Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development

NICHD Early Child Care Research Network*

and Greg J. Duncan

Running Head: Modeling the Impacts of Child Care Quality

*Authors' footnote: This work is a collaborative work between the National Institute of Child Health and Human Development (NICHD) Study of Early Child care and Greg J. Duncan, Northwestern University. This study is directed by a Steering Committee and supported by NICHD through a cooperative agreement (U10), which calls for scientific collaboration between the grantees and the NICHD staff. Participating investigators, listed in alphabetical order, are: Virginia Allhusen, University of California, Irvine; Jay Belsky, Birkbeck University of London; Cathryn L. Booth, University of Washington; Robert Bradley, University of Arkansas, Little Rock; Celia A. Brownell, University of Pittsburgh; Margaret Burchinal, University of North Carolina, Chapel Hill; Bettye Caldwell, University of Arkansas; Susan B. Campbell, University of Pittsburgh; K. Alison Clarke-Stewart, University of California, Irvine; Martha Cox, University of North Carolina, Chapel Hill; Sarah L. Friedman, NICHD, Bethesda, Maryland; Kathyrn Hirsh-Pasek, Temple University; Aletha Huston, University of Texas, Austin; Deborah J. Johnson, Michigan State University; Jean F. Kelly, University of Washington; Kathleen McCartney, Harvard University; Nancy Marshall, Wellesley College Marion O'Brien, University of Kansas; Margaret Tresch Owen, University of Texas, Dallas; Chris Payne, University of North Carolina, Greensboro; Deborah Phillips, Georgetown University; Robert Pianta, University of Virginia; Suzanne M. Randolph, University of Maryland, College Park; Wendy Robeson, Wellesley College; Susan Spieker, University of Washington; Deborah Lowe Vandell, University of Wisconsin, Madison; Marsha Weinraub, Temple University. We express our appreciation to the study coordinators at each site who supervised the data collection, the research assistants who collected the data, and especially the families and child-care providers who welcomed us into their homes and workplaces and cooperated willingly with our repeated requests for information. Correspondence concerning this article should be addressed to NICHD Early Child Care Research Network, Office of Extramural Policy, Office of the Director, NICHD, 6100 Executive Blvd., 4A01, Rockville, Maryland 20852.

An earlier version of this paper was presented at the biennial meetings of the Society for Research on Child Development, Minneapolis, April 2001. We would like to thank Christina

Gibson, Mark Appelbaum and members of the Brookings Roundtable on Children for comments on a prior draft, but in no way hold them responsible for remaining errors.

Abstract

Using observational measures of child care quality from the NICHD Study of Early Child Care, we compared several statistical methods for controlling for selection bias in relating children's cognitive and achievement outcomes to the quality and type of their child care arrangements prior to school entry. The methods included multiple regression models focused on the level of cognitive development and achievement at 54 months and both simple change and residualized change models focused on change in skills from 24 and 54 months. The impact of possible selection bias was examined by comparing coefficients across models and within models as additional selection factors were added. We also examined possible bias in estimation by testing model assumptions. Since definitive conclusions regarding bias rest on untestable assumptions in our models, we were unable to establish empirically which model best adjusted for omitted-variable bias. However, we were able to establish a fairly small range of estimates of the impacts of observational measures of child care quality gathered when children were 6-24 and 36-54 months of age. Results suggested that a one-standard-deviation increase in child care quality between 36 and 54 months was associated with increases in standardized cognitive test scores of between one-half and one-and-one-half points. Quality between 6 and 24 months had an independent and additive impact on age 54-month cognitive development that was of similar magnitude. These impacts were estimated over the entire range of child care quality and socioeconomic circumstances observed in the sample.

Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development

Knowing whether and how much the quality of child care experienced by infants, toddlers and preschool children affects their cognitive development is crucial for formulating policies regarding regulation of child-care settings, family leave, public funding of child care through tax credits or subsidies for low-income families, and infant-care exemptions from welfare-to-work requirements. Research on this issue by developmental psychologists and economists has led to varying conclusions about the magnitude of quality effects, in part because investigators use different conceptual and analytic models. Our purpose in this paper is to compare various analytic approaches to the quality-outcome connections using data from the NICHD Study of Early Child Care (NICHD SECC) and an interdisciplinary team consisting of an economist, Greg Duncan, and the developmental psychologists of the NICHD SECC.

The study has followed a large and diverse sample of children born in ten sites around the United States. Longitudinal data gathered in the study are used to link the type and quality of child care settings experienced by the children between age 6 and 54 months to their cognitive development at 54 months. We exploit three special features of these data. First, the study assesses quality using an intensive observational measure of caregiver-child interactions that is specific to each study child (rather than assessed at the arrangement or classroom level). Second, data are gathered on an unusually rich set of maternal and child characteristics that are likely to capture many plausible sources of selection bias. And third, the availability of measures of children's cognitive ability, family environment, and child care quality between ages 6 and 54 months enables us to estimate change models of child care quality. We begin with a brief review of the literature and then outline our analytic approach for estimating child care impacts.

Background

Poor quality child care is of concern because infants and preschoolers need responsive and stimulating interactions with adults – parents and other caregivers - to enhance social, cognitive, and language development in early childhood (Bronfrenbrenner & Morris, 1998; Sameroff, 1983). In particular, infants and young children best learn language and cognitive skills during interactions with adults that involve ample talking, turn-taking,and focused attention on the child (Kantz & Snow, 2000; Tomasello & Farrar, 1986). Ideally, such interactions also provide young children with the secure attachments and positive role models needed to develop social skills (Bradley et al., 1989; Howes, 2000). The extent to which children's development relates to their child care experiences has been widely studied, but conclusions drawn vary across disciplines.

Many psychologists and educators conclude that the quality of children's nonparental care affects cognitive and language skills based on evidence from experimental and observational studies (c.f., Clarke-Stewart, 1989; Lamb, 1998). The random-assignment experimental designs used in evaluations of center-based care (e.g., Barnett, 1995; Campbell & Ramey, 1994; IHDP, 1990; Schweinhart, Weikart, & Larner, 1986) provide definitive evidence that expensive, high-quality child care can enhance the cognitive and language development of children from impoverished families. The cognitive advantage linked to this level of quality child care has been tracked to early adulthood for the most intensive program activity (Campbell et al., 2001), and was linked to increased levels of academic achievement and rates of employment for the two most intensive programs (Campbell et al., 2002; Schweinhart et al., 1986; Yoshikawa, 1995). Children from less intensive programs did not maintain cognitive advantages a few years after leaving child care for public school, but were more likely than control children to complete high school and less likely to be placed in special education or retained in grade (Lazar & Darlington, 1982).

Psychologists have also found that child care quality is associated with cognitive and language skills in non-experimental studies (Vandell & Wolfe, 2000). Using standard measures of child care quality, researchers found that child care quality was related to language and cognitive development, even after they controlled for family selection factors such as socioeconomic status, maternal education, parenting, or family structure in other large multi-site studies: the Chicago Study (Clarke-Stewart, Gruber, & Fitzgerald, 1994); the Child Care and Family Study (Kontos, Howes, Shim, & Galinsky, 1995); the Cost, Quality, and Outcomes Study (Peisner-Feinberg, & Burchinal, 1997); and in smaller single-site studies (Burchinal et al., 1996; 2000; Dunn, 1993; McCartney, 1984; Phillips, McCartney, & Scarr, 1987; Schliecker, White & Jacobs, 1991). Published analyses of the current study -- the NICHD Study of Early Child Care -- demonstrate that observed child care quality measured from age 6 months onward was positively related to cognitive and language development at ages 2, 3, and 4½ years, even in the presence of substantial controls for selection factors (NICHD Early Child Care Research Network [ECCRN], 2000; 2002; in press).

In most of these observational studies, quality of child care was modestly related to children's development. Quality of care often accounted for less than 5% of the variance in children’s developmental outcomes in analyses that adjusted for family selection factors. Child care quality was more strongly related to outcomes for poor children than for middle class children in many studies (Baydar & Brooks-Gunn, 1991; Bryant, Burchinal, Lau, & Sparling, 1994; Burchinal, Peisner-Feinberg, Bryant, & Clifford, 2000; Caughy et al., 1994; Peisner-Feinberg & Burchinal, 1997; Vandell & Corasaniti, 1990), but not all (NICHD ECCRN, 2000). However, none of these studies implemented the most rigorous analytic methods designed to address biased estimation of quality effects resulting from parental selection of child care.

Finally, regulatable aspects of child care such as caregiver education and training, adult-child ratios, and group size have also been related to cognitive development in non-experimental studies. Findings from the National Child care Staffing Study (Howes, Phillips, & Whitebook, 1992), the Cost, Quality, and Outcomes Study (Howes, 1997), the Carolina Otitis Media Project (Burchinal et al, 2000) as well as the NICHD Study of Early Child care (NICHD ECCRN, 1999a; in press), show that children attending programs in which caregivers had more education and training, and in which child-staff ratios were smaller, performed better across a range of measures.

Not all psychologists and economists are convinced that child care quality is related to child outcomes because of potential selection effects in observational studies (Blau, 1999; Scarr, 1998). Both developmentalists and economists worry that the quality impacts estimated in the nonexperimental studies can misstate true causal impacts (Blau, 1999; McCartney et al., 1984; Burchinal et al., 1995), but have used different statistical methods to address the bias problem. These concerns, spelled out in greater detail below, focus on bias from failure to adjust for characteristics of families or children that are associated with both quality of care and children’s cognitive development. Family selection is an important issue to consider in child care research, because the type and quality of child care are related to demographic and family characteristics that predict child outcomes (c.f., Lamb, 1998). Children from more advantaged families, in which parents have more education and income, are more likely to experience center-based child care as well as higher-quality care (Phillips et al., 1987). In addition, parents are more likely to select high-quality care if they provide more responsive parenting and have less authoritarian child-rearing beliefs. Parents also appear to make decisions about child care based on child characteristics. Parents are more likely to select center care for preschoolers than infants (Lamb, 1998; Singer, Fuller, Keiley, & Wolf, 1998).

The associations between family and child care characteristics, especially child care quality, often are quite modest, and some studies question the degree of parental choice in the type and quality of care. Further, some have questioned parents' ability to judge child care quality (Helburn, 1994). Parents almost uniformly report their child is receiving high quality care. However, in one of the few studies directly comparing parental and observational ratings, parental ratings were only modestly related to observed quality, despite very strong endorsement of the dimensions of care measured by the observational scale and close agreement between the parent and observational rating forms (Cryer & Burchinal, 1996). Two large child care studies reported only modest (i.e., in the .10 to .25 range) correlations between family characteristics and child care quality measures (Peisner-Feinberg & Burchinal, 1997; NICHD ECCRN, 1996). Our empirical efforts are focused on contrasting methods for assessing bias from both measured and unmeasured characteristics of families and children.

