

Systematic Differences in How Mothers Assess Children and Implications for Developmental Research

Cory Koedel
University of Missouri

Teerachat Techapaisarnjaroenkit*
University of Missouri

October 2012

We examine the agreeability between mothers and caregivers in assessing the non-cognitive development of young children. Our analysis extends the standard agreeability framework to consider systematic directional differences between mothers and caregivers across maternal subgroups. Using caregiver assessments as a benchmark, we find that minority mothers provide consistently more-favorable evaluations of their children relative to white mothers. Holding race constant, mothers who raise their children in less-favorable family structures also provide more-favorable evaluations. These patterns in the data cannot be explained by any obvious source. We consider several possible explanations, and discuss implications.

* We thank Irma Arteaga, Xiaoyan (Laura) Chen, Colleen Heflin and Peter Mueser for useful comments and suggestions. The usual disclaimers apply.

There is considerable interest among researchers and policymakers in the non-cognitive development of young children. This interest stems from two recurring findings throughout the research literature: (1) the early-childhood years are formative years in the human development process, and (2) individuals who develop stronger non-cognitive skills are more successful as adults. In synthesizing the findings from his work and the work of others, Heckman (2011) writes “inequality in early childhood experiences and learning produces inequality in ability, achievement, health and adult success” (p. 32). But research on children’s non-cognitive development is challenging. A key issue is that children’s mothers are often the only source of information about their development, particularly when children are very young.

We examine the agreeability between maternal and caregiver assessments of non-cognitive development for young children. Examples of previous work on the agreeability between parental and non-parental assessments include Achenbach et al. (1987), Cai et al. (2004), Chilcoat and Breslau (1997), Crane et al. (2011), Kinard (1995), Mitsis et al. (2000), Offord et al. (1996), Ritchers and Pellegrini (1989), Ritchers (1992), Winsler and Wallace (2002) and Youngstrom et al. (2000). Our study builds on this prior work and makes two primary contributions. First, we add to the relatively sparse literature on cross-informant agreeability for preschool-aged children (in the above list, only Cai et al. (2004), Crane et al. (2011) and Winsler and Wallace (2002) study preschoolers). Second, we extend our analysis beyond the initial question of agreeability and examine *systematic directional differences* between maternal and non-maternal assessments for different types of mothers (other studies that consider directional differences between maternal and non-maternal assessments include Youngstrom et al. (2000) and Murray et al. (2007)). We compare maternal subgroups divided by race and family-structure status, although our analytic framework can be extended to consider other divisions. For our racial analysis we compare African American, Hispanic and white mothers. Our

family-structure analysis compares mothers from six different family-structure groups, which we describe in more detail below.

We follow the general methodological approach established in prior research for evaluating assessor agreeability; that is, we compare maternal and non-maternal assessments for the same children. We evaluate over 4,000 age-4 children from different racial and socioeconomic backgrounds. Consistent with previous work in this area we find considerable disagreement between mothers and caregivers. In our racial analysis we find that minority mothers are much more optimistic relative to caregivers than white mothers. We consider several explanations for this finding drawing on previous research. We also identify systematic directional differences in the agreeability between maternal and caregiver assessments across mothers who raise their children in different family structures. Mothers from less-favorable family structures, within race, are consistently more optimistic about their children's development. Unlike the differences in maternal assessments by race, the literature offers much less guidance to explain these findings, although we consider several possibilities.

We conclude with a discussion about the implications of our findings for future research and policy. First, we show that our results can inform research design, and the interpretation of findings, in studies of early-childhood development. Using an application taken from the family-structure literature to illustrate, we show that research results can depend in systematic ways on whether mothers or outside caregivers provide the non-cognitive assessment information for children. We also show the potential social consequences of our findings using objective scores for parents and children from a parent-child interaction task (the two-bags task – see Najarian et al., 2010). We show that mothers who provide relatively more-favorable assessments of their children (again, using caregiver assessments as a benchmark) get lower scores from raters during an observed child-parent interaction. Their children also receive lower scores.

I. Data and Method

Data

Our data are from the Early Childhood Longitudinal Survey, Birth Cohort (ECLS-B), provided by the National Center for Education Statistics (NCES). The ECLS-B began tracking children at birth and administered surveys when the children were aged 9-months, 2-years, 4-years and in kindergarten. We use data primarily from the age-4 survey. During this survey each child who was enrolled in non-relative childcare received two assessments of her non-cognitive development: one from her mother, and one from her childcare provider.

Table 1 lists fourteen questions that were asked to mothers and caregivers about the children. We use factor analysis to divide the questions into four subgroups: (1) Behavioral/General, (2) Eagerness and Independence, (3) Maturity and (4) Peer Interactions. The subgroup divisions are shown in the table. The loading weights from the factor analysis differ between the maternal and caregiver assessments. A typical approach in such circumstances is to use uniform weights within question subgroups. We take this approach in the present analysis to maintain consistency in outcome metrics across mothers and caregivers. In addition to evaluating the four subgroups independently, we also estimate models based on outcome measures that weight all questions across all subgroups uniformly.¹

We re-order the responses to each question so that they convey the same information directionally. Specifically, we reorder the responses where necessary so that for all questions a “1” indicates the least-favorable response and a “5” indicates the most-favorable response. For example, for the first question in Table 1 a response of “1” (never) is the most-favorable response, so we re-

¹ There were also three other questions that did not load consistently across the maternal- and caregiver-assessment data. Our primary analysis excludes these three questions due the loading ambiguity, but our findings are qualitatively robust to their inclusion in reasonable ways (e.g., we can include them in the subgroup suggested by either the maternal- or caregiver-assessment factor analysis and we obtain similar results).

code it as a “5”. For question eight, a response of “1” (never) is the least-favorable response, so it remains as initially coded.

We can only compare maternal and caregiver assessments for children who were enrolled in non-relative childcare during the age-4 ECLS-B survey. The first two columns in Table 2 show how this group of children compares to the full ECLS-B sample of African American, Hispanic and white children (combined for brevity).² Unsurprisingly, our estimation sample is a relatively advantaged group. On the one hand, the selectivity of our sample should not influence our ability to gain inference because the key condition for identification is not violated (namely, that we have maternal and caregiver assessments for the same children). However, the selectivity highlights an important practical limitation: in studies where disadvantaged children are of primary interest, outside assessments of non-cognitive development are less likely to be available because children are less likely to be enrolled in non-relative childcare. This suggests that in some instances the use of maternal assessments in research will be unavoidable, which points to the importance of better understanding the ways in which maternal assessments systematically differ from each other, and outside assessments.

When we move to our evaluation of family structure, we impose additional restrictions on the dataset based on children’s family structures over the course of the first three waves of ECLS-B survey. First, we start with children who can be categorized into one of two types of family structures at the time of the 9-month survey. The first structure is an intact family, which we define as a family where the biological mother and father both live in the same household with the child. The second structure is a female-headed household, which we define as a household where the mother is raising the child without a male cohabitant.

² Children who are coded as multi-race in the ECLS-B are excluded from our analysis.

