, and 922 again at age 3. We expect at least 800 completed cases in our age-4 lab 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, 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 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 age-4 lab 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 are releasing data and documentation from our study to the research community approximately 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.

Age-4 Resting EEG Hypotheses and Analysis Plan.

Following our publication of Age-1 resting EEG treatment impacts (Troller-Renfree et al., 2022), we amended our Age-4 resting EEG analysis plan to include primary and secondary

hypotheses. The original preregistration of Age-1 EEG data included hypotheses across multiple frequency bands. However, due to participant refusal of EEG, the rejection of artifact-laden EEG files, and the high correlation between EEG bands as well as the expected effect size and consistency of the hypothesized effects, we were left with inadequate statistical power for multiple hypothesis testing across bands. As we have uniform, directional hypotheses for all three mid- to high-frequency bands, we have updated this analysis plan to instead include an index of mid- to high-frequency power (described below; primary hypothesis) as well as more traditional neuroscientific investigation of power within bands (described below; secondary hypothesis). Please see the history of preregistrations, including analysis plans, to see a history of how Age-1 EEG findings altered our preregistered analyses.

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 aggregates 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. 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.

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. Covariates will include all preregistered covariates as well as the number of artifact-free epochs contributed by each participant.

In addition, to explore all regional-frequency effects, we will perform separate mixed-ANOVAs 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 be 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.

Changes to Control Variables Following the Posting of the Phase I Analysis Plan (finalized on July 3, 2025)

Beginning with Age 4, we have added controls for child age at the time of the data collection to adjust for age-related developmental differences that may affect both preregistered child and maternal outcomes. In the case of all preregistered outcomes gathered in lab assessments or maternal interviews, we will also control for an indicator for the interviewer/assessor for the specified outcome. In the case of non-survey and non-assessment measures (e.g. tract-level neighborhood poverty rate, school record test scores) we will adjust only for child age at the time of the interview. Child age controls for age-related developmental differences that may affect both preregistered child and maternal outcomes, while the indicator for the interviewer/assessor has been added to account for any idiosyncratic assessor component to variation in preregistered child and maternal measures based on lab or survey assessments.

These two measures were not initially preregistered for analyses of child outcomes (Receptive One-Word Picture Vocabulary Test, Minnesota Executive Function Scale, Child Behavior Checklist, Resting Brain Activity, Matrices, Reading House, and Developmental Condition Diagnosis, age/sex adjusted BMI percentile and maternal report of child health) and maternal outcomes (depression, anxiety, and Body Mass Index (BMI)).

Baby's First Years:
Statistical Analysis Plan for Phase Two (Data Collections at ages 6 and 8)
July 3, 2025

Project Summary

In the Baby's First Years (BFY) study, one thousand infants born to mothers with incomes falling 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 recruiting issues 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 a subset of participating families in each of the four sites to participate in an in-depth qualitative study, but do not elaborate on those plans in this document.)

Mothers were recruited in postpartum wards of 12 participating hospitals shortly after giving birth and, after consenting to participate in a longitudinal child development study, were administered a 30-minute baseline interview. They were then informed about the opportunity to receive a cash gift. The "high-cash gift" treatment group mothers (40% of all mothers) are receiving unconditioned cash payments of $333 per month ($3,996 per year) via debit card for 76 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 76 months. The 40/60 randomization assignment is stratified by site, but not by hospitals within each of the four sites.

BFY was originally designed to study the effects of poverty reduction via a 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. Subsequently, payments were extended a second time, such that they will now be provided to participants for a total of 6 years, 4 months. Interviews conducted at child ages 1, 2 and 3 provide information about family functioning as well as several maternal reports of developmentally-appropriate measures of children's cognitive and behavioral development. An in-person visit at universities at the child age 4 follow-up provided high-quality measures of child well-being as well as maternal responses to a briefer questionnaire.

We have worked with state and local officials to ensure to the extent feasible that the study's cash gifts are not considered countable income for the purposes of determining benefit levels from social assistance programs. Conditional on participants' consent and our success in securing agreements with state and county agencies, we are also collecting state and local

administrative data on parental employment, utilization of public benefits such as Medicaid and Supplemental Nutrition Assistance Programs (SNAP), and any involvement in child protective services.

The compensation difference between families in the high and low cash gift groups has the potential to boost family incomes by $3,756 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, a total sample size of 800 at ages 6 and 8 years, 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 ages 6 and 8 in-person visits at universities we will administer validated, reliable and developmentally sensitive measures of language, executive functioning and socioemotional skills. We will also collect the same measure of resting EEG that we did at age 4 (see below).

The family process measures that we will gather are based on two theories of change—via investments and stress—that might change in response to receipt of the cash gifts. *Investment pathway*: Additional income enables 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 income 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. As detailed in the Baby's First Years clinicaltrials.gov's Study Record Versions, we originally preregistered hypotheses with clinicaltrials.gov within a month after recruitment began (May, 2018) and have preregistered updates prior to each data collection wave.

Overview of Hypothesis Testing. Our Phase 2 key aims are to evaluate the impacts of the high- versus low-cash gifts on validated, reliable, and developmentally-sensitive measures of reading and math achievement, self-regulation, and socioemotional functioning, as well as lower rates of special education and grade retention at child ages 6 and 8 – this is Aim 1 in our Phase 2 NICHD application. Aim 2 calls for measurement of developmentally-sensitive electroencephalographic-based measures of brain functioning at child ages 6 and 8. Aim 3 is focused on hypotheses and data related to family expenditures, neighborhood quality, parent stress and mental health, parenting practices, and children's time spent in nonparental care, all gathered at child ages 6 and 8.

The study's 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 families and children in the high-cash gift group if they had not been offered the additional $313/month cash gift. Therefore, differences in outcomes for the high- compared with the low-cash gift group (after random assignment) can be interpreted as an estimate of the causal effect of the $313/month cash gift (regardless of whether treatment-group participants actually expend all of the funds.) These are commonly known as intent-to-treat effects. In contrast to prior pre-registrations, this child age 6 and 8 preregistration provides directional hypotheses for all of our pre-registered primary and almost all of our secondary hypotheses and will assess them using one-tailed tests (see discussion below).

Before conducting these main analyses, all measures will be examined for psychometric equivalence (configural invariance) across race/ethnicity and whether Spanish or English is a primary language spoken at home.

All pre-registered hypotheses focus on full-sample impacts, although we will also estimate in exploratory analyses moderation of impacts by child gender, race/ethnicity (Black/African American vs. Latino families), presence of older siblings in the family at birth, and depth of poverty at birth (income to needs $\leq .5$ or not).

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-cash gift group as in (1).

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

"Take-up" of the cash gifts is virtually universal, with fewer than six families never having used their debit cards by child age 3. 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, birth order and interview/assessor ID. For child outcomes we also control for child's age in months at the time of the survey/assessment.

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 gathered at ages 6 and 8. To investigate family process impacts, we will apply our estimation strategy to maternal and family measures gathered at child ages 6 and 8.

Attrition. The greatest threat to internal validity is 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 were very high for the ages 1-3 data collections and varied across measures at age 4. Of the original 1,000 recruited participants, we secured interviews with 931 at age 1, 922 at age 2, and 922 again at age 3. By age 4, 5 mothers and 5 children had died and we gathered information on between n=634 (for child EEG) to n=882 (for maternal reports of behavior problems) of the remaining 990 families.

For ages 6 and 8 data collections, we will conduct sensitivity checks to evaluate whether missing data might be biasing estimates using weighting strategies and with multiple imputation. We will use two types of weights created by the Toolkit for Weighting and Analysis of Nonequivalent Groups (TWANG) (Ridgeway et al. 2022) in our regression analyses. Broadly speaking, TWANG uses generalized boosted models to flexibly estimate propensity scores and analytic weights. Generalized boosted modeling is a flexible and nonparametric estimation method that has been shown to outperform other algorithms for propensity score estimation with respect to bias (McCaffrey et al., 2004). These models will include all baseline control variables and the child's age.

First, we will create inverse probability of treatment weights, which are intended to provide estimates of the average treatment effect on the treated (ATT). In this approach, participants from the low-cash gift group analytic sample are weighted by the likelihood of being in the high-cash gift group analytic sample given their baseline observed characteristics, thereby creating a weighted sample in which the low-cash and high-cash gift groups have similar baseline characteristics. This should reduce any bias that is arising from non-equivalence on baseline covariates in the treatment and control analytic samples.

Second, we will create a set of non-response weights intended to adjust for missing data. These weights adjust regression estimates using a weight corresponding to the inverse probability of providing enough usable data to be included in our analytic sample. Weighting in this way produces an analysis sample with characteristics similar to the full BFY baseline sample.

We will also address the problem of missing data using 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 observed information,

data are missing at random. Finally, because we have permission to collect administrative data from over 75% of mothers, we will be able to compare survey respondents and survey non-respondents on 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 cash gift money on ages 6 and 8 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 mediational 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 exogenous (experimental) variation in income 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 $23,788 over the course of the 76 months. The annual equivalent of 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 age-6 lab 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, a sample size of 800 at age 6 or 8, divided 40%/60% between high and low payment groups, provides 80% statistical power to detect a .18 sd impact on cognitive functioning and family processes at $p < .05$ in a one-tailed test and a .20 sd impact at $p < .05$ in a two-tailed test. For our planned EEG measurements (see below), we anticipate a sample size that is closer to 700 than 800. A sample size of 700 at age 6 or 8, divided 40%/60% between high and low payment groups, provides 80% statistical power to detect a .19 sd impact on cognitive functioning and family processes at $p < .05$ in a one-tailed test and a .22 sd impact at $p < .05$ in a two-tailed test.

Multiple comparisons. One strength of this study is the collection of survey, child assessment, neuroscience 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) and Westfall, Tobias and Wolfinger (2011). Pre-registered clusters of measures are identified with grey bars in appendix tables.

Data release. We are releasing data and documentation from the study to the research community through the Inter-university Consortium for Political and Social Research approximately 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.

Additional Details for Age 6 and 8 Resting EEG Hypotheses and Analyses

Consistent with our approach at age 4, we are updating our analysis plan for the ages 6 and 8 collection with primary and secondary hypotheses that have been informed by results from our age-1 resting EEG treatment impacts (Troller-Renfree et al., 2022) and, in the case of age 6, our ongoing but as yet unpublished analyses of the age-4 resting EEG data. As with other outcomes at ages 6 and 8, we now provide directional hypotheses for all our pre-registered primary and secondary hypotheses and will assess them using one-tailed tests (see discussion below).

Our analysis plan includes preregistration of an index of mid- to high-frequency power (described below; primary hypothesis) as well as a more traditional neuroscientific investigation of power within bands (described below; secondary hypotheses). Please see the history of preregistrations, including analysis plans, for more detail.

Ages 6 and 8 primary hypothesis and analyses: Because of limitations in power expected with multiple testing adjustments, we are preregistering a single composite index of mid-to-high-frequency whole-brain power summed across alpha, beta, and gamma bands (defined as between 7 and 45 Hz). This frequency composite index sums absolute power across all single-Hz interval in the Alpha, Beta, and Gamma bands.