Analytic Approach

Our goal is to estimate the causal impact of child care quality on children's cognitive development. By this we have in mind what Raudenbush (2001) and others have termed the "Rubin-Rosenbaum-Holland" theory of causation. When applied to our situation, the causal impact of an increment in child care quality for a given child is the difference between two potential outcomes: a child's cognitive development in the presence of an increment to existing child care quality and that same child's cognitive development in the absence of such an increment. The policy "experiment" we seek to address is the extent to which children's cognitive development would improve with increments to child care quality as it currently exists, but no other concurrent changes in the circumstances of the child or family.

Note that this definition of causation is child-specific, can vary from child to child, is never observed directly, and amounts to a missing-data problem. An experiment in which children are allocated randomly to either quality increments or no such increments ensures that the unobserved counterfactual is missing at random. Such random-assignment experiments are expensive and ethically difficult to implement, so statistical methods have been developed to address biases resulting from the fact that family selection of child care quality is not a random process.

In the context of nonexperimental data such as ours, the task of causal inference requires a carefully specified model of the developmental process under investigation that includes all relevant predictors and correctly specifies all main effects and interactions among predictors. We consider three analytic approaches to estimating the relations between child care quality and a child's cognitive development. One is a "level" model relating age 54-month cognitive development to a child's past history of child care quality and characteristics of the family, child, and child care experiences. "Change" and "residualized change" analyses relate the change in a child's development at two distinct points in early childhood to the family and child care experiences between the 24- and 54–month assessments. These models provide varying levels of adjustment for observed and unobserved selection factors, and, correspondingly, make differing assumptions about the associations among the predictors and outcome variables.

The general cognitive model assumed in all analyses in this paper represents cognitive development at the end of the preschool period as the product of child, parent, family, and child care characteristics. We assume that these characteristics can vary over time, and divide the early childhood period into the infant and preschool periods. A simple version of this model, proposed by Blau (1999), assumes that child i's cognitive development (Y_{it}) at the point of school entry (t = age 54 months in our data) is an additive function of the child's history of the blocks of variables representing the quality and quantity of home (HOME) and child care (CARE), plus time-invariant child (CHILD) and maternal/family (FAM) characteristics. In most of our empirical work we are able to distinguish between inputs before and after the child's second birthday, and do so in our model with early (E) and late (L) notation:[1]

(1) $Y_{it} = a_1 + ß_{1E}\ CARE_{iE} + ß_{2E}\ HOME_{iE} + ß_{1L}\ CARE_{iL} + ß_{2L}\ HOME_{iL} + ß_3\ CHILD_i + ß_4\ FAM_i + e_{it}$

Our particular interest is in estimating the impact of the quality of both early ($ß_{1E}$) and later ($ß_{1L}$) child care on cognitive development. Threats to unbiased estimation of these impacts include unmeasured characteristics of the child, mother or family environment – elements of $CHILD_i$ and FAM_i in equation (1). Such bias will arise to the extent that unmeasured variables are correlated with both choice of child care quality and children's cognitive development. The following bias stories provide an indication of the diverse nature of sources and directions of bias, and of the kinds of factors that need to be measured if the empirical estimation of (1) is to provide an unbiased estimate of the causal impact of CARE on cognitive development (Duncan & Gibson, 2000).

Suppose that parents who make sacrifices to obtain quality child care for their children also promote their children's development in other ways. If this high level of concern for their children's development is excluded from the estimated model, then its omission will impart an upward bias to the estimated impact of child care quality on children's cognitive development. A similar upward bias occurs if parents who, perhaps owing to unmeasured family circumstances, lack of information, lack of access to affordable care and other important resources, or personal characteristics, are unable to arrange for good quality care for children and less able to promote their children's healthy development in other ways.

Upward bias in the estimated impact of child care quality may also arise from the child's own characteristics if, for example, a child with a difficult temperament bites or fights in a child

[1] An alternative to this kind of longer-run "accumulation of inputs" model is an incremental one in which current-period development is a function of last period's development, inputs to development between the last period and the current one, and other family and child-specific characteristics:

(1a) $Y_{it} = a_1 + \eta_1\ Y_{it-1} + \eta_2\ CARE_{it} + \eta_3\ HOME_{it} + \eta_4\ CHILD_i + \eta_5\ FAM_i + \varepsilon_{it}$

with $CARE_{it}$ and $HOME_{it}$ representing child care and home inputs between time t-1 and t. Such an incremental model better captures the dynamic nature of developmental processes but, as shown by Blau, is completely consistent with (1).

care situation and is expelled from higher-quality child care centers. If a problematic temperament also adversely affects the child's cognitive development, then failure to control for child temperament will impart an upward bias in the quality estimates as well.

Downward bias may result if parents take actions on behalf of their children. For example, a difficult-to-measure developmental delay or behavioral problem in early childhood might motivate a parent to seek out unusually high-quality care to address the problem. Failure to adjust for child characteristics prior to entry into child care in this case will impart a downward bias to the estimated impact of quality on child outcomes.

One approach to the bias problem, used in earlier analyses of these data (e.g., NICHD ECCRN, in press), is to estimate multiple regression models relating final level of cognitive ability to the child's history of child care and other inputs, controlling for a number of selection factors and taking care to include only selection factors that are themselves unlikely to have been influenced by child care arrangements. We term these approaches to unbiased estimation of equation (1) as "level" models, since we are relating the final level of pre-school cognitive ability to the history of child care quality and other inputs.

Our data provide us with an unusually extensive set of selection-factor variables, most of which are measured prior to the child's entry into child care settings, including: maternal verbal ability, positive regard for the child, depressive symptoms, agreeableness, extroversion, neuroticism, child-rearing attitudes, separation anxiety; parenting sensitivity, non-intrusiveness, and involvement; as well as the child's own temperament. We estimate models that include more conventional measures of socioeconomic status (e.g., maternal education) and progressively more of these typically-unmeasured child-care selection factors to gauge the extent to which quality impacts vary with the extent of both conventional and more extensive controls for selection factors (Altonji and Taber, 2000). Altonji and Taber argue that if one assumes that observable controls are a random subset of all possible controls, then the changes in the key quality coefficients across models that add more observed covariates is informative of the possible bias that remains when all observables have been included. Small coefficient changes suggest little additional bias from unobservables; larger coefficient changes suggest that more bias is possible.

We also attempt to control for unobservable sources of bias by estimating change models, where change in cognitive ability spans the period from ages 24 to 54 months. These models are based on the assumption that unobserved variables have similar impacts on both early and later outcomes, and error in assessment at both ages is random. To motivate our change model, suppose that an analogous relationship to (1) describes a child's cognitive development at age 24 months (denoted by s) as an additive function of his or her history (from birth to time s) of the quality and quantity of home (HOME) and child care (CARE) inputs, plus time-invariant child (CHILD) and maternal/family (FAM) influences. In this case there are only early (E) inputs:

(2) $$Y_{is} = a_2 + ß_{1E}\ CARE_{iE} + ß_{2E}\ HOME_{iE} + ß_3\ CHILD_i + ß_4\ FAM_i + e_{is}$$

Note that this equation includes many of the same parameters ($ß_{1E}$, $ß_{2E}$, $ß_3$, and $ß_4$) as the level model, equation (1). This reflects the strong assumption that the impacts of increments to $CARE_E$, $HOME_E$, CHILD and FAM on cognitive development are identical, conditional on covariates, at ages s (24 months in our data) and t (54 months). This assumption is dubious in light of the possible changing nature of the impacts of family and child characteristics and our differential ability to measure cognitive development at those two points. We ignore such problems for the time being, but return to them below.

A simple difference model of equations (1) and (2), using "Δ" to denote the t-s difference, is:

(3) $\Delta Y_i = \Delta a + ß_{1L} CARE_{iL} + ß_{2L} HOME_{iL} + \Delta e_i$

The obvious advantage of (3) over (2) or (1) is that the biases associated with unmeasured child and maternal/family characteristics have been eliminated by the differencing. A disadvantage is that (3) differences out the quantity and quality of early child care and home environments and thus provides no way of estimating the developmental impacts of these factors.

It is crucial to note a common point of confusion over "change" model (3): even though the dependent variable is a change score, the independent variables are the average levels of child care quality and other inputs between the two measurement points. Furthermore, the key $ß_{1L}$ parameter in change equation (3) corresponds precisely to the $ß_{1L}$ "level" parameter in (1). Thus, the differencing in (3) is merely a method to secure less biased estimates of the impacts of $CARE_{iL}$ on the final preschool level of cognitive development.

If the $ß_{1E}$, $ß_{2E}$, $ß_3$ or $ß_4$ parameters do change over time and thus differ between (1) and (2), then, continuing to use "Δ" to denote change and applying some simple algebra, the difference between (1) and (2) becomes:

(4) $\Delta Y_i = \Delta a + ß_{1L} CARE_{iL} + ß_{2L} HOME_{iL} + \Delta ß_{1E} CARE_{iE} + \Delta ß_{2E} HOME_{iE} + \Delta ß_3 CHILD_i + \Delta ß_4 FAM_i + \Delta e_i$

In this case, the role of early inputs and of the time-invariant child and family effects are not differenced out of the change equation since these factors are assumed to have a differential impact on cognitive development at ages two and five. Equation (4) thus provides a rationale for including early inputs and parent and child characteristics in the change equation, but also suggests that omitted elements of $CARE_{iE}$, $HOME_{iE}$, $CHILD_i$ and FAM_i may bias impacts of $CARE_{iL}$ estimated in (4).

Note, however, that the conditions for omitted-variable bias are quite different in change model (4) compared with level model (1). In (1), bias arises when elements of $CHILD_i$ or FAM_i have significant impacts on cognitive development and are correlated with child care quality. In (4), bias arises when unmeasured elements of $CHILD_i$ or FAM_i have significantly *different* impacts on development between the early and late periods and are correlated with child care quality. Note also that the coefficients on $CARE_{iE}$, $HOME_{iE}$, $CHILD_i$ and FAM_i in equation (4) have a very different interpretation from their corresponding coefficients in either (1) or (2): they represent the change in the impact of these measures on development at the two measurement points.

Equations (1)-(4) fail to account for the possibility that the impacts of quality may be nonlinear, and that quality and quantity of child care may interact, or depend on the gender or early cognitive ability of the child, or on the education level of the mother. All of these possibilities are explored in our empirical work.

Psychologists and educators are reluctant to rely on simple change scores over "level" scores because of their greater measurement error (lower reliability). Typically, change scores are substantially less reliable than the original two scores when their "true" scores are moderately to highly correlated (Cronbach & Furby, 1967). One consequence of analyzing a less reliable outcome measure is increased standard errors for parameter estimates in a change equation such as (3) as compared with level equations (1) or (2).[2]

Measurement error in dependent variables may or may not bias regression coefficients. If measurement error is uncorrelated with right-hand side variables and the error term, the change model parameter estimates will be unbiased (Allison, 1990; Gujarati, 1995). However, bias may result if the measurement error in the change model is correlated with the true levels of the dependent and independent variables at time 1 or 2 (Bound et al., 1994). The most likely causes of such bias are failure to include relevant interactions, differences in the impact of the omitted variable at the two time points, or use of assessment tools in which error is related to ability.