Starting with these two baseline family structures we identify six family-structure subgroups based on changes to family structure over time, or lack thereof, between the 9-month and 4-year surveys. The six subgroups are (where again, we define intact families as families where the biological mother and father cohabit): (1) persistent intact families, (2) initially intact families at the time of the 9-month survey where the parents separated before the age-4 survey, (3) initially intact families where a separation occurred and the mother began cohabitating with a new male, all before the age-4 survey, (4) persistent female-headed households, (5) initially female-headed households at the time of the 9-month survey where a new cohabitating male was introduced prior to the age-4 survey, and (6) initially female-headed households where a new cohabitating male was introduced and a separation occurred, all before the age-4 survey. For initially female-headed households, we drop cases where the incoming male in categories (5) and (6) is the biological father, but this occurs only in a very small fraction of families.³

The final column in Table 2 provides summary statistics for our family-structure analytic sample. The table shows the share of families that display each structure. We report our findings for all family-structure groups, but focus on the largest groups when we present our results: groups (1), (2) and (4). Children from families where the structure does not fit into any of the above-listed groups are excluded from the family-structure portion of our analysis.

Empirical Strategy

We estimate models that compare maternal and caregiver responses within children using child fixed effects. We use the following empirical model, beginning with our analysis by race:

$$Y_{ij} = \delta_{iq} + D_j \alpha_1 + (D_j * R_i) \alpha_2 + \varepsilon_{ij} \quad (1)$$

³ We acknowledge that these family-structure designations are imperfect, mostly in the sense that they are too broad. For example, Brown (2004) shows that children whose biological parents cohabit but are not married are worse off, on average, than children with married biological parents; but our family-structure categories treat these family types as being the same. Brown's findings are based on an analysis of older children, but they may generalize to younger children as well. We do not evaluate family structure in greater depth because it is not the primary focus of our paper. That said, the general lessons from our study will carry over to research that uses more-detailed definitions of family structure.

In (1), Y_{ij} measures the non-cognitive well-being of child i as measured by question-subgroup q , assessed by person j , where $j = m$ (mother) or c (caregiver). δ_{iq} is a vector of child fixed effects for that question subgroup, D_j is an indicator variable equal to one if $j = m$ and zero otherwise, and R_i is a vector of indicator variables for race. Because the data are ordinal, we code a binary variable set to one for each outcome whenever the response is most-favorable, and take the average of the binary variables by group as our outcome measure.⁴

The model does not include any child-, parent- or caregiver-level controls because these controls are absorbed by the child-by-question fixed effects. The only within-child variability in the model comes from which respondent is providing the assessment (mother or caregiver). Our estimates of α_1 measure the baseline disparity between the maternal and caregiver assessments, and our estimates of α_2 measure the differential mother-caregiver response gaps by race. We omit white children from the vector R_i , which means that the maternal-caregiver response gaps for African American and Hispanic children are estimated relative to the gaps for white children (whose response gaps are reflected by the baseline coefficient, α_1). Recall that we re-sort the responses to the questions so that higher-order responses always indicate more-positive outcomes. Therefore, positive coefficients for African American and Hispanic mothers indicate that these mothers provide more favorable assessments relative to caregivers than their white counterparts.

We follow a similar analytic approach to compare mothers who raise their children in different family structures. The family-structure models are estimated as follows:

$$Y_{ij} = \gamma_{iq} + D_j\beta_1 + (D_j * F_i)\beta_2 + e_{ij} \quad (2)$$

⁴ We also estimated models where we averaged the ordinal variables to create outcome measures instead of using the binary measures. Although it is not technically correct to average the ordinal variables, as a practical matter we typically obtained similar results from models for the averaged outcomes (although the results are weaker statistically in general). See Table 4, and Appendix Tables A.1 and A.2, for information about the similarity between our binary measures and the average ordinal scores in the data for each non-cognitive question. As would be expected, higher average ordinal scores are associated with a larger fraction of most-positive responses.

Equation (2) is analogous to equation (1) except that we replace the vector of race indicators with a vector of family-structure indicators. The omitted group includes mothers from persistent intact families – that is, families where the mother and father cohabitated continuously between the 9-months and 4-years ECLS-B surveys.

Our estimates of β_2 reflect differences in the gaps between the maternal and caregiver assessments for mothers from different family structures. Positive values indicate that mothers from less-favorable family structures are relatively more optimistic than mothers from persistent intact families, compared to caregivers. We estimate equation (2) separately for African American, Hispanic and white children. We divide the family-structure analysis by race for two reasons. First, our results from equation (1) reveal that maternal-caregiver assessment gaps are not consistent across races (see below). Second, the composition of family structure varies dramatically by race, particularly between whites/Hispanics and African Americans, as shown Table 3 (also see Cancian and Reed, 2009; Martin et al., 2010).

We extend our family-structure analysis with an application to research. Based on the general literature relating differences in family structure to early-childhood development, we model developmental outcomes as follows:⁵

$$Y_i = \lambda_0 + X_i\lambda_1 + F_i\lambda_2 + u_i \quad (3)$$

In (3), Y_i is a non-cognitive outcome measure for child i based on one of four question subgroups (or the all-question average); X_i a vector of observable information about the child and his or her family, and F_i is the vector of family-structure indicators. The controls that we include in the vector X_i are for maternal and paternal (biological) education, whether the mother’s first language was English, whether the child was born to a teenage mother, whether the child is a twin (oversample

⁵ Our model is based on prior studies of family-structure effects. There are many similar models estimated in the literature – see, for example, Carlson and Corcoran (2001) and Painter and Levine (2000, 2004).

population in the ECLS-B), the child's number of siblings, the number of family members who are 18 or younger living in the household, child gender, and child age.⁶ Similarly to equation (2), we estimate equation (3) separately by race and omit persistent intact families as the comparison group. We estimate the model and compile results separately using the maternal and caregiver assessments. Of interest in model (3) is how our findings differ depending on whether we use the maternal or caregiver assessments to measure outcomes.

II. Results

Table 4 presents an example of the basic summary statistics that underlie the regression analysis. For each survey question in the data, the table reports the average of the ordinal outcome as reported by mothers and caregivers, and the share of responses that indicate the most-positive outcome. For brevity, Table 4 shows results for white children only – we replicate the table separately for African American and Hispanic children in Appendix A.

Table 5 shows results from the initial regressions as described by model (1). The rows of the table are for different models by question subgroup, and the columns show estimates for the maternal-response effects. The maternal-assessment effects for white mothers can be read from the table by the estimates in column (1). For example, the upper-left estimate of -0.105 in Table 5 indicates that the maternal assessments from white mothers, on average, are considerably less likely to indicate a most-favorable response relative to the caregiver assessments for the same children. For African American and Hispanic mothers, the maternal-assessment effects are equal to the sum of the estimate in the first column and the relevant estimate by race.

Table 5 reveals large and statistically significant differences in the maternal-caregiver response gaps by race. The estimates reveal that minority mothers provide more-favorable

⁶ We do not include family income in the model because it is partly determined by family structure; however, as a practical matter our key findings are not qualitatively affected if we include income controls. In fact, our findings are qualitatively similar when we estimate several variants of the specification in equation (3). Because the correct specification of (3) is not central to our analysis, we omit these robustness results for brevity.

assessments of their children relative to white mothers in all developmental areas. The response gaps are most-similar between races along the peer-interaction dimension, but they differ suggestively even there.