Because this approach is based on an estimated intent-to-treat difference in a single index score, there is no need for multiple-testing adjustments. Covariates will include all covariates preregistered at baseline as well as the number of artifact-free epochs contributed by each participant. Power values above the 99th percentile of their respective power band distribution will be truncated from above at the 99th percentile. We hypothesize that, in the eyes-closed condition, the high-cash gift group will show more power in this composite relative to the low-cash gift group. One-tailed t-tests will be used to test this directional hypothesis.

Ages 6 and 8 secondary hypotheses and analyses: Our primary EEG-based hypothesis (detailed above) creates a single composite based on data gathered in the eyes-closed condition. Here we preregister the corresponding analyses of data from the eyes-open condition as secondary hypotheses, with the same directional hypotheses as the eyes-closed condition. We expect effects to be smaller in the eyes-open than in the eyes-closed condition, but do not pre-register hypotheses related to condition by group differences as they will be tested in separate models, are not of immediate theoretical interest, and we are likely to lack the statistical power to detect them.

A second set of secondary analyses are based on a different estimation approach. Consistent with another RCT examining brain activity (Debnath et al., 2020; Marshall et al., 2004; Vanderwert et al., 2010, 2016), we will analyze band-specific effects using mixed-effects models, with group as the between-subjects factor and region as within-subjects factor, to examine ITT impacts on each band of relative power in the eyes-closed condition. Bands will be defined as follows: Theta 3-6 Hz, Alpha 7-12 Hz, Beta 13-20 Hz, Gamma 21-45 Hz. Again, covariates will include all preregistered covariates as well as the number of artifact-free epochs contributed by each participant. Power values above the 99th percentile of their respective power band distribution in absolute power will be truncated from above at the 99th percentile. Our secondary hypotheses for each band are as follows: Theta (high-cash<low-cash), Alpha (high-cash>low-cash), Beta (high-cash>low-cash), and Gamma (high-cash>low-cash). One-tailed t-tests will be used given directional hypotheses. No further multiple-testing adjustments beyond the use of multi-level models will be used (Gelman et al., 2012).

The mixed effects models will also enable us to rank the four power bands by the (absolute) size of their impacts. Although we are not preregistering hypotheses regarding these rankings, we can record our expectations about them. Based on past results from this project (Ages 1 and 4), a review of the relevant literature, and the expertise of our neuroscience team, we expect that the cash gift impact will be largest in the Alpha band at both ages 6 and 8. Following Alpha, associations between SES and Theta are the most reported in the literature and, as such, we expect that the impact of the cash gift will second largest in the Theta band, although we have seen no indication of this difference at Ages 1 and 4. Finally, some group differences in Beta and Gamma were detected at Age 1, but preliminary analyses suggest these are not detectible at age 4. Continuous, correlational relations between SES and Beta and Gamma are less consistent in the broader literature (perhaps as these bands are easily influenced by artifacts, as we saw some evidence of at Age 4), leading us to expect that impact of the cash gift will be smallest for these bands at Ages 6 and 8. In supplemental, exploratory analyses, we will report the same analyses for absolute power and in the eyes-open condition.

A few changes in our EEG analysis plans between age 4 and ages 6 and 8 are worth noting. First, as with other study outcomes, we now provide directional hypotheses for all analyses. As such, we now also propose the use of one-tail analyses to test these hypotheses. This switch to one-tailed tests was motivated by two factors. First, one-tailed tests provide more

statistical power for testing our hypotheses. Second, although we had directional hypotheses in the past, these were not always clearly preregistered and, as such, two-tailed tests were used. However, since our initial preregistration, more evidence has emerged linking SES to EEG (Brito et al., 2020, 2022; Cantiani et al., 2019; Maguire & Schneider, 2019; Rockers et al., 2023) as well as our own analyses (Troller-Renfree et al., 2022), which supports our making directional hypotheses for all bands. However, we are still unsure of which brain regions are likely to drive these whole-brain effects as the literature is mixed with significant findings inconsistently observed in almost every region of the brain. As such, we make no region-specific hypotheses.

Additional Details for Age 6 and Age 8 Epigenetic Processing and Analysis

We are preprocessing age 4 genome-wide DNA-methylation on a newly released array (EPIC v2) - see OSF link https://osf.io/ahv2p and will then upload a more detailed age-6 and age-8 DNA-methylation preprocessing pipeline on OSF, which will probably deviate from age 4 given new bioinformatic tools and lessons learned. We will re-preprocess age-4, age-6 and age-8 DNA-methylation together to minimize technical artifacts. We will exclude DNA-methylation sites and samples failing technical criteria (e.g., low detection p threshold, sex mismatch) based on methylation-wide data *before* these preregistered methylation profile scores are computed. We will follow preregistered intent-to-treat analyses described in this document and add the following variables as covariates to DNA-methylation analyses: cell composition, technical DNAm factor (plate), child's age in months at the time of age-6 and age-8 data collection.

We hypothesize the following with regard to *child epigenetic outcomes*:

1) We hypothesize slower pace of aging, as measured by DunedinPACE, in the high-cash gift group, relative to the low-cash gift group. Methylation pace of aging was developed from DNA-methylation analysis of Pace of Aging in the Dunedin Study birth cohort (Belsky et al., 2022). Pace of Aging is a composite phenotype derived from analysis of longitudinal change in 18 biomarkers of organ-system integrity (Belsky et al., 2015). Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. We will also report GrimAge Acceleration, which we consider an exploratory analysis. GrimAge represents a DNA-methylation metric designed to predict morbidity and mortality (Lu et al 2019). Briefly, the initial phase entailed the computation of models incorporating physiological indicators, age, sex, and smoking history, with the objective of optimizing mortality prediction within the Framingham Heart Study Offspring cohort (Lu et al., 2019). GrimAge will be based on principal components of DNA-methylation to bolster reliability (Higgins-Chen et al., 2022) and residualized for chronological age derived from sample receipt age to reflect accelerated biological age.

2) We hypothesize higher Epigenetic-g in the high-cash group relative to the low-cash gift group. 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. If a higher quality measure of epigenetic profile of cognitive functioning becomes available at the time of analysis, we will substitute that instead.

We hypothesize the following with regard to *maternal epigenetic outcomes*:

1) We hypothesize slower pace of aging, as measured by DunedinPACE (Belsky et al., 2022), in the high-cash gift group, relative to the low-cash gift group. We will also report GrimAge Acceleration (Lu et al 2019; described above) and PhenoAge Acceleration as exploratory analyses. PhenoAge is conceptualized on the foundation of physiological markers and chronological age, which are subsequently employed to model a novel sample derived from DNA methylation, culminating in the establishment of a definitive DNA methylation clock (Levine et al., 2018). This metric exemplifies the age, measured in years, at which the average mortality risk in the NHANES III cohort aligns with the mortality risk as forecasted by the PhenoAge algorithm. PhenoAge will be based on principal components of DNA-methylation to bolster reliability (Higgins-Chen et al., 2022) and residualized for chronological age derived from sample receipt age to reflect accelerated biological age.

Changes to Control Variables Following the Posting of the Phase II Analysis Plan (finalized on July 3, 2025)

As with Age 4 analyses, for our Age 6 and 8 preregistered analyses, we have added controls for child age at the time of the data collection to adjust for age-related developmental differences that may affect both preregistered child and maternal outcomes. In the case of all preregistered outcomes gathered in lab assessments or maternal interviews, we will also control for an indicator for the interviewer/assessor for the specified outcome. In the case of non-survey and non-assessment measures (e.g. tract-level neighborhood poverty rate, school record test scores) we will adjust only for child age at the time of the interview.

These measures were not initially preregistered for analyses of child outcomes (Receptive One-Word Picture Vocabulary Test, Minnesota Executive Function Scale, Child Behavior Checklist, Resting Brain Activity, Matrices, Reading House, and Developmental Condition Diagnosis, age/sex adjusted BMI percentile and maternal report of child health) and maternal outcomes (depression, anxiety, and Body Mass Index (BMI)). Child age controls for age-related developmental differences that may affect both preregistered child and maternal outcomes, while

the indicator for the interviewer/assessor has been added to account for any idiosyncratic assessor component to variation in preregistered child and maternal measures based on lab or survey assessments.

July 1, 2026 Addendum

This addendum to our July 3, 2025 Phase 2 Analysis Plan includes changes to our preregistered analyses of age 8 EEG data. Analysis plans for the epigenetic data we will be gathering will be posted on the OSF website at a future date.

Additional Details for Age 8 Resting EEG Hypotheses and Analyses

Our primary and secondary hypotheses and analytic approach for age 8 remain identical to those preregistered for ages 6 and 8 prior to age-6 data collection: our primary outcome is a single composite index of absolute mid-to-high-frequency power (combining alpha, beta, and gamma, likely 7–45 Hz [see below about band definition] in the eyes-closed condition), and our secondary outcomes are the four bands of relative power (theta, alpha, beta, and gamma), each tested with one-tailed tests given our directional hypotheses.

The pattern of results from resting EEG data observed at age 4 (Troller-Renfree et al., 2026) and 6 (Rico-Pico et al., in prep), together with literature published since our initial age 6 preregistration, have led us to develop new expectations for the results from our age 8 EEG analyses. We summarize both prior findings from BFY and the literature that have led to our change in expectations (even though we have preserved our original pre-registered outcomes).

For our primary preregistered outcome, the mid-to-high-frequency absolute power composite, at age 6 analyses showed group differences in the direction we expected with a p-value <.10. At age 4, the estimate was in the expected direction, but not statistically significant. At age 1, this composite was significant ($p<.05$) but not preregistered.

Across data collected at ages 1, 4, and 6, we have observed consistent cash gift group differences with more power observed in the high cash gift group in absolute, but not relative alpha power. Specifically, across all three measurement occasions, group differences favoring children in the high-cash group in absolute alpha power are statistically significant at $p<.10$ or $p<.05$ (Troller-Renfree et al., 2022, 2026).

Group differences in absolute and relative theta have been inconsistent. At age 1, absolute theta showed essentially no group difference (effect size = .02, p = .83), while relative theta was lower in the high-cash gift group (the predicted direction; effect size = −.21) though the group differences were p = .07). At age 4, neither absolute nor relative theta differed meaningfully between groups (effect sizes of .06 and −.10, respectively, both ns). By age 6, absolute theta power was higher in the high-cash gift group in both the eyes-closed and eyes-open blocks — the opposite of our preregistered direction — with effect sizes (.16 –.18) comparable in magnitude to those observed for alpha, while relative theta power showed no meaningful group difference (effect size of approximately 0). This pattern of results raised two further questions: whether alpha is better captured using an absolute or relative power measure, and whether a one-tailed test remains appropriate for theta given the pattern of results observed at age 6.

Considerations and expectations based on recent literature.

