NBER WORKING PAPER SERIES

THE EFFECT OF A MONTHLY UNCONDITIONAL CASH TRANSFER ON CHILDREN'S DEVELOPMENT AT FOUR YEARS OF AGE: A RANDOMIZED CONTROLLED TRIAL IN THE U.S.

Kimberly Noble
Katherine Magnuson
Greg Duncan
Lisa A. Gennetian
Hirokazu Yoshikawa
Nathan A. Fox
Sarah Halpern-Meekin
Sonya Troller-Renfree
Sangdo Han
Shannon Egan-Dailey
Timothy D. Nelson
Jennifer Mize Nelson
Sarah Black
Michael Georgieff
Debra Karhson

Working Paper 33844
http://www.nber.org/papers/w33844

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2025

Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01HD087384). The content is solely the responsibility of the authors and does not represent the official views of NIH. Additional support was provided by the US Department of Health and Human Services, Office of Planning, Research and Evaluation; Andrew and Julie Klingenstein Family Fund; Annie E. Casey Foundation; Arnold Ventures; Arrow Impact; BCBS of Louisiana Foundation; Bezos Family Foundation, Bill and Melinda Gates Foundation; Bill Hammack and Janice Parmelee, Brady Education Fund; Chan Zuckerberg Initiative (Silicon Valley Community Foundation); Charles and Lynn Schusterman Family Philanthropies; Child Welfare Fund; Esther A. and Joseph Klingenstein Fund; Ford Foundation; Greater New Orleans Foundation; Heising-Simons Foundation; Holland Foundation; Jacobs Foundation; JPB Foundation; J-PAL North America; Lozier Foundation; NYC Mayor's Office for Economic Opportunity; Office of Behavioral and Social Sciences Research-Office of the Director, NIH; Perigee Fund; Robert Wood Johnson Foundation; Robin Hood Foundation; Sherwood Foundation; Valhalla Foundation; Weitz Family Foundation; W.K. Kellogg Foundation; and three anonymous donors. We thank the University of Michigan Survey Research Center for their outstanding work on study recruitment, data collection, and outreach. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Effect of a Monthly Unconditional Cash Transfer on Children's Development at Four Years of Age: A Randomized Controlled Trial in the U.S.
Kimberly Noble, Katherine Magnuson, Greg Duncan, Lisa A. Gennetian, Hirokazu Yoshikawa, Nathan A. Fox, Sarah Halpern-Meekin, Sonya Troller-Renfree, Sangdo Han, Shannon Egan-Dailey, Timothy D. Nelson, Jennifer Mize Nelson, Sarah Black, Michael Georgieff, and Debra Karhson
NBER Working Paper No. 33844
May 2025
JEL No. I3, J13

ABSTRACT

Developmental differences between children growing up in poverty and their higher-income peers are frequently reported. However, the extent to which such differences are caused by differences in family income is unclear. To study the causal role of income on children's development, the Baby's First Years randomized control trial provided families with monthly unconditional cash transfers. One thousand racially and ethnically diverse mothers with incomes below the U.S. federal poverty line were recruited from postpartum wards in 2018-19, and randomized to receive either $333/month or$20/month for the first several years of their children's lives. After the first four years of the intervention (n=891), we find no statistically significant impacts of the cash transfers on four preregistered primary outcomes (language, executive function, social-emotional problems, and high-frequency brain activity) nor on three secondary outcomes (visual processing/spatial perception, pre-literacy, maternal reports of developmental diagnoses). Possible explanations for these results are discussed.

Kimberly Noble
Columbia University
Teachers College
kgn2106@tc.columbia.edu

Katherine Magnuson
University of Wisconsin - Madison
kmagnuson@wisc.edu

Greg Duncan
University of California, Irvine
School of Education
gduncan@uci.edu

Lisa A. Gennetian
Duke University
Sanford School of Public Policy
and NBER
lisa.gennetian@duke.edu

Hirokazu Yoshikawa
New York University
Steinhardt School of Culture,
Education, and Human Development
hiro.yoshikawa@nyu.edu

Nathan A. Fox
University of Maryland, College Park
fox@umd.edu

Sarah Halpern-Meekin
University of Wisconsin - Madison
sarah.halpernmeekin@wisc.edu

Sonya Troller-Renfree
Columbia University
svt2110@tc.columbia.edu

Sangdo Han
Washington University in St. Louis
wshan23@wisc.edu

Shannon Egan-Dailey
Duke University
shannon.dailey@duke.edu

Timothy D. Nelson
University of Nebraska–Lincoln
tnelson3@unl.edu

Jennifer Mize Nelson
University of Nebraska–Lincoln
jnelson18@unl.edu

Sarah Black
Kent State University
sblack38@kent.edu

Michael Georgieff
University of Minnesota
georg001@umn.edu

Debra Karhson
University of New Orleans
dkarhson@uno.edu

A randomized controlled trials registry entry is available at:
https://clinicaltrials.gov/ct2/show/NCT03593356?term=NCT03593356&draw=2&rank=1

Introduction

Early life experience has a profound influence on the developing child (Boyce et al., 2021), and economic resources shape many early experiences (Garcia Coll et al., 1996; Sroufe et al, 1992). Higher family income has been associated with higher scores on assessments of children's language, executive functioning and social-emotional development, as well as differences in brain structure and function (Noble, Hart & Sperber, 2021; Pace et al., 2017; Reardon, 2011). Further, adults raised in households with low incomes are likely to experience a range of unfavorable labor market and health outcomes (NASEM, 2023).

Whether family income itself causes differences in children's development is not clear. Identifying causal impacts is complex, because it is difficult to isolate income from its many correlates, including parental education, community socioeconomic factors, and structural racism (Duncan et al., 2017; Iruka et al., 2022). Evidence on the causal effects of income on child development comes from recent evaluations of programs providing either unconditional cash transfers – cash without constraints – or conditional cash transfers, which condition payments on desired behaviors (Shah & Gennetian, 2024). Only one RCT evaluation has estimated the effects of a monthly unconditional cash transfer to mothers on their children's outcomes. Based in Ecuador, the study found no main effects on early childhood development (Fernald & Hidsrobo, 2011).

Most RCT-based evaluations of conditional cash transfers come from low- and middle-income countries, with payments rewarding behaviors such as school attendance or doctor visits. These studies often document significant, but selective, improvements in children's development, education, and health (Bastagli et al., 2016). Of course, it is hard to know whether impacts of these programs should be attributed to the conditioned behaviors or to the cash itself.

In the U.S., quasi-experimental studies have taken advantage of roll-outs or expansions of programs such as the Earned Income Tax Credit (EITC) and Supplemental Nutrition Assistance Program (SNAP). Taken together, findings suggest that among low-income school-aged children, higher cash benefits predict higher test scores, educational attainment, and adult earnings, as well as better health (NASEM, 2023; McInnis et al., 2024). But the EITC conditions payments on parental employment, and SNAP assistance can be used only to purchase food. Thus, these studies cannot isolate the causal impacts of income per se on children's development. Moreover, most prior studies focus on older children or adolescents, providing no information on children's development during the sensitive early childhood period. The present study uses data from Baby's First Years (BFY), a multi-site randomized controlled trial (Noble, Magnuson, et al., 2021), to estimate the causal impact of four years of monthly unconditional cash transfers on preregistered measures of early childhood development.

Economic models argue that higher income enables parents to invest resources in their children's development (Attanasio et al., 2022). The Family Stress theory argues that family poverty causes material hardship, which in turn causes parental psychological distress and compromises the quality of family relationships and parent-child interactions (Masarik & Conger, 2017). Past work (Gennetian et al., 2024; Magnuson et al., 2024) has shown that the BFY cash transfers have increased family's net income and economic resources, and have led to increased parental investments in children (e.g., greater expenditures on books and toys; more time spent reading and storytelling). However, the transfers did not affect other household expenditures, decrease material hardship or parental stress, or improve mothers' mental health or family relationships.

The first year of monthly cash transfers provided some preliminary evidence of impacts on infants' brain activity (Troller-Renfree et al., 2021), but no group differences in maternal reports of children's health or development were detected across the first three years of life (Hart et al., 2024; Sperber et al., 2023). The present study extends these findings by examining the impacts of four years of monthly unconditional cash transfers on direct assessments of children's cognition and brain activity, and on maternal reports of behavior and developmental diagnoses.

Methods

Participants

Between May 2018 and June 2019, 1,000 mothers were recruited shortly after giving birth in 12 postpartum wards across 4 U.S. metropolitan areas: New York, the greater Omaha metropolitan area, New Orleans, and Minneapolis/St. Paul. Recruitment took place over a 12-month period. During recruitment days, all mothers in the postpartum wards who had given birth within the last two days were approached for possible interest in participating in the study, unless nursing staff indicated to the research team that approaching a particular mother was medically or otherwise inappropriate.

Mothers who expressed interest were administered a short screening interview. Those reporting household incomes below the U.S. federal poverty threshold in the prior calendar year were then offered the opportunity to participate in a multi-year longitudinal study of child development and family life. The full set of eligibility criteria included: 1) being of legal age to provide consent; 2) reporting a household income below the federal poverty line in the calendar year prior to the birth (including the newborn in the household); 3) proficiency in speaking English or Spanish; 4) a singleton birth; 5) infant not admitted to the neonatal intensive care unit; 6) residing in the state of recruitment and likely to continue to reside in the state for the next 12 months, per self-report, and 7) infant to be discharged to the mother's custody. We placed no restrictions on birth parity.

Study Procedures

Written informed consent to participate in a longitudinal study was obtained prior to administration of a baseline interview. At the completion of the interview, mothers were told of the additional opportunity to receive a monthly, unconditional cash transfer (hereafter referred to as a "cash gift"). They were told that they could spend the money however they wished, and that receipt of the funds did not require continued participation in the study. Mothers expressing interest in receiving the cash gifts were randomly assigned to either the high-cash gift group ($333/month; n=400) or the low-cash gift group ($20/month; n=600). To the extent possible, state agencies and legislation ensured that the cash gifts affected neither mothers' eligibility for nor the amount of most public benefits.

As detailed in the Supplemental Materials, randomization occurred within each of the metropolitan areas, with 60% of mothers randomized to the low-cash gift group and 40% to the high-cash gift group. Although neither the participant nor the interviewer was blind to the gift amount, the randomization process was designed so that interviewers could not influence the assigned cash gift level. In subsequent rounds of data collection, interviewers were not reminded (or, in the case of different interviewers, informed) of participants' treatment status.

When mothers were recruited into the study, the cash gifts were promised for the first 40 months of the children's lives. Subsequently, the gifts were extended twice, for a total of 76 months. Following recruitment, mothers were invited to participate in annual in-person or telephone-based data collections around the time of their child's birthday. The present study reports results after families received the monthly cash gifts for about 48 months.

The cash gifts are disbursed via an electronic MasterCard debit card (co-branded with "4MyBaby"; Gennetian et al., 2024). Debit cards were activated in the hospital, with subsequent deposits occurring monthly on the day of the child's birth date. Mothers were notified of their monthly disbursements with a text message. Use of the cash gift has been virtually universal; only five families had not withdrawn funds from the debit card three years post-randomization (Gennetian et al., 2024). A programming error by the company administering the debit cards led 24 families to receive one-time overpayments (see Supplemental Materials).

As shown in the CONSORT diagram in Figure 1, a total of 13,482 mothers were identified for possible recruitment during the year-long recruitment period. Many were not assessed for eligibility because they had been discharged or refused to consent to the screening interview. Roughly half of the remainder did not meet the low-income recruitment criterion. Figure 1 shows how we arrived at the n=1,000-mother baseline recruitment sample.