Next, in Table 6 we compare mothers who raise their children in different family structures. The three panels in Table 6 (A-B-C) are for white, African American and Hispanic children. Again, the rows of the table are for different models by question category, and the columns show estimates for the baseline maternal-response effects and differences in the maternal-response effects by family-structure group. Recall that there are three common family structures in the data: persistent intact families, intact families where the mother and father separate (but there is no new male cohabitant), and persistent female-headed families (see Table 2). The estimates that compare these three groups to each other are bolded in the tables.

The comparison mothers in each model in Table 6 are from persistent intact families. Analogously to Table 5, the estimates for these mothers are reflected in column 1, and the estimates for mothers from other family-structure types can be identified from the table by summing the estimates from column (1) and the relevant family-type column. A positive estimate in columns (2) through (6) indicates that mothers from the specified family-structure type are more-positive about their children when compared to the caregiver, relative to mothers from persistent intact families when compared to the caregiver. For white children the patterns in Table 6 are the strongest and most-precisely estimated (panel A). Almost all of the point estimates that compare mothers from the less-favorable family structures to mothers from persistent intact families are positive. For the main family-structure comparisons, many of the estimates are large and statistically significant.

Our estimates for African American and Hispanic mothers in Tables 6B and 6C are less precise, and therefore less compelling. This is largely due to the fact that our samples for minority children are much smaller than for whites (see Table 3). Nonetheless, for Hispanic mothers (panel

C) our point estimates are qualitatively consistent with the findings for white mothers. For African American mothers our estimates are somewhat consistent with the results for white mothers for persistent female-headed households, but not for intact families that experience a separation.

III. Sensitivity Analysis and Robustness

First we consider three potential confounding issues that may explain our findings from Table 5 (by race). First, minority mothers could send their children to childcare at earlier or later ages, in which case caregivers could be more or less familiar with the children at the time of the age-4 ECLS-B survey (although the predicted effect of caregiver-child familiarity on the assessments is unclear *ex ante*). Second, caregivers who watch wealthier children could be generally more optimistic than caregivers who watch lower-income children. Because minority children are more likely to be from low-income families, this could also explain the maternal-caregiver response gaps (i.e., they would be driven by differences in caregiver responses). Third, our results could be driven by racial incongruence between caregivers and minority children. Racial incongruence has been shown to be important in other contexts (Dee 2004, 2005; Grissom and Keiser, 2011) and empirically, minority children are more likely to have a racially incongruent caregiver.⁷ Racial incongruence may lead to caregiver assessments that are too negative, which could explain our findings in Table 5 even if there were no real differences in how mothers assess their children across races.

We investigate the potential for each of the above explanations to influence our findings in turn. To address the familiarity issue we estimate the model only for children who were reported to have been in care in the same facility for at least 6 months at the time of the age-4 ECLS-B survey. To address the income issue, we exclude high-income families from the analysis (i.e., we drop all families with household income above \$50,000). To address the racial-incongruence issue, we augment the model to directly control for incongruence between children and their caregivers. Our

⁷ For white, African American and Hispanic children, racial incongruence between the caregiver and child occurs for 10, 42 and 57 percent of observations, respectively.

results from these supplementary analyses are reported in Appendix B. In all cases, the differences between white and minority mothers remain similar to what we report in Table 5.

We perform a similar sensitivity analysis for our family-structure findings, focusing on white children only (for brevity, and due to the small samples of minority children). Specifically, we replicate our results in Table 6 for the subsample of children who were in care in the same facility for at least 6 months, and for the subsample of children with household income less than or equal to \$50,000.⁸ The motivation for these data divisions is similar to what we describe above. We report our findings in Appendix C, and again, our estimates in the appendix are consistent with our estimates from the main models as shown in Table 6.

Overall, our sensitivity analyses do not uncover any obvious confounding issues that can account for our findings. One factor that may be important in explaining our results by race are differences in children's at-home responsibilities (e.g., see Summerhill, 2001). We are not aware of any research on differences in at-home responsibilities for children who are raised in different family structures; still, this may be an important explanation for our family-structure findings as well. Another possibility is that mothers from different racial backgrounds and/or family structures have different assessment standards, even holding the assessment criteria fixed. Both of these possibilities merit attention in future research; we cannot definitively test the extent to which our results are driven by these competing factors with the ECLS-B data.

IV. Implications

Research Implications

We extend our family-structure analysis to consider the research implications of our findings. In particular, in Table 7 we show results from model (3) as outlined in the previous section, where

⁸ As noted above, racial incongruence is rare for white children. Nonetheless, in results omitted for brevity we verify that differences in racial incongruence across white children from different family structures are not driving our findings in Table 6A.

we estimate the model separately using outcome data from each respondent (mother and caregiver). We present results for white children only. The rows in Table 7 compare the different family-structure types and the columns are for different models. For example, consider the first two entries in the upper-left panel of the table: -0.018 and -0.040. The first number shows that if we use the maternal-assessment data, we find that the average value of the top-score variable for children from families that experience a separation is 1.8 percentage points lower than for children from persistent intact families. If we use the caregiver-assessment data and perform the same analysis, these children look marginally worse off – the gap is 4.0 percentage points. More generally, whenever the estimate in the “caregiver” column is more negative than the estimate in the “mother” column, it means that the caregiver assessments suggest that less-favorable family structures have more adverse consequences for children than the corresponding assessments from mothers themselves.

The pattern of the estimates in Table 7 shows that when we apply the same model to evaluate the same intervention for the same children, our choice of whether to use maternal or caregiver assessments can systematically impact the results. Using maternal assessments instead of caregiver assessments consistently produces smaller estimates of the associations between non-cognitive development and family structure. That is, the coefficients based on the caregiver assessments are more negative than the analogous coefficients based on the maternal responses, reflecting a larger gap between children from intact families and children being raised in less-favorable family structures. In some instances the maternal assessments reveal no differences, on average, between children; whereas the caregiver assessments identify a statistically significant gap (e.g., see the gaps between children from intact families and children from persistently female-headed families in the “Behavioral/General” category).

Social Implications

Next we briefly examine whether the maternal-caregiver response gaps that we identify above are related to outside assessments of children and their mothers. We use data from the two-bags task, a semistructured play task, for this portion of our analysis. Scores on the two-bags task are from trained raters who watch children and their parents interact, then evaluate mothers and children along several dimensions (see Najarian et al. (2010) for additional details about this assessment). We use factor analysis to produce a single score from the two-bags task for each mother and child.

We incorporate the scores from the two-bags task into our analysis by estimating the following models:

$$Y_{ij} = \lambda_{iq} + D_j \tau_1 + (D_j * Q_i^M) \tau_2 + \eta_{ij} \quad (4)$$

$$Y_{ij} = \pi_{iq} + D_j \phi_1 + (D_j * Q_i) \phi_2 + \xi_{ij} \quad (5)$$

Equations (4) and (5) are analogous to equations (1) and (2) above. That is, Y_{ij} again measures the non-cognitive well-being of child i as measured by question-subgroup q , assessed by person j , where $j = m$ (mother) or c (caregiver), and D_j is also defined as above. The key difference is that we now look for systematic differences across mothers and children who receive different ratings on the two-bags task. Specifically, in equation (4) we divide mothers into three groups based on our factor analysis of their two-bags-test scores: the top 30 percent, the middle 40 percent, and the bottom 30 percent. The variable Q_i^M indicates the placement of child i 's mother into one of these three groups. Similarly, the variable Q_i in equation (5) divides *children* into three groups in the same way based on their two-bags-task ratings.