Any repeated measures analysis is prone to bias when the ability to measure the outcome or the metric of the measure of that outcome varies over time (Bryk & Weisburg, 1977; Labouvie, 1980). This is of special concern in our case because omitted variables are likely to be more strongly associated with well-measured cognitive development at 54 months than with poorly measured cognitive development at 24 months. We are especially concerned because infant cognitive assessments provide more reliable assessment of children with lower than higher scores (McCall, 1977). Prediction of subsequent individual differences in cognitive skills is much better during the preschool years than during infancy (McCall, 1977). We examine our data for evidence of such bias for the level and change models.

The "residualized change" model (Cronbach & Furby, 1970; Labouvie, 1980; Diggle, Liang & Zeger, 1994) is an alternative to a simple change model. In it, the later outcome assessment is used as dependent variable and the early assessment as a right-hand variable. As such, it is a variant of change model (3) in which the implicit coefficient on the time s (for us, 24 months) outcome is an estimated parameter rather than, as is the case for change equation (3), effectively fixed at 1.0. The residualized change version of equation (3) is:

(5) $Y_{it} = a_3 + \delta_1 Y_{is} + \delta_2 \text{ CARE}_{iL} + \delta_3 \text{ HOME}_{iL} + \varepsilon_{it}.$

The residualized change version of equation (4) adds CARE_{iE}, HOME_{iE}, CHILD_i and FAM_i to this equation.

This approach can provide considerably more power to detect associations when outcomes are highly correlated (Cronbach & Furby, 1970), and can be used when the earlier and later assessment of the outcome are not measured identically over time. On the other hand, including initial level on the right-hand side of equation (5) fails to take into account random variability in initial scores, and therefore almost certainly builds in a biasing correlation between it and the error term of (5). Furthermore, there is no longer a direct relationship between an explicit model of development and the parameters estimated in (5), rendering the interpretation of δ_2 problematic. Finally, this model will result in biased estimation when the unobserved variables have differential impact on the early and late cognitive development.

In summary, there are a variety of analytic methods for addressing selection bias. The level model can adjust for observed but not unobserved selected factors. The simple change model adjusts for observed and unobserved selection factors, but has less power and imposes

[2] Our own empirical work, presented below, shows that increases in these standard errors are relatively modest.

additional strong assumptions about the data than the level model. The residualized change model also adjusts for observed and unobserved selection factors. It should provide more power and bias than the change model, but less power and bias than the level model. Results emerging from the level, change, and residualized change models are contrasted using data from the NICHD Study of Early Child Care.

Methods

Participants

Our data come from the NICHD Study of Early Child Care, which recruited mothers from hospitals near the following locations throughout 1991: Little Rock, Arkansas; Irvine, California; Lawrence, Kansas; Boston, Massachusetts; Philadelphia, Pennsylvania; Pittsburgh, Pennsylvania; Charlottesville, Virginia; Morganton, North Carolina; Seattle, Washington; and Madison, Wisconsin. The sample plan was not intended to provide a representative national sample, but was designed to represent healthy births to non-teen parents at the selected hospitals. Potential participants were selected from among 8,986 mothers giving birth during selected 24-hour sampling periods.

Planned exclusions and unplanned attrition. The sample of 8,986 mothers was reduced to 5,416 mothers eligible for a phone call two weeks after the birth owing to both unplanned attrition (438 cases; mostly refusals) and planned sample exclusions (3,142 cases; mother <18, multiple births, mother not fluent in English, family expects to move, medical complications, baby being put up for adoption, family lives too far away, family in another study, family lives in an unsafe neighborhood in one site [Philadelphia]).

A conditional sub-sampling plan was next imposed to assist in eventually reaching a planned enrollment target of around 1,200 families. The subsampling attempted to ensure that single parent, low-maternal education and minority distributional targets were met, while maintaining random case selection at this stage of the process. The use of targets led to somewhat differential selection probabilities for enrolled cases.[3] All told, the subsampling process reduced the number of families for which two-week phone calls were attempted from 5,416 to 3,015. This reduction should not be considered attrition-related nonresponse.

Unplanned (1,153 cases; refusals and lack of success with contacts at three different times of the day), planned (151 cases; baby in hospital more than 7 days, planning to move within 3 years and 185 cases; not contacted because enrollment quota was achieved before that family's name appeared on the contact sheets) reasons further reduced the sample from 3,015 screened mothers to the 1,364 mothers who provided information at the one-month interview. The 1,153 refusals and no-contact cases constitute the bulk of the study's nonresponse cases. Further attrition between the one-month and six-month interviews reduced the 1,364 enrollments by 37 cases. Modest attrition reduced the sample of families providing information for the 15-, 24-, 36-month and subsequent interviews.

---

[3] We intend to gauge the impact of both attrition and differential sampling when weights have been developed for this purpose.

Duncan and Gibson (2000) construct unplanned-attrition-related response rates for the sample by counting as attrition the "refusals" and "unable to locates" but not the planned exclusions from the sample. They find that the attrition-related cumulative response rate as of the one-month interview to be 52.2%. This response rate, together with the plan to sample at 10 sites chosen due to scientific merit of the investigators, rather than from a national sampling frame, created a sample that cannot be viewed as statistically representative of any a priori-defined population. Nevertheless, the sample is large and economically, geographically, and ethnically diverse, especially for an observational child care study.

Procedures

Child outcomes. Our outcome measure at 15 and 24 months is the Bayley Mental Developmental Index, which is normed to have a mean of 100 and standard deviation of 15. The Bayley test (Bayley, 1969) used at 15 months was based on 1969 norms, whereas the Bayley test used at 24 months was based on a 1993 revision of the test (Bayley, 1993). At 54 months we created cognitive and achievement composite scores. The cognitive score was computed as the mean of four scales scores (alpha = .83.46 < r < .70): the Woodcock-Johnson Picture Vocabulary and Memory for Sentences tests and the Preschool Language Scale Expressive and Receptive tests. The achievement scores was computed as the mean of three scale scores (alpha=.72; .36 < r < .58): Woodcock-Johnson Applied Problems (mathematical skill), Letter-Word Identification (reading skill), and Incomplete Words Scales (phonological knowledge). These scale scores are normed to have a mean of 100 and a standard deviation of 15. The 54-month cognitive composite score technically is not an intelligence score, but both language and memory are specific cognitive processes that are so highly correlated with IQ scores that they can be used to approximate cognitive scores (Neisser et al., 1996).

Child care type and quality. In all cases, our child care measures include: average quality; mean hours of care per week; proportion of time in center care at least 10 hours/week; whether exclusive mother care; and whether quality of care data are missing (yes=1). Child care quality was observed in all but mother-provided settings. Observational assessments of caregiver-child interaction were obtained for children who were in 10 or more hours per week of nonmaternal care. At least one such assessment was obtained for 91% of the sample; and at least two assessments were obtained for 779 children. Observations were conducted during two half-day visits scheduled within two-week intervals at 6, 15, 24, and 36 months and during one half-day visit at 4 ½ years. At each visit, observers completed two 44-minute cycles of the Observational Record of the Caregiving Environment (ORCE), during which they first coded the frequency of specific caregiver behaviors and then rated the quality of the caregiving. Positive caregiving composites were calculated for each age level observed by averaging the ratings (see NICHD ECCRN, 1996, 2000). At the end of each observation counts were made of the numbers of children and of child care providers in the setting from which measures of group size and adult-child ratio were derived. In addition, information on the educational levels of the child care providers were obtained as part of a caregiver interview.

At 6, 15, and 24 months, quality scores were the mean of five 4-point subscales (sensitivity to child's nondistress signals, stimulation of child's development, positive regard toward child, detachment [reflected], and flatness of affect [reflected]). Cronbach alphas for the composite ranged from .87 to .89. At 36 months, these five scales plus two additional subscales, "fosters child's exploration" and "intrusive" [reflected], were included in the composite (alpha = .83). At 4 ½ years, the positive caregiving composite was the mean of 4-point ratings of caregivers' sensitivity/responsivity, stimulation of cognitive development, intrusiveness [reflected], and detachment [reflected] (alpha = .72). To ensure that observers at the ten sites were making comparable ratings, all observers were certified before beginning data collection and tested for observer drift every 3-4 months. Reliability estimates were computed for both the master-coded videotapes and live observations using Pearson correlations and the repeated-measures ANOVA formulation. Reliability exceeded .80 at all ages.

Child care type and quantity were collected during interviews with the mother beginning when the child was one month of age. Every three months when the child was between 3 and 36 months, mothers were asked about the type and hours per week spent in up to three care arrangments. From 42 to 54 months, mothers were interviewed every four months. To be consistent with prior reports (NICHD ECCRN, 2002), center care was indicated if the mother reported the child spent at least 10 hours per week in center care, and total hours of care was tallied across all arrangements.

Our demographic controls included study site, child gender, child ethnic group (Non-Hispanic African American, Non-Hispanic European American, Hispanic American, or other), maternal years of education at birth, average family income-to-poverty threshold ratio from 6-54 months, and the percent of measurement occasions when a partner lived in the household (1-54 months). These were included because each has been related to child care experiences (Pungello & Kurtz-Costes, 1999).

A great advantage of the NICHD Study data is its wealth of measures of maternal, family and child conditions measured prior to the child's first entry into child care. <u>Child difficult temperament</u> was measured by a 55-item Infant Temperament Questionnaire (Medoff-Cooper, Carey & McDevitt, 1993) completed by mother. A composite measure, difficult temperament, was formed from the subscales, approach, activity, intensity, mood, and adaptability. <u>Maternal sensitivity</u> (positive, nonintrusive, responsive and supportive maternal care) was coded from videotaped 15-minute mother-child observations during semi-structured play at 6 months. The maternal sensitivity score was a composite of 4-point ratings of sensitivity to nondistress, intrusiveness (reverse scored), and positive regard. Videotapes from all sites were coded at one location (see NICHD ECCRN, 1999b for details).

Quality of home environment was measured with the Infant/Toddler version of the Home Observation for Measurement of the Environment (HOME, Caldwell & Bradley, 1984), which is an assessment of the overall quality of the physical and social resources available to the child in the family context. Maternal <u>depressive symptoms</u> were measured using the Center for Epidemiological Studies Depression Scale (CES-D, Radloff, 1977) administered at 6 months. <u>Maternal personality</u> was measured with a composite of the neuroticism, extroversion, and agreeableness scales from the NEO Five-Factor Inventory, a short form of the NEO Personality Inventory (Costa and McCrae 1985). Maternal <u>separation anxiety</u> was assessed using Subscale I of the Separation Anxiety Scale (Hock, Gnezda & McBride, 1983). High scores indicate high levels of maternal worry, sadness, and guilt during separation from her child, and adherence to beliefs about the value of exclusive maternal care. Maternal <u>attitudes and beliefs about child rearing</u> were measured with a 30-item questionnaire probing mothers' ideas about raising children (Schaefer & Edgerton, 1985). High scores indicate authoritarian child rearing attitudes and beliefs. Maternal <u>beliefs about costs and benefits of maternal employment</u> for children were measured using the Attitudes Toward Maternal Employment questionnaire (Greenberger, Goldberg, Crawford, & Granger, 1988). High scores indicate positive beliefs. <u>Maternal</u>

vocabulary was assessed by the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981), which was administered to mothers when their children were 36 months old. Each of these family characteristics was included because they had been linked theoretically or empirically to both child outcomes and family selection of child care arrangements (Pungello & Kurtz-Costes, 1999).