Age-4 data collection began in July of 2022 and was completed in August of 2023. For this data collection, 984 of the original 1,000 mother-infant pairs remained eligible (there were five maternal deaths, five child deaths, two maternal-child separations, and four instances of maternal incarceration). Additional mothers were not locatable or had withdrawn from the study. Among eligible families, 94.7% of the high-cash gift group and 87.8% of the low-cash gift group provided at least some data for the age-4 data collection. As explained below, the number of valid observations for our seven preregistered assessments of child development ranges from n=637 to n=886. All study procedures were approved by the Institutional Review Boards of Teachers College, Columbia University, and the New York State Psychiatric Institute.

Measures

We preregistered four primary child outcomes: language, executive function, social-emotional development, and resting high-frequency brain activity, as well as three secondary child outcomes: visual processing/spatial perception, pre-literacy skills, and diagnosis of developmental conditions.

Language

Receptive vocabulary was measured using the Receptive One Word Picture Vocabulary Test (ROWPVT; Martin & Brownell, 2011). The monolingual (English) or bilingual (English/Spanish) versions were administered as appropriate; see the Supplemental Materials for details of how this was determined. In the ROWPVT, the experimenter names a word, and the child is instructed to point to the corresponding picture among four choices. For the bilingual administration, the experimenter begins using the child's dominant language, and can switch between languages as needed (e.g., by prompting in the other language if the child provides an incorrect response). Because the monolingual and bilingual versions of the ROWPVT are not normed on the same populations, we derived a "conceptual score" for our primary outcome by summing the raw scores on all individual items that appear on both versions of the test. There are 173 overlapping items on the ROWPVT forms. However, all items are not administered (as the test is normed through adulthood), so to create the conceptual scores, we excluded the 17 non-overlapping items and rescored the assessments. In addition, we report age-standardized receptive vocabulary scores for monolingual and bilingual children as supplemental outcomes (Martin & Brownell, 2011; Supplemental Table 1). These measures have a population average of 100 and standard deviation of 15.

There were 700 valid receptive vocabulary scores. An additional 31 participants attempted this assessment but were excluded from analysis for the following reasons: child's behavior or responses did not yield a valid score (*n*=25); technical error (*n*=4); experimenter error (*n*=1); primary language neither English nor Spanish (*n*=1).

Executive Function

Executive function was measured using the Minnesota Executive Function Scale (MEFS; Carlson, 2017; Carlson & Zelazo, 2014), a direct, tablet-based assessment of executive functioning. Children alternated between sorting pictures by shape and sorting by color with increasing levels of difficulty. We use the MEFS standard score, which is an age-standardized, norm-referenced score with a population average of 100 and standard deviation of 15.

There were 783 valid MEFS scores. An additional 31 participants attempted this assessment but were excluded from analysis for the following reasons: child unable/unwilling to complete assessment (*n*=24); experimenter error (*n*=4); technical error (*n*=3).

Social-Emotional Development

Social-emotional development was measured via maternal report using a shortened version of the Child Behavior Checklist (CBCL; Achenbach et al., 2000). Four subscales of the CBCL were administered: Anxiety/Depression, Aggressive Behavior, Attention Problems, and Emotionally Reactive. For our main preregistered analyses, we calculated a total score by summing raw scores on the four subscales (maximum possible score = 82). CBCL total scores were found to have high internal consistency (41 items; $\alpha = 0.92$). The Emotionally Reactive subscale consists of 9 items ($\alpha = .74$); the Anxious/Depressed subscale consists of 8 items ($\alpha = .62$); the Attention Problems subscale consists of 5 items ($\alpha = .66$); and the Aggressive Behavior subscale consists of 19 items ($\alpha = .88$). In Supplemental Table 1, we report analyses of T-scores and cutoff scores for clinical concern for each subscale. There were 882 valid scores.

Resting Brain Function

Using procedures described in the Supplemental Materials, resting brain activity was measured using high-density electroencephalography (EEG). Because of limitations in power expected with multiple testing adjustments, we followed the approach in Troller-Renfree et al. (2021) and preregistered a single composite of mid-to-high-frequency whole-brain power, summing across the individual single-Hz bins that comprise the alpha, beta, and gamma bands, from 7 to 45 Hz.

We have EEG composite scores for 637 children. An additional 187 participants were excluded for the following reasons: mother did not consent to EEG collection (n=70), child or parent refused to participate after consenting (n=86), technical problems/experimenter error or ill-fitting cap (n=13), visit ended before EEG collection (n=4), a developmental condition prohibiting collection (n=1), poor data quality (n=4), not enough usable data (n=9).

Visual Processing/Spatial Perception

Visual processing and abstract spatial perception were measured by the Matrices subtest of the Wechsler Nonverbal Scale of Ability (Wechsler & Naglieri, 2006). This is a nonverbal assessment where the child is asked to look at an incomplete matrix of images and select the missing image. Age-normed T-scores are available for participants at least 48 months of age; however, most children in our sample were 46 or 47 months old at the time of assessment.[1] Thus, we report treatment impact regression results using raw scores and controls for the child's age at the time of assessment. We have valid scores for 769 children. An additional 19 participants were excluded from analysis because the child's behavior or responses did not yield a valid score.[2]

Pre-Literacy Skills

Pre-literacy skills were measured using The Reading House (Hutton et al., 2019; Hutton et al., 2021). In this assessment, the child is given a board book (*The Reading House*), and the researcher guides the child through a series of questions and prompts to assess emergent literacy

[1] Our initial invitation to participate in the age-4 data collection had been issued when the children were 46 months old, as we had anticipated that locating and persuading participants to take part would be a lengthy process for some.

[2] Originally, IQ scores on the Wechsler Nonverbal Scale of Ability, which are calculated using both Matrices and Recognition subtests, were preregistered as a primary outcome. We began data collection with both, but preliminary analysis of the first 71 participants showed that 21% of participants scored at the floor of the Recognition assessment. We therefore dropped the Recognition subtest from our data collection instrument, precluding us from calculating IQ for subsequent participants. We changed our preregistration to consider the Matrices subtest to be a secondary outcome.

skills (range of possible scores: 0-14). We use raw scores for our main preregistered analyses and provide a supplemental analysis of reading level in Supplemental Table 1.

There were 754 valid Reading House scores. An additional 25 participants attempted this assessment but were excluded from analysis for the following reasons: child's behavior or responses did not yield a valid score (*n*=22); experimenter error (*n*=1); primary language not English or Spanish (*n*=1).

Diagnosis of Developmental Condition

Diagnosis of developmental condition (*n*=881) was collected via maternal report. During the maternal survey, mothers were asked, "Has [child's name] been diagnosed with any developmental condition, like speech delay, autism, or ADHD?" (yes/no). Responses to this question constitute our dichotomous preregistered outcome of diagnosis of developmental condition. Supplemental impact results for speech delay and autism are shown in Supplemental Table 1.

Statistical Power

Our review of the quasi-experimental literature on income effects in childhood led us to expect that our $313 (=$333-$20) monthly payment differential between our high-cash and low-cash gift groups would result in at least a .20 SD differences in the child outcomes (Duncan et al., 2011). For this reason, the study was planned to have a sufficient initial sample size so that, after 20% expected attrition, we would be able to detect an effect size of approximately 0.20 SD. An initial sample size of n=1,000 (and expected n=800 at age 4), divided 40%/60% between high and low cash gift groups, provided 80% statistical power to detect a .20 SD age-4 impact at p <.05 in a two-tailed test. The use of covariates was expected to slightly increase statistical power. Using Bloom's (1995) conversion of standard errors into minimum detectable effect sizes, our regression results confirmed the accuracy of our original power calculations. Because the seven outcomes are conceptually distinct, we did not preregister multiple outcome adjustments.

Statistical Analysis

Following preregistered analytic protocols, intent-to-treat estimates of the cash gift differential on child outcomes were estimated by regressing each preregistered outcome on an indicator for being in the high-cash gift group and a set of covariates. In the case of dichotomous outcomes, a logistic regression was fit to the data, and results are reported as marginal changes in the probability of affirmative responses. All analyses included preregistered covariates measured at baseline (see Table 1 and Supplemental Table 2) as well as the child's age in months. An indicator of the staff member administering the child assessment was also included, except as noted for logistic regressions when they created problems with perfect prediction (see Table notes for details). All analyses were conducted using Stata (Version 18).

Results

Participants

Table 1 presents baseline descriptive statistics by treatment status for all participants in the age-4 analysis sample (n=891). Forty-one percent of mothers self-identified as Hispanic, and 40% self-identified as non-Hispanic Black. Approximately 9% of the sample self-identified as White. On average, mothers were about 27 years old, had completed close to 12 years of schooling, and had between 1 and 2 older children at the time of the birth. Thirty-eight percent reported living with the biological father of the baby at the time of the birth. Virtually all infants were of normal birth weight (Mean=7.1 pounds, SD=1.0) and were born at term (Mean=39.1 weeks, SD=1.2 weeks).

With regard to balance across the two treatment groups, the equivalence of the two groups was assessed with a joint test of orthogonality using a probit model with robust standard errors and site-level fixed effects. Results indicate that we are unable to reject the null hypothesis of joint equivalence (p=.268; bottom panel of Table 1), suggesting the two groups did not statistically differ in baseline characteristics. The standardized mean group difference averaged across all baseline variables was 0.06, well below the maximum recommended value of 0.15. However, the standardized mean group differences for mothers' marital status and self-reported health did exceed the 0.15 threshold. Additional rows at the bottom of Table 1 show sample equivalence across the various sample subsets used in the child outcome analyses. We conclude that, as would be expected from randomization, the high-cash and low-cash gift groups that participated in age-4 data collection were very similar at baseline, across a host of characteristics.

Descriptive information on the primary child development outcomes is presented in the correlation matrix shown in Table 2, Table 3, and Supplemental Table 1. On average, children in both groups were performing approximately 0.2-0.3 SD below the normed average for receptive vocabulary, and about a third to a half-SD below average for executive function. Approximately 14% of children in both groups were reported to have received a developmental diagnosis, consistent with national averages (Hagerman & Houtrow, 2021). Children in both groups were rated as slightly worse than national norms on anxiety/depression, aggressive behavior, attention problems and emotional reactivity.

Table 2 shows that, as expected, cognitive assessments were generally positively correlated with each other, and higher cognitive scores were generally associated with fewer behavior problems and lower likelihood of diagnosis with a developmental condition. Although the high-frequency brain activity composite is not correlated with cognition or behavior, individual frequency bands of brain activity were (Troller-Renfree et al., 2024).

Impacts of the Cash Gift on Primary and Secondary Child Outcomes

Our primary questions were whether children whose mothers received the BFY high-cash gift performed better on direct assessments of language and cognition and had lower scores on maternal reports of problem behavior and developmental delays, compared with children whose mothers received the low-cash gifts. Results in Table 3 show that none of the standardized assessments or maternal reports differed statistically between the high- and low-cash gift groups (columns 3 and 4). Only two of the six estimated coefficients associated with directional hypotheses were in the hypothesized direction (Figure 2).[3]

Supplemental Analyses

We conducted a number of exploratory supplementary analyses to assess the sensitivity of these findings to alternative specifications.

Impacts on Subscales

In some cases, the outcomes we use combined scores across differing forms of the assessment or included multiple subscales or items. In Supplemental Table 1, we provide descriptive statistics and regression results for i) the monolingual and bilingual ROWPVT scale, ii) subscales and "clinical concern" thresholds for the Child Behavior Checklist, iii) a dichotomous indicator of above- or below-average scores on the Reading House, and iv) separate impact estimates for reports of speech delay and autism.