Equation (4) is designed to answer the following question: Do high-scoring mothers on the two-bags test assess their children in ways that systematically differ from how low-scoring mothers

assess their children? The output from equation (5) answers a complementary question: Do high-scoring *children* on the two-bags test have mothers who assess them in ways that systematically differ from the maternal assessments received by their low-scoring peers.

Our findings from equation (4) and (5) are reported in Tables 8 and 9, respectively. The omitted comparison group in each table is the middle group – that is, the middle 40 percent of mothers and children based on their two-bags scores.⁹ Correspondingly, the response gap between mothers and caregivers for this middle group is reflected by the main “mother response effect” in each row of each table, consistent with our previous analysis; while the coefficients in the last two columns of each table reflect differential assessments for high- and low-scoring mothers (Table 8) or children (Table 9) relative to the middle group.

The estimates in Tables 8 and 9 indicate that differences in maternal assessment behaviors are reflected in the two-bags scores. In particular, Table 8 shows that higher-scoring mothers generally provide less-favorable assessments of their children. Complementing this result, Table 9 shows that higher-scoring *children* on the two-bags test have mothers who rate them less favorably.

V. Concluding Remarks

We use a simple statistical framework to assess systematic differences in maternal-caregiver assessment gaps across mothers that differ by race and family-structure status. Differences emerge along both dimensions. It has been well-documented that maternal and non-maternal assessments disagree for the same children; however, our analysis highlights the importance of systematic directional disagreement across observably-different maternal groups, an issue that has received much less attention in research.

Our findings have implications for research design and interpretation in early-childhood development. We illustrate this point using an example based on the family-structure research

⁹ Tables 8 and 9 show estimates based on our analysis of the entire analytic sample, but in an omitted analysis we confirm that these findings are maintained qualitatively if we focus on white children only.

literature, where we show that researchers may draw different conclusions depending on whether they use maternal or caregiver assessments. We also show that systematic differences in how mothers assess their children are important in substantive ways. Specifically, we use independent assessments from raters, who observe mothers and children interacting in a semistructured play task, to show that maternal-assessment differences are reflected in rater-provided scores for mothers and children. Our results show that mothers who provide comparatively less-favorable assessments of their children score higher on the play task, as do their children.

References

- Achenbach, T. M., S. H. McConaughy and C. T. Howell. 1987. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 41, 213–232.
- Borghans, Lex, Angela Lee Duckworth, James J. Heckman and Bas ter Weel. 2008. The Economics and Psychology of Personality Traits. *Journal of Human Resources*, 43(4): 972-1059.
- Briggs-Gowan, M. J., A. S. Carter and M. Schwab-Stone. 1996. Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Child Psychology*, 24, 749–765
- Bronte-Tinkew, Jancinta, Allison Horowitz and Mindy E. Scott. 2009. Fathering With Multiple Partners: Links to Children’s Well-Being in Early Childhood. *Journal of Marriage and Family*, 71(3): 608 – 631.
- Brown, Susan L. 2004. Family Structure and Child Well-Being: The Significance of Parental Cohabitation. *Journal of Marriage and Family*, 66(2): 351-367.
- Cai, Xinsheng, Ann P. Kaiser and Terry B. Hancock. 2004. Parent and Teacher Agreement on Child Behavior Checklist Items in a Sample of Preschoolers From Low-Income and Predominantly African American Families. *Journal of Clinical Child and Adolescent Psychology*, 33(2): 303-312.
- Cancian, M. and D. Reed. 2009. “Changes in Family Structure, Childbearing, and Employment: Implications for the Level and Trend in Poverty.” In M. Cancian and S. Danziger (Eds), *Changing Poverty*, NY: Russel Sage.
- Carlson, Marcia J. and Mary E. Corcoran. 2001. Family Structure and Children's Behavioral and Cognitive Outcomes. *Journal of Marriage and Family*, 63(3): 779-792.
- Cherlin, Andrew J., Frank F. Furstenberg Jr., P. Lindsay Chase-Lansdale, Kathleen E. Kiernan, Philip K. Robins, Donna Ruane Morrison, and Julien O. Teitler. 1991. Longitudinal studies of effects of divorce on children in Great Britain and the United States. *Science Magazine*, 252 (5011): 1386-1389.
- Chilcoat, Howard D. and Naomi Breslau. 1997. Does Psychiatric History Bias Mothers’ Reports? An Application of a New Analytic Approach. *Journal of the American Academy of Child and Adolescent Psychiatry*, 36(7): 971-979.
- Crane, Jennifer, Melissa S. Mincic and Adam Winsler. Parent-Teacher Agreement and Reliability on the Devereux Early Childhood Assessment (DECA) in English and Spanish for Ethnically Diverse Children Living in Poverty. *Early Education and Development*, 22(3): 520-547.
- Dee, Thomas.S. 2004. Teachers, Race, and Student Achievement in a Randomized Experiment. *Review of Economics and Statistics*, 86(1): 195-210.

- . 2005. A Teacher Like Me: Does Race, Ethnicity or Gender Matter? *American Economic Review*, 95(2): 158-165
- Duncan, Greg J., Katherine A. Magnuson and Jens Ludwig. 2004. The Endogeneity Problem in Developmental Studies. *Research in Human Development*, 1(1&2): 59-80.
- El Nokali, Nermeen E., Heather J. Bachman, and Elizabeth Votruba-Drzal. 2010. Parent Involvement and Children's Academic and Social Development in Elementary School. *Child Development*, 81(3): 988-1005.
- Grissom, Jason A., and Lael Keiser. 2011. A Supervisor Like Me: Race, Representation, and the Satisfaction and Turnover Decisions of Public Sector Employees. *Journal of Policy Analysis and Management* 30(3): 557-580.
- Heckman, James J. 2011. "The Economics of Inequality: The Value of Early Childhood Education." *American Educator*, Spring 2011: 31-47.
- Kinard, E. Milling. 1995. Mother and Teacher Assessments of Behavior Problems in Abused Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 34(8): 1043-1053.
- Martin, J. A., B. E. Hamilton P. D. Sutton and S. J. Ventura. 2010. "Births: Final data for 2008." *National vital statistics reports*, 59(1). Hyattsville, MD: National Center for Health Statistics.
- McCloskey, Laura A., Aurelio José Figueredo, and Mary P. Koss. 1995. The Effects of Systemic Family Violence on Children's Mental Health. *Child Development*, 66(5): 1239-1261.
- McMunn, Anne M., James Y. Nazroo, Michael G. Marmot, Richard Boreham and Robert Goodman. 2001. Children's Emotional and Behavioral Well-Being and the Family Environment: Findings from the Health Survey for England. *Social Science & Medicine*, 53: 423-440.
- Mitsis, E. M., K. E. McKay, K. P. Schulz, J. H. Newcorn and J. Halperin. 2000. Parent-teacher concordance for DSM-IV attention- deficit/hyperactivity disorder in a clinic-referred sample. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 308-313.
- Murray, Desiree W., Scott H. Kollins, Kristina K. Hardy, Howard B. Abikoff, James M. Swanson, Charles Cunningham, Benedetto Vitiello, Mark A. Riddle, Mark Davies, Laurence L. Greenhill, James T. McCracken, James J. McGough, Kelly Posner, Anne Skrobala, Tim Wigal, Sharon B. Wigal, Jaswinder K. Ghuman and Shirley Z. Chuang. 2007. Parent Versus Teacher Ratings of Attention-Deficit/Hyperactivity Disorder Symptoms in the Preschoolers with Attention-Deficit/Hyperactivity Disorder Treatment Study (PATS). *Journal of Child and Adolescent Psychopharmacology*, 17(5): 605-619.
- Najarian, Michelle, Kyle Snow, Jean Lennon, Susan Kinsey, Gail Mulligan. 2010. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Preschool – Kindergarten Psychometric Report. Institute for Education Sciences.
- Offord, David R., Michael H. Boyle, Yvonne Racine, Peter Szatmari, Jan E. Fleming, M ark Sanford, Ellen L. Lipman. 1996. Integrating Assessment Data from Multiple Informants. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35(8): 1078-1085.