Results

The analysis plan involved fitting a series of level, change, and residualized change models to 54 month cognitive and achievement composite scores. For these analyses, we transformed the ORCE quality scores to have a mean of 0 and standard deviation of 1 so coefficients can be interpreted as estimated changes in outcomes associated with a standard deviation change in quality. All models included early and/or later ORCE quality scores, average hours of care per week, proportion of time in center care, and dummy variables that indicated whether child received exclusive maternal care or had missing quality data. The variables representing early quality of care were computed as the mean of observed quality at 6, 15, and 24 months, and early hours of care and amount of center care were averaged from maternal reports collected every 3 months from 3 to 36 months. The later quality variable was computed from observed quality at 36 and 54 months, and later hours and amount of care variables from maternal report from interviews every 4 months from 42 to 54 months. Missing quality data occurred for children whose care setting should have been observed, but was not. Our missing data procedures include the use of a dummy variable indicating who has missing data and assigning the mean value on the quality variable for those same individuals. We also included a dummy variable for exclusive maternal care. The resulting model produces coefficients for quality that describes the linear association between quality and cognitive development based only on data from individuals observed in care. The coefficient for maternal care describes the mean difference in cognitive outcomes for children in care and not in care, and the coefficient for missing care describes the mean difference in cognitive outcomes for children observed and not observed in care. Missing value dummy variables were created for all covariates and individuals with missing values were assigned the mean score on that measure. These missing value variables were added to analyses to ensure that no child was deleted from the analysis due to missing child care variables or covariates.

Within each set of analyses (level, change, residualized change), four models were fit that correspond to models used in previous studies relating child care quality to cognitive development. The first model included only site as a covariate. The second model added ethnicity, maternal education, and gender. The third model added early and later family income/poverty threshold, maternal sensitivity and HOME enrichment scores at 6 months, maternal depression, maternal vocabulary, proportion of time with a partner in the household, and missing value dummy variables for each covariate. The final model added the 1-month childrearing attitudes score, 1-month maternal separation anxiety, 1-month rating of benefits of work, 6-month temperament rating, and 6-month maternal psychological adjustment ratings. Covariates were selected if we viewed them as reflecting potential selection bias and as exogenous influences. Any covariates that could be causally influenced by child care experiences (e.g., subsequent maternal sensitivity) were excluded as potentially endogenous influences.

Table 1 presents descriptive statistics for the longitudinal data used in these analyses. The columns in which variables are listed correspond to the interviewing wave in which they were gathered. With average maternal education between a high school and college degree and income-to-poverty threshold ratios slightly above 3.0, the sample is somewhat more affluent than the nation as a whole. Mean ORCE ratings are close to 3.0 on its 4-point scale. A score of less than 2 on the ORCE indicates low quality care. When averaged between 6 and 24 months of age or 36 and 54 months, very few children were observed in low quality care (n's = 43 and 31, respectively). The child care providers in the sample had, on average, over 13 years of education (*SD*'s ranged from 2.09 to 2.25 years). As expected, group sizes and child-adult ratios increased with age and were characterized by substantial variation given the range of types of care observed in this study. The increasing use of center-based care with age follows a familiar pattern.

<u>Level model analysis of cognitive development</u>. Table 2 presents regression results from "level" models in which the cognitive score at 24 months or 54 months is regressed on a linear measure of ORCE quality ratings and other controls. In Model 1, the child's 24-month cognitive score is related to child-care quality between 6 and 24 months and a full set of birth-to-24-month parental and child controls. It produces an estimated impact of 1.7 for a one-standard deviation increase in ORCE quality. This can viewed as an effect size of 1.66/14.6 = .11, when the sample standard deviation on the MDI (14.6) was used, or as an effect size of .13 when the pooled standard deviation from the model was used.

Models 2 through 5 provide estimates of "level" models that relate cognitive development at 54 months to child care quality averaged over both the 6-24 and 36-54 month intervals, with progressively more comprehensive control variables included in the model. In all cases the child care measures listed in the first column are included in the models. Interestingly, the coefficients on the early ORCE quality measure are generally similar to the coefficients on the later ORCE quality measure.

Looking across the cognitive outcome results for Models 2 through 5, it can be seen that controls for maternal education and the child's sex and gender in Model 3 cut the two ORCE child-care quality coefficients in Model 2 roughly in half. The addition of many more selection factors in Models 4 and 5 produces proportionately smaller reductions in the estimated impacts of ORCE-based quality. This suggests that a readily available maternal measure – completed schooling – accounts for most but not all of the adjustment to the quality impact estimate obtained in a regression with a much more extensive set of observable maternal and home characteristics.

The most inclusive "level" model (Model 5) suggests that a one standard deviation increase in either early or later child care quality raises cognitive scores by 1.4 and 1.2 points, which in turn suggest that higher quality across the entire period is associated with 2.6 point increase. Both of these coefficients are statistically significant at conventional levels. Each coefficient amounts to an effect size of about .08 to .09 (1.3/15.1 = .086), while both combined produce an effect size of about .17 for a one standard deviation change in ORCE-based quality maintained across the entire four-year period of measurement. Using the standard deviation from the analysis model, the effect sizes are only slightly bigger (.09 and .10).

Also noteworthy in the 54-month cognitive outcome regression models in Table 2 is that the proportion of time spent in center care between ages 27 and 54 months (but not earlier) appears to have a positive impact on cognitive development, with the difference between 100% of time vs. no time in such care associated with a 4.1-point increase in cognitive test scores in

Model 5 (effect size = .27). In addition, the cases with missing quality assessments (primarily because the child care provider refused to be observed) in the early but not later period have lower scores. Prior analyses indicated that child care settings with missing data in infancy probably provided lower quality care because caregivers tended to have less education and training (NICHD ECCRN, 1996).

Change model analysis of cognitive development. Even the inclusion of the extensive set of characteristics in Model 5 does not rule out the possibility that yet more controls might reduce the estimated impact further. If one accepts their assumptions, change models remove the biasing impacts of all persistent characteristics, both observed and not.

Table 3 details results from various formulations of simple and residualized change models for both cognitive (top half of the table) and achievement (bottom half of the table) outcomes. The first such model (Model 6) includes only the earlier and later child care variables listed in the first column of Table 3. Recalling that the interpretation of a child-care quality coefficient in our change model is identical to its interpretation in our level model, it produces an estimated impact of child care quality between 36 and 54 months on the cognitive outcome (1.17 points) that is similar to estimated impacts of 36- to 54-month quality in the most comprehensive level models, and is statistically significantly different from zero ($p<.05$).

The addition of site, gender and ethnicity reduces the ORCE's coefficient further, and the additional of extensive family characteristics in Model 8 reduces it to a statistically and substantively insignificant .59, which translates into an effect size of .04. [5] Taken together, alternative formulations of the simple change model in the left half of Table 3 suggest that ORCE-based quality between 36 and 54 months has a null to very modest impact on cognitive development**.**

Estimates from various residualized change models (in which 54-month cognitive development is regressed on 24-month cognitive development plus the quality measures) are presented in the right half of Table 3. The quality impacts retain statistical significance in all of the models and imply that a one standard deviation increase in ORCE-based quality is associated with between .9 and 1.7 points change in cognitive development, with an effect size for the most comprehensive model equal to .06.

Level models of academic achievement. Table 2 also presents regression results from "level" models in which academic achievement at 54 months is regressed on a linear measure of

[5] The .04 effect size estimate is the ratio of .59 to 15.1. We continue to use the standard deviation of 54-month cognitive score to maintain comparability with the effect sizes reported for the level models. The effect size is .05 if the estimated standard deviation from the analysis model is used. Only the effect sizes from the sample standard deviation are presented from this point on. Effect sizes based on the estimated standard deviation under the analysis model was always slightly larger due to the variance accounted for when all the covariates are considered.

ORCE quality ratings and other controls. In the case of site-only-covariate Model 2, the coefficients on quality imply that a one-standard deviation change in ORCE quality in either early or late childhood is associated with a 1.0- to 1.3-point increase in children's achievement scores. Controls for additional family and child characteristics reduced the ORCE child care quality coefficients in Model 2 only slightly. As with the level analyses of cognitive scores, the addition of many more selection factors in Models 4 and 5 produced smaller reductions in the estimated impacts of ORCE-based quality.

The most inclusive "level" model (Model 5) suggests that a one standard deviation increase in either early or later child care quality raises achievement scores by .9 to 1.1 points (suggesting that higher quality across the entire four-year period is associated with a 2.0 point increase). As before, the proportion of time spent in center care between ages 27 and 54 months (but not earlier) is positively related to cognitive development, with the difference between 100% of time vs. no time in such care associated with a 3.1-point increase in cognitive test scores in model 5 (effect size of .27).

Change models of academic achievement. The bottom half of Table 3 details results from analyzing the 54 month achievement outcome with the simple and residualized change models. We included the change model analyses as a point of comparison, although we recognize that the 24-month cognitive assessment and 54-month achievement measures cannot be regarded as repeated measures. The first such model (Model 6) includes only child care variables, and suggests that a one standard deviation change in quality is significantly associated with a .38 point change in achievement scores at 54 months. This estimate is about one-third that from the most restrictive level model (model 5), and is substantively and statistically nonsignificant.

The residualized change models yielded quality coefficients that were somewhat larger than those from the corresponding simple change models, but smaller than those from the level models. Quality impacts retain statistical significance in all of the models. The most comprehensive model suggests that that a standard-deviation increase in ORCE-based quality is associated with a .84-point change in cognitive development, which implies an effect size of .07.

Extensions

Nonlinear relations of quality to outcomes. Our worry that a linear ORCE might miss important thresholds in the impacts of child care quality on cognitive development led us to estimate the models presented in Table 4. The relatively small set of children in the lowest quality category (ORCE scores between 1.0 and 2.0) serves as the reference group, and we estimate impacts of half-unit ORCE intervals above 2.0.[6] We present estimates from both level and change models that include the full set of covariates. The coefficients listed in Table 4 represent the difference in adjusted means for children in the very lowest quality care and the group indicated in the first column of that row. For example, in the first column we see that children in the highest quality care (row labeled 6-24 ORCE quality 3.5-4) scored an average of 5.68 points higher on the 24 months cognitive test than did the children in lowest quality care.