As with the analyses of preregistered outcomes, none of the cash gift group differences for these alternative outcomes is statistically significant, with one exception. With regard to clinically concerning levels of anxiety and depression, the estimate was sizable and statistically

[3] We do not formulate a directional hypothesis for diagnosis of developmental conditions, because of two offsetting possibilities: i) the high-cash gift group may have had better access to services, which may lead to higher rates of diagnosis, and/or ii) the cash gifts may have led to higher scores on cognitive assessments and therefore lower rates of developmental diagnoses.

significant (effect size= .33, p<.04), but the direction of the estimated effect ran contrary to expectations.

Sensitivity to Robust Regression

To ensure that results were not sensitive to the inclusion of possible outlying observations, we repeated the estimations in Table 3 using the robust regression command in Stata (Berk, 1990). Supplemental Table 3 shows that the findings are similar to those presented in Table 3.

An Alternative Measure of Self-Regulation

We gathered information on two experimenter-reported subscales of the Preschool Self-Regulation Assessment – Attentive/Impulse Control and Positive Emotion – and conducted exploratory analyses of impacts on these measures. Estimated effect sizes are very small and not statistically significant (Supplemental Table 4).

Impacts Using Various Nonresponse Adjustments

Differential response rates favoring the high-cash gift group characterized all seven of the outcomes included in our impact analyses (Figure 1). Supplemental Table 5 shows the sensitivity of the results to various adjustments for nonresponse. To address whether results were affected by attrition and any imbalances in baseline characteristics, we adjusted regression analyses using two types of weights created by the Toolkit for Weighting and Analysis of Nonequivalent Groups (TWANG) (Ridgeway et al. 2022). Broadly speaking, TWANG uses generalized boosted models to flexibly estimate propensity scores and analytic weights (McCaffrey et al., 2004). These models included all baseline control variables and the child's age.

First, we constructed inverse probability of treatment weights to address imbalances that might have been introduced by differential response rates across the groups. 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.

Additionally, we created a set of nonresponse weights that adjust regression estimates for 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. Results in Supplemental Table 5 show that neither weighting adjustment substantively changed the results: none of the coefficients is statistically significant and effect sizes are similar to those reported in Table 3.

Multiple Imputation

We next estimated the impact of the BFY high-cash gift on preregistered outcomes using multiple imputation to account for missing data. We imputed 20 datasets using chained equations, or MICE, linear regression, or logistic regression, and predictive mean matching. As shown in Supplemental Table 6, multiple imputation produces a pattern of results similar to those shown in Table 3.

Supplemental Analysis of Children's Brain Activity

Additional analyses of the EEG data are detailed in a complementary manuscript (Troller-Renfree et al., 2024). At age 4, consistent with procedures used to analyze EEG data in another adversity-reduction RCT (Debnath et al., 2019), we estimated three multi-level models (MLMs) with cash gift group, brain region, and group-by-brain region interactions predicting whole-brain alpha, beta, and gamma power. The individual whole-brain bands were not preregistered as outcomes at age 4, because they were part of the primary preregistered EEG composite. However, we expected differences in the same direction as were preregistered for the composite (i.e., greater power in each band in the high-cash gift group).

Past research (Benasich et al., 2008; Brito et al., 2016; Brito et al., 2020; Tomalski et all, 2013), plus our own prior findings at 12 months of age (Troller-Renfree et al., 2022), led us to preregister frontal gamma power as a secondary outcome at age 4, with the hypothesis of greater frontal gamma power in the high-cash than low-cash gift group. Counter to our predictions, and in contrast to first-year results, the MLM of gamma power showed that group differences in the gamma band at four years of age were neither present across the whole brain, nor were they stronger or weaker over any particular brain region (Troller-Renfree et al., 2024). Confirmatory intent-to-treat regression analysis of frontal gamma power confirmed this null result (effect size = -.09, p=.27).

In the other MLM models at age 4, we found some evidence of higher whole-brain power in the alpha band among preschoolers in the high-cash gift group relative to those in the low-cash gift group (Troller-Renfree et al., 2024). The magnitude of this group difference in alpha power at age 4 was similar to that found after the first year of cash gifts (Troller-Renfree et al., 2022). However, in contrast to the first-year results, there were no group differences in beta power. Exploratory analyses revealed no group-by-region interactions in either alpha or beta power at four years of age.

Taken together, the EEG findings suggest that monthly unconditional cash transfers may have selective impacts on preschoolers' brain activity, with possibly different impacts across brain frequency bands throughout early childhood (i.e., impacts on beta and gamma power in infancy, and impacts on alpha power in the preschool years). Whether this suggests the possibility of future impacts of the high-cash gift on cognition and behavior is discussed in Troller-Renfree et al., 2024. However, we caution that the group differences in the preregistered primary and secondary outcomes at age 4 were not statistically significant. In addition, while results were in the expected direction, whole-brain alpha power was not a preregistered outcome, and thus the impact of unconditional cash transfers on brain activity in this frequency band needs further study and replication.

Discussion

The Baby's First Years study tests whether monthly unconditional cash transfers to low-income mothers beginning shortly after birth affect children's development. This paper reports results after the first 4 years of the planned 6-year RCT, at a point when mothers in the high-cash gift group had received about $16,000 in cash gifts and mothers in the low-cash gift group had received less than $1,000. We found no evidence of group differences on preregistered primary (language, executive function, social-emotional development, composite of high-frequency brain activity) or secondary (visual processing/spatial perception, pre-literacy skills, diagnosis of developmental conditions) outcomes (Figure 2).

We consider several possible explanations for these null results. First, we consider implementation elements of the cash gifts and/or study design that may have hindered our ability to detect true impacts of cash transfers. These include lack of intervention fidelity, differential attrition, and measurement error.

With respect to intervention fidelity, the cash gifts were disbursed monthly to virtually all families in both groups across the four-year period, suggesting that results were not downwardly biased by failure to receive the cash gifts. With respect to differential attrition, preregistered outcomes were gathered from 637 to 886 children, depending on the measure. For all measures, response rates were higher for the high-cash gift group compared with the low-cash gift group. However, results were robust to various approaches for using baseline demographic characteristics to model the possible biases associated with differential attrition and missing data.

Nevertheless, because it was not possible to measure the same child outcomes at baseline to model missing outcomes, it remains possible that unmeasured group differences arising from differential attrition may bias our estimates.

Whenever possible we used validated, standardized measures for primary outcomes. However, children had little-to-no prior exposure to similar assessment settings; subjective reports by the research staff suggest that maintaining children's attention posed challenges, as has been reported among children who experienced the pandemic in early childhood (Kuehn et al., 2024). While these considerations apply to children in both groups, they may have increased measurement error. Whether this would bias the estimated impacts is unclear.

We next consider conceptual explanations for why the cash gifts may not have affected children's development at age 4. These include the size and duration of the monthly cash gifts, the pandemic's effects on family life and economic conditions, and consideration of whether increased income alone is sufficient to affect early childhood development.

It is possible that the monthly cash gifts were simply not large enough to affect early child development. The annual BFY high-cash gift transfers amounted to 18% of the average mother's income at baseline. Over the first three years, this led to an average 14% impact on net income (see Gennetian et al. 2024). Nearly all (91%) of the study families reported household incomes (including the gifts) that were less than twice the official poverty line when children were four years of age, and the two groups did not report different levels of material hardship (Magnuson et al., under review). Moreover, inflation was unusually high and led to about an 18% drop in purchasing power over the study period (U.S. Bureau of Labor Statistics, 2024). On the other hand, the BFY cash transfers were as large as the average EITC and welfare reform transfers that boosted child school achievement in prior studies (Barr et al., 2022; Dahl and Lochner, 2012; Duncan et al., 2011).

It may be that a $4,000 annual cash transfer requires more than four years to produce substantive effects on child development, or that early cash support will not affect children's outcomes until later. Although we found some suggestive evidence of neuroplasticity after only one year of cash gifts (Troller-Renfree et al., 2022), and mixed evidence of impacts on preschoolers' brain activity (Troller-Renfree et al., 2024), it is possible that it takes longer than four years to see effects on cognition and behavior. Because the BFY cash gifts will continue until children's 6th birthdays, future data will shed light on whether a longer duration of payments – and longer follow-up of families – will affect children's development.

Another possible explanation for null impacts centers on the COVID-19 pandemic, which may have affected the impact of the cash gifts on children's development. The pandemic, which took place roughly between BFY children's first and third birthdays, may have increased caregiving demands, isolation, and a range of other challenges. Over 70% of BFY mothers reported major changes in their own behaviors in response to the pandemic conditions (Sauval et al., 2024). Surprisingly, though, mothers' subjective wellbeing and mental health did not worsen during the height of the pandemic (Premo et al. 2023).

The economic context of families also fluctuated during the pandemic, with 60% of BFY mothers reported losing income "because of the coronavirus" (Sauval et al., 2024). However, 72% of mothers reported receiving pandemic-related government payments when their children were two years of age (Gennetian et al, 2024). According to a tax simulation program (Feenberg & Coutts, 1993), BFY mothers were eligible in 2021 for approximately $15,000 from the expanded child tax credit, economic impact payments, and the EITC – an amount that was about double what they could have received from these sources in 2020. Families likely also benefited from other pandemic policy changes, including expansions in unemployment insurance, food assistance, and eviction moratoria. These changes, plus the concurrent spike in inflation, may have reduced the relative value of an additional modest monthly cash transfer. Arguing against a pandemic disruption explanation, positive impacts on the time and money parents spent on their child did not substantively fluctuate before, during or after the pandemic (Gennetian et al., 2024), nor did null impacts on maternal-reported child health and development (Hart et al., 2024; Sperber et al., 2023).

Finally, the lack of impacts on age-4 child outcomes raises the possibility that income alone may not affect children's early development. Although prior theory and quasi-experimental research suggested that income alone may improve children's developmental outcomes (NASEM, 2019), prior studies have not been able to isolate the impact of income itself (i.e., rule out the possibility that income only matters in combination with other incentives or policy levers). Our strong test of the impacts of four years of unconditional cash transfers on the development of young children living in low-income families finds consistent null results, which may indicate that cash income alone does not have a causal effect on young children's development in the contemporary policy context.

Additional research is needed to shed light on these various possibilities and to replicate these findings. Data from future study waves – which will include direct assessments of child development after a full six years of monthly unconditional cash support – will provide the opportunity to test for emergent impacts later in life.

Supplemental Materials

Randomization Procedures

Randomization occurred within each of the four sites. The first step in the randomization process was to create four rosters of 250 rows each, with 150 rows designated as “low cash gifts” and 100 designated as “high cash gifts”. Each of the four 250-row rosters was then randomly ordered. Rows were assigned consecutively-numbered cash gift IDs. The resulting roster data on cash gift condition and cash gift IDs were then stored on the survey contractor’s server.

As the May, 2018 to June, 2019 recruitment period in hospitals during the immediate postnatal period proceeded, it became clear that IRB and other institutional issues in one site (the Twin Cities) would lead to fewer than 250 recruited participants. This led to a roughly equal increase in the number of roster rows in the other three sites. To accomplish this, additional roster rows were created in each of these sites using the same randomization procedure. When aggregated, the 1,000-row roster matched exactly the 40%/60% distribution of cash gifts across all possible respondents.

The second step was to create a web-based application that, when interfacing with the interviewers’ questionnaire software program, could access these rosters, determine the high-cash vs. low-cash gift condition to be offered to each participant, record that the condition was offered, and return the gift value for the interviewer to share with the participant.