We note two considerations raised by recent literature and by our own results that support a change in our expectations. First, with respect to absolute versus relative power: a new systematic review of the pediatric resting EEG literature indicates that absolute power is used more frequently to measure resting state analyses than relative power (Gray et al., 2026). Moreover, a survey of domain experts found that they would be more likely to preregister absolute than relative power if designing a hypothetical RCT study like the design of BFY (Troller-Renfree et al., 2025). Relative power has commonly been used in developmental research because it offers a way to remove or hold constant non-neural and non-task-related variability (e.g., state effects like fatigue) and individual differences in power, although recent work suggests this correction is fairly modest (Sandre & Troller-Renfree, 2026). In addition, relative power has some significant weaknesses relative to absolute power in the context of an RCT, and this has shifted our thinking about which outcomes are preferred. First, the very framework of an RCT means that individual differences do not bias analyses by design. This means that the stated strength of relative power is less important. More important, however, is the fact that single-band relative power measures are calculated with total absolute power as the denominator. This means that a single band measure reflects changes in all bands, not just the band of interest. This is especially problematic if there are changes in the largest bands because these changes may result in differences in the smaller bands being harder to detect. This may the case in our age-6 data. We observed an unexpected significantly higher absolute theta at age 6 in the high cash gift group compared with the low cash gift group. This pattern is relevant because theta is largest component of the denominator used to compute relative power. Thus, differences in the denominator by treatment group, in turn, affect estimates of the difference by treatment of relative power in alpha, beta, and gamma bands (because it introduces differences in the denominator). Based on these considerations, we judge absolute power to be a better way of measuring band specific treatment group differences in the BFY study.

Second, prior to age 6, we preregistered a one-tailed test predicting lower relative theta power in the high-cash gift group relative to the low-cash gift group. This directional test was not significant at age 6; indeed, the observed difference ran in the opposite direction, with the high-cash gift group showing higher relative theta power, at a magnitude comparable to the alpha effect, and the corresponding two-tailed test would have reached conventional significance. Our original preregistration for ages 6 and 8 calls for a one-tailed, directional test (high-cash<low-cash). Given the age 6 pattern, we now think it is plausible that group differences in relative theta will favor the high-cash group at age 8. For the sake of the field and to increase data transparency will report any findings from two-tailed tests (in all bands) as a part of our analyses of age 8 data.

Age 8 band definition. At Age 8, frequency bands will be provisionally defined as follows: Theta 4–6 Hz, Alpha 7–11 Hz, Beta 13–19 Hz, and Gamma 21–45 Hz. However,

considering the development of these frequency ranges during childhood (e.g., Freschl et al., 2022), we will blind ourselves to treatment condition and examine the presence of oscillatory peaks (e.g., theta and alpha peaks) to accurately adjust upper and lower boundaries of the frequency bands. We will use the Specparam toolbox (Donoghue et al., 2020) to isolate oscillatory activity from the aperiodic background. We will then examine the oscillatory power spectrum to determine the general number of peaks and extract their frequencies from the data provided by Specparam, classifying them in the standard ranges. Based on this information, we will analyze the distributions of peak frequencies and widths (i.e., how many frequency bins the peak occupies) to determine a range that includes the average and limits for at least 90% of participants. At age 8, we anticipate a predominant and single alpha peak (see Cellier et al., 2021) with minimal appearance of the theta peak. In such a case, we will use alpha as the reference to adjust the upper and lower boundaries of theta and beta, respectively, while keeping the rest of the ranges to the standard defined above. In any case, the changes in band boundaries are expected to be minimal (no more than 1 Hz).

References

Belsky, D. W., Caspi, A., Houts, R., Cohen, H. J., Corcoran, D. L., Danese, A., Harrington, H., Israel, S., Levine, M. E., Schaefer, J. D., Sugden, K., Williams, B., Yashin, A. I., Poulton, R., & Moffitt, T. E. (2015). Quantification of biological aging in young adults. Proceedings of the National Academy of Sciences, 112(30), E4104–E4110. https://doi.org/10.1073/pnas.1506264112

Belsky, W. D. et al (2020). Quantification of the pace of biological aging in humans through blood tests, 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

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

Brito, N. H., Troller-Renfree, S. V., Leon-Santos, A., Isler, J. R., Fifer, W. P., & Noble, K. G. (2020). Associations among the home language environment and neural activity during infancy. *Developmental Cognitive Neuroscience*, 43, 100780. https://doi.org/10.1016/j.dcn.2020.100780

Brito, N. H., Werchan, D., Brandes-Aitken, A., Yoshikawa, H., Greaves, A., & Zhang, M. (2022). Paid maternal leave is associated with infant brain function at 3 months of age. *Child Development*, 93(4), 1030–1043. https://doi.org/10.1111/cdev.13765

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

Cantiani, C., Piazza, C., Mornati, G., Molteni, M., & Riva, V. (2019). Oscillatory gamma activity mediates the pathway from socioeconomic status to language acquisition in infancy. *Infant Behavior and Development*, 57, 101384. https://doi.org/10.1016/j.infbeh.2019.101384

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

Freschl, J., Al Azizi, L., Balboa, L., Kaldy, Z., & Blaser, E. (2022). The development of peak alpha frequency from infancy to adolescence and its role in visual temporal processing: A meta-analysis. *Developmental Cognitive Neuroscience*, *57*, 101146.

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal Of Research on Educational Effectiveness*, 5(2), 189-211.

Gray, K. N., & Troller-Renfree, S. V. (2026, April 6). Pediatric resting EEG collection, preprocessing, and analysis: A systematic review. Retrieved from osf.io/u7q32

Higgins-Chen, A. T., Thrush, K. L., Wang, Y., Minteer, C. J., Kuo, P.-L., Wang, M., Niimi, P., Sturm, G., Lin, J., Moore, A. Z., Bandinelli, S., Vinkers, C. H., Vermetten, E., Rutten, B. P. F., Geuze, E., Okhuijsen-Pfeifer, C., van der Horst, M. Z., Schreiter, S., Gutwinski, S., … Levine, M. E. (2022). A computational solution for bolstering reliability of epigenetic clocks: Implications for clinical trials and longitudinal tracking. Nature Aging, 2(7), 644–661. https://doi.org/10.1038/s43587-022-00248-2

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

Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., & Horvath, S. (2018). An epigenetic biomarker of

aging for lifespan and healthspan. *Aging*, *10*(4), 573–591. https://doi.org/10.18632/aging.101414

Lu, A, Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging*, *11*(2), 303–327. https://doi.org/10.18632/aging.101684

Maguire, M. J., & Schneider, J. M. (2019). Socioeconomic status related differences in resting state EEG activity correspond to differences in vocabulary and working memory in grade school. *Brain and Cognition*, 137. https://doi.org/10.1016/j.bandc.2019.103619

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

McCaffrey, D. F., Ridgeway, G., & Morral, A. R. (2004). Propensity score estimation with boosted regression for evaluating causal effects in observational studies. *Psychological methods*, *9*(4), 403.

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

Rockers, P. C., Leppänen, J. M., Tarullo, A., Coetzee, L., Fink, G., Hamer, D. H., Yousafzai, A. K., & Evans, D. (2023). Evaluation of a community health worker home visit intervention to improve child development in South Africa: A cluster-randomized controlled trial. *PLOS Medicine*, 20(4), e1004222. https://doi.org/10.1371/journal.pmed.1004222

Ridgeway, C. L., & Markus, H. R. (2022). The significance of status: what it is and how it shapes inequality. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, *8*(7), 1-25.

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

Sandre, A., & Troller-Renfree, S. V. (2026). Power struggles: Absolute vs. relative EEG power in developmental neuroscience. *Developmental Cognitive Neuroscience*, 101698.

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

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., ... & Noble, K. G. (2022). The impact of a poverty reduction intervention on infant brain activity. *Proceedings of the National Academy of Sciences*, *119*(5), e2115649119.

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., ... & Noble, K. G. (2026). The impact of a monthly unconditional cash transfer on child brain activity: A 4-year follow-up. *Developmental Cognitive Neuroscience*, 101673.

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 and notes (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) We will administer the monolingual (English) or bilingual (English/Spanish) versions as appropriate. Because the two versions of the test are not co-normed, the primary outcome will be a derived "conceptual score," or sum of the raw scores on all individual items that appear on both versions of the test.
	Martin & Brownell, 2011	Internal consistency: English version: .96 to .97; Bilingual version: .94 to .98	6, 8		Expressive One-Word Picture Vocabulary (EOWPVT). Monolingual and Bilingual. Score range monolingual version: 0-185, bilingual version: 0-180; higher scores indicate better performance. Because the two versions of the test are not co-normed, the primary outcome will be a derived "conceptual score," or sum of the raw scores on all individual items that appear on both versions of the test. Expect higher scores in high- than low-cash gift group.
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)
Academic Achievement: Reading					
Reading	McGrew & Woodcock, (2018)	Test-retest reliability ages 5-19: 0.92	6, 8		Woodcock Johnson IV Test of Achievement: Letter-Word Identification. Score range: 0-78; higher scores indicate better performance. Expect higher scores in high- than low-cash gift group.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Reading comprehension	McGrew & Woodcock, (2018)		Originally planned for age 8 but dropped due to visit time constraints		Woodcock Johnson IV Test of Achievement: Passage Comprehension. Score renge: 0-52. Expect higher scores in high- than low-cash gift group. Dropped this test prior to A8. Reason: This test is either short (if children cannot read) or long (if they can) but the median time is relatively long (10 mins). It is highly correlated with WJ Letter Word (.93) and children started to lose effort during piloting. To make the overall visit length work we had to cut this measure.
Academic Achievement: Math					
Math	McGrew & Woodcock, (2018)	Test-retest reliability ages 5–19 = 0.91	6, 8		Woodcock Johnson IV Tests of Achievement: Applied Problems. Score range: 0-56; higher scores indicate better performance. Expect higher scores in high- than low-cash gift group.
Executive Function and Behavioral Regulation					
Executive Function	Diamond & Taylor, 1996; Weiland& Yoshikawa, 2013; Bierman et al., 2008			Originally registered for age 4 and then dropped	Intended to be measured by the pencil tap test. This item was preregistered as an age-4 secondary outcome but was dropped on September 13, 2022, due to evidence of floor effects, and numerous reports from research staff that children were not understanding the instructions.
Executive Function: Minnesota Executive Function Scale	Carlson, 2017; Carlson, & Zelazo 2014	MEFS: validity .92 test-retest .93	4, 6, dropped at 8		Measured by the Minnesota Executive Function Scale (MEFS). Score range:0-100; higher scores indicate better performance. Expect higher scores in high- than low-cash gift group. We dropped this measure at A8 because a. less variability than the Flanker, b. piloting data quality concerns (children sometimes seem confused by the instructions) and c. we had to cut some measures to keep overall visit length acceptable.
Executive Function: NIH Toolbox Flanker Inhibitory Control and Attention Test	Gershon, Wagster, Hendrie, Fox, Cook, & Nowinski, 2013	Test-retest: English 3-15 yo: ICC=0.95; 20-85 yo: ICC=0.83. Spanish 18-85 years old: ρ=.65	6, 8		Measured by the NIH Toolbox Flanker task. Score range: 0-30; higher scores indicate better performance.Expect higher scores in high- than low-cash gift group.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Working Memory	Wechsler (2012)		6		Measured by the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) picture memory subtest. Picture Memory score range: 0-35, higher scores indicate better performance. Expect higher scores in high- than low-cash gift group.We pre-register three scores (MEFS, Flanker, Picture Memory), with a plan to do a confirmatory factor analysis and pre-register the impact on the common factor.
Working Memory	Wechsler (2011)		Originally planned for age 8 but dropped due to visit time constraints		Measured by Wechsler Intelligence Scale for Children: 5th Edition (WISC-V) subtest Digit Span. Score Range 0-54. Higher scores indicate better performance. Expect higher scores in high- than low-cash gift group. We will pre-register three scores (MEFS, Flanker, Digit Span), with a plan to do a confirmatory factor analysis and pre-register the impact on the common factor. To make the overall visit length work we had to cut this measure.
Socio-Emotional Processing					
Social-Emotional Problems	Briggs-Gowan et al., 2004	internal consistency .65-.79 test-retest reliability .87		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. At age 3, we will estimate the statistical significance of the entire family of related measures in the Child Socio-Emotional Processing outcome cluster measured during the same wave using step-down resampling methods for multiple testing (see statistical analysis plan for more details; Westfall and Young, 1993).