Painter, G. and D. I. Levine. 2000. "Daddies, Family Structure and Youths' Outcomes: Which Correlations are Causal?" *The Journal of Human Resources*, 35(3): 524-549.

--. 2004. "Daddies, Devotion, and Dollars: How Do They Matter for Youth?" *American Journal of Economics and Sociology*, 63(4): 813-850.

Peterson, James L. and Nicholas Zill, 1986. Marital Disruption, Parent-Child Relationships, and Behavior Problems in Children. *Journal of Marriage and Family*, 48(2): 295-307.

Petterson, Stephen M. and Alison B. Albers. 2001. Effects of Poverty and Maternal Depression on Early Child Development. *Child Development*, 72(6): 1794 – 1813.

Pettit, Gregory S., Robert D. Laird, Kenneth A. Dodge, John E. Bates, and Michael M. Criss. 2001. Antecedents and Behavior-Problem Outcomes of Parental Monitoring and Psychological Control in Early Adolescence. *Child Development*, 72(2): 583–598.

Richters, J. 1992. Depressed mothers as informants about their children: A critical review of the evidence for distortion. *Psychological Bulletin*, 112: 485-499.

Richters J. and D. Pellegrini. 1989. Depressed Mothers' Judgments about Their Children: An Examination of the Depression-Distortion Hypothesis. *Child Development*, 60(5): 168-1075.

Roosa, Mark W., Shiyong Deng, Ehri Ryu, Ginger L. Burrell, Jenn-Yun Tein, Sarah Jones, Vera Lopez and Sakina Crowder. 2005. Family and Child Characteristics Linking Neighborhood Context and Child Externalizing Behavior. *Journal of Marriage and Family*, 67(2): 515-529.

Summerill, Lynette. 2001. The Chore of Chores. *ASU Research*, Fall 2001: 10-11.

Watanabe, Sarah E., Deborah A. Phillips, Taryn W. Morrissey, Kathleen McCartney and Kristen Bub. 2011. Double Jeopardy: Poorer Social-Emotional Outcomes for Children in the NICHD SECCYD Experiencing Home and Child-Care Environments That Confer Risk. *Child Development*, 82(1): 48-65.

Winsler, Adam and Gregory L. Wallace. 2002. Behavior Problems and Social Skills in Preschool Children: Parent-Teacher Agreement and Relations with Classroom Observations. *Early Education and Development* 13(1), 41-58.

Youngstrom, E., R. Loeber and M. Stouthamer-Loeber. 2000. Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology*, 68, 438–450.

Table 1. Questions from ECLS-B Survey Used to Measure Non-Cognitive Development.

Behavioral / General Well-Being

1. Child seems unhappy
2. Child worries about things
3. Child is physically aggressive
4. Child has temper tantrums
5. Child annoys other children

Eagerness and Independence

6. Child shows eagerness to learn new things
7. Child works or plays independently (without the need for adult direction)

Maturity

8. Child pays attention well
9. Child is overly active
10. Child has difficulty concentrating

Interactions with Peers

11. Child makes friends easily
 12. Child stands up for others rights
 13. Child comforts others
 14. Child tries to understand others
-

Notes: The “most favorable” response is the one that suggests the best outcome for children. For example, for Question 1, the most-favorable response is “never”; for Question 8, it is “very often.”

Table 2. Weighted Sample Means for Our Analytic Sample from the Age-4 ECLS-B Survey.

	All African American, Hispanic and white Children	Estimation Sample for Racial Analysis	Estimation Sample for Family Structure Analysis
<i>Basic Demographics</i>			
White	0.58	0.61	0.62
African American	0.15	0.15	0.15
Hispanic	0.27	0.24	0.23
Female	0.48	0.48	0.48
<i>Family Income (9-month Survey)</i>			
Less than 20,000	0.25	0.22	0.21
20,001– 35,000	0.24	0.22	0.22
35,001-50,000	0.15	0.15	0.15
50,001-75,000	0.16	0.17	0.18
75,001-100,000	0.10	0.11	0.12
More than 100,000	0.10	0.12	0.13
<i>Maternal Education (9-month Survey)</i>			
Less than High School	0.17	0.14	0.13
High School	0.30	0.28	0.17
Some College	0.28	0.29	0.29
Bachelor’s Degree	0.17	0.20	0.22
Higher than Bachelor’s Degree	0.07	0.09	0.09
Did not know or Refused	0.00	0.00	0.00
Teen Mom	0.12	0.11	0.10
Mother Native English Speaker	0.18	0.15	0.15
<i>Family Structure Group</i>			
(1) Persistent intact family	0.69	0.68	0.74
(2) Initial intact family with separation	0.07	0.07	0.08
(3) Initial intact family with separation and new male cohabitant	0.02	0.02	0.02
(4) Persistent female-headed family	0.11	0.12	0.13
(5) Initial female-headed with new male cohabitant	0.02	0.02	0.03
(6) Initial female-headed with new male cohabitant and subsequent separation	0.00	0.00	0.00
Family structure is unknown or does not fit any of the above-listed categories	0.08	0.08	0.00
Unweighted Sample Size	7,005	4,662	4,225

Note: We exclude all children from our analysis who are categorized as multi-race or any race that is not African American, Hispanic or white.

Table 3. Family Structure Classifications by Race. Family-Structure Analytic Sample.

	Sample Means		
	White	African American	Hispanic
<i>All</i>			
Persistent intact family	0.84	0.36	0.71
Initial intact family with separation	0.07	0.10	0.10
Initial intact family with separation and new male cohabitant	0.02	0.01	0.02
Persistent female-headed family	0.04	0.47	0.14
Initial female-headed with new male cohabitant	0.02	0.05	0.02
Initial female-headed with new male cohabitant and subsequent separation	0.00	0.02	0.00
N	2529	790	906
<i>Low Income</i>			
Persistent intact family	0.63	0.25	0.65
Initial intact family with separation	0.17	0.11	0.12
Initial intact family with separation and new male cohabitant	0.04	0.01	0.03
Persistent female-headed family	0.11	0.56	0.17
Initial female-headed with new male cohabitant	0.04	0.06	0.03
Initial female-headed with new male cohabitant and subsequent separation	0.01	0.02	0.00
N	874	621	641
<i>Low Maternal Education</i>			
Persistent intact family	0.66	0.21	0.69
Initial intact family with separation	0.14	0.09	0.10
Initial intact family with separation and new male cohabitant	0.05	0.01	0.02
Persistent female-headed family	0.09	0.60	0.17
Initial female-headed with new male cohabitant	0.05	0.06	0.02
Initial female-headed with new male cohabitant and subsequent separation	0.01	0.03	0.01
N	660	430	521

Notes: The low-income group includes all families where household income was reported to be \$50,000 annually or less at the time of the age-4 follow-up survey (similar family-structure discrepancies exist between races if we use an income cutoff of \$35,000). The low-maternal-education group includes families where the mother's education does not exceed a high-school diploma.