The interviews themselves were conducted using the Blaise-based Computer-Assisted Personal Interview program (version 4.8). After a participant was successfully recruited in the hospital and agreed to receive a cash gift, the Blaise instrument accessed the web application with a pre-loaded link. The web-based application collected information on the site and the IDs of both the respondent and interviewer from the Blaise software for validation purposes. Once that information was processed, the web application accessed the randomized roster for the specified site, retrieved the next available cash gift ID and its amount, and recorded for which respondent, on which date, and by which interviewer it has been claimed. These requests for cash gift assignments could have come from any of the hospitals within the site and were processed in the order in which they were received by the web-based application. Cash gift amounts, cash gift ID and respondent ID were displayed (via Blaise) to the interviewer for confirmation, along with the randomly generated monthly gift amount. The interviewer confirmed the information and proceeded with the interview by announcing the cash gift amount to the respondent and setting up a debit card with that amount so that the participant could use it right away.

These procedures ensured a randomization process in which the interviewers could not influence the assigned amount. At the same time, the procedures unblinded both the participant and the interviewer to the gift amount during the recruitment period. In subsequent rounds of data collection, interviewers were not reminded (or, in the case of different interviewers, informed) of participants' treatment status during follow-up assessments. Of course, participants were reminded of the cash gift amount on a monthly basis.

EEG methods

As has been described elsewhere (Troller-Renfree et al., 2024), EEG data were collected using four lab-based EEG systems provided by Magstim EGI, Eugene, OR. All EEG data were recoded using a 128-channel Geodesic Sensor Net (Magstim EGI, Eugene, OR) with eye and face electrodes removed. As our sample is both racially and ethnically diverse, modified geodesic sensor nets with longer pedestals were made available, and specialized net application protocols were used to increase inclusivity in our recordings (see Adams et al., 2024; Troller-Renfree et al., 2021 for more information). Following cap application, impedances were kept below 50 kΩ whenever possible. The sampling rate was 1000 Hz and data were referenced online to Cz. During the recording, children sat on a chair approximately 60cm from the monitor while they watched silent non-social videos of toys currently being used as a part of data collection for the Health Brains and Child Development Study (HBCD, see Fox et al., 2024 for more information). Recordings lasted approximately 3 minutes. Data were analyzed offline on virtual machines; data processors received no treatment group information.

EEG Processing

Identical to the methods reported in Troller-Renfree et al. (2024), EEG preprocessing was conducted with EEGlab (Delorme & Makeig, 2004), custom MATLAB scripts (MathWorks, Natick, MA), and the MADE processing pipeline (Debnath et al., 2020). First, as the MADE pipeline utilizes independent components analysis (ICA) for systematic artifact removal, all EEG data from resting and task-related recordings were visually inspected, and any non-brain data (e.g., recording after cap removal) were then merged to maximize data for ICA. The bottom row of electrodes was then deleted from merged, continuous EEG data and then high-pass filtered at 0.3 Hz and low-pass filtered at 50 Hz. Next, the EEGLAB plug-in FASTER (Nolan et al., 2010) was used to identify bad channels prior to undergoing the standard MADE ICA removal procedure (Debnath et al., 2020). Artifactual independent components were removed from the original dataset by using the Adjusted-ADJUST algorithm (Leach et al., 2020).

Following ICA, resting EEG data were identified and segmented into 1s epochs which were allowed to overlap by 50%. After the segmentation procedure, residual ocular artifacts were removed by completely rejecting any epochs in which the ocular channels (electrodes 1, 8, 14, 21, 25, and 32) voltages exceeded ±100 μV. Next, bad channels were identified by determining whether channel voltages exceeded ±100 μV. Channels containing artifacts were interpolated at the epoch level (Perrin et al., 1989). However, if more than 10% of the channels (not considering globally rejected channels) were interpolated, the entire epoch was removed from further processing.

Finally, cleaned data were decomposed via a Fast Fourier Transform (FFT) with a Hanning window used to deconstruct the EEG data into 1-Hz frequency bins. Spectral power (µV2) of all frequency bins between 7 and 45 was then summed to create the primary preregistered composite of brain activity.

Determining Language of Assessment

Children were assessed in English and/or Spanish, depending on their language preferences. In order to determine which language(s) was/were most appropriate, mothers were asked before the laboratory visit which language their child was most comfortable speaking at home and whether they spoke more than one language. If mothers reported that the child was a monolingual English speaker, all measures were administered in English. If it appeared that a child preferred English but spoke some Spanish, the bilingual ROWPVT was administered, and all other measures were administered in English. If a child was a monolingual Spanish speaker, or spoke some English but preferred Spanish, the bilingual ROWPVT was administered, and all other measures were administered in Spanish. Research staff further adapted the language of assessment depending on the child's apparent preference during the visit.

Self-Regulation

In addition to the preregistered primary and secondary outcomes described above, we report an exploratory analysis of experimenter-reported self-regulation in the Supplemental Materials. Self-regulation (n=843) was measured using the Preschool Self-Regulation Assessment (PSRA; Smith-Donald, et al., 2007) Assessor Report. The PSRA Assessor Report assesses research staff observations of children's emotions, attention, and behavior during an assessment, and was originally designed to accompany a structured set of tasks. In this study, we did not employ the PSRA tasks; instead, assessors completed the PSRA assessor report measure (28 items) after administration of the child cognitive battery described above. The PSRA assessor report generates two scores: Attention/Impulse Control and Positive Emotion. Cronbach's alpha was 0.965 for the Attentive/Impulse Control and 0.867 for Positive Emotion. Impact results for these two measures are shown in Supplemental Table 3.

Cash gift overpayments

On two occasions in the four-year treatment period (when children were ages 2 and 4), the financial institution distributing BFY's monthly cash gifts provided erroneously high payments to a total of 24 families. Eight of the families were in the low-cash gift group and 16 were in the high-cash gift group. The mean overpayment was $1293 [SD=$1234]. Corresponding means and standard deviations for the low-cash and high-cash gift groups were $93 [$49] and $1894 [$1086]. For the age-4 data collection, 22 families that ever received an overpayment of the cash gift provided data for at least one child health outcome, including 14 from the high-cash gift group and 8 from the low-cash gift group.

We investigated the possible impact of overpayments on our treatment results in two ways (Supplemental Table 7). First, we added a dichotomous control variable to the impact regressions shown in Table 3. The resulting effect sizes and significance levels (top panel of Supplemental Table 7) are virtually identical to their Table 3 counterparts. Second, we reran Table 2 regressions leaving out the 22 cases with overpayments. Here again, the resulting effect sizes and significance levels (bottom panel of Supplemental Table 4) are virtually identical to their Table 3 counterparts.

References

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. https://doi.org/10.1542/pir.21.8.265

Attanasio, O., Cattan, S., & Meghir, C. (2022). Early childhood development, human capital, and poverty. *Annual Review of Economics*, *14*(1), 853-892. https://doi.org/10.1146/annurev-economics-092821-053234

Barr, A., Eggleston, J., & Smith, A. A. (2022). Investing in infants: The lasting effects of cash transfers to new families. *The Quarterly Journal of Economics*, *137*(4), 2539-2583. https://doi.org/10.1093/qje/qjac023

Bastagli, F., Hagen-Zanker, J., Harman, L., Barca, V., Sturge, G., Schmidt, T., & Pellerano, L. (2016). Cash transfers: what does the evidence say. A rigorous review of programme impact and the role of design and implementation features. London: ODI, 1(7), 1.

Benasich, A. A., Gou, Z., Choudhury, N., & Harris, K. D. (2008). Early cognitive and language skills are linked to resting frontal gamma power across the first 3 years. Behavioural brain research, 195(2), 215-222. https://doi.org/10.1016/j.bbr.2008.08.049

Berk, R. (1990). A primer on robust regression. In J. Fox & J. S. Long (Eds.), *Modern Methods of Data Analysis*. Sage.

Bloom, H. S. (1995). Minimum detectable effects: A simple way to report the statistical power of experimental designs. *Evaluation Review*, 19(5), 547-556. https://doi.org/10.1177/0193841X9501900504

Boyce, T.W., Levitt, P., Martinez, F.D., McEwen, B.S., and Shonkoff, J.P. (2021). Genes, Environments, and Time: The Biology of Adversity and Resilience. *Pediatrics*, 147(2). https://doi.org/10.1542/peds.2020-1651

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., Fifer, W. P., Myers, M. M., Elliott, A. J., & Noble, K. G. (2016). Associations among family socioeconomic status, EEG power at birth, and cognitive skills during infancy. *Developmental Cognitive Neuroscience*, *19*, 144-151. https://doi.org/10.1016/j.dcn.2016.03.004

Carlson, S. M., & Zelazo, P. D. (2014). *Minnesota Executive Function Scale*. Reflection Sciences, Inc.

Carlson, S. M. (2017). *Minnesota executive function scale: Technical Report, v. 2*. Reflection Sciences, Inc.

Coll, C. G., Crnic, K., Lamberty, G., Wasik, B. H., Jenkins, R., Garcia, H. V., & McAdoo, H. P. (1996). An integrative model for the study of developmental competencies in minority children. *Child development*, *67*(5), 1891-1914. https://doi.org/10.1111/j.1467-8624.1996.tb01834.x

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

Duncan, G. J., Magnuson, K., & Votruba-Drzal, E. (2017). Moving beyond correlations in assessing the consequences of poverty. *Annual Review of Psychology*, *68*, 413-434. https://doi.org/10.1146/annurev-psych-010416-044224

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. https://doi.org/10.1037/a0023875

Feenberg, D., & Coutts, E. (1993). An introduction to the TAXSIM model. *Journal of Policy Analysis and Management*, 12(1), 189-194.

Fernald, L. C. H., & Hidrobo, M. (2011). Effect of Ecuador's cash transfer program (Bono de Desarrollo Humano) on child development in infants and toddlers: A randomized effectiveness trial. *Social Science & Medicine*, 72(9), 1437–1446. https://doi.org/10.1016/j.socscimed.2011.03.005

Gennetian, L. A., Duncan, G. J., Fox, N. A., Halpern-Meekin, S., Magnuson, K., Noble, K. G., & Yoshikawa, H. (2024). Effects of a monthly unconditional cash transfer starting at birth on family investments among US families with low income. *Nature Human Behaviour*, 1-16. https://doi.org/10.1038/s41562-024-01915-7

Gennetian, L. A., Halpern-Meekin, S., Meyer, L., Fox, N., Magnuson, K., Noble, K., & Yoshikawa, H. (2023). Cash to US families at scale: Behavioral insights on implementation from the Baby's first years study. In D. Soman, J. Zhao, & S. Datta (Eds.), *Using cash transfers to build an inclusive society: A behavioral Lens.* University of Toronto Press. http://dx.doi.org/10.2139/ssrn.4286345

Hagerman, T. K., & Houtrow, A. J. (2021). Variability in prevalence estimates of disability among children in the National Survey of Children's Health. *JAMA Pediatrics*, 175(3), 307-310. https://doi.org/10.1001/jamapediatrics.2020.5073

Hart, E. R., Gennetian, L. A., Sperber, J. F., Penalva, R., Magnuson, K., Duncan, G. J., Halpern-Meekin, S., Yoshikawa, H., Fox, N., & Noble, K. (2024). The Effect of Unconditional Cash Transfers on Maternal Assessments of Children's Early Language and Socioemotional Development: Experimental Evidence from U.S. Families Residing in Poverty. *Developmental Psychology*. https://doi.org/10.1037/dev0001824

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). https://doi.org/10.1542/peds.2020-1641

Hutton, J. S., Justice, L., Huang, G., Kerr, A., DeWitt, T., & Ittenbach, R. F. (2019). The Reading House: a children's book for emergent literacy screening during well-child visits. *Pediatrics*, *143*(6). https://doi.org/10.1542/peds.2018-3843

Iruka, I. U., Gardner-Neblett, N., Telfer, N. A., Ibekwe-Okafor, N., Curenton, S. M., Sims, J., Sansbury, A. B., & Neblett, E. W. (2022). Effects of racism on child development:

Advancing antiracist developmental science. *Annual Review of Developmental Psychology*, *4*(1), 109-132. https://doi.org/10.1146/annurev-devpsych-121020-031339

Kuehn, L. M., Jones, A., Helmkamp, L., Knudtson, M., Domek, G. J., & Allison, M. A. (2024). Socioemotional development of infants and toddlers during the COVID-19 pandemic. *JAMA Pediatrics*, 178(2), 151-159. https://doi.org/10.1001/jamapediatrics.2023.5684

Magnuson, K., Duncan, G., Yoshikawa, H., Yoo, P., Han, S., Gennetian, L. A., Halpern-Meekin, S., Fox, N., & Noble, K. (2024). Can a Poverty Reduction Intervention Improve Maternal Wellbeing and Family Processes among Families with Young Children? An Experimental Analysis. *SSRN.* https://ssrn.com/abstract=4188131

Masarik, A. S., & Conger, R. D. (2017). Stress and child development: A review of the Family Stress Model. *Current Opinion in Psychology*, 13, 85-90. https://doi.org/10.1016/j.copsyc.2016.05.008

Martin, N. A., & Brownell, R. (2011). *ROWPVT-4: Receptive One-Word Picture Vocabulary Test* (4th ed.).