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Social-Emotional Behavior	Roggman et al., 2013; Griffen & Friedman, 2007; Belsky, 2007			Originally registered for age 1 but unable to be coded	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)
Behavior/Emotional Problems	Achenbach, McConaughy, Ivanova, & Rescorla (2011)	internal consistency .80 - .92, test-retest reliability .81-.85	6, 8		Measured by the maternally-reported Brief Problem Monitor (BPM), assesses attentional, behavioral, and internalizing problems in children. 19 questions and answer choices. Score range: 0-38; higher scores indicate more behavioral problems. Expect lower scores in high- than low-cash gift group. 3 answer choices: 0. Not true as far as you know, 1. Somewhat true, 2. Very true
Visual Processing and Abstract Spatial Perception					
Visual processing and abstract spatial perception*	Wechsler & Naglieri, 2006	internal consistency .88 test-retest reliability .77	Originally registered for age 4 but not able to be calculated	4, for matrices subtest only	The Wechsler Nonverbal Scale of Ability was originally pre-registered as a Primary Outcome. The IQ score is calculated using two subtests -- Matrices and Recognition -- and we began our fieldwork on July 9, 2022 with both. On the basis of preliminary analysis of the first 71 cases, we discovered that 21% of participants scored at the floor of the Recognition assessment. We therefore dropped the Recognition subtest from our data collection instrument on September 30 2022, precluding us from calculating IQ in subsequent participants. Scores on the Matrices subtest, which measures visual processing and abstract spatial perception (not IQ per se), are now registered as an age-4 secondary outcome.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Fluid Reasoning					
Fluid Reasoning*	Wechsler(2012)		6		Measured by the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) fluid reasoning index, assessed through a composite of two tasks: picture concepts (score range: 0-27; higher scores indicate better performance) and matrix reasoning (score range: 0-26; higher scores indicate better performance). Expect higher scores in high- than low-cash gift group. Note: The Picture Concept subtest had to be dropped on 9/17/2024 due to floor effects.
Perceptual Reasoning					
Perceptual Reasoning*	Wechsler (2011)		8		Measured by the Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II) perceptual reasoning index, assessed through a composite of two tasks: block design (score range: up to 8 years old: 0-57; > 9 yo: 0-71; higher scores indicate better performance) and matrix reasoning (up to 8 years old: 0-24; > 9 yo: 0-30; higher scores indicate better performance). Expect higher scores in high- than low-cash gift group.
Pre-Literacy					
Pre-Literacy	Hutton et al., 2019; Hutton et al., 2021			4	Measured by The Reading House

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
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; Troller-Renfree et al. 2022	n/a	4	4	Measured by high-density in-lab electroencephalography (EEG) **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. **Age-4 secondary**: We hypothesize 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. Note: The original preregistration of EEG data collected when children were 12 months old included hypotheses across multiple frequency bands. Please see the history of preregistrations, including analysis plans.
Age-6 and Age-8 Resting Brain Function	Brito et al., 2020; Brito et al., 2022; Cantiani et al., 2019; Debnath et al., 2020; Gelman et al., 2012; Maguire et al., 2019; Marshall et al., 2004; Rockers et al., 2023, Troller-Renfree et al., 2022; Vanderwert et al., 2010; Vanderwert et al., 2016		6, 8		Measured by high-density in-lab electroencephalography (EEG). Primary hypothesis: Because of limitations in power expected with multiple testing adjustments, we are preregistering a single composite index of mid-to-high-frequency whole-brain power summed across alpha, beta, and gamma bands (defined as between 7 and 45 Hz). This frequency composite index sums absolute power across all single-Hz interval in the Alpha, Beta, and Gamma bands. We hypothesize that, in the eyes-closed condition, the high-cash gift group will show more power in this composite relative to the low-cash gift group. One-tailed t-tests will be used to test this directional hypothesis. See the Statistical Analysis Plan for additional details.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Age-6 and Age-8 Resting Brain Function, continued	Brito et al., 2020; Brito et al., 2022; Cantiani et al., 2019; Debnath et al., 2020; Gelman et al., 2012; Maguire et al., 2019; Marshall et al., 2004; Rockers et al., 2023, Troller-Renfree et al., 2022; Vanderwert et al., 2010; Vanderwert et al., 2016,			6, 8	Secondary hypotheses: Our primary hypothesis (detailed above) is based on data gathered in the eyes-closed condition. Corresponding analyses of data from the eyes-open condition will be pre-registered as one of our secondary hypotheses, with the same directional hypotheses as the eyes-closed condition. We expect effects to be smaller in the eyes- open than in the eyes closed condition but do not pre-register hypotheses related to condition by group differences as they will be tested in separate models, are not of immediate theoretical interest, and we are likely to lack the power to detect them. A second set of secondary analyses are based on a different estimation approach. Consistent with another RCT examining brain activity (Debnath et al., 2020; Marshall et al., 2004; Vanderwert et al., 2010, 2016), we will analyze band-specific effects using mixed-effects models, with group as the between-subjects factor and region as within-subjects factor, to examine ITT impacts on each band of relative power in the eyes-closed condition. Our secondary hypotheses for each band are as follows: Theta (high-cash<low-cash), Alpha (high-cash>low-cash), Beta (high-cash>low-cash), and Gamma (high-cash>low-cash). One-tailed t-tests will be used given directional hypotheses. No further multiple-testing adjustments beyond the use of multi-level models will be used (Gelman et al., 2012). See the Statistical Analysis Plan for additional details.
Task-Related Brain Function					
Auditory Discrimination Brain Function*	Choeur et al., 2000; Garcia-Sierra et al., 2011; Kuhl et al., 2005	n/a		4 dropped	Our original plan was to measure the mismatch negativity (MMN) ERP with larger differences between standard and deviant stimulus in high-cash gift group compared to the low-cash gift group. While the auditory oddball task was administered, an MMN waveform was not produced.Instead, a lateralized MMR was produced, which has been theorized to be a correlary of the MMN (McWeeny & Norton, 2020), but not a fully overlapping construct. Due to differences in timing, amplitude, and topography, we are not reporting MMR amplitudes as an MMN. As the MMN was not elicited, we will not deposit MMN impacts and analyses of the MMR will be reported as exploratory.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Health: BMI					
Body Mass Index (BMI)	Kuczmarski, 2000	n/a		4, 6, 8	Measured by CDC BMI percentile scales **Age 6/8:** We expect to see a reduced percentage of overweight or obese (greater than or equal to 85th percentile) children in the high-cash gift group compared to the low-cash gift group. We will report mean percentile scores of the two groups in descriptive analyses.
Health: Physiological Stress					
Physiological Stress: Hair cortisol	Ursache et al., 2017; Meyer et al., 2014; Davenport et al., 2006	n/a		Originally registered for age 4 but unable to collect	Our original plan was to measure physiological stress using hair cortisol concentration. The first several months of data collection revealed large racial and ethnic differences in willingness to provide a hair sample, due to both cultural and practical reasons. Because of the large amounts of non-random missing data, which would both compromise our statistical power and limit the generalizability of any findings, we dropped hair cortisol from our data collection procedures on October 25, 2022.
Physiological Stress: Nail cortisol	Phillips et al., 2021		6 dropped at 8		Child nail samples will be collected to yield a measure of the concentration of cortisol in pg/mg (picograms per milligram). To ensure cortisol levels are within an expected range, values above 500 and equal to or less than 0 will be assigned a missing value. To account for potential outliers, values below 500 will be truncated at the 99th percentile. Cortisol values will be log-transformed. We hypothesize the high-cash gift group will have lower cortisol values when compared to the low-cash gift group. If both fingernail and toenail are collected we will control for whether fingernail or toenail. Age 8: We dropped this measure as it was not feasible to collect enough nail samples.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Health: Sleep					
Sleep problems	Yu et al., 2012	reliability .9	3	1, 2	Measured by PROMIS Sleep Disturbance- Short Form adapted from ECHO; For ages 1 and 2, additive index of the following items with 5-point answer (0: never; 1: almost never; 2: sometimes; 3: almost always, 4: always): 1. difficulty falling asleep 2. sleeping through night (reverse coded) 3. problem with sleep 4. problem sleeping For Age 3, item 1 was not included in the survey
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	Child's overall health item source: Idler & Benyamini, 1997	n/a		4, 6, 8	Additive index of the following two items (reverse score 1, then add 1 and 2): 1. Child's overall health? 5 answer choices: excellent, very good, good, fair, or poor. Score range: 1-5; higher scores indicate better health. Expect higher scores in high- than low-cash gift group. 2.About how many times in the last year was child sick? 4 answer categories: 0-1 times, 2-3 times, 4-6 times, 7+. Score range: 1-4; higher scores indicate increased occurrences of illness. Expect lower scores in high- than low-cash gift group.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Diagnosis of Health Condition					
Diagnosis of Health Condition	BFY PIs			6, 8	Has child been diagnosed with any chronic health condition? Yes/No. We will estimate whether there are group differences, but do not formulate a directional hypothesis, because of two offsetting possibilities: i) the high-cash gift group may have better access to services, which may lead to higher rates of diagnosis and/or ii) the cash gifts may lead to fewer chronic health conditions. For children reported to have been diagnosed with a chronic health condition, we will ask for descriptive purposes whether the child was diagnosed with asthma and/or something else. We will explore the presence of group differences in specific diagnoses, but do not anticipate having the statistical power to detect significant group differences.
Diagnosis of Developmental Condition					
Diagnosis of Developmental Condition	Study PIs	n/a		4, 6, 8	Dichotomous measure 1:yes 0:no, constructed from the questionnaire categories: Diagnosis of Dev. Condition: speech delay, autism, ADHD, something else (this Q is part of Health Qs) We will estimate whether there are group differences, but do not formulate a directional hypothesis, because of two offsetting possibilities: i) the high-cash gift group may have better access to services, which may lead to higher rates of diagnosis and/or ii) the cash gifts may lead to fewer developmental conditions