Table 4. Mother and Caregiver Responses by Question, Whites.

	Average Ordinal Response		Share of Responses that are Most Positive	
	Mothers	Caregivers	Mothers	Caregivers
Question 1	4.19	4.24	0.43	0.33
Question 2	3.60	3.84	0.16	0.29
Question 3	3.78	4.23	0.18	0.51
Question 4	3.32	4.07	0.07	0.43
Question 5	3.78	4.06	0.15	0.39
Question 6	4.41	4.17	0.53	0.43
Question 7	4.09	4.03	0.33	0.33
Question 8	3.64	3.81	0.12	0.25
Question 9	3.37	3.75	0.12	0.29
Question 10	3.52	3.65	0.09	0.21
Question 11	4.26	4.16	0.42	0.40
Question 12	3.57	3.40	0.16	0.16
Question 13	3.71	3.39	0.18	0.16
Question 14	3.62	3.31	0.16	0.13
Simple Average	3.78	3.87	0.22	0.31

Note: Maternal and caregiver responses are re-ordered where necessary so that higher values correspond to more-positive outcomes throughout, as described in the text.

Table 5. Differential Response Gaps Between Mothers and Caregivers by Race.

	Relative to White Mothers:		
	Mother Response Effect	African American Mother Differential Response	Hispanic Mother Differential Response
<u>Models</u>			
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>			
All Outcomes	-0.105 (0.006)**	0.060 (0.013)**	0.050 (0.011)**
Behavioral / General	-0.232 (0.009)**	0.087 (0.019)**	0.081 (0.016)**
Eagerness and Independence	0.053 (0.011)**	0.056 (0.022)*	0.011 (0.022)
Maturity	-0.142 (0.008)**	0.046 (0.017)**	0.071 (0.016)**
Peer Interactions	0.017 (0.008)†	0.030 (0.018)†	0.016 (0.016)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: Robust standard errors are in parentheses. ECLS-B sample weights are used in all regressions.

Table 6A. Differential Response Gaps Between Mothers and Caregivers by Family Structure Category. White Children.

	Mother Response Effect	Initially Intact with Separation	Initially Intact with Separation, Followed by New Male Cohabitant	Persistent Female-Headed Household	Initially Female-Headed Household with Male Cohabitant Entrant	Initially Female-Headed Household with Male Cohabitant Entrant, followed by Separation
Relative to Mothers from Persistent Intact Families:						
<u>Models</u>						
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>						
All Outcomes	-0.116 (0.007)**	0.055 (0.022)*	0.117 (0.047)*	0.077 (0.027)**	0.003 (0.050)	0.108 (0.061)†
Behavioral / General	-0.251 (0.009)**	0.083 (0.031)**	0.219 (0.076)**	0.146 (0.088)**	0.065 (0.067)	0.076 (0.107)
Eagerness and Independence	0.054 (0.013)**	0.002 (0.042)	0.005 (0.076)	-0.001 (0.059)	-0.022 (0.076)	0.196 (0.196)
Maturity	-0.159 (0.009)**	0.076 (0.031)*	0.125 (0.055)*	0.111 (0.032)**	0.050 (0.060)	0.200 (0.102)*
Peer Interactions	0.016 (0.009)	0.012 (0.031)	0.034 (0.066)	0.022 (0.047)	-0.049 (0.060)	0.109 (0.111)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: Robust standard errors are in parentheses. ECLS-B sample weights are used in all regressions. See Section I of the text for more information about how the family-structure categories were constructed.

Table 6B. Differential Response Gaps Between Mothers and Caregivers by Family Structure Category. African American Children.

	Mother Response Effect	Initially Intact with Separation	Initially Intact with Separation, Followed by New Male Cohabitant	Persistent Female-Headed Household	Initially Female-Headed Household with Male Cohabitant Entrant	Initially Female-Headed Household with Male Cohabitant Entrant, followed by Separation
Relative to Mothers from Persistent Intact Families:						
<u>Models</u>						
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>						
All Outcomes	-0.069 (0.020)**	-0.034 (0.046)	-0.007 (0.139)	0.047 (0.027)†	0.010 (0.055)	0.147 (0.062)*
Behavioral / General	-0.199 (0.029)**	-0.001 (0.062)	0.085 (0.172)	0.098 (0.040)*	0.019 (0.081)	0.183 (0.111)†
Eagerness and Independence	0.153 (0.035)	-0.068 (0.078)	-0.091 (0.124)	-0.050 (0.047)	-0.166 (0.094)†	0.055 (0.128)
Maturity	-0.138 (0.027)**	0.002 (0.060)	-0.070 (0.149)	0.086 (0.035)*	0.018 (0.067)	0.193 (0.048)**
Peer Interactions	0.057 (0.029)*	-0.087 (0.059)	-0.199 (0.279)	-0.028 (0.037)	0.012 (0.076)	0.110 (0.129)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: See Notes from Table 6A.

Table 6C. Differential Response Gaps Between Mothers and Caregivers by Family Structure Category. Hispanic Children.

	Mother Response Effect	Initially Intact with Separation	Initially Intact with Separation, Followed by New Male Cohabitant	Persistent Female-Headed Household	Initially Female-Headed Household with Male Cohabitant Entrant	Initially Female-Headed Household with Male Cohabitant Entrant, followed by Separation
Relative to Mothers from Persistent Intact Families:						
<u>Models</u>						
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>						
All Outcomes	-0.064 (0.013)**	0.043 (0.033)	0.005 (0.054)	0.021 (0.032)	-0.010 (0.085)	0.137 (0.092)
Behavioral / General	-0.167 (0.018)**	0.064 (0.052)	0.023 (0.076)	0.069 (0.043)	-0.033 (0.104)	0.217 (0.100)*
Eagerness and Independence	0.068 (0.024)**	0.007 (0.066)	-0.013 (0.140)	-0.023 (0.060)	-0.092 (0.142)	-0.193 (0.246)
Maturity	-0.086 (0.018)**	0.008 (0.048)	0.049 (0.083)	0.061 (0.044)	-0.031 (0.095)	0.253 (0.171)
Peer Interactions	0.034 (0.018)†	0.046 (0.046)	-0.034 (0.080)	-0.013 (0.049)	0.028 (0.098)	-0.097 (0.067)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: See Notes from Table 6A.

Table 7. Associations Between Family Structure and Children’s Non-Cognitive Development Using Mother and Caregiver Responses. Number of Negative (Negative Significant) Estimates by Family Structure Category. Whites.