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. https://doi.org/10.1037/1082-989X.9.4.403

McInnis, N., Michelmore, K., & Pilkauskas, N. (2024). The Intergenerational Transmission of Poverty and Public Assistance: Evidence from the Earned Income Tax Credit. *Journal of Human Resources,* 0422-1224R2. https://doi.org/10.3368/jhr.0422-12241R2

National Academies of Sciences, Engineering, and Medicine (2019). *A Roadmap to Reducing Child Poverty*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25246

National Academies of Sciences, Engineering, and Medicine. (2023). *Reducing Intergenerational Poverty*. Washington, DC: The National Academies Press. https://doi.org/10.17226/27058

Noble, K. G., & Giebler, M. A. (2020). The neuroscience of socioeconomic inequality. *Current Opinion in Behavioral Sciences*, *36*, 23-28. https://doi.org/10.1016/j.cobeha.2020.05.007

Noble, K. G., Hart, E. R., & Sperber, J. F. (2021). Socioeconomic disparities and neuroplasticity: Moving toward adaptation, intersectionality, and inclusion. *American Psychologist*, *76*(9), 1486. https://doi.org/10.1037/amp0000934

Noble, K. G., Magnuson, K., Gennetian, L. A., Duncan, G. J., Yoshikawa, H., Fox, N. A., & Halpern-Meekin, S. (2021). Baby's first years: design of a randomized controlled trial of poverty reduction in the United States. *Pediatrics*, 148(4). https://doi.org/10.1542/peds.2020-049702

Pace, A., Luo, R., Hirsh-Pasek, K., & Golinkoff, R. M. (2017). Identifying pathways between socioeconomic status and language development. *Annual Review of Linguistics*, 3(1), 285–308. https://doi.org/10.1146/annurev-linguistics-011516-034226

Premo, E. M., Magnuson, K. A., Lorenzo, N. E., Fox, N. A., & Noble, K. G. (2023). Mental health and sleep quality of low-income mothers of one-year-olds during the COVID-19 pandemic. *Infant Mental Health Journal*, *44*(4), 572-586. https://doi.org/10.1002/imhj.22074

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither Opportunity*. Russell Sage Foundation.

Ridgeway, G., McCaffrey, D. F., Morral, A. R., Cefalu, M., Burgette, L. F., Pane, J. D., & Griffin, B. A. (2022). *Toolkit for weighting and analysis of nonequivalent groups: a tutorial for the R TWANG package*. Santa Monica, Calif: Rand.

Sauval, M., Duncan, G. J., Gennetian, L. A., Magnuson, K. A., Fox, N. A., Noble, K. G., & Yoshikawa, H. (2024). Unconditional cash transfers and maternal employment: Evidence from the Baby's First Years study. *Journal of Public Economics*, 236, 105-159. https://doi.org/10.1016/j.jpubeco.2024.105159

Sroufe, L. A., Cooper, R. G., DeHart, G. B., Marshall, M. E., & Bronfenbrenner, U. E. (1992). *Child development: Its nature and course*. McGraw-Hill Book Company.

Shah, H., & Gennetian, L. A. (2024). Unconditional cash transfers for families with children in the US: a scoping review. *Review of Economics of the Household,* 22(2), 415-450. https://doi.org/10.1007/s11150-023-09672-8

Sperber, J. F., Gennetian, L. A., Hart, E. R., Kunin-Batson, A., Magnuson, K., Duncan, G. J., Yoshikawa, H., Fox, N. A., Halpern-Meekin, S., & Noble, K. G. (2023). Unconditional Cash Transfers and Maternal Assessments of Children's Health, Nutrition, and Sleep: A Randomized Clinical Trial. *JAMA Network Open*, *6*(9), e2335237-e2335237. https://doi.org/10.1001/jamanetworkopen.2023.35237

Tomalski, P., Moore, D. G., Ribeiro, H., Axelsson, E. L., Murphy, E., Karmiloff-Smith, A., Johnson, M. H., & Kushnerenko, E. (2013). Socioeconomic status and functional brain development–associations in early infancy. *Developmental Science*, *16*(5), 676-687. https://doi.org/10.1111/desc.12079

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., Halpern-Meekin, S., Fox, N., & 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

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., … Noble, K. G. (2024, December 19). The impact of a monthly unconditional cash transfer on child brain activity: A 4-year follow-up. Retrieved from osf.io/preprints/psyarxiv/dw2em

Troller-Renfree, S. V., Morales, S., Leach, S. C., Bowers, M. E., Debnath, R., Fifer, W. P., Fox, N., & Noble, K. G. (2021). Feasibility of assessing brain activity using mobile, in-home collection of electroencephalography: Methods and analysis. *Developmental psychobiology*, 63(6), e22128. https://doi.org/10.1002/dev.22128

U.S. Bureau of Labor Statistics. (2024). *CPI Inflation Calculator*. U.S. Deparetment of Labor.

Wechsler, D., & Naglieri, J. A. (2006). Wechsler nonverbal scale of ability. https://doi.org/10.1037/t15176-000

References for Supplemental Materials

Adams, E. J., Scott, M. E., Amarante, M., Ramírez, C. A., Rowley, S. J., Noble, K. G., & Troller-Renfree, S. V. (2024). Fostering inclusion in EEG measures of pediatric brain activity. *Npj Science of Learning*, 9(1), 1–8. https://doi.org/10.1038/s41539-024-00240-y

Debnath, R., Buzzell, G. A., Morales, S., Bowers, M. E., Leach, S. C., & Fox, N. A. (2020). The Maryland Analysis of Developmental EEG (MADE) Pipeline. *Psychophysiology*, 57(6), e13580. https://doi.org/10.1111/psyp.13580

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods, 134(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Fox, N. A., Pérez-Edgar, K., Morales, S., Brito, N. H., Campbell, A. M., Cavanagh, J. F., Gabard-Durnam, L. J., Hudac, C. M., Key, A. P., Larson-Prior, L. J., Pedapati, E. V., Norton, E. S., Reetzke, R., Roberts, T. P., Rutter, T. M., Scott, L. S., Shuffrey, L. C., Antúnez, M., Boylan, M. R., ... & the HEALthy Brain and Child Development (HBCD) Study EEG Workgroup. (2024). The development and structure of the HEALthy Brain and Child Development (HBCD) Study EEG protocol. *Developmental Cognitive Neuroscience*, 101-447. https://doi.org/10.1016/j.dcn.2024.101447

Leach, S. C., Morales, S., Bowers, M. E., Buzzell, G. A., Debnath, R., Beall, D., & Fox, N. A. (2020). Adjusting ADJUST: Optimizing the ADJUST algorithm for pediatric data using geodesic nets. *Psychophysiology*, 57(8). https://doi.org/10.1111/psyp.13566

Nolan, H., Whelan, R., & Reilly, R. B. (2010). FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. *Journal of Neuroscience Methods,* 192(1), 152–162. https://doi.org/10.1016/j.jneumeth.2010.07.015

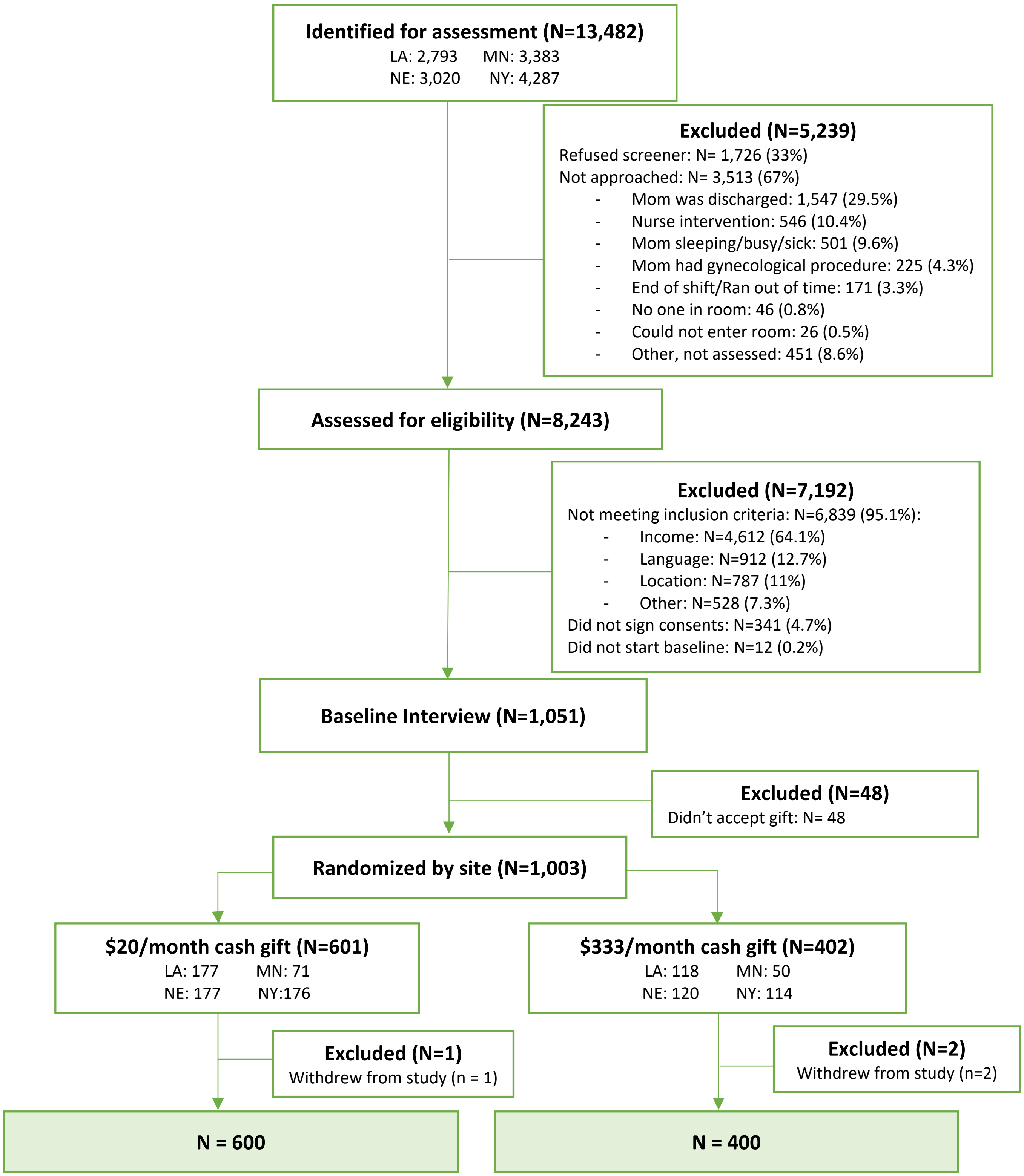
Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72(2), 184–187. https://doi.org/10.1016/0013-4694(89)90180-6