[6] We would normally not want to use such a small group as a reference category, but in this case the advantage for hypothesis testing is compelling: coefficients on the included dummy categories show differences in cognitive scores for higher-quality care relative to the lowest quality care.

In the case of early care, the impact estimates of which come from "level" models, an *F*-test on the entire set of quality dummies suggests that quality does indeed matter. The pattern of coefficients presented in the first two columns suggests a possible threshold such that cognitive, benefits emerge when quality is above 3.0 on the ORCE scale – a level of care enjoyed by slightly more than 50% of children experiencing child care in the sample. Pairwise comparisons of 24-month cognitive scores revealed that children in the highest two quality groups scored significantly higher than children in the lowest two quality groups. The children in the highest quality group also outscored children cared for exclusively by their mother as well as those in unobserved formal care.

In the case of care between ages 36 and 54 months and the cognitive outcomes, there is more of a suggestion that care above the 2.5 threshold (rather than the 3.0 threshold seen above) produces cognitive benefits. Pairwise comparisons of the 54-month cognitive scores in the level and residualized change models indicated that children in the highest two 36-54 month quality groups scored significantly higher than children in the lowest quality groups, and children in highest quality group scored higher than children in the next lowest group (2.0-2.5 ORCE scores). Comparisons in the change model showed significantly higher cognitive composite scores in the highest quality group than in the lowest quality group, and comparisons in the residualized change model showed significantly higher scores for children in the highest quality group than in the lowest two quality groups. In summary, all three analyses indicated that children in highest quality care scored higher than children in the lowest quality care, and two of the three analyses also indicated that children in high quality care scores higher than children in moderately low quality care.

The pattern of results was less clear from these analyses of the achievement outcome at 54 months. Surprisingly, the few children who experienced the lowest quality care (n=26) showed nonsignificantly higher levels of academic skills than children in moderate quality care. Pairwise comparisons of achievement means in the level model analyses and residualized change models yielded significant differences between children in the two highest quality groups and children and children in the middle and second-lowest quality groups. Comparisons involving the lowest quality groups did not yield significant pairwise differences but significant overall differences for all but the simple change models.

The safest conclusion from this analysis is that there are consistently significant differences between child care quality and cognitive development and that the effect size when comparing extreme groups (a three standard deviation quality difference) tends to be moderate (level model: 6.55/15.5=.42; change model 5.47/15.5= .35; residualized change model: 5.67/15.5=.37). The precise nature of the child care quality/academic achievement association is less clear. All in all, it appears that our data are not up to the task of revealing thresholds in the impacts of child care quality due to the very small number of children in our sample who experience very low quality care.

<u>Moderators of relations of quality to outcomes.</u> We next tested for various interactions between: early and later child care quality; child care quality and child gender; child care quality and mother's schooling level; and child care quality and type of care (center vs. not). Again, we estimated level, change, and residualized change models. Only one statistically significant interaction emerged - a gender x quality in the change model analysis of achievement.

Child's initial cognitive level as moderator. We also tested whether the impacts of child care quality depended on the initial cognitive ability of the child. There are two reasons to believe this might be the case. On a substantive level, it could well be that children with early cognitive deficits might profit the most from sensitive and stimulating interactions with caregivers. On a methodological level there was evidence that the Bayley scores used in our change model are likely more reliable and thus better predictors of later performance for infants who score in the lower range than for those who score in the normal to high range.

To investigate these possibilities, we created a dummy variable indicating whether the child had scored in a problematic cognitive range at 15 months, defined as scores less than or equal to half a standard deviation below the mean.[7] This dummy variable and its interactions with all child care and family variables were added to the final level (Model 5 in Table 2), change (Model 10 in Table 3), and residualized change (second Model 10 in Table 3) models.

The resulting coefficients are shown in Table 5. The first two rows in Table 5 repeat coefficients shown in Tables 2 and 3. The next rows show the corresponding coefficients for children who did and did not score in the low range on the Bayley MDI at 15 months. The coefficients for later child care quality were significantly different for these two groups of children in the change model analyses of cognitive ($F(1,940)=4.16$, $p=.04$) and achievement outcomes ($F(1,924)=6.09$, $p=.01$), but not in the level model analyses of cognitive ($F(1,972)=0.00$, $p=.99$), achievement outcomes ($F(1,955)=1.15$, $p=.28$), the residualized change model analyses of cognitive ($F(1,939)=.99$, $p=.32$) or achievement outcomes ($F(1,925)=2.01$, $p=.16$). Later quality is related to cognitive development in a very similar manner for both groups of children in the level model, but is positively and significantly related only for children who initially scored low in the change model analysis. Similar results were observed when achievement scores were analyzed.

The change model both make the assumption that error in assessments of outcomes at both time points is random, whereas the level and residualized change model makes that assumption only about the later assessment. Accordingly, these analyses suggest that the lower reliability of measuring cognitive ability for children scoring higher on the test at the earlier assessment and not the later assessment may be biasing estimates from our change models, especially since this interaction was not observed in the level analysis that do not include the 24-month cognitive score in the model. Alternatively, these results may reflect the fact that children with early cognitive deficits profit the most from high-quality care. Further work on these issues is an important research priority.

Other quality measures. Finally, we estimated quality impacts using alternative measures of quality: staff-to-child ratio and group size both as reported by the mother and as observed in the child care setting; and child care-setting reported teacher education and training (Table 6). We employed the range of "level" and "change" models listed at the bottom of the table. The first model included type of observed care (center yes or no) and hours of child care as covariates. The second model added maternal education, gender, and ethnicity, and the final model included all of the family and child covariates included in the models described above.

[7] We chose the 15-month measure since it preceded the 24-month measure used in our level and change models.

Teacher education showed relatively consistent associations with children's 54-month achievement outcomes (i.e., math and reading skill, and phonological knowledge) in the change but not level models. Group size also showed consistent, although weaker, negative associations with 54-month cognitive outcomes in all change model analyses, but not in any level model analyses. Ratio showed no consistent pattern of association.

We find these results puzzling. In the ORCE-based results, simple change models generally produced the smallest and least consistently significant quality coefficients, which led us to speculate that faulty model assumptions may be leading to a systematic underestimation of quality effects. But Table 6 shows that the simple change models generally produce the largest and most consistently significant coefficients for these alternative measures of quality. Further research on the range of quality measures available in the NICHD SECC is another priority.

Discussion

This paper employed a variety of approaches to address the problem of unobservable sources of bias in estimates of the impacts of child care quality on children's cognitive development. We found considerable evidence of bias from models that do little to control for child care selection factors. Not surprisingly, the various approaches to adjusting for selection bias yielded somewhat different conclusions, and we were not able to definitively determine which model provided the least biased estimates of the impact of child care quality on child outcomes. We were, however, able to establish a fairly narrow range of effects sizes (.04 to .08) on cognitive ability and achievement for a one standard-deviation increment in child care quality sustained between 24 and 54 months of age. Child care quality increments sustained during the first two years of life may have similar positive impacts, although we were unable to establish the robustness of these estimates using our full range of methods.

Until recently, most of the existing literature relies on far fewer covariates than some psychologists and economists feel are necessary to provide the needed bias adjustments (Duncan & Gibson, 2000; Scarr, 1998). Compared with "level" models that controlled only for site, our level models that also controlled for a few demographic and child characteristics (notably, mother's schooling and the child's gender and ethnicity) reduced our estimates of the impact of 36-54-month quality on age 54-month cognitive development by close to half. Adding covariates included in various prior NICHD ECCRN analyses (e.g., family income/poverty, maternal sensitivity, HOME scores) reduced the coefficients by an additional 14%. Adding all baseline maternal and child characteristics available in the NICHD study data had virtually no additional effect on the quality impact estimates. Coefficient reductions were smaller in the case of the achievement than cognitive outcomes, and smaller in the residualized change than in the level model analyses.

Our change models further reduced the coefficient on child care quality, which is consistent with the possibility that unobserved selection factors that are correlated with both choice of child care quality and children's cognitive development may still be biasing coefficients even in the most complete level models. For cognitive skills, our least comprehensive change model yielded similar coefficients to those reported for the "level" models; the most comprehensive change model reduced the quality coefficients to statistical nonsignificance.

It is important to appreciate that the three types of models we employ vary in the number of assumptions on which the model depends and in the degree to which the model adjusts for selection bias due to omitted variables. All three models assume that included covariates are exogenous (i.e. are not changed by quality of care in a way that affects how they relate to the outcome measure). Efforts were made to include only family and child measures that would not be influenced by child care experiences, although it is possible that some covariates might be at least partially endogenous. If so, estimates of the quality impact from the models with extensive covariates may be biased.

Most of our models assumed that the relation between child care quality and child outcomes is linear. This assumption was supported in follow-up analyses for the cognitive outcomes, but the absence of large numbers of children in very low quality care leaves open the possibility that the impacts of quality increments are different for children at the low end of the child care quality distribution. Follow-up analyses also suggested no important interactions between child care quality and family characteristics.

The change and residualized change models, as distinct from the level models, made additional assumptions as they attempted to adjust for omitted variable bias. For example, estimates of the quality impact from the residualized change model will be biased if important selection variables are not included and they have a similar impact across time. The simple change model provides the strongest adjustment for omitted variable bias by subtracting the earlier assessment from the outcome of interest. However, the assumptions of this model will be violated if omitted variables correlated with child care quality have differential impacts on child outcomes at different ages. Our attempts to test whether model assumptions were met in these change analyses suggested that there might be reason for concern. Keeping these strengths and weaknesses of the various models in mind, it is not surprising that the quality effect sizes from the residualized change analyses generally fell between those from the level and simple change models. This analytic approach is incapable of eliminating bias, but may provide more accurate estimates when change model assumptions are not met in the data.

The Role of Selection Bias

What do these analyses imply about bias due to omitted variables in analyses relating child care quality and child outcomes? A starting point for answering this question is to note that the range of coefficients estimates for the impact of a one-standard deviation increment in child care quality from the least sophisticated level model to the most complete change models is large for both cognitive outcomes (raw score coefficients range from 2.72 to .59) and achievement outcomes (1.33 to .43). A small set of easily-measured covariates (mostly importantly maternal schooling) reduces the upperbound of this range considerably (from 2.72 to 1.57 in the case of cognitive outcomes and from 1.33 to 1.17 in the case of achievement outcomes). Adding an extensive set of covariates for the level and residualized change models produces small additional coefficient changes, which suggests that there might be relatively little additional bias from unobserved variables (Altonji & Taber, 2000).

Adding the earlier and time-invariant covariates in the change models markedly reduced the child care quality coefficients. These results indicate that concern regarding the possible impact of omitted variables is legitimate, but raises additional methodological concerns. Adding these variables can change the quality coefficients only if they have a stronger shared association between child care quality and cognitive outcomes at one age than at the other age (i.e., if the impact was comparable then adding their impact would subtract out when the change score was computed). This raises questions about the underlying assumption that each omitted variable also has a comparable effect at both times. Testing this assumption provided some evidence suggesting that estimates of ORCE-based child care quality coming from our change model were biased downward. On the other hand, our change models of impacts of alternative measures of child care quality generally produced larger estimates than did corresponding level and residualized change models. All in all, our analyses do not clearly answer the question about the degree to which omitted variable bias has led to faulty conclusions regarding whether and how much child care quality affects children's cognitive outcomes.