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Special Services (IEP)	PSID, 2021			6, 8	Maternal report of whether the child has an IEP or receives special educational services using questions adapted from the School Enrollment and Expectations section of the PSID Child Development Supplement. 1: yes; 0: no We will estimate whether there are group differences, but do not formulate a directional hypothesis, because of two offsetting possibilities: i) the high-cash gift group may have better access to services, which may lead to higher rates of diagnosis and receipt of special education and/or ii) the cash gifts may lead to higher school achievement and therefore lower need for special education and individual education plans.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Child Epigenetic Pace of Aging					
Methylation pace of aging	Belsky et al., 2015; Belsky et al., 2020; Belsky et al., 2022; Higgins-Chen et al., 2022; Levine et al., 2018, Lu et al., 2019	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. 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. If a higher quality measure of epigenetic aging at the time of analysis becomes available, we we will substitute that instead. OSF preregistration link: https://osf.io/ahv2p/overview
				6	Expect slower DunedinPACE- pace of aging in high-cash group: 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. Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. We will also report GrimAge Acceleration, which we consider an exploratory analysis. GrimAge represents a DNA-methylation metric designed to predict morbidity and mortality. Briefly, the initial phase entailed the computation of models incorporating physiological indicators, age, sex, and smoking history, with the objective of optimizing mortality prediction within the Framingham Heart Study Offspring cohort. GrimAge will be based on principal components of DNA-methylation to bolster reliability and residualized for chronological age derived from sample receipt age to reflect accelerated biological age.OSF link: https://osf.io/75f2s/overview

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
				8	Expect slower DunedinPACE- pace of aging in high-cash group: Methylation pace of aging was developed from DNA-methylation analysis of Pace of Aging in the Dunedin Study birth cohort (Belsky et al., 2022). Pace of Aging is a composite phenotype derived from analysis of longitudinal change in 18 biomarkers of organ-system integrity (Belsky et al., 2015). Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. We will also report GrimAge Acceleration as exploratory analyses. GrimAge represents a DNA-methylation metric designed to predict morbidity and mortality (Lu et al., 2019). Briefly, the initial phase entailed the computation of models incorporating physiological indicators, age, sex, and smoking history, with the objective of optimizing mortality prediction within the Framingham Heart Study Offspring cohort (Lu et al., 2019). GrimAge will be based on principal components of DNA-methylation to bolster reliability (Higgins-Chen et al., 2022) and residualized for chronological age derived from sample receipt age to reflect accelerated biological age. OSF preregistration link to come.
Child DNA Methylation					
DNA methylation	McCartney et al, 2022	n/a		4, 6, 8	Expect higher Epigenetic-g in high-cash group: 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. 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. If a higher quality - and validated - measure of epigenetic profile of cognitive functioning becomes available at the time of analysis, we will substitute that instead. OSF preregistration links for ages 4 and 6: https://osf.io/ahv2p/overview and https://osf.io/75f2s/overview

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
Child Nutrition					
Consumption of healthy foods	Los Angeles County WIC Survey, 2017			2, 6, 8	Additive index of the number of times per day consumed the following items: 1. eat fruits 2. eat vegetables Index score: 1: zero, 2:1 time, 3: 2 times, 4: 3 times, 4: 5 times, 6: 5 or more times. Score range: 1-12; higher scores indicate more comsuption of healthy foods. Age 6/8: Expect higher scores in high- than low-cash gift group.
Consumption of unhealthy foods	Los Angeles County WIC Survey, 2017; Hunsberger et al., 2012			2, 6, 8	Age 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 Age 6/8: . Additive index of the number of times per day consumed the following items: 1. salty snacks such as potato chips, Doritos, Fritos, tortilla chips; 2. sweets or sweetened foods, such as sweetened cereals, fruit bars, Pop-Tarts, donuts, cookies, or candies 3. drink sweetened beverages such as juice, soda, chocolate milk, or other sweet drinks? Index score: 1. zero, 2. 1 time, 3. 2 times, 4. 3 times, 5. 4 times, 6. 5 or more times. Score range: 1-18; higher scores indicate more comsuption of unhealthy foods. Expect lower scores in high- than low-cash gift group.
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

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
School Achievement					
Retention	PSID, 2021			6, 8 dropped at 6	Maternal report of whether the child has repeated a grade, adapted from the School Enrollment and Expectations section of the PSID Child Development Supplement. Item: Since kindergarten, has child ever repeated a grade or been held back because the school recommended it? (1.yes 0.no) Expect fewer retentions in high cash group. We originally preregistered a retention question for age 6 but were not able to collect this information owing to a questionnaire sequence error.
School Behavior					
Suspensions	SIPP, 2023			6, 8 dropped at 6	Maternal report of the number of school suspensions, adapted from the Survey of Income and Program Participation. Item: Has child ever been suspended or expelled from school since the beginning of kindergarten? (no/yes) Expect fewer suspensions in high cash group. We originally preregistered a suspension question for age 6 but realized that not all children will have attended school for very long at the time of those interviews. That is not an issue with 8-year old children.
Child engagement in school	Ehrle & Moore, 1999	Cronbach's alpha=.76		6, 8	Maternal report of child's engagement in school using an item adapted from the Survey of Income and Program Participation. "How often would you say that (CHILD) cares about doing well in school"? 4 answer choices: 1. none of the time; 2. some of the time, 3. most of the time; 4. all of the time. Score range: 1-4; higher scores indicate more engagement in school. Expect more engagement in high cash group.
Attendance	BFY PIs			8	During the last 12 months, how many times has (CHILD) been absent from school for any reason? Answer is estimated number of days. Expect fewer absences in high cash group.

Domains (in gray) and sub-domains	Measure source	Psychometrics	Age preregistered Primary Outcome	Age preregistered Secondary Outcome	Measures and notes (All measures between grey lines measured in the same wave will be subject to multiple testing adjustments)
School Quality					
School Quality	Stanford data matched to name of school and city/district			8	Gathered from school names and city/district. We expect an increase in school quality for the high cash group. The measure we select will be a function of what data will be available at the time of Age 8 wave.

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 collection. Assessment target time is around the child's birthday for each wave, i.e. 12 months, 24 months, 36, months, 48 months, 72 months, and 96 months.

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 were 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 later ages, so it is removed from the pre-registration table.

*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"; then "IQ". Presently re-named to domain tested (not IQ per se)
-Previously "Language Related Brain Function"; presently "Auditory Discrimination Brain Function".

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. Martin, N., & Brownell, R. (2011). Expressive one-word picture vocabulary test (4th ed.). Novato: Academic Therapy Publications.
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.
Academic Achievement	
Reading	McGrew, K. S., & Woodcock, R. W. (2018). Woodcock-Johnson IV Tests of Achievement. Rolling Meadows, IL: Riverside Publishing.
Reading Comprehension	McGrew, K. S., & Woodcock, R. W. (2018). Woodcock-Johnson IV Tests of Achievement. Rolling Meadows, IL: Riverside Publishing.
Math	McGrew, K. S., & Woodcock, R. W. (2018). Woodcock-Johnson IV Tests of Achievement. Rolling Meadows, IL: Riverside Publishing.
Executive Function & Self-Regulation	
Pencil Tap	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 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
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.
Flanker	National Institutes of Health and Northwestern University (2006-2023). NIH Toolbox® for Assessment of Neurological and Behavioral Function Administrator's Manual. NIHToolbox.org. Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH Toolbox for assessment of neurological and behavioral function. Neurology, 80(11 Suppl 3), S2–S6. https://doi.org/10.1212/WNL.0b013e3182872e5f Weintraub S, Bauer PJ, Zelazo PD, Wallner-Allen K, Dikmen SS, Heaton RK, Tulsky DS, Slotkin J, Blitz DL, Carlozzi NE, Havlik RJ, Beaumont JL, Mungas D, Manly JJ, Borosh BG, Nowinski CJ, Gershon RC. I. NIH Toolbox Cognition Battery (CB): introduction and pediatric data. Monogr Soc Res Child Dev. 2013 Aug;78(4):1-15. doi: 10.1111/mono.12031. PMID: 23952199; PMCID: PMC3954750.
Working Memory	Wechsler, D. (2012). Wechsler Preschool and Primary Scale of Intelligence Fourth Edition (WPPSI-IV). San Antonio, TX: The Psychological Corporation.
	Wechsler, D. (2014). Wechsler Intelligence Scale for Children—Fifth Edition (WISC-V). Bloomington, MN: Pearson.
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. Achenbach, T. M., McConaughy, S. H., Ivanova, M. Y., & Rescorla, L. A. (2011). Manual for the ASEBA Brief Problem Monitor (BPM). Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.: ASEBA.
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.
Visual Processing and Abstract Spatial Perception	
Visual processing and abstract spatial perception	Wechsler, D., Naglieri, J. A. (2006). Wechsler Nonverbal Scale of Ability. San Antonio, TX: Pearson.
Fluid and Perceptual Reasoning	
Fluid Reasoning	Wechsler, D. (2012). Wechsler Preschool and Primary Scale of Intelligence Fourth Edition (WPPSI-IV). San Antonio, TX: The Psychological Corporation.
Perceptual Reasoning	Wechsler, D. (2011). Wechsler Abbreviated Scale of Intelligence–Second Edition (WASI-II). San Antonio, TX: NCS Pearson.
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.

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., Halpern-Meekin, S., Fox, N. A., & Noble, K. G. (2022). The impact of a poverty reduction intervention on infant brain activity. Proceedings of the National Academy of Sciences, 119(5), e2115649119. https://doi.org/10.1073/pnas.2115649119

Age-6 Resting Brain Function

Brito, N. H., Troller-Renfree, S. V., Leon-Santos, A., Isler, J. R., Fifer, W. P., & Noble, K. G. (2020). Associations among the home language environment and neural activity during infancy. Developmental Cognitive Neuroscience, 43, 100780. https://doi.org/10.1016/j.dcn.2020.100780

Brito, N. H., Werchan, D., Brandes-Aitken, A., Yoshikawa, H., Greaves, A., & Zhang, M. (2022). Paid maternal leave is associated with infant brain function at 3 months of age. Child Development, 93(4), 1030–1043. https://doi.org/10.1111/cdev.13765

Cantiani, C., Piazza, C., Mornati, G., Molteni, M., & Riva, V. (2019). Oscillatory gamma activity mediates the pathway from socioeconomic status to language acquisition in infancy. Infant Behavior and Development, 57, 101384. https://doi.org/10.1016/j.infbeh.2019.101384

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

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal Of Research on Educational Effectiveness* , *5* (2), 189-211.

Maguire, M. J., & Schneider, J. M. (2019). Socioeconomic status related differences in resting state EEG activity correspond to differences in vocabulary and working memory in grade school. Brain and Cognition, 137. https://doi.org/10.1016/j.bandc.2019.103619

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.