Smith-Donald, R., Raver, C. C., Hayes, T., & Richardson, B. (2007). Preliminary construct and concurrent validity of the Preschool Self-regulation Assessment (PSRA) for field-based research. *Early Childhood Research Quarterly*, 22(2), 173-187. https://doi.org/10.1016/j.ecresq.2007.01.002

Troller-Renfree, S. V., Costanzo, M. A., Duncan, G. J., Magnuson, K., Gennetian, L. A., Yoshikawa, H., … Noble, K. G. (2024, December 19). The impact of a monthly unconditional cash transfer on child brain activity: A 4-year follow-up. Retrieved from osf.io/preprints/psyarxiv/dw2em

Troller-Renfree, S. V., Morales, S., Leach, S. C., Bowers, M. E., Debnath, R., Fifer, W. P., Fox, N. A., & Noble, K. G. (2021). Feasibility of assessing brain activity using mobile, in-home collection of electroencephalography: Methods and analysis. *Developmental psychobiology*, 63(6), e22128. https://doi.org/10.1002/dev.22128

Figure 1. Consort Diagram



Age-4 Child Outcomes Data Collection

Low-cash gift group (N=600)

Excluded at previous waves (N=5)
Child deceased: N=1
Mother deceased: N=4

Excluded at Age 4 (N=2)
Child deceased: N=1
Mother deceased: N=1

Ineligible for Age-4 follow-up:	**N=4**
Mother not with child:	N=2
Mother incarcerated:	N=2
Eligible for Age-4 follow-up	**N=589**
Nonresponse: unavailable, not found	N=43
Nonresponse: refused data collection	N=26
Nonresponse: study withdrawal	N=4
Nonresponse: data collection interrupted	N=0
Age-4 Data collected	N=516
Any Available Age-4 Data	**N=518**
Valid Maternal report measures	
CBCL	513
Child health	515
Valid Lab-based child assessments	
Matrices	443
Reading House	434
One-Word PVT	404
MEFS	449
EEG (resting)	359
Experimenter report measure	
PSRA	484

High-cash gift group (N=400)

Excluded at previous waves (N=3)
Child deceased: N=3

Excluded at Age 4 (N=0)

Ineligible for Age-4 follow-up:	**N=2**
Mother not with child:	N=0
Mother incarcerated:	N=2
Eligible for Age-4 follow-up	**N=395**
Nonresponse: unavailable, not found	N=9
Nonresponse: refused data collection	N=12
Nonresponse: study withdrawal	N=0
Nonresponse: data collection interrupted	N=0
Age-4 Data collected	N=374
Any Available Age-4 Data	**N=374**
Valid Maternal report measures	
CBCL	369
Child health	369
Valid Lab-based child assessments	
Matrices	326
Reading House	320
One-Word PVT	296
MEFS	334
EEG (resting)	278
Experimenter report measure	
PSRA	359

Table 1. Sample Characteristics and Baseline Balance for High and Low Cash Gift Groups in the Age-4 analytic sample (N=890)

	Total Sample		Low-Cash Gift		High-Cash Gift		Std Mean Difference		
	Mean (sd)	N	Mean (sd)	N	Mean (sd)	N	Hedges' g	Cox's Index	p-value
CHILD									
Child is female	0.490	890	0.499	517	0.477	373		-0.053	0.522
Child weight at birth (pounds)	7.119 (1.032)	888	7.139 (1.044)	516	7.092 (1.016)	372	-0.045		0.507
Child gestational age (weeks)	39.069 (1.221)	886	39.106 (1.206)	513	39.018 (1.242)	373	-0.072		0.302
MOTHER									
Mother age at birth (years)	27.181 (5.880)	890	27.002 (5.929)	517	27.429 (5.809)	373	0.073		0.246
Mother education (years)	11.900 (2.843)	882	11.912 (2.775)	511	11.884 (2.939)	371	-0.010		0.911
Mother white, non-Hispanic	0.093	890	0.101	517	0.083	373		-0.131	0.302
Mother Black, non-Hispanic	0.422	890	0.402	517	0.450	373		0.119	0.116
Mother race/ethnicity: multiple, non-Hispanic	0.037	890	0.041	517	0.032	373		-0.156	0.476
Mother race/ethnicity: other or unknown	0.035	890	0.048	517	0.016	373		-0.685	0.004
Mother Hispanic	0.412	890	0.408	517	0.418	373		0.025	0.553
Mother never married	0.452	890	0.420	517	0.496	373		0.186	0.021
Mother single, living with partner	0.236	890	0.253	517	0.212	373		-0.139	0.150
Mother married	0.221	890	0.222	517	0.220	373		-0.007	0.954
Mother divorced/separated	0.039	890	0.046	517	0.029	373		-0.290	0.189
Mother marital status: other or unknown	0.052	890	0.058	517	0.043	373		-0.191	0.277
Mother health is good or better	0.900	890	0.884	517	0.922	373		0.266	0.055

Mother depression (CESD)	0.684 (0.457)	890	0.687 (0.463)	517	0.681 (0.450)	373	-0.013		0.810
Cigarettes per week during pregnancy	3.830 (16.082)	883	4.240 (18.741)	512	3.263 (11.440)	371	-0.061		0.291
Alcohol drinks per week during pregnancy	0.099 (1.317)	887	0.151 (1.694)	515	0.027 (0.394)	372	-0.094		0.110
Number of children born to mother	2.463 (1.405)	890	2.412 (1.399)	517	2.534 (1.413)	373	0.086		0.217
Number of adults in household	2.056 (0.972)	890	2.083 (0.980)	517	2.019 (0.960)	373	-0.066		0.338
Biological father lives in household	0.376	890	0.400	517	0.343	373		-0.148	0.084
Household combined income	21847.73 (19628.38)	834	22566.33 (21824.87)	487	20839.20 (16023.33)	347	-0.088		0.198
Household income unknown	0.063	890	0.058	517	0.070	373		0.122	0.485
Household net worth	-1976.51 (24337.30)	791	-1802.07 (30130.01)	458	-2216.41 (12636.57)	333	-0.017		0.796
Household net worth unknown	0.111	890	0.114	517	0.107	373		-0.043	0.776

Joint tests on the equality of means across all baseline characteristics for all of the samples used in the paper:

All participants providing any data in the analyses: $\chi^2(30)=32.157$, p-value= 0.268, n=886

Joint tests on the equality of means across all baseline characteristics for all participants providing data in analyses of:

Receptive One-Word Picture Vocabulary Test $\chi^2(30)=32.087$, p-value= 0.271, n=696

Minnesota Executive Function Scale $\chi^2(30)=28.302$, p-value= 0.449, n=779

Child Behavior Checklist $\chi^2(30)=32.080$, p-value= 0.271, n=879

Resting Brain Activity $\chi^2(30)=22.767$, p-value= 0.745, n=637

Matrices $\chi^2(30)=32.594$, p-value= 0.251, n=765

Reading House $\chi^2(30)=29.852$, p-value= 0.370, n=750

Developmental Condition Diagnosis $\chi^2(30)=31.575$, p-value= 0.292, n=877

Notes: p-values were derived from a series of OLS bivariate regressions in which each respective baseline characteristic was regressed on the treatment status indicator using robust standard errors and site-level fixed effects. Standardized mean differences were calculated using Hedges' g for continuous variables and Cox's Index for dichotomous variables. If there were more than 10 missing cases for a covariate, missing data dummies were included in the table and the joint test. If there were less than 10 cases missing, missing data dummies were not included in the table but were included in the joint test. The joint tests of orthogonality were conducted using a probit model with robust standard errors and site-level fixed effects. Across all joint tests, 3 to 4 observations dropped because of a perfect predictor issue from the missing indicator of child gestational age.

Table 2. Correlation Table of Preregistered Child Measures in Age-4 Sample

	1	2	3	4	5	6
1. Receptive One-Word Picture Vocabulary Test (ROWPVT)	-					
2. Minnesota Executive Function Scale (MEFS)	0.354** 675	-				
3. Child Behavior Checklist (CBCL)	-0.071+ 694	-0.073* 777	-			
4. Resting Brain Activity	0.007 551	0.011 613	-0.023 632	-		
5. Matrices	0.409** 675	0.313** 744	-0.098** 763	-0.058 606	-	
6. Reading House	0.518** 662	0.242** 723	-0.043 748	0.005 597	0.305** 737	-
7. Developmental Condition Diagnosis	-0.232** 692	-0.175** 775	0.214** 877	0.015 631	-0.171** 761	-0.086* 746

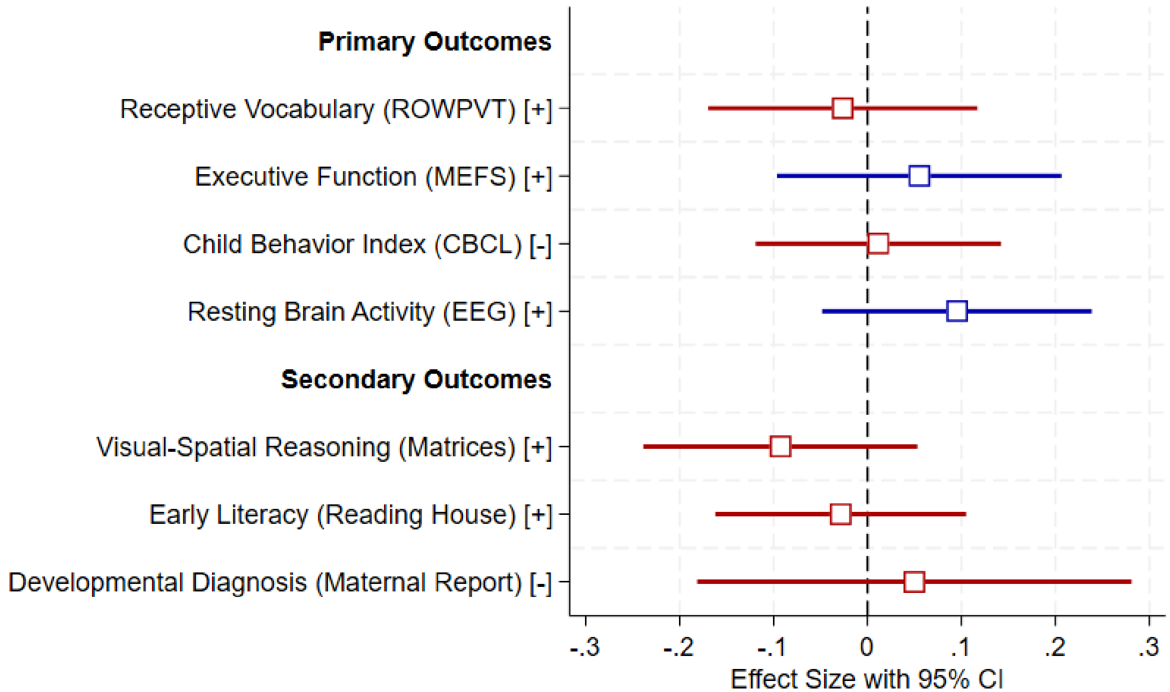
Note: Each rows presents the pairwise correlation and the statistical significance level of that correlation. The number of observations is provided in the row below the outcome. The developmental condition diagnosis outcome is a binary measure.
+ p<0.10; * p<0.05; ** p<0.01