	All (17 Questions)		Behavioral / General		Eagerness and Independence		Maturity		Peer Interactions	
	Mother	Caregiver	Mother	Caregiver	Mother	Caregiver	Mother	Caregiver	Mother	Caregiver
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>										
Intact – Separate	-0.018 (0.018)	-0.040 (0.027)	-0.053 (0.022)*	-0.061 (0.034)†	-0.033 (0.039)	-0.038 (0.039)	-0.003 (0.023)	-0.056 (0.036)	0.019 (0.036)	0.021 (0.034)
Intact – Separate – New	0.028	-0.0031	0.053	-0.117	-0.044	0.111	-0.008	-0.065	0.058	0.038
Male Cohabitant	(0.026)	(0.041)	(0.046)	(0.067)†	(0.068)	(0.062)†	(0.030)	(0.057)	(0.054)	(0.048)
Persistent Female Head	-0.007 (0.021)	-0.054 (0.032)†	-0.016 (0.032)	-0.105 (0.046)*	0.013 (0.053)	0.023 (0.053)	-0.041 (0.022)†	-0.092 (0.038)*	0.049 (0.050)	0.055 (0.056)
Female Head – New Male	-0.032	0.080	-0.052	0.025	-0.030	0.113	-0.044	-0.020	0.024	0.175
Cohabitant	(0.028)	(0.061)	(0.041)	(0.073)	(0.061)	(0.082)	(0.026)†	(0.061)	(0.060)	(0.079)*
Female Head – New Male	-0.113	-0.100	-0.095	-0.080	-0.144	-0.082	-0.099	-0.252	-0.070	-0.045
Cohabitant – Separate	(0.082)	(0.064)	(0.129)	(0.114)	(0.143)	(0.093)	(0.024)**	(0.039)**	(0.135)	(0.069)

Notes: Robust standard errors are in parentheses. ECLS-B sample weights are used in all regressions. See Section I of the text for more information about how the family-structure categories were constructed.

Table 8. Differential Response Gaps between Mothers and Caregivers by Parental Two-Bags Score

	Relative to Middle-Scoring Mother:		
	Mother Response Effect	High-Scoring Mother	Low-Scoring Mother
<u>Models</u>			
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>			
All Outcomes	-0.095 (0.007)**	0.000 (0.011)	0.028 (0.011)*
Behavioral / General	-0.213 (0.010)**	-0.010 (0.015)	0.045 (0.016)**
Eagerness and Independence	0.067 (0.013)**	0.010 (0.019)	-0.018 (0.020)
Maturity	-0.138 (0.009)**	-0.007 (0.014)	0.044 (0.015)**
Peer Interactions	0.024 (0.010)*	0.007 (0.014)	0.000 (0.015)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: Robust standard errors are in parentheses. ECLS-B sample weights are used in all regressions.

Table 9. Differential Response Gaps between Mothers and Caregivers by Type of Children (From an index generated by factor analysis two bags scores)

	Relative to Middle-Scoring Child:		
	Mother Response Effect	High-Scoring Child	Low-Scoring Child
<u>Models</u>			
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>			
All Outcomes	-0.082 (0.008)**	-0.023 (0.011)*	0.017 (0.011)
Behavioral / General	-0.197 (0.011)**	-0.035 (0.015)*	0.027 (0.016)†
Eagerness and Independence	0.074 (0.014)**	-0.009 (0.019)	-0.017 (0.021)
Maturity	-0.135 (0.011)**	-0.016 (0.010)	0.046 (0.015)**
Peer Interactions	0.035 (0.011)**	-0.014 (0.015)	0.010 (0.015)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: Robust standard errors are in parentheses. ECLS-B sample weights are used in all regressions.

Appendix A Supplementary Tables

Appendix Table A.1. Mother and Caregiver Responses by Question, African American children.

	Average Ordinal Response		Share of Responses that are Most Positive	
	Mothers	Caregivers	Mothers	Caregivers
Question 1	4.35	4.22	0.49	0.48
Question 2	4.09	4.18	0.43	0.46
Question 3	3.61	4.04	0.21	0.45
Question 4	3.33	4.06	0.16	0.46
Question 5	3.71	3.98	0.23	0.39
Question 6	4.25	3.93	0.49	0.34
Question 7	4.00	3.85	0.34	0.28
Question 8	3.52	3.58	0.16	0.21
Question 9	2.86	3.55	0.10	0.28
Question 10	3.39	3.41	0.13	0.21
Question 11	4.27	4.16	0.47	0.41
Question 12	3.65	3.37	0.23	0.19
Question 13	3.60	3.31	0.21	0.17
Question 14	3.67	3.26	0.20	0.15
Simple Average	3.74	3.78	0.28	0.32

Note: Maternal and caregiver responses are re-ordered where necessary so that higher values correspond to more-positive outcomes throughout, as described in the text.

Appendix Table A.2. Mother and Caregiver Responses by Question, Hispanic children.

	Average Ordinal Response		Share of Responses that are Most Positive	
	Mothers	Caregivers	Mothers	Caregivers
Question 1	4.26	4.26	0.45	0.47
Question 2	3.55	3.79	0.23	0.29
Question 3	3.80	4.21	0.27	0.51
Question 4	3.28	4.01	0.12	0.44
Question 5	3.89	4.05	0.28	0.40
Question 6	4.31	4.08	0.46	0.39
Question 7	3.99	3.96	0.34	0.29
Question 8	3.66	3.76	0.18	0.23
Question 9	2.96	3.51	0.11	0.24
Question 10	3.57	3.49	0.19	0.22
Question 11	4.24	4.16	0.45	0.41
Question 12	3.69	3.40	0.21	0.17
Question 13	3.55	3.36	0.17	0.16
Question 14	3.60	3.30	0.18	0.14
Simple Average	3.74	3.81	0.26	0.31

Note: Maternal and caregiver responses are re-ordered where necessary so that higher values correspond to more-positive outcomes throughout, as described in the text.

Appendix B Robustness of Findings by Race

Robustness Check 1: Restricting the sample to children who were in care at the same facility for at least 6 months prior to the Age-4 ECLS-B survey

Table B.1 shows results from model (1) where we restrict our sample to children who were reported to have been in care in the same facility for at least 6 months at the time of the age-4 survey. The differences in the maternal-caregiver response gaps for white and African American children are nearly identical to those reported in Table 5. Between white and Hispanic children the differences are qualitatively similar to what we report in Table 5, but the coefficients are smaller.

Robustness Check 2: Restricting the sample to low-income children based on household income as reported by parents on the Age-4 ECLS-B survey

Table B.2 shows results from models for children who come from families where household income was less than or equal to \$50,000 annually. This cutoff was selected to preserve a reasonably-sized sample of white children, who make up the comparison group. The differences between white and minority mothers are again very similar to what we report in Table 5 in the main text.

Robustness Check 3: Controlling for racial incongruence

There is considerable general evidence to support the notion that racial incongruence affects social interactions (e.g., Dee 2004, 2005; Grissom and Keiser, 2011). Racial incongruence between children and their caregivers may influence the estimated gaps between maternal and caregiver assessments by race because it is more common for minority children (see footnote 11 in text).

We investigate the extent to which our findings in Table 5 are driven by differences in racial incongruence by augmenting equation (1) from the main text to include a direct incongruence control:

$$Y_{ij} = \pi_{iq} + D_j \tau_1 + (D_j * NC_i) \tau_2 + (D_j * R_i) \tau_3 + \eta_{ij} \quad (\text{B.1})$$

Equation (B.1) is analogous to equation (1) with the addition of the control $(D_j * NC_i)$, where NC_i is an indicator equal to one if the child and caregiver are not of the same race. This control captures the effect of racial incongruence on the maternal-caregiver assessment gaps. Our findings in Table B.3 again mirror those in Table 5.