Child Care and Cognitive Development

What about the child care quality effect sizes implied by the coefficients in our models? While there is no way to know definitively which model provides the best estimates, almost all of them employing a reasonable number of covariates suggest modest quality effect sizes – i.e., a .04 to .08 standard-deviation increment to cognitive ability and achievement for a one standard-deviation increment in child care quality sustained between 24 and 54 months of age. Effect sizes for standard-deviation child care quality increments prior to age 24 months were around .08, although in this case we were unable to use change models to investigate the robustness of this estimate. Previous analyses from the NICHD Study of Early Child Care (NICHD ECCRN, 1999c), which relied on level models and did not include as many controls for selection bias, reported a considerably higher effect size of .19 for a one standard deviation change in quality sustained over a 1- to 3-year period on the school readiness scores of 3-year olds.

Models run separately on children with low early MDI scores produced larger effect size, which ranged from .07 to .18. This higher range either challenges the change model assumptions (it was not found in the level models) or indicates that child care quality has a bigger impact for these developmentally at-risk children, as has been found in the early intervention literature (cf. Ramey, Bryant, & Suerez, 1985), and the change models provided the best estimates of those effects.

Although our primary focus has been on ORCE-based quality impacts, it is important to point out that time in center-based child care in the third and fourth years of life had the most consistently significant associations with both cognitive and achievement outcomes across all of our various models. Our most complete models produced effect size estimates for spending all vs. none of this developmental period in center-based care ranged from .09 to .27 for cognitive outcomes and .22 to .33 for achievement outcomes. In contrast, center-based care earlier in childhood did not have consistently significant associations with cognitive and achievement outcomes. Comparable findings regarding center care for cognitive outcomes have previously been reported with the NICHD data (NICHD SECC, 2000)

Putting the Effect Sizes in Context

McCartney and Rosenthal (2000) urged investigators to place effect sizes from any given analysis in the context of other relevant effect sizes as a means of assessing their practical significance. In nonexperimental studies, for example, a .10 effect size is judged small (Cohen, 1988), but still could have both statistical and policy significance. The most complete models reported in this paper suggest that the effect sizes of a one-standard deviation increase in child care quality between 24 and 54 months of age ranges from .04 to .08 for the full sample of children and from .07 to .18 for initially-low MDI children. How do these compare with the range of effect sizes from (a) other naturalistic studies of variation in child care quality, (b) analyses that combine child care and early intervention programs, and (c) experimental studies of high-quality early intervention programs, welfare reform interventions, and elementary class size interventions?

Taking into account the duration of the quality increments, the effect sizes obtained in the present analysis for a 2.5-year quality enhancement between 24 and 54 months are considerably less than those obtained in other non-experimental studies of variation in child care quality. For example, the Cost, Quality, and Outcome study (CQO Study), which encompassed a broader range of quality than did the NICHD study, reported an effect size for a nine-month increment in child care quality of approximately .20 on a measure of language and math skills at entry to kindergarten in a sample of over 700 predominantly middle-class children (Peisner-Feinberg et al., 2001). A recent, secondary analysis combining results from the CQO study, a study of Head Start programs in North Carolina, and study of public preschool programs in North Carolina (Burchinal et al., 2001) reports effect sizes for one standard deviation change in quality sustained over a 6- to 9-month period of .21-.24 for pre-reading and math skills.

Our effect sizes report are also substantially smaller than those reported in experimental studies of early preschool intervention programs offering levels of quality that routinely exceed those of more typical community-based child care programs and focusing on children at risk due to both economic and developmental factors (Layzer, Goodson & Moss, 1993; Phillips et al., 1994). For example, treatment effect sizes on IQ were 2.0 at age 5 in the two-year Milwaukee Project; 1.0 at 3 years and .75 at age 5 for the three and five year treatment of the Abecedarian Project, .60 for the 1-2 year Perry Preschool Project, and .50 at age 5 for five-year Project CARE (Ramey, Bryant, & Suerez, 1985). These projects provided high quality care to disadvantaged children randomly assigned to treatment for a minimum of two years. The much greater intensity and duration of these treatments, as well as their target population of at-risk children, are the most likely explanations for their larger effect sizes. In addition, the effect sizes in these randomized experimental studies may be more precise because extensive covariates are not needed to adjust for potential selection bias.

The range of child care quality effect sizes obtained in the current study is also in the lower half of the range of effect sizes (.01 to .31) for parent and teacher reports on early elementary school achievement reported in recent studies of the effects of welfare reform experiments on child outcomes for children who were preschool-aged at random assignment (Morris et al., 2001). Effect size estimates ranging from .15 to .35 were obtained in experimental studies of the impacts of class-size reduction on elementary children's math and reading scores (Word et al., 1990; Krueger, 1999).

One crucial policy question is whether our estimated effects are economically significant, which depends on a comparison of the benefits associated with a .04 to .08 standard-deviation increase in test scores and the cost of producing a standard-deviation (half-point) change in the ORCE for three years**.** This is an important policy issue because quality child care seems to provide a more effective means of improving children's cognitive outcomes, especially for at-risk children, than other types of widely used programs such as home-visit programs (Lazar and St. Pierre, 1996). It is commonly assumed that the cost differences between high and low-quality care are substantial because quality is associated with caregiver education and wages (Phillips, et al., 1994), but parents in this sample reported relatively modest differences in the fees they paid for low and high quality care.[9] If quality could somehow be increased by one standard deviation for the estimated increment in wages ($15 per week, or about $2,000 for the interval between 24 and 54 months), then our results suggest that the quality impacts may be cost effective, particularly for children with low initial cognitive ability. Of course, it is hard to believe that merely increasing child care workers' wages by this amount would produce a one-standard-deviation increment in quality. Perhaps an even more promising line of intervention research is to focus on increasing the use of center-based care for 3 and 4 year-olds, given our findings of consistently larger effect sizes associated with that kind of change. A final judgment on the economic significance of the effect sizes we estimate requires more information than we have at hand.

Limitations of the Study

There are important qualifications to these conclusions. First, even our most comprehensive level models omit potentially important variables, and we do not know how including them would affect our estimates. The change model provides less biased estimates with somewhat larger standard errors when model assumptions are met, but possibly more biased estimates when the impact of the omitted variables vary over time (Wang & Burchinal, 2001). In our case, it is likely that at least some omitted variables operate differently over time. For example, entitlement programs such as Head Start or pre-kindergarten programs may be primarily responsible for placing low income children in higher quality care at older ages, but differential parental selection of higher quality care among more educated parents may be primarily responsible for placing children in higher quality care at younger ages. If this is the case, then the change model may underestimate the impact of quality.

Second, few children were observed in low-quality settings, and quality impacts may be larger for children in the worst child care settings. Our attempts to isolate the impacts of improving upon the lowest quality observed were not successful, perhaps because the very worst settings are underrepresented in the data, rendering it difficult to assess the benefits of improving upon them. Another consequence of the truncated distribution of child care quality is that we may be underestimating effect sizes. Efforts to simulate nationally representative data with our sample suggest that quality of care observed in this study is somewhat higher and less variable than overall quality of care in the nation (NICHD ECRRN, 2000).

---

[9] When we used age 54-month data to regress the ORCE on the hourly wages of staff, observed ratio and observed group size, we found that both hourly wages and observed ratios were significant predictors of ORCE for child care centers but not child care homes. For centers, the weekly wage cost associated with a one-standard deviation increase in ORCE was about $15. This does not imply that raising wages by this amount would cause a one standard deviation increase in the ORCE, but does provide some information on the possible costs associated with increasing the ORCE ratings.

Third, center care was coded only if the child spent at least 10 hours in that setting. Follow-up analyses indicated similar quality and center effects when any center care was coded. Finally, our quality measure does not explicitly reflect educational dimensions of the child care settings, which may be particularly important for the outcomes assessed in this study. Most of our conclusions relate to child care quality measured by the ORCE during the time when the child is between 24 and 54 month old. The ORCE focuses on the sensitivity/responsivity and affective quality of caregiver-child dyadic interactions. As such, it may be missing dimensions of quality more directly related to the educational content of the child care setting and to cognitive-achievement outcomes for children, such as specific learning-focused exchanges, the curriculum, or available learning materials. The fact that being in center-based care, independently of quality, has a consistent positive relation to cognitive outcomes suggests that there may be features in the structure and organization of child care centers, and the typically stronger educational qualifications of center-based providers, that are important influences not captured in the ORCE.

Conclusions

There has been substantial debate about the adequacy of common-used statistical methods and non-experimental data for securing unbiased estimates of the effects of child care quality on child outcomes (e.g., Duncan & Gibson, 2000). The current study used a variety of methods for controlling for selection bias, derived from both the economics and developmental literatures, to gain a better understanding of this important scientific and policy issue. Our findings do not identify <u>the</u> best method, although they do provide a fairly narrow range of likely effect sizes. They suggest that while the prevailing developmental literature on child care may have overestimated the developmental consequences of child care quality, studies that control for demographic, family and maternal variables produce estimates that are close to the high end of our range of preferred estimates. There remains a compelling need for experimental studies of child care that encompass the typical range of quality in community-based arrangements.

At the same time, policy decisions continue to be made regarding the desirability of investing in efforts to improve the quality of child care for the nation's children. Our findings suggest that such efforts are likely to improve children's cognitive and achievement outcomes, although we cannot establish whether the benefits of such efforts exceed their costs. We find more consistent support for the likely benefits of providing more opportunities for preschoolers to attend center-based child care. Finally, our results also suggest that efforts to improve quality may be especially important for young children with low initial cognitive skills. All of our conclusions, while often consistent with the existing research, require replication using even stronger research designs and samples with strong representation of children in low-SES families.

References

Allison, P. (1990). Change scores as dependent variables in regression analysis. In C.C. Clogg (Ed.), Sociological methodology, (pp. 93-114). Oxford: Basil Blackwell.

Altonji, J., Elder, T., & Taber, C. (2000). Selection on observable and unobservable variables: Assessing the effectiveness of Catholic schools. (Mimeographed) Northwestern University.

Barnett, W.S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. The Future of Children, 5(3), 25-50.

Baydar, N., & Brooks-Gunn, J. (1991). Effects of maternal employment and child care arrangements on preschoolers' cognitive and behavioral outcomes: Evidence from the children of the national longitudinal survey of youth. Developmental Psychology, 27, 932-945.

Bayley, N. (1969). Bayley Scales of Infant Development. New York: Psychological Corp.

Bayley, N. (1993). Bayley II Scales of Infant Development. New York: Psychological Corp.