Rockers, P. C., Leppänen, J. M., Tarullo, A., Coetzee, L., Fink, G., Hamer, D. H., Yousafzai, A. K., & Evans, D. (2023). Evaluation of a community health worker home visit intervention to improve child development in South Africa: A cluster-randomized controlled trial. PLOS Medicine, 20(4), e1004222. https://doi.org/10.1371/journal.pmed.1004222

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., Halpern-Meekin, S., Fox, N. A., & Noble, K. G. (2022). The impact of a poverty reduction intervention on infant brain activity. Proceedings of the National Academy of Sciences, 119(5), e2115649119. https://doi.org/10.1073/pnas.2115649119

	Vanderwert, R. E., Marshall, P. J., Nelson, C. A., 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), e11415. https://doi.org/10.1371/journal.pone.0011415 Vanderwert, R. E., Zeanah, C. H., Fox, N. A., Nelson, C. A., & III. (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
Task-Related Brain	
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. Journal of visualized experiments: JoVE, (83). Davenport, M. D., Tiefenbacher, S., Lutz, C. K., Novak, M. A., & Meyer, J. S. (2006). Analysis of endogenous cortisol concentrations in the hair of rhesus macaques. General and comparative endocrinology, 147(3), 255-261. Phillips, R., Kraeuter, A. K., McDermott, B., Lupien, S., & Sarnyai, Z. (2021). Human nail cortisol as a retrospective biomarker of chronic stress: A systematic review. *Psychoneuroendocrinology* , *123* , 104903.
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	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 Health Condition	
Diagnosis of Health Condition	BFY Study PIs
Diagnosis of Developmental Condition	
Diagnosis of Developmental Condition	BFY Study PIs
Special Services (IEP)	Adapted from the *School Enrollment and Expectations* (SEE) section of the 2021 Panel Study of Income Dynamics (PSID) Child Development Supplement: Beaule, A. Campbell, F., Insolera, N., Juska, P., McAloon-Fernandez, R., McGonagle, K., Mushtaq, M., Simmert, B., & Warra, J. (2023). PSID-2021 Main Interview User Manual: Release 2023. Institute for Social Research, University of Michigan. https://psidonline.isr.umich.edu/data/Documentation/UserGuide2021.pdf
Child Epigenetic Pace of Aging	
Methylation pace of aging	Belsky, D. W., Caspi, A., Houts, R., Cohen, H. J., Corcoran, D. L., Danese, A., Harrington, H., Israel, S., Levine, M. E., Schaefer, J. D., Sugden, K., Williams, B., Yashin, A. I., Poulton, R., & Moffitt, T. E. (2015). Quantification of biological aging in young adults. Proceedings of the National Academy of Sciences, 112(30), E4104–E4110. https://doi.org/10.1073/pnas.1506264112 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 Higgins-Chen, A. T., Thrush, K. L., Wang, Y., Minteer, C. J., Kuo, P.-L., Wang, M., Niimi, P., Sturm, G., Lin, J., Moore, A. Z., Bandinelli, S., Vinkers, C. H., Vermetten, E., Rutten, B. P. F., Geuze, E., Okhuijsen-Pfeifer, C., van der Horst, M. Z., Schreiter, S., Gutwinski, S., … Levine, M. E. (2022). A computational solution for bolstering reliability of epigenetic clocks: Implications for clinical trials and longitudinal tracking. Nature Aging, 2(7), 644–661. https://doi.org/10.1038/s43587-022-00248-2

	Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., & Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging* , *10* (4), 573–591. https://doi.org/10.18632/aging.101414 Lu, A, Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* , *11* (2), 303–327. https://doi.org/10.18632/aging.101684
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 Hunsberger M, O'Malley J, Block T, Norris JC. Relative validation of Block Kids Food Screener for dietary assessment in children and adolescents. Matern Child Nutr. 2012:1–11. doi:10.1111/j.1740-8709.2012.00446.x.
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.

Appendix Table 12: 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, 6, 8	**Age 1-4:** 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).*changes to "in the past 12 months" for surveys at ages 2 through 4 9. needed medical or dental care and did not get it? (0=no; 1=yes) *item 9 dropped at age 4 owing to survey time constraint. **Age 6,8:** Additive index of dichotomous variables (higher score=more stress); expect less stress for high-cash group. In the past 12 months: 1. have you ever missed a rent or mortgage payment? (1: Yes; 0: No) 2. did you ever miss a payment for oil, gas, water, or electricity? (1: Yes; 0: No/Not applicable, do not pay or does not have these utilities) 3. was your gas, water, or electricity ever shut off for nonpayment? (1: Yes; 0: No) 4. have you ever been forced to leave or were evicted from your home? (1: Yes; 0: No) 5. were you ever hungry, but didn't eat because you couldn't afford enough food? (1: Yes; 0: No) 6. did you receive free food or meals? (1: Yes; 0: No)
Maternal Hardship	BFY PIs			6, 8	One item: 1.In the last 12 months, have you experienced an emergency need for extra money that would amount to $500 or more? (1: Yes; 0: No) 1a. [If yes] Were you able to manage it? (Would you say: 1. Easily with no worries; 2. With some stress; 3. With a lot of difficulty; 4. Not at all, couldn't manage it) Expect less hardship for high-cash group.We will create an indicator for " Had an emergency need for extra money of $500 and more dollars and managed it with some stress or more" — Hardship Indicator= (item #1=yes and item #2 >= 2)
Household Poverty rate	US Census Bureau			1, 2, 3, 4, 6, 8	Measured using the Census Bureau's poverty thresholds by size of family and number of children. Age 6/8: Expect lower poverty rate for high-cash group.
Index of food insecurity*	Economic Research Service, USDA, 2012			1, 2, 3, 4	Additive index of 6 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. If yes to (3), how often? (0: only one or two months; 1: almost every month or some months) 5. Eat less than should (0:No; 1: Yes) 6. Hungry^{+} (0:No; 1: Yes)

Appendix Table 12: 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)
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 9. LIHEAP / heat/AC assistance* *Indicates benefits that were not asked about at age 3. Note: Age 4 benefit index was not pre-registered because of the availability of administrative records for some of the benefits
Mother's Labor Market and Education Participation					
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 education attainment	Current Population Survey			6,8	Age 6: Dichotomous variable indicating that mother participated in education in the past 12 months plus any degrees received. Expect higher educational attainment in high-cash gift group. Age 8: Lifetime grades/degrees completed.
Mother's training attainment	Current Population Survey			6, dropped at age 8	Age 6: Dichotomous variable indicating that mother participated in job training activities in the past 12 months. Expect more training in high-cash gift group. Age 8: Dropped owing to combination of training and education to one question of lifetime grades/degrees (above)

Appendix Table 12: 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)
Mother's Labor Market Participation	Current Population Survey			4, 6, 8	Dichotomous variable indicating whether mother is participating in the labor market using the item "do you currently work for pay?" We don't register group differences on employment and earnings impacts because of two offsetting arguments: i) income effects from our payments (our clawback arrangements with states preclude substitution effects) could lead mothers to opt for fewer work hours and possibly dropping out of the labor force altogether; or ii) our payment may enable the mothers in our study to engage in training activities and/or afford center-based child care that eventually boost earning and employment.
Maternal Earnings	PSID			4, 6, 8	Mother's Earnings in the previous calendar year. We don't register group differences on employment and earnings impacts because of two offsetting arguments: i) income effects from our payments (our clawback arrangements with states preclude substitution effects) could lead mothers to opt for fewer work hours and possibly dropping out of the labor force altogether; or ii) our payment may enable the mothers in our study to engage in training activities and/or afford center-based child care that eventually boost earnings
Child-Focused Expenditures					
Index of child-focused expenditures (since birth)	Lugo-Gil, Yoshikawa, 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	Lugo-Gil, Yoshikawa, 2006, and Schild et al., 2003			1, 2, 3, 4, 6, 8	Continuous dollar amount of age-relevant items. Past 30 days, total $ amount spent on... **Age 1:** 1. books 2. toys 3. clothes 4. diapers 5. videos; **Age 2-4:** 1. books 2. toys 3. clothes 4. activities 5. videos. **Age 6/8:** 1. clothing 2. electronic devices 3. toys 4. school materials 5. activities. 6. medical expenses 7. child care 8. travel Expect higher spending for high-cash group.
Index of expenditures on all children in the household, including target child	Schild et al., 2003			6, 8	Past 30 days, total $ amount spent on all children in household. Expect higher spending for high-cash group. 1. clothes 2. electronic devices 3. toys 4. school materials 5. activities 6. medical expenses 7. child care 8.travel
Cost of paid child care	National Study of Early Care and Education			1, 2, 3, 4	Out of pocket spending on child care last month. 1. altogether, about how much money did you spend out-of-pocket on all of [CHILDNAME]'s child care arrangements last month? Note: Age 4: dropped wording "out-of-pocket"

Appendix Table 12: 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)
Use of center-based care	National Study of Early Care and Education			1	1. Has child spent any time in childcare or day care? (Y/N)
				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 9 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	Dichotomous indicator of whether the mother has ever been homeless or in a group shelter (age 1 "since birth", age 2-3 "in the past 12 months"): 0: No 1: Yes
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, 6, 8	# of residents below poverty line in census tract divided by total number of residents in census tract. Age 6/8: Expect less neighborhood poverty in the high-cash group.

Appendix Table 12: 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)
Family and Maternal Perceived Stress					
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 (originally also registered for age 4 and then dropped)	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) Note: Index dropped from age 4 survey owing to time constraints

Appendix Table 12: 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 Happiness and Optimism					
Global happiness	The General Social Survey from NORC			1, 2, 3, 6, 8	One-item with 3-point response scale. **Age 1-3:** "Taken altogether, how happy are you these days?" (0: not happy; 1: pretty happy; 2: very happy) One-item with 3-point response scale. Expect higher score for high-cash group. **Age 6/8:** "Taken all together, how would you say things are these days, would you say that you are 1. very happy 2. pretty happy or 3. not too happy?" Expect higher score for high-cash group.
Maternal Agency	Snyder et al., 1991	alpha: .86 test-retest: .81		1, 2, 3	The Adult 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	At age 1, we attempted to collect maternal hair cortisol for all in-person visits, prior to the onset of the pandemic (when data collection became limited to phone-based survey administration only). This resulted in a hair sample being collected from 409 of the 605 mothers who participated in an in-person visit, with large racial and ethnic differences in willingness to provide a sample. At age-4, we attempted to improve collection rates following focus groups and the development of informational videos. However, the first several months of data collection again revealed large racial and ethnic differences in willingness to provide a hair sample, due to both cultural and practical reasons. Because of the large amounts of non-random missing data, which would both compromise our statistical power and limit the generalizability of any findings, we dropped hair cortisol from the age-4 data collection procedures on October 25, 2022.
Maternal Executive Function					
Executive Function	Carlson, 2017; Carlson & Zelazo 2014; Reflection Sciences 2021	ICC=0.69		4, 6 dropped at 8	Minnesota Executive Function Scale (MEFS). Age 6: We expect a higher value of the maternal MEFS standardized score for the high-cash gift group mothers than the low-cash gift mothers. Higher scores indicate better performance. Age 8: Maternal MEFS will not be administered at Age 8 because of time and burden concerns on mothers. The measure has also not proven to work well in prior data collections.