Appendix Table B.1. Differential Response Gaps Between Mothers and Caregivers by Race. Sample in Care for 6+ Months Only.

	Baseline Mother Response Effect	Relative to White Mothers (baseline):	
		African American Mother Differential Response	Hispanic Mother Differential Response
<u>Models</u>			
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>			
All Outcomes	-0.096 (0.007)**	0.059 (0.014)**	0.030 (0.013)*
Behavioral / General	-0.196 (0.010)**	0.099 (0.021)**	0.050 (0.018)**
Eagerness and Independence	0.039 (0.013)**	0.045 (0.026)†	0.007 (0.025)
Maturity	-0.130 (0.009)**	0.050 (0.019)**	0.043 (0.018)*
Peer Interactions	-0.001 (0.010)	0.017 (0.021)	0.013 (0.018)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: Robust standard errors are in parentheses. ECLS-B sample weights are used in all regressions.

Appendix Table B.2. Differential Response Gaps Between Mothers and Caregivers by Race. Sample with Household Income \leq \$50,000.

	Baseline Mother Response Effect	Relative to White Mothers (baseline):	
		African American Mother Differential Response	Hispanic Mother Differential Response
<u>Models</u>			
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>			
All Outcomes	-0.078 (0.010)**	0.049 (0.016)**	0.033 (0.016)*
Behavioral / General	-0.198 (0.014)	0.086 (0.024)**	0.064 (0.022)**
Eagerness and Independence	0.052 (0.019)**	0.035 (0.029)	0.005 (0.029)
Maturity	-0.116 (0.013)**	0.043 (0.021)*	0.062 (0.021)**
Peer Interactions	-0.038 (0.014)**	0.009 (0.023)	-0.002 (0.022)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: See notes from Appendix Table B.1.

Appendix Table B.3. Differential Response Gaps Between Mothers and Caregivers by Race after Controlling for Racial Incongruence Between the Child and Caregiver.

	Baseline Mother Response Effect	Effect of Racial Incongruence Between Child and Caregiver	Relative to White Mothers (baseline):	
			African American Mother Differential Response	Hispanic Mother Differential Response
<u>Models</u>				
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>				
All Outcomes	-0.109 (0.005)**	0.016 (0.007)*	0.058 (0.011)**	0.045 (0.010)**
Behavioral / General	-0.233 (0.014)**	0.016 (0.011)	0.082 (0.016)**	0.073 (0.014)**
Eagerness and Independence	0.034 (0.010)**	0.030 (0.014)*	0.064 (0.019)**	0.014 (0.019)
Maturity	-0.143 (0.007)**	-0.010 (0.011)	0.051 (0.015)**	0.078 (0.014)**
Peer Interactions	0.009 (0.007)	0.035 (0.010)**	0.024 (0.015)	0.004 (0.014)

** Indicates statistical significance at the 1 percent level.

* Indicates statistical significance at the 5 percent level.

† Indicates statistical significance at the 10 percent level.

Notes: See notes from Appendix Table B.1.

Appendix C

Are Violations to the Fair Benchmark Assumption Driving our Findings by Family-Structure Status?

Robustness Check 1: Restricting the sample to children who were in care at the same facility for at least 6 months prior to the Age-4 ECLS-B survey (white children only)

Appendix Table C.1 shows results from model (2) where we restrict our sample to children who were reported to have been in care for at least 6 months at the time of the age-4 survey. We report estimates for white children only because our minority samples are too small support further divisions of the data. The estimates in Table C.1 are entirely consistent with what we report in Table 6A in the main text.

Robustness Check 2: Restricting the sample to low-income children based on household income as reported by parents on the Age-4 ECLS-B survey (white children only)

Table C.2 shows results from models for children who come from families where household income was less than or equal to \$50,000 annually. Our sample of white children declines significantly when we focus only on low-income children, which increases our standard errors, but our estimates in Table C.2 are qualitatively consistent with what we report in Table 6A in most instances.

Appendix Table C.1. Differential Response Gaps Between Mothers and Caregivers by Family Structure Category. Sample in Care for 6+ Months Only. White Children.

	Baseline Mother Response Effect	Initially Intact with Separation	Initially Intact with Separation, Followed by New Male Cohabitant	Persistent Female- Headed Household	Initially Female- Headed Household with Male Cohabitant Entrant	Initially Female-Headed Household with Male Cohabitant Entrant, followed by Separation
<i>Models</i>						
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>						
All Outcomes	-0.110 (0.008)**	0.055 (0.024)*	0.158 (0.047)**	0.069 (0.028)*	-0.017 (0.063)	0.130 (0.059)*
Behavioral / General	-0.218 (0.011)**	0.079 (0.033)*	0.230 (0.077)**	0.125 (0.042)**	0.042 (0.085)	0.084 (0.056)
Eagerness and Independence	0.045 (0.016)**	-0.019 (0.045)	-0.001 (0.075)	-0.013 (0.067)	-0.084 (0.087)	0.122 (0.361)**
Maturity	-0.152 (0.011)**	0.104 (0.033)**	0.193 (0.054)**	0.082 (0.036)*	0.037 (0.073)	0.152 (0.011)**
Peer Interactions	-0.006 (0.011)	0.019 (0.037)	0.112 (0.069)	0.056 (0.052)	-0.074 (0.080)	0.172 (0.069)*

** Indicates statistical significance at the 1 percent level. * Indicates statistical significance at the 5 percent level. † Indicates statistical significance at the 10 percent level.

Notes Robust standard errors are in parentheses. ECLS-B sample weights are used in all regressions. See Section I of the text for more information about how the family-structure categories were constructed.

Appendix Table C.2. Differential Response Gaps Between Mothers and Caregivers by Family Structure Category. Sample with Income ≤ \$50,000. White Children.

	Baseline Mother Response Effect	Initially Intact with Separation	Initially Intact with Separation, Followed by New Male Cohabitant	Relative to Mothers from Persistent Intact Families (baseline): Persistent Female- Headed Household	Initially Female- Headed Household with Male Cohabitant Entrant	Initially Female-Headed Household with Male Cohabitant Entrant, followed by Separation
<u>Models</u>						
<i>Outcomes: Average of Binary Indicators for Most-Positive Response</i>						
All Outcomes	-0.091 (0.014)**	0.021 (0.028)	0.150 (0.055)**	0.049 (0.033)	-0.019 (0.048)	0.083 (0.077)
Behavioral / General	-0.228 (0.019)**	0.045 (0.040)	0.310 (0.087)**	0.131 (0.045)**	-0.001 (0.082)	0.054 (0.113)
Eagerness and Independence	0.054 (0.025)*	-0.008 (0.053)	-0.013 (0.093)	-0.003 (0.069)	-0.040 (0.092)	0.196 (0.205)
Maturity	-0.142 (0.018)**	0.044 (0.038)	0.151 (0.060)*	0.097 (0.040)*	0.049 (0.074)	0.184 (0.107)†
Peer Interactions	0.052 (0.019)**	-0.019 (0.038)	0.036 (0.085)	-0.031 (0.055)	-0.124 (0.077)	0.073 (0.117)

** Indicates statistical significance at the 1 percent level. * Indicates statistical significance at the 5 percent level. † Indicates statistical significance at the 10 percent level.

Notes: See notes from Appendix Table C.1.