Blau, D. (1999). The effects of child care characteristics on child development, Journal of Human Resources, 34, 786-822.

Bound, J., Brown, C., Duncan, G. & Rodgers, W. (1994), Evidence on the validity of cross-sectional and longitudinal labor market data, Journal of Labor Economics, 12, 345-368.

Bradley, R.H., Caldwell, B.M., Rock, S.L., Ramey, C.T., Barnard, K.E., Gray, C., Hammond, M.A., Mitchell, S., Gottfried. A.W., Siegel, L., & Johnson, D.L. (1989). Home environment and cognitive development in the first three years: A collaborative study involving six sites and three ethnic groups in North America. Developmental Psychology, 25, 217-235.

Bracken, B.A. (1984). Bracken basic concept scales. San Antonio, TX: Psychological Corporation.

Bronfrenbrenner, U., & Morris, P.A. (1998). The ecology of developmental processes. In W. Damon, & R. M. Lerner (Eds.), Handbook of child psychology: Vol. 1 Theoretical models of human development 5th ed, (pp. 993-1028). New York: Wiley & Sons, Inc.

Bryant, D.M., Burchinal, M.R., Lau, L.B.,& Sparling, J.J. (1994). Family classroom correlates of Head Start children's developmental outcomes. Early Childhood Research Quarterly, 9, 289-309

Bryk, A., & Weisburg, H. (1977). Use of nonequivalent control group design when subjects are growing. Psychological Bulletin, 84, 950-962.

Byrk & Raudenbush, (1992). Heirarchical linear models: applications and data analysis methods. Newbury Park, California: Sage Publications.

Burchinal, M.R., Ramey, S.L., Reid, M.K., & Jaccard, J. (1995) Early child care experiences and their association with family and child characteristics during middle school. Early Childhood Research Quarterly, 10, 33-61.

Burchinal, M.R., Roberts, J.E., Nabors, L., & Bryant, D.M., (1996). Quality of center infant care and infant cognitive and language development. Child Development, 67, 606-620.

Burchinal, M.R., Peisner-Feinberg, E., Bryant, D.M., & Clifford, R. (2000). Children's social and cognitive development and child care quality: Testing for differential associations related to poverty, gender, or ethnicity. Journal of Applied Developmental Sciences, 4, 149-165.

Burchinal, M.R., Roberts, J.E. Riggins, R., Zeisel. S., Neebe, E., & Bryant, M. (2000). Relating quality of center child care to early cognitive and language development longitudinally. Child Development, 71, 339-357.

Caldwell, B. M., & Bradley, R. H. (1984). Home observation for measurement of the environment. Little Rock: University of Arkansas Press.

Campbell, F. & Ramey, C.T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. Child Development, 65, 684-698.

Campbell, F.A., Pungello, E.P., Miller-Johnson, S., Burchinal, M.R., & Ramey, C. (2001). The development of cognitive and academic abilities: Growth curves from an early intervention educational experiment. Developmental Psychology, 37, 231-242.

Campbell, F. A., Ramey, C. T., Pungello, E. P., Sparling, J., & Miller-Johnson, S. (in press). Early childhood education: Young adult outcomes from the Abecedarian Project. Applied Developmental Science.

Caughy, M.O., DiPietro, J.A., & Strobino, D.M. (1994). Day-care participation as a protective factor in the cognitive development of low-income children. Child Development, 65, 457-471.

Clarke-Stewart, K., (1989). Infant day care: maligned or malignant? American Psychologist, 44, 266-273.

Clarke-Stewart, K.A., Gruber, C.P., & Fitzgerold, L.M. 1994. Children at home and in day care. Hillsdale, NJ: Lawrence Erlbuam.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale NJ: Lawrence Erlbaum.

Costa, P.T., & McCrae, R.R. (1985). The NEO personality inventory manual. Odessa, FL: Psychological Assessment Resources.

Cronbach, L.J., & Furby, L. (1970). How should we measure "change" - Or should we? Psychological Bulletin,74, 68-80.

Diggle, P., Liang, K., & Zeger, S. (1994). Fitting transition models. Analysis of longitudinal data. (pp. 192-194). Oxford: Clarendon Press.

Duncan, G.J., & Gibson, C., Selection and attrition in the NICHD Childcare Study's analyses of the impacts of childcare quality on child outcomes, mimeo, Institute for Policy Research, Northwestern University, July 2000

Duncan, G., Magnuson, K., & Ludwig, J. (2000) The endogeneity problem in developmental studies. (Mimeographed) Northwestern University.

Dunn, L. (1993). Proximal and distal features of day care quality and children's development. Early Childhood Research Quarterly, 8, 167-192.

Dunn, L.M., & Dunn, L.M. (1981). Peabody picture vocabulary test-revised. Circle Pines, MN: American Guidance Service.

Greenberger, E., Goldberg, W.A., Crawford, T.J., & Granger, J. (1998). Beliefs about the consequences of maternal employment for children. Psychology of Women Quarterly, 12, 35-59.

Gujarati, D.M. (1995). Basic Econometrics, McGraw-Hill, New York

Hock, E., Gnezda, M.T., & MacBride, S., (1983). The measurement of maternal seperation anxiety . Paper presented at the biennial meeting of the Society for Research in Child Development, Detroit.

Howes, C. (1997). Children's experiences in center-based child care as a function of teacher background and adult:child ratio. Merrill Palmer Quarterly, 43, 404-425.

Howes, C. (2000). Social development, family, and attachment relationships of infants and toddlers: Research to practice. In D. Cryer & T. Harms (Eds.), Infants and toddlers in out-of-home care (pp. 87-116). Baltimore: Brookes Publishing.

Howes, C., Phillips, D.A., & Whitebook, M. (1992). Thresholds of quality: Implications for the social development of children in center-based child care. Child Development, 63, 449-460.

IHDP. (1990). Enhancing the outcomes of low-birth-weight, premature infants - A multisite, randomized trial. Journal American Medical Association, 263, 3035-3042.

Jennrich, R.I., & Schluchter, (1986). Unbalanced repeated-measures models with structured covariance matrices, Biometrics, 42, 805-820.

Katz, J.R., & Snow, C.E. (2000). Language development in early childhood: The role of social interaction. In D. Cryer & T. Harms (Eds.), Infants and toddlers in out-of-home care (pp. 49-86). Baltimore: Brookes Publishing.

Kontos, S., Howes, C., Shinn M., & Galinsky, E. (1995). Quality in family child care and relative care. New York: Teachers College Press.

Krueger, Alan. B. (1999). Experimental estimates of education production functions. Quarterly Journal of Economics, 114, 497-532.

Labouvie, E.W. (1980). Measurement of individual differences in intraindividual changes. Pychological Bulletin, 88, 54-59.

Laird, N.M., & Ware, J.H. (1982). Random-effects models for longitudinal data. Biometrics, 38, 963-974.

Lamb, M. (1998). Nonparental child care: Contest, quality, correlates, and consequences. In I. Sigel & A. Renninger (Eds.), W. Damon (Series Ed.), Handbook of child psychology: Vol 4. Child psychology in practice. New York: Wiley.

Layzer, J.I., Goodson, B.D., & Moss, M. (1993). Life in preschool: Final report, Vol. 1. Observational Study of Early Childhood Programs. Washington, DC: U.S. Dept of Ed.

Lazar, I., & Darlington, R. (1982). Lasting effects of early education: A report from the Consortium for Longitudinal Studies. Monographs of the Society for Research in Child Development, 47 2-3, Serial No. 195.

McCall, R.B. (1977). Childhood IQs as predictors of adult educational and occupational status. Science, 1977, 482-483.

McCartney, K. (1984). The effect of quality of day care environment upon children's language development. Developmental Psychology, 20, 244-260.

McCartney, K., & Rosenthal, R. (2000). Effect size, practical importance, and social policy for children. Child Development, 71, 173-180.

Medoff-Cooper, B., Carey, W.B., & McDevitt, S. (1993). The Early Infancy Temperament Questionnaire. Journal of Developmental and Behavioral Pediatrics, 14, 230-235.

Morris, P. A., Huston, A. C., Duncan, G. J., Crosby, D. A., and Bos, J. M. (2001). How welfare and work policies affect children: A synthesis of research. NY, New York: Manpower Demonstration Research Corporation.

Neisser, U., Boodoo, G., Bouchard, T.J., Boykin, A.W., Brody, N., Ceci, S.J., Halpern, D.F., Loehlink, J.C., Perloff, R., Steinberg, R., & Urbina, S. (1996). Intelligence: Knowns and unknowns. American Psychologist, 51, 77-101.

NICHD Early Child Care Research Network. (1996). Characteristics of infant child care: Factors contributing to positive caregiving. Early Childhood Research Quarterly, 1, 269-306.

NICHD Early Child Care Research Network. (2000). Characteristics and quality of child care for toddlers and preschoolers. Journal of Applied Developmental Sciences, 4, 116-135

NICHD Early Child Care Research Network. (1997). Familial factors associated with characteristics of nonmaternal care for infants. Journal of Marriage and the Family, 59, 389-408.

NICHD Early Child Care Network. (1998). Early child care and self-control, compliance and problem behavior at 24 and 36 months. Child Development, 69, 1145-1170.

NICHD Early Child Care Network. (1999a). Child outcomes when child-care center classes meet recommended standards for quality. American Journal of Public Health, 89, 1072-1077.

NICHD Early Child Care Research Network (1999b). Child care and mother-child interactions in the first three years of life. Developmental Psychology, 35(6), 1399-1413.

NICHD Early Child Care Network. (1999c, April). Effect sizes from the NICHD study of early child care. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Albuquerque, NM.

NICHD Early Child Care Research Network. (in press). Early child care and children's development prior to school entry. American Education Research Journal.

NICHD Early Child Care Network. (2000). Poor children in working families: Parent employment and poverty as predictors of developmental outcomes. (Mimeographed)

NICHD Early Child Care Research Network.(In Press). Structure>Process>Outcome: Direct and indirect Effects of Caregiving Quality on Young Children's Development. Psychological Science.

NICHD Early Child Care Network. (2000). The relation of child care to cognitive and language development Child Development, 71, 960-980

Peisner-Feinberg, E., & Burchinal, M. (1997). Concurrent relations between child care quality and child outcomes: The study of cost, quality, and outcomes in child care centers. Merrill-Palmer Quarterly, 43, 451-477.

Peisner-Feinberg, E.S., Burchinal, M.R., Clifford, R.M., Culkin, M.L., Howes, C., Kagan, S.L., & Yazejiam, N., (2001). The relation of preschool child care quality to children's cognitive and social developmental trajectories through second grade. Child Development, 72, 1534-1553.

Phillips, D.A., McCartney, K., & Scarr, S. (1987). Child care quality and children's social development. Developmental Psychology, 23, 537-543.

Phillips, D., Voran, M., Kisker, E., Howes, C., & Whitebook, M. (1994). Child care for children in poverty: Opportunity or inequity, Child Development, 65, 472-492.