Appendix Table 12: 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 Attentional Resources					
Scarcity-Primed Inhibitory Control and Attention	Gershon et al., 2013; Slotkin et al., 2012; Zelazo et al., 2013	Test-retest: English 3-15 years: ICC=0.95; 20-85 years: ICC=0.83. Spanish 18-85 years: ρ=.65	6 dropped at 8		NIH Toolbox Flanker Inhibitory Control and Attention Test. Age-corrected standardized score. We expect a higher value of the maternal Flanker score for the high-cash gift group mothers than the low-cash gift mothers. Higher scores indicate better performance. Prior to the Flanker administration, respondent/mother will be asked to listen and reflect for a minute on these questions: "Imagine that an unforeseen event requires of you an immediate $1,000 expense. Are there ways in which you may be able to come up with that amount of money on a very short notice? How would you go about it? How stressful would it be to manage this?" Age 8: Maternal Flanker was not in the original study design as a measure to capture family processes as proposed in the NIH renewal grant. It was administered at Age 6 as a potential expansion of capturing maternal attention. We will not be administering it at Age 8.
Maternal Mental Health					
Index of maternal depression	Kroenke & Spitzer, 2002			1, 2, 3, 4, 6, 8	PHQ-8: additive index of 8 items (0: not at all; 1: several days; 2: more than half of days; 3: every day) Age 6/8: We expect less depression in the high- as opposed to low-cash group. 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, 6, 8	GAD-7: additive index of 7 items (0: not at all; 1:several days; 2: more than half the days; 3: nealy every day) Age 6/8: We expect less anxiety in the high- as opposed to low-cash group. 1. nervous and anxious 2. no control worrying 3. worrying too much 4. no relaxing 5. restless 6. annoyed and irritable 7. afraid

Appendix Table 12: 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 Substance abuse[x]					
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 Wellbeing			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: rarely; 3: sometimes; 4: often; 5: always)

Appendix Table 12: 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	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	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, 6, 8	One item with 5-point response scale "overall, how would you describe your health…" (1:poor - 5:excellent) Age 6/8: Expect better health for high-cash group
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, 6, 8	**Age 4:** Measured by CDC BMI percentile scales **Age 6, 8:** Dichotomous measure based on CDC BMI percentile scales We expect to see a reduced percentage of overweight or obese (BMI greater than or equal to 25) mothers in the high-cash gift group compared to the low-cash gift group. We will report mean BMI of the two groups in descriptive analyses.

Appendix Table 12: 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)
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
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 the total of four sub-scales (affection, responsiveness, encouragement and teaching) with responses ranging from 0: absent, 1: barely, 2: clearly. The total composite score is preregistered. Exploratory analyses will examine differences across the subscales, and factor analysis will be used to confirm the extent to which the four subscales best fit the data. Parent child interaction task and script adapted from the NICHD Study of Early Child Care and Youth Development.
Index of mother's positive parenting behaviors	(Aran et al., 2022; Biringen, 2008; Biringen et al., 2014; Darling & Steinberg, 2017; Leyva et al., 2017; Leyva et al., 2019)	ICCs for inter-rater reliability range between .76-.92		6 dropped at 8	Block play and a grocery shopping game (12 minutes). Grocery shopping game task and script adapted from Leyva et al., 2017 and Leyva et al., 2019. Coding: Measured using Emotional Availability scale (EAS) - 4th Edition middle childhood/youth version (Biringen et al., 2008; Biringen et al., 2014) coding of parenting behaviors from a total of four maternal sub-scales (sensitivity, structuring, non-intrusiveness, and non-hostility) with responses scored on a 7-point scale across the two PCI tasks. The additive total composite score will be pre-registered. We hypothesize a higher composite score for the high-cash gift group relative to the low-cash gift group. Prior studies have reported a unidimensional structure of the maternal dimensions of the EAS data (Aran et al., 2022); although most evidence includes samples of younger children. Earlier work by the authors of the EA scales, using a prior version of the EAS, suggested two distinct factors of maternal style, an affective and a control one (see Biringen et al., 2014). Similarly, literature on parenting practices with older children and adolescents support the idea of distinct parenting dimensions related to parental support (i.e., warmth, responsiveness) and parental control (i.e., intrusiveness) (Darling & Steinberg, 2017). Given the limited empirical evidence on the dimensionality of the EAS with children aged six, we will first conduct confirmatory factor analysis in the full sample using SEM to assess whether the measure is best represented by a single latent factor, as expected. If the goodness-of-fit statistics (RMSEA, CFI, TFI, SRMR) do not indicate support of a single factor structure, we will then perform split-half exploratory factor analysis (principal component analysis with oblique rotation) to determine how many factors underlie the EA parenting measure. Following the results of our exploratory analysis, we will conduct confirmatory factor analysis with a revised number of factors and assess fit statistics. Factor scores derived from this factor analysis will be used in exploratory analyses. We dropped parent child interaction at age 8 primarily for cost reasons but also agreed that it was less age-appropriate.

Appendix Table 12: 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 language quantity	(Anderson et al., 2021; Rowe, 2008; Rowe, 2012)			6 dropped at 8	Measured by the total number of word tokens (adult word count) in maternal speech during the parent-child interaction tasks (Anderson et al 2021; Rowe, 2008; Rowe, 2012). Coded through transcripts at the utterance level. We hypothesize a higher adult word count for mothers in the high-cash gift group relative to the low-cash gift group. We dropped parent child interaction at age 8 primarily for cost reasons but also agreed that it was less age-appropriate.
Maternal language quality	(Anderson et al., 2021; Cristofaro & Tamis-LeMonda, 2012; Luo et al., 2022; Rowe, 2008; Rowe, 2012)			6 dropped at 8	Measured through three indicators of quality: Mean length of utterance (MLU; average number of morphemes per utterance), word types (number of different word roots produced) as indicators of language complexity and diversity, respectively (Anderson et al 2021; Rowe, 2008; Rowe, 2012), and the proportion of utterances that are wh- questions –What, Where, When, Which, Why, Who, and How (referential and inferential questions)– (Luo et al., 2022; Cristofaro & Tamis-LeMonda, 2012) in maternal speech during the parent-child interaction tasks. Language quality variables will be coded through transcripts at the utterance level. We will standardize each variable and run confirmatory factor analysis in the full sample to assess whether the measure is best represented by a single latent factor of language quality. If fit statistics (RMSEA, CFI, TFI, SRMR) do not indicate support of a single factor structure, we will use each individual measure for analyses. We hypothesize higher language quality scores, either the composite factor or the individual scores, for the high-cash gift group relative to the low-cash gift group. We dropped parent child interaction at age 8 primarily for cost reasons but also agreed that it was less age-appropriate.

Appendix Table 12: 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)
Epigenetic Pace of Aging					
Methylation pace of aging	Belsky et al., 2015; Belsky et al., 2020; Belsky et al., 2022; Higgins-Chen et al., 2022; Levine et al., 2018, Lu et al., 2019			4, 6	**Age 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. OSF preregistration link: https://osf.io/ahv2p/overview **Age 6:** Expect slower DunedinPACE- pace of aging in high-cash group: Methylation pace of aging was developed from DNA-methylation analysis of Pace of Aging in the Dunedin Study birth cohort (Belsky et al.,2022).. Pace of Aging is a composite phenotype derived from analysis of longitudinal change in 18 biomarkers of organ-system integrity (Belsky et al.,2015). Increments of methylation pace of aging correspond to "years" of physiological change occurring per 12-months of chronological time. We will also report GrimAge Acceleration and PhenoAge Acceleration as exploratory analyses. GrimAge represents a DNA-methylation metric designed to predict morbidity and mortality (Lu et al., 2019). Briefly, the initial phase entailed the computation of models incorporating physiological indicators, age, sex, and smoking history, with the objective of optimizing mortality prediction within the Framingham Heart Study Offspring cohort (Lu et al., 2019).GrimAge will be based on principal components of DNA-methylation to bolster reliability (Higgins-Chen et al.,2022) and residualized for chronological age derived from sample receipt age to reflect accelerated biological age. PhenoAge is conceptualized on the foundation of physiological markers and chronological age, which are subsequently employed to model a novel sample derived from DNA methylation, culminating in the establishment of a definitive DNA methylation clock (Levine et al., 2018). This metric exemplifies the age, measured in years, at which the average mortality risk in the NHANES III cohort aligns with the mortality risk as forecasted by the PhenoAge algorithm. PhenoAge will be based on principal components of DNA-methylation to bolster reliability (Higgins-Chen et al.,2022) and residualized for chronological age derived from sample receipt age to reflect accelerated biological age. Please also see OSF https://osf.io/ju7ra/overview **Age 8:** We hypothesize slower pace of aging, as measured by DunedinPACE (Belsky et al., 2022), in the high-cash gift group, relative to the low-cash gift group. We will also report GrimAge Acceleration (Lu et al 2019; described above). OSF preregistration link to come.

Appendix Table 12: 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 DNA Methylation					
DNA methylation	McCartney et al, 2022			4	**Age 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. OSF preregistration link: https://osf.io/ahv2p/overview
Frequency of Parent Child Activity					
Self-Report of Parent-child activities	Rodriguez & Tamis-LeMond a 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
				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
				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. Activities are: read books, tell stories, play game/build something, pretend play, learning activities, screen activities. 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).
	Rodriguez & Tamis-LeMond a 2011; BFY PIs			6, 8	Additive index of activities where the number of days reported doing the activity are multiplied by the number of minutes on a given day. We excpect higher amount for high cash group. Activities are: read books, tell stories, play together, pretend play, physical games outside, learning activities, watch tv or videos. Age 8 additional items: household chores, nature activities, and playing together building things like Legos/dominos/construction set. 1. How many days did you participate in [activity]? (1: no days; 2: 1-2 days; 3: 3-5 days; 4: 6-7 days) 1a. On those days, how many minutes do you do typically do this per day? (1: 15 minutes or less; 2: 15-30 minutes; 3: more than 30 minutes).