Table 3. Descriptive Statistics and ITT estimates of the Cash-Gift Effect on Preregistered Child Measures at Age 4

	Low Cash Gift Group mean (sd)	High Cash Gift Group mean (sd)	OLS* (se) w/FE	OLS* (se) w/FE w/covariates	Effect Size	p-value	N
PRIMARY OUTCOMES							
ROWPVT conceptual score (+)	40.998 (16.204)	39.199 (15.517)	-1.809 (1.204)	-0.427 (1.179)	-0.026	0.717	700
MEFS standard score (+)	93.746 (10.480)	94.554 (10.596)	0.775 (0.762)	0.578 (0.807)	0.055	0.474	783
CBCL index (-)	22.175 (12.928)	22.393 (12.504)	0.209 (0.863)	0.139 (0.860)	0.011	0.872	882
Resting Brain Activity (+)	2.920 (1.369)	3.051 (1.349)	0.131 (0.107)	0.133 (0.100)	0.097	0.184	637
SECONDARY OUTCOMES							
Matrices raw score (+)	7.625 (3.359)	7.304 (3.308)	-0.313 (0.243)	-0.311 (0.250)	-0.093	0.213	769
Reading House raw score (+)	3.016 (2.242)	2.806 (2.174)	-0.206 (0.160)	-0.058 (0.153)	-0.026	0.705	754
Developmental Condition Diagnosis	0.137 (0.344)	0.146 (0.354)	0.010 (0.023)	0.010 (0.022)	0.052	0.655	878

Note: This table shows descriptive characteristics and estimated ITT effects for all preregistered child outcomes. Hypothesized directions of effects are given in parentheses in the first column. Effect sizes and p-value come from regressions with site fixed-effects and, in the fourth column, controls for the covariates listed in Supplemental Table 2. The effect sizes of continuous outcomes are computed by dividing the treatment effect by the standard deviation of the low-cash sample. ROWPVT=Receptive One-Word Picture Vocabulary Test. MEFS=Minnesota Executive Function Scale. CBCL=Child Behavior Checklist. *In the case of the developmental condition binary model, a logistic regression was run, with coefficients and standard errors expressed as marginal changes in probabilities. To prevent a perfect predictor issue in the logistic regression, four missing indicators of baseline covariates were dropped from the main model: missing indicator of alcohol drinks during pregnancy, smoking cigarettes during pregnancy, child's weight at birth, and child's gestational age. The inclusion and exclusion of missing indicators barely affected results. The effect size of impacts on the developmental condition binary outcome (Developmental Condition Diagnosis) are expressed as Cohen's d, which is converted from the log odds estimate using: $\text{Log odds ratio} \times \frac{\pi}{\sqrt{3}}$ (Bornstein et al., 2021).

Figure 2. Effect Sizes and 95% Confidence Intervals for Pre-Registered Child Outcomes



Notes: Data are taken from the fourth and fifth columns of Table 3, with markers reflecting effect sizes and confidence intervals derived from the unstandardized standard error estimates from the models that include site fixed effects and baseline control measures. Hypothesized direction of effects are given in brackets after the variable labels. Blue lines and markers reflect point estimates in hypothesized directions, while red lines and market reflect the opposite.

Supplemental Table 1. Descriptive Statistics and Estimated Cash-Gift Effect on Child Subscale Measures at Age 4 sample

	Low Cash Gift Group mean (sd)	High Cash Gift Group mean (sd)	OLS* (se) w/FE	OLS* (se) w/FE w/covariates	Effect Size	p-value	N
LANGUAGE							
ROWPVT monolingual standard score (+)	96.398 (14.624)	96.844 (14.301)	0.461 (1.411)	1.334 (1.555)	0.091	0.391	428
ROWPVT bilingual standard score (+)	97.396 (17.157)	94.483 (15.955)	-2.532 (1.992)	-1.775 (1.993)	-0.103	0.374	272
BEHAVIOR PROBLEMS							
CBCL: Anxiety/Depression (-)	56.620 (7.258)	57.273 (7.412)	0.647 (0.499)	0.505 (0.510)	0.070	0.323	883
CBCL: Aggressive Behavior (-)	55.735 (8.147)	55.526 (7.050)	-0.208 (0.514)	-0.202 (0.541)	-0.025	0.709	882
CBCL: Attention Problems (-)	57.150 (7.520)	56.892 (7.027)	-0.269 (0.491)	-0.021 (0.476)	-0.003	0.965	883
CBCL: Emotionally Reactive (-)	55.895 (7.770)	56.491 (7.990)	0.588 (0.539)	0.621 (0.545)	0.080	0.254	882
CBCL: Anxiety/Depression - borderline clinical concern (-)	0.146 (0.354)	0.189 (0.392)	0.042 (0.025)	0.040 (0.026)	0.173	0.113	876
CBCL: Aggressive Behavior - borderline clinical concern (-)	0.136 (0.344)	0.117 (0.321)	-0.020 (0.023)	-0.012 (0.023)	-0.061	0.616	875
CBCL: Attention Problems - borderline clinical concern (-)	0.197 (0.398)	0.168 (0.374)	-0.030 (0.027)	-0.023 (0.026)	-0.095	0.373	874
CBCL: Emotionally Reactive - borderline clinical concern (-)	0.181 (0.386)	0.198 (0.399)	0.017 (0.027)	0.021 (0.026)	0.084	0.418	875
CBCL: Anxiety/Depression - clinical concern (-)	0.043 (0.203)	0.076 (0.265)	0.032 (0.016)	0.030 (0.015)	0.332	0.048	862

CBCL: Aggressive Behavior - clinical concern (-)	0.066 (0.249)	0.046 (0.210)	-0.021 (0.017)	-0.018 (0.018)	-0.195	0.314	869
CBCL: Attention Problems - clinical concern (-)	0.105 (0.307)	0.097 (0.297)	-0.009 (0.021)	0.007 (0.020)	0.048	0.730	874
CBCL: Emotionally Reactive - clinical concern (-)	0.049 (0.216)	0.054 (0.227)	0.005 (0.015)	0.003 (0.015)	0.042	0.825	857
PRE-LITERACY							
Reading House level: Below average vs. average and above (-)	0.382 (0.487)	0.344 (0.476)	-0.038 (0.035)	-0.034 (0.034)	-0.095	0.304	754
DEVELOPMENTAL CONDITIONS							
Developmental Diagnosis: Speech Delay	0.089 (0.285)	0.102 (0.303)	0.014 (0.019)	0.021 (0.016)	0.153	0.202	887
Developmental Diagnosis: Autism	0.041 (0.198)	0.051 (0.220)	0.011 (0.014)	0.012 (0.015)	0.183	0.424	890

Note: This table shows descriptive characteristics and estimated ITT effects for subscales from a subset of the preregistered child outcomes. Hypothesized directions of effects are given in parentheses in the first column. Effect sizes and p-value come from regressions with site fixed-effects and, in the fourth column, controls for the covariates listed in Supplemental Table 2. The effect sizes of continuous outcomes are computed by dividing the treatment effect by the standard deviation of the low-cash sample. ROWPVT=Receptive One-Word Picture Vocabulary Test. CBCL=Child Behavior Checklist. ROWPVT standard scores have a normed mean of 100 with an SD of 15 (normed separately in monolingual and bilingual samples). CBCL T scores are calculated using percentiles, as follows: a T score of 50 is assigned to all raw scores that were at or below approximately the 50th percentile of the normative sample. T scores from 51 to 70 are assigned according to the percentiles of the normative sample. T scores from 71 to 100 are assigned in relation to equal intervals of the raw scores that were above the 98th percentile in the normative sample. * For the CBCL--clinical concern, pre-literacy and developmental conditions models, a logistic regression was run, with coefficients and standard errors expressed as marginal changes in probabilities. To prevent a perfect predictor issue in logistic regressions, some missing indicators of baseline covariates were dropped from the main model: missing indicator of alcohol drinks during pregnancy, smoking cigarettes during pregnancy, child's weight at birth, and child's gestational age. The inclusion and exclusion of missing indicators does not affect results. The effect size of impacts estimated in the logistics regressions are expressed as Cohen's d, which is converted from the log odds coefficients using: $\text{Log odds ratio} \times \frac{\pi}{\sqrt{3}}$ (Bornstein et al., 2021).

Supplemental Table 2. Covariate list

Measure	Explanation
Child is female	Dichotomous, based on maternal report at baseline
Child weight at birth	Birthweight in pounds, based on maternal report of birthweight in pounds and ounces at baseline
Child gestational age	Child's gestational age at birth in weeks, based on a comparison of due date from maternal report at baseline and date of birth as recorded in hospital records and verbally confirmed by mother in the baseline interview
Mother age at birth	Mother age in years at time of child's birth, based on maternal report at baseline
Mother education	Coded into six mutually exclusive indicators (less than high school degree; high school degree; some college; associate's degree; bachelor's degree or more; unknown), based on maternal report at baseline
Mother race/ethnicity	Coded into six mutually exclusive indicators (Non-Hispanic, White; Non-Hispanic, Black or African American; Others including Asian or Pacific Islander, and American Indian, Eskimo, or Aleut; and reporting multiple races; whether the mother considered herself to be Hispanic or Latino; and unknown), based on maternal report of race and ethnicity at baseline
Mother marital status	Coded into six mutually exclusive indicators (single and never married; single and cohabitating; married; divorces or separated; other, including widowed; unknown), based on maternal report at baseline
Mother health	Dichotomous measure of whether mother's health is good or not, based on maternal report of health status as "excellent", "very good", or "good" as opposed to "fair" or "poor" at baseline
Mother depression	Sum of the Center for Epidemiologic Studies Depression Scale (CESD) items, based on maternal report at baseline
Cigarettes per week during pregnancy	Number of cigarettes smoked per week during pregnancy, based on maternal report at baseline
Alcohol drinks per week during pregnancy	Number of alcoholic drinks consumed per week during pregnancy, based on maternal report at baseline
Number of children born to mother	Number of children born to the same mother, based on maternal report at baseline
Number of adults in household	Number of adults in the household, based on maternal report at baseline
Biological father lives in household	Dichotomous measure of whether a biological father lives in the household, based on maternal report at baseline
Household combined income	Combined household income from previous calendar year (see table note), coded into six mutually exclusive indicators (less than $10,000; $10,000 to $14,999; $15,000 to $19,999; $20,000 to

	$29,999; equal to or more than $30,000; and unknown), based on maternal report at baseline
Household net worth	Coded into six mutually exclusive indicators (debt is equal to or more than $5,000; debt is $0 to $4,999; breaking even; leftover is $0 to $4,999; leftover is equal to or more than $5,000; unknown), based on maternal report at baseline
Child age at interview[a]	Child age in months at time of age-4 assessment
Language of assessment[a]	Language (English or Spanish) that age-4 child assessment or maternal survey (as relevant) was administered in (see Appendix for protocol used to determine which language to use)
Interviewer identifier[a]	Dichotomous indicators of interviewer/assessor, based on a coded identifier for the individual interviewer who assess children or conducted the maternal survey at age 4. These indicators are included in regression models involving continuous outcomes but not, owing to reasons of multicollinearity, for the dichotomous outcomes.
Usable epochs[a]	Number of usable epochs that remained after age 4 EEG data cleaning

Note: Household combined income is calculated to sum the total household income, government income (social program benefits such as the Supplemental Security Income and the Unemployment Insurance), and other income (business profit, aid from relatives, and child support benefits). Outcomes of resting brain activity and EEG power only include usable epochs as a covariate.
[a]Measure was not preregistered as a baseline control variable. All other listed covariates were measured prior to random assignment.

Supplemental Table 3. Descriptive Statistics and Estimated Cash-Gift Effect on Preregistered Child Measures at Age 4 Using Robust Regression

	Low Cash Gift Group mean (sd)	High Cash Gift Group mean (sd)	OLS (se) w/FE	OLS (se) w/FE w/covariates	Effect Size	p-value	N
PRIMARY OUTCOMES							
ROWPVT conceptual score	40.998 (16.204)	39.199 (15.517)	-2.003 (1.182)	-0.531 (1.084)	-0.033	0.625	700
MEFS standard score	93.746 (10.480)	94.554 (10.596)	0.788 (0.549)	0.836 (0.567)	0.080	0.141	783
Child Behavior Checklist	22.175 (12.928)	22.393 (12.504)	0.575 (0.873)	0.512 (0.895)	0.040	0.567	882
Resting Brain Activity	2.920 (1.369)	3.051 (1.349)	0.125 (0.091)	0.094 (0.093)	0.069	0.313	637
SECONDARY OUTCOMES							
Matrices raw score	7.625 (3.359)	7.304 (3.308)	-0.268 (0.242)	-0.278 (0.256)	-0.083	0.279	769
Reading House raw score	3.016 (2.242)	2.806 (2.174)	-0.221 (0.150)	-0.088 (0.153)	-0.039	0.563	754
Developmental Condition Diagnosis	0.137 (0.344)	0.146 (0.354)	0.010 (0.023)	0.010 (0.022)	0.052	0.655	878

Note: This table shows descriptive characteristics and estimated ITT effects for all preregistered child outcomes. Effect sizes and p-values come from robust regressions with site fixed-effects and, in the fourth column, controls for covariates, listed in Supplemental Table 2. The effect sizes of continuous outcomes are computed by dividing the treatment effect by the standard deviation of the low-cash sample. ROWPVT=Receptive One-Word Picture Vocabulary Test. MEFS=Minnesota Executive Function Scale. CBCL=Child Behavior Checklist. The effect size of impacts on the developmental condition binary outcome (Developmental Condition Diagnosis) are expressed as Cohen's d, which is converted from the log odds estimated with a logistic regression using: $\text{Log odds ratio} \times \frac{\pi}{\sqrt{3}}$.

Supplemental Table 4. Descriptive Statistics and Estimated Cash-Gift Effect on non-preregistered Child Measures at Age 4

	Low Cash Gift Group mean (sd)	High Cash Gift Group mean (sd)	OLS (se) w/FE	OLS (se) w/FE w/covariates	Effect Size	p-value	N
PSRA: Attentive/Impulse Control	36.314 (13.764)	36.462 (13.937)	0.126 (0.965)	0.313 (0.952)	0.023	0.742	833
PSRA: Positive Emotion	17.653 (4.972)	17.638 (5.429)	0.014 (0.355)	0.022 (0.327)	0.004	0.946	833

Note: This table shows descriptive characteristics and estimated ITT effects for two non-preregistered child outcomes. Effect sizes and p-value come from regressions with site fixed-effects and, in the fourth column, controls for covariates, listed in Supplemental Table 2. The effect sizes are computed by dividing the treatment effect by the standard deviation of the low-cash sample. PSRA=Preschool Self-regulation Assessment.

Supplemental Table 5. Cash-gift treatment effect size estimates for base and covariate-adjusted models, applying no weights, inverse probability of treatment weights to correct for possible bias.

	Unweighted (taken From Table 3)		Corrected for imbalance between high and low cash groups at Age 4		Corrected for imbalance between Age-4 and baseline samples		
	Effect Size	p-value	Effect Size	p-value	Effect Size	p-value	Unweighted N
ROWPVT conceptual score	-0.026	0.717	-0.064	0.442	-0.026	0.721	700
MEFS standard score	0.055	0.474	0.022	0.802	0.055	0.472	783
CBCL index	0.011	0.872	0.013	0.867	0.010	0.876	882
Resting Brain Activity	0.097	0.184	0.064	0.427	0.097	0.184	637
Matrices raw score	-0.093	0.213	-0.099	0.215	-0.093	0.213	769
Reading House raw score	-0.026	0.705	-0.004	0.958	-0.026	0.708	754
Developmental Condition Diagnosis	0.052	0.655	0.050	0.714	0.052	0.675	878

Note: This table repeats Table 3 ITT effects in the first two columns and then shows ITT effect estimates weighted in two different ways to adjust for possible biases in the sample. Estimates in the third and fourth columns are estimated using Inverse Probability of Treatment Weights, which are intended to weight the high cash gift group to be as similar as possible to the low cash gift group. NRW estimates in the fifth and sixth columns are based on inverse probability weights intended to adjust for missing data and make the sample with available data as similar as possible to the baseline sample. The effect sizes of continuous outcomes are computed by dividing the treatment effect by the standard deviation of the low-cash sample. The effect size of impacts on the developmental condition binary outcome (Developmental Condition Diagnosis) are expressed as Cohen's d, which is converted from the log odds estimated with a logistic regression using: $\text{Log odds ratio} \times \frac{\pi}{\sqrt{3}}$. ROWPVT=Receptive One-Word Picture Vocabulary Test. MEFS=Minnesota Executive Function Scale. CBCL=Child Behavior Checklist.

Supplemental Table 6. Descriptive Statistics and Estimated Cash-Gift Effect on Preregistered Child Measures at Age 4 Using Multiple Imputation

	Low Cash Gift Group mean	High Cash Gift Group mean	OLS (se) w/FE	OLS (se) w/FE w/covariates	Effect Size	p-value	N
ROWPVT conceptual score	37.969	37.793	-0.159 (1.383)	-0.930 (1.133)	-0.057	0.412	995
MEFS standard score	90.026	92.369	2.338 (1.020)	0.809 (0.941)	0.077	0.391	995
CBCL index	21.536	22.037	0.478 (0.855)	0.093 (0.824)	0.007	0.910	995
Resting Brain Activity	2.982	3.085	0.108 (0.117)	0.105 (0.107)	0.077	0.326	995
Matrices raw score	7.321	6.998	-0.324 (0.281)	-0.397 (0.257)	-0.118	0.124	995
Reading House raw score	2.788	2.711	-0.074 (0.164)	-0.033 (0.155)	-0.015	0.833	995
Developmental Condition Diagnosis	0.150	0.151	0.001 (0.025)	0.008 (0.024)	0.039	0.750	995

Note: This table shows descriptive characteristics and estimated ITT effects for all preregistered child outcomes calculated with multiple imputation by chained equations, or MICE, using linear regression, logistic regression, and predictive mean matching and imputing 20 datasets. Five children were excluded from the analyses after the multiple imputation because they had died before the age 4 survey. To facilitate comparison to effect sizes in Table 3, the effect size of continuous outcome used the sample control group's standard deviation. The effect size of impacts on the developmental condition binary outcome (Developmental Condition Diagnosis) are expressed as Cohen's d, which is converted from the log odds estimated with a logistic regression using: $\text{Log odds ratio} \times \frac{\pi}{\sqrt{3}}$. ROWPVT=Receptive One-Word Picture Vocabulary Test. MEFS=Minnesota Executive Function Scale. CBCL=Child Behavior Checklist.

Supplemental Table 7. Sensitivity of Cash Gift Impacts on preregistered Child Measure at Age 4 to Overpayments

	Low Cash Gift Group mean (sd)	High Cash Gift Group mean (sd)	OLS (se) w/FE	OLS (se) w/FE w/covariates	Effect Size	p-value	N
Panel A. Including overpayments as a covariate							
PRIMARY OUTCOMES							
ROWPVT conceptual score	40.998	39.199	-1.878	-0.492	-0.030	0.676	700
	(16.204)	(15.517)	(1.198)	(1.176)			
MEFS standard score	93.746	94.554	0.690	0.491	0.047	0.544	783
	(10.480)	(10.596)	(0.763)	(0.809)			
CBCL index	22.175	22.393	0.236	0.139	0.011	0.872	882
	(12.928)	(12.504)	(0.865)	(0.860)			
Resting Brain Activity	2.920	3.051	0.137	0.137	0.100	0.176	637
	(1.369)	(1.349)	(0.108)	(0.101)			
SECONDARY OUTCOMES							
Matrices raw score	7.625	7.304	-0.346	-0.347	-0.103	0.164	769
	(3.359)	(3.308)	(0.243)	(0.249)			
Reading House raw score	3.016	2.806	-0.225	-0.075	-0.033	0.624	754
	(2.242)	(2.174)	(0.159)	(0.153)			
Developmental Condition	0.137	0.146	0.011	0.011	0.058	0.627	878
Diagnosis	(0.344)	(0.354)	(0.023)	(0.023)			
Panel B. Dropping overpayment cases							
PRIMARY OUTCOMES							
ROWPVT conceptual score	40.899	39.147	-1.761	-0.480	-0.029	0.688	682
	(16.297)	(15.287)	(1.216)	(1.196)			
MEFS standard score	93.630	94.429	0.768	0.547	0.052	0.508	763
	(10.483)	(10.695)	(0.777)	(0.825)			

CBCL index	22.275 (12.935)	22.338 (12.521)	0.043 (0.875)	-0.077 (0.867)	-0.006	0.929	860
Resting Brain Activity	2.920 (1.371)	3.057 (1.353)	0.143 (0.109)	0.138 (0.102)	0.101	0.176	621
SECONDARY OUTCOMES							
Matrices raw score	7.586 (3.347)	7.248 (3.306)	-0.331 (0.246)	-0.346 (0.253)	-0.104	0.171	749
Reading House raw score	3.007 (2.240)	2.769 (2.124)	-0.236 (0.161)	-0.096 (0.154)	-0.043	0.535	733
Developmental Condition Diagnosis	0.137 (0.344)	0.152 (0.359)	0.014 (0.024)	0.014 (0.025)	0.071	0.576	856

Note: This table shows descriptive characteristics and estimated ITT effects for all preregistered child outcomes considering overpayment. Overpayments of the cash gift occurred for 22 families providing data in the age 4 sample (14 high-cash gift and 8 low-cash gift families). Panel A retains all observations in which an overpayment of the cash gift was ever received. For regression results, receipt of the cash gift is incorporated in all regression models. Results presented in Panel B drop all observations in which an overpayment of the cash gift was ever received, n=22 for the age 4 sample providing any data for preregistered child outcomes. Effect sizes and p-value come from regressions with site fixed-effects and, in the fourth column, controls for covariates, listed in Supplemental Table 2. To prevent a perfect predictor issue in logistic regression, some missing indicators of baseline covariates were removed from the main model: missing indicator of alcohol drinks during pregnancy, smoking cigarettes during pregnancy, child's weight at birth, and child's gestational age. The inclusion and exclusion of missing indicators did not affect results. If the missing indicator was statistically significant predictor in a Linear Probability Model, it was retained in the logistic regression model which resulted in 3 cases being dropped because of perfect prediction. The effect sizes of continuous outcomes are computed by dividing the treatment effect by the standard deviation of the low-cash sample. The effect size of impacts on the developmental condition binary outcome (Developmental Condition Diagnosis) are expressed as Cohen's d, which is converted from the log odds estimated with a logistic regression using: $\text{Log odds ratio} \times \frac{\pi}{\sqrt{3}}$ (Bornstein et al., 2021). ROWPVT=Receptive One-Word Picture Vocabulary Test. MEFS=Minnesota Executive Function Scale. CBCL=Child Behavior Checklist.