Appendix Table 12: 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)
Child meal and sleep routine index	Study PIs			4, 6, 8	Additive index of 2 survey items (higher score=more routines); Age 6/8: expect higher score for high-cash group. In past week... **Age 4:** 1. eat meals together (0: 0 days; 1: 1+ days) **Age 6/8:** 1. eat meals together (0: 0 days; 1: 1-2 days; 2: 3-5 days; 3: 6-7 days **Ages 4/6/8:** 2. had regular bedtime (0: no; 1: yes) 2a. If yes, indicate bedtime
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).
Parent-Child Conflict Tactics Scale (CTSPC)	Straus et al., 1998	alpha = .70 to .75		6, dropped at 8	Subscales: Nonviolent Discipline, Psychological Aggression, and Corporal punishment and 1 item from the Severe Assault (physical abuse) subscale. Total 14 items about conflict strategies with child and harsh discipline are asked about the past year. Min: 0 Max: 98 We expect a reduction of harsh discipline in the high-cash group compared to the low-cash group. 1.Once 2.Twice 3.3-5 times 4.6-10 times 5.11-20 times 6.More than 20 times 7.This has never happened. In the past year have you.... 1. Explained to [CHILDNAME] why something he or she did was wrong? 2. Put [CHILDNAME] in "time out" (or sent [CHILDNAME] to his or her room)? 3. Shook [CHILDNAME] ? 4. Hit [CHILDNAME]on the bottom with something like a belt, hairbrush, a stick, or some other hard object? 5. Gave him/her something else to do instead of what he or she was doing? 6. Shouted, yelled, or screamed at [CHILDNAME]? 7.Spanked him or her bottom on the bottom with your bare hand? 8.Swore or cursed at him or her? 9. Said you would send him or her away or would kick him or her out of the house? 10. Threatened to spank or hit him or her but did not actually do it? 11. Slapped him or her on the hand, arm, or leg? 12. Took away privileges or grounded him or her? 13. Pinched him or her? 14. Called him or her dumb or lazy or some other name like that? Age 8: We dropped this measure at age 8 for timing/length and because there was little variation in mothers' responses

Appendix Table 12: 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)
Parent School Involvement					
Parent School Involvement	Tourangeau et al., 2019			8	We expect a positive effect on this additive index. During this school year, have you or any other adult members of your household… 1. attended an open house or a back-to-school night or similar type of event at [CHILDNAMEF]'s school? (1: Yes; 0: No) 2. attended a meeting of PTA, PTO, or Parent-Teacher Organization at [CHILDNAMEF]'s school? (1: Yes; 0: No) 3. gone to a regularly scheduled parent-teacher conference with [CHILDNAMEF]'s teacher or meeting with [CHILDNAMEF]'s teacher? (1: Yes; 0: No) 4. served as a volunteer in [CHILDNAMEF]'s classroom or elsewhere in the school? (1: Yes; 0: No) 5. Reached out or contacted [CHILDNAMEF]'s teacher or school for any reason having to do with [CHILDNAMEF]'s (1: Yes; 0: No) 6. Participated in fundraising for [CHILDNAMEF]'s school? (1: Yes; 0: No) 7. Attended a school or class event, such as a play, sports event, or science fair at [CHILDNAMEF]'s school? (1: Yes; 0: No)
Enrichment Activities					
Child enrichment	Tourangeau et al., 2019			6, 8	We expect a positive effect on this additive index. In the past month, has anyone in your family... 1. visited a library with [CHILD] (1: Yes; 0: No;) 2. gone to a play, concert, or other live show with [CHILD]? (1: Yes; 0: No) 3. visited an art gallery, museum, or historical site with [CHILD]? (1: Yes; 0: No) 4. visited a zoo, aquarium, or petting farm with [CHILD]? (1: Yes; 0: No) Additional items added at age 8: 5. visited a street fair, festival, parade or neighborhood celebration with [CHILDNAMEF]? (1: Yes; 0: No) 6. visited a park, amusement part, playground, pool, lake, or beach with [CHILDNAMEF]? (1: Yes; 0: No)
Child lessons, sports, etc. activity participation	Adapted from Tourangeau et al., 2019			6, 8	We expect a positive effect on this additive index In the past year, outside of school hours, has [CHILD] ever participated in ... 1. organized athletic activities, like basketball, soccer, baseball, dance, or gymnastics? (1: Yes; 0: No) 2. academic activities, like tutoring, or math or language lessons? (1: Yes; 0: No) 3. art or carfts classes or lessons, for example, painting, drawing, or sculpture? (1: Yes; 0: No) 4. music classes or lessons like piano, other instruments, singing or choirs? (1: Yes; 0: No)

Notes. A 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. The targeted age for each data collection wave is around the child's birthday, i.e. at 12 months, 24 months, 36 months, 48 months, 72 months, and 96 months.

Appendix Table 12: 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)

+ 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 6 from the index of food insufficiency ("hungry"), and item 11 from the relationship quality index ("partner threatened or hurt your child/children? "). These indices were therfore comprised of one less item at age 1.

×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 age 4. 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, spanking discipline strategy, anxiety.

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		
Household Economic Hardship		
Index of economic stress	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica* , 75(1), 83-119.	http://www2.nber.org/mtopublic/
Maternal Hardship	BFY Study PIs	
Index of food insecurity	https://www.ers.usda.gov/media/8282/short2012.pdf	
Household poverty rate	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.	
Social Services Receipt		
Number of Benefits received by mother	BFY 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	Current Population Survey, retrieved from: https://www.census.gov/programs-surveys/cps/technical-documentation/questionnaires.html	
Mother's education and training attainment		
Maternal 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.	Schild, J., Collyer, S. M., Garner, T., Kaushal, N., Lee, J., Waldfogel, J., & Wimer, C. T. (2023). Effects of the Expanded Child Tax Credit on Household Spending: Estimates Based on US Consumer Expenditure Survey Data (No. w31412). National Bureau of Economic Research. https://www.bls.gov/osmr/research-papers/2023/pdf/ec230010.pdf
Cost of paid child care	National Study of Early Care and Education	
Use of center-based care		
Housing and Neighborhoods		
Index of perceptions of neighborhood safety		
Index of housing quality	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica* , 75(1), 83-119.	
Residential mobility		

Homelessness		
Neighborhood poverty		
Family and Maternal Perceived Stress		
Perceived stress	Cohen, S., & Williamson, G. M. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan & S. Oskamp (Eds.), The social psychology of health: Claremont Symposium on Applied Social Psychology (pp. 31– 67). Newbury Park, CA: Sage. 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, University of Chicago, 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, D. W., Caspi, A., Houts, R., Cohen, H. J., Corcoran, D. L., Danese, A., Harrington, H., Israel, S., Levine, M. E., Schaefer, J. D., Sugden, K., Williams, B., Yashin, A. I., Poulton, R., & Moffitt, T. E. (2015). Quantification of biological aging in young adults. Proceedings of the National Academy of Sciences, 112(30), E4104–E4110. https://doi.org/10.1073/pnas.1506264112	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	Higgins-Chen, A. T., Thrush, K. L., Wang, Y., Minteer, C. J., Kuo, P.-L., Wang, M., Niimi, P., Sturm, G., Lin, J., Moore, A. Z., Bandinelli, S., Vinkers, C. H., Vermetten, E., Rutten, B. P. F., Geuze, E., Okhuijsen-Pfeifer, C., van der Horst, M. Z., Schreiter, S., Gutwinski, S., … Levine, M. E. (2022). A computational solution for bolstering reliability of epigenetic clocks: Implications for clinical trials and longitudinal tracking. Nature Aging, 2(7), 644–661. https://doi.org/10.1038/s43587-022-00248-2
	Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., & Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging* , *10* (4), 573–591. https://doi.org/10.18632/aging.101414	Lu, A, Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* , *11* (2), 303–327. https://doi.org/10.18632/aging.101684
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 Cognitive ResourcesResources		
MEFS	Carlson, S. M., & Zelazo, P. D. (2014). Minnesota Executive Function Scale: Test Manual. St. Paul, MN: Reflection Sciences, Inc. Reflection Sciences (2021). Minnesota executive function scale technical report. Reflection Sciences. https://reflectionsciences.com/wp-content/uploads/MEFS-Technical-Report-July-2021.pdf	Carlson, S. M. (2017). *Minnesota Executive Function Scale* : *Technical Report, v. 2* . St. Paul, MN: Reflection Sciences, Inc.

Flanker	Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH Toolbox for assessment of neurological and behavioral function. Neurology, 80(11 Suppl 3), S2–S6. https://doi.org/10.1212/WNL.0b013e3182872e5f Zelazo PD, et al. (2013) II. NIH Toolbox cognition battery (CB): Measuring executive function and attention. Monogr Soc Res Child Dev 78:16–33.	Slotkin J, et al. (2012) NIH Toolbox Scoring and Interpretation Guide (Northwestern University and National Institutes of Health, Chicago). Zelazo PD, et al. (2013) II. NIH Toolbox cognition battery (CB): Measuring executive function and attention. Monogr Soc Res Child Dev 78:16–33.
Maternal Mental Health		
Index of maternal depression	University and National Institutes of Health, Chicago).	
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 **Opioid use**	Kling, J.R., Liebman, J.B., Katz, L.F. (2007). Experimental analysis of neighborhood effects. *Econometrica* , 75(1), 83-119.	
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		
Frequency of Arguing	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	
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		
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*
	Biringen, Z., Derscheid, D., Vliegen, N., Closson, L., & Easterbrooks, M. A. (2014). Emotional availability (EA): Theoretical background, empirical research using the EA Scales, and clinical applications. Developmental Review, 34(2), 114–167.	Biringen, Z. (2008). The Emotional Availability (EA) Scales Manual. 4.1 edition, MiddleChildhood/Adolescent Version. Boulder, CO: emotionalavailability.com
	Leyva, D., Tamis-LeMonda, C. S., Yoshikawa, H., Jimenez-Robbins, C., & Malachowski, L. (2017). Grocery games: How ethnically diverse low-income mothers support children's reading and mathematics. Early Childhood Research Quarterly, 40, 63–76. https://doi.org/10.1016/j.ecresq.2017.01.001	Aran, P., Lewis, A. J., Watson, S. J., MacMillan, K. K., Power, J., & Galbally, M. (2022). A psychometric study of the Emotional Availability Scales: Construct validity and measurement invariance between depressed and nondepressed mother-infant dyads. Psychological assessment, 34(1), 70–81. https://doi.org/10.1037/pas0001067
	Darling, N., & Steinberg, L. (2017). Parenting style as context: An integrative model. *Psychological Bulletin, 113* (3), 487-496.	Leyva, D., Tamis-LeMonda, C. S., & Yoshikawa, H. (2019). What parents bring to the table: Maternal behaviors in a grocery game and first graders' literacy and math skills in a low-income sample. The Elementary School Journal, 119(4), 629– 650. https://doi.org/10.1086/703104
Maternal language quantity	Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development* , *92* (2), 484-501.	Rowe, M. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35,* 185-205. *doi:10.1017/S0305000907008343*

	Rowe, M. (2012). A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development. *Child Development, 83* (5), 1762-1774.	
Maternal language quality	Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development* , *92* (2), 484-501.	Rowe, M. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35,* 185-205. *doi:10.1017/S0305000907008343*
	Rowe, M. (2012). A Longitudinal Investigation of the Role of Quantity and Quality of Child-Directed Speech in Vocabulary Development. *Child Development, 83* (5), 1762-1774.	
	Cristofaro, T. N., & Tamis-LeMonda, C. S. (2012). Mother-child conversations at 36 months and at pre-kindergarten: Relations to children's school readiness. Journal of Early Childhood Literacy, 12(1), 68–97. https://doi.org/10.1177/1468798411416879	Luo, R., Masek, L. R., Alper, R. M., & Hirsh-Pasek, K. (2022). Maternal question use and child language outcomes: The moderating role of children's vocabulary skills and socioeconomic status. Early Childhood Research Quarterly, 59, 109–120. https://doi.org/10.1016/j.ecresq.2021.11.007
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.	BFY Study PIs
Child meal and sleep routine index	BFY Study PIs	
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.	