Pungello, E., & Kurtz-Costes, B (1999) Developmental Psychology

Radloff, L. (1977). The CES-D scale: A self-report depression scale for research in the general population. Applied Psychological Measurement, 1, 385-410.

Ramey, C.T., Bryant, D.M., & Suarez, T.M. (1985). Preschool compensatory education and the modifiability of intelligence: A critical review. In D. K. Detterman (Ed.), Current topics in human intelligence. Vol. 1. Research methodology (pp. 247-296). Norwood, NJ: Ablex Publishing Corporation.

Raudenbush, S.W. (2001). Toward a coherent framework for comparing trajectories of individual change. In L. Collins & A. Sayers (Eds.), New Methods for the Analysis of Change (pp. 35-64). Washington D.C.: American Psychological Association.

Reynell, J. (1991). Reynell developmental language scales (US Edition). Los Angeles: Western Psychological Service.

Rogosa, D., Brandt, D., & Zimowski, M. (1982). A growth curve approach to measurement of change. Psychological Bulletin, 92, 726-748.

Sameroff, A.J. (1983). Developmental systems: Contexts and evolution. In W. Kessen (Ed.), Handbook of child development: Vol. 1 History, theories, and methods. (pp. 237-294). New York: Wiley.

Scarr, S. (1998). American child care today. American Psychologist, 53, 95-108.

Schaefer, E.S., & Edgerton, M. (1985). Parent and child correlates of parental modernity. In I. E. Sigel (Ed.), Parental belief systems: The psychological consequences for children. Hillsdale, NJ: Erlbaum.

Schliecker, E., White, D.R., & Jacobs, E. (1991). The role of day care quality in the prediction of children's vocabulary. Canadian Journal of Behavioral Science, 23, 12-24.

Schweinhart, L.J., Weikart, D.P., & Larner, M.B. (1986). Consequences of three preschool curriculum models through age 15. Early Childhood Research Quarterly, 1, 15-45.

Singer, J.D., Fuller, B., Keiley, M.K., & Wolf, A. (1998). Early child-care selection: Variation by geographic location, maternal characteristics, and family structure. Developmental Psychology, 34, 1129-1144.

Tomasello, M., & Farrar, J. (1986). Joint attention and early language. Child Development, 57, 1454-1463.

Vandell, D.L., & Corasaniti, M.A. (1990). Child care and the family: Complex contributors to child development. In K. McCartney (Ed.), New directions in child development. San Francisco: Jossey-Bass.

Wang, C., & Burchinal, M. (2001). Modeling behavioral science data with unobserved covariates. Unpublished manuscript, University of North Carolina

Woodcock, R.W., & Johnson, M.B. (1990). Tests of achievement, WJ-R. Examiner's manual. Allen, TX: DLM Teaching Resources.

Yoshikawa, H. (1995). Long-term effects of early childhood programs on social outcomes and delinquency. The Future of Children, 5(3), 51-75.

Zimmerman, I.L., Steiner, & Pond, R.E. (1979). Preschool language scale. San Antonio : The Psychological Corporation.

Table 1: Description of sample and analysis variables

	1 month			6 months			15 months			24 months			36 months			54 months		
	N	Mean	sd	N	Mean	Sd	N	mean	sd	n	mean	sd	n	mean	sd	n	mean	sd
54m Cognitive Score							1180	108.6	14.1	1162	92.1	14.6				1078	98.2	15.1
PLS Auditory Compreh.																1064	98.3	19.9
PLS Expressive																1055	100.6	20.0
WJ Picture Vocabulary																1060	101.0	14.7
WJ Memory for Sentences																1054	93.0	18.6
54m Achievement Score																1056	99.7	11.7
WJ Applied Problems																1053	103.3	15.9
WJ Letter-Word Ident.																1056	99.3	13.9
WJ Incomplete Words																1050	96.7	13.6
Family/Mother Characteristics																		
Maternal Education	1363	14.2	2.5															
Household-partner	1364	.85		1279	.86		1243	.85		1207	.85		1216	.83		1084	.83	
Income/Pov Threshold				1270	3.7	3.1	1234	3.7	3.2	1187	3.7	3.0	1208	3.6	3.1			
Ethnicity – white	1364	.79																
African-American	1364	.11																
Hispanic	1364	.06																
Other	1364	.04																
Maternal Depression	1363	11.4	9.0	1278	9.0	8.3	1241	9.0	8.2	1119	9.4	8.6	1202	9.2	8.3	1077	9.8	8.7
Mat Agreeableness				1272	46.3	5.3												
Mat Extroversion				1272	42.5	5.8												
Mat neuroticism				1272	29.8	7.2												
M-Separation Anxiety	1351	70.3	13.3															
M-Benefits of Work	1363	19.2	3.2															
M-IQ (PPVT)													1167	99.0	18.4			
Child Characteristics																		
Gender (male=1)	1353	.50																
Temperament				1279	3.2	.4												
Home/Parenting																		
HOME Total				1279	36.5	4.6												

Mat Sensitivity				1272	9.2	1.8												
Progressive Attitude	1358	32.7	3.5															
Child Care																		
Quality: ORCE rating				593	2.98	.57	656	2.93	.57	669	2.81	.56	707	2.80	.46	854	2.98	.56
Hours/week				1320	22.4	20.6	1276	24.6	20.5	1251	25.2	20.5	1230	26.3	20.1	1057	35.2	17.5
Any Center Care (1=yes)				1364	.09		1269	.12		1239	.20		1229	.36		1136	.74	
Center Care, 10+ hrs/wk				1364	.08		1269	.11		1239	.17		1229	.27		1136	.54	
Exclusive Maternal Care				1316	.43		1269	.37		1239	.34		1229	.31		1136	.20	
Caregiver Education				583	13.35	2.25	607	13.34	2.21	584	13.5	2.09	636	13.8	2.14	737	15.1	2.4
Observed child/adult ratio				591	2.77	2.02	608	2.99	1.90	668	3.70	2.23	707	4.86	2.97	854	6.90	3.59
Observed group size				591	3.29	3.21	668	3.77	3.28	669	5.06	4.51	707	7.34	5.99	854	12.71	7.37

Table 2: "Level Models": Regression Coefficients And Standard Errors For Various Regression Models Relating Continuous Measures Of Child care Quality To Children's Outcomes At 24 And 54 Months

Independent variables	24-month Cognitive score	Level model: Cognitive score				Level model: Achievement score			
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 2	Model 3	Model 4	Model 5
N	1162	1078	1078	1078	1078	1056	1056	1056	1056
R^2	.34***	.18***	.36***	.42***	.43***	.13***	.29***	.34***	.35***
Coefficients	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)
ORCE Quality 6-24 mo. Average (z score)	1.66*** (.48)	3.35*** (.55)	1.64** (.50)	1.34** (.48)	1.36** (.49)	2.26*** (.45)	.96* (.41)	.85* (.41)	.90* (.41)
ORCE Quality 36-54 mo. Average (z score)		2.72*** (.50)	1.57*** (.44)	1.19** (.43)	1.17** (.43)	2.15*** (.40)	1.33*** (.36)	1.17** (.36)	1.12** (.36)
(1,0) Mother care 6-24 mo.	-.27 (1.15)	-1.50 (1.38)	-1.35 (1.23)	-1.84 (1.18)	-1.51 (1.19)	-.87 (1.11)	-.81 (1.01)	-.98 (.98)	-.64 (.99)
(1,0) Mother care 36-54 mo.=1		-4.41* (1.74)	-1.84 (1.55)	-.46 (1.49)	-.33 (1.49)	-3.16* (1.41)	-1.25 (1.29)	-.45 (1.26)	-.42 (1.26)
(1,0) Quality 6-24 mo. Missing	-.64 (1.40)	-6.88*** (1.61)	-4.11** (1.44)	-3.93** (1.40)	-4.19** (1.40)	-3.67* (1.31)	-1.61 (1.19)	-1.89 (1.18)	-2.15 (1.18)
(1,0) Quality 36-54 mo. missing		-1.57 (1.69)	-1.19 (1.50)	.06 (1.44)	-.02 (1.45)	-1.79 (1.37)	-1.48 (1.24)	-.71 (1.21)	-.69 (1.21)
Hours of care 3-24 mo. average	.04 (.03)	.14** (.04)	.07 (.04)	.06 (.04)	.06 (.04)	.06 (.03)	.01 (.03)	-.01 (.03)	-.01 (.03)
Hours of care 27-54 mo. Average		-.15*** (.05)	-.07 (.04)	-.07 (.04)	-.06 (.04)	-.06 (.04)	.01 (.04)	.02 (.03)	.02 (.03)
Prop. Center care 3-24m	5.41*** (1.58)	1.20 (2.27)	-1.29 (2.02)	-1.39 (1.93)	-1.09 (1.94)	-.47 (1.83)	-2.30 (1.67)	-2.07 (1.62)	-2.07 (1.63)
Prop. Center care 27-54m		7.19*** (1.65)	5.28*** (1.48)	4.34** (1.43)	4.10** (1.44)	5.15*** (1.33)	3.60*** (1.22)	3.21** (1.20)	3.13** (1.21)

Notes: ORCE quality was rescaled to have a mean of 0 and standard deviation of 1.

*p<.05 ** p<.01 ***p<.001

Model 1: *Covariates:* site, gender, ethnicity, maternal education, family income/poverty threshold (6-24m), 6m maternal sensitivity, 6m HOME total score, maternal depression (6-24m), maternal Peabody Picture Vocabulary Test score, % time partner in household (6-24), 1m child temperament, 6m maternal psychological adjustment (PI agree, extroversion, neuroticism), 1m maternal child rearing attitudes, 1m maternal separation anxiety, 1m maternal attitudes about benefits of work, missing value variables. *Child care variables:* mean of ORCE child care quality at 6, 15, 24m; mean of hours of care/week 3-24m; proportion of time in center care at least 10 hours/week 3-24m; exclusive mother care 6-24m (yes or no); missing quality of care data 6-24m.

Model 2: *Covariates*: site only. *Child care variables*: mean of ORCE child care quality at 6-24m; mean of hours of care/week 3-24m; proportion of time in center care at least 10 hours/week 3-24m; exclusive mother care 6-24m (yes or no); missing quality of care data 6-24m; mean of ORCE child care quality at 36-54m; mean of hours of care/week 27-54m; proportion of time in center care at least 10 hours/week 27-54m; exclusive mother care 27-54 (yes or no); missing quality of care data 27-54m.

Model 3 adds to Model 2 covariates: gender, ethnicity, and maternal education, missing value variables.

Model 4 adds to Model 3 covariates: income/poverty thresholds averaged between 6-24 months & 36-54 months, 6m maternal sensitivity, 6m HOME total score, maternal depression averaged between 6-24 months & 36-54 months, maternal PPVT, % time partner in household for 6-24 months & 36-54 months, missing value variables.

Model 5 adds to Model 4 covariates: 6m child temperament, 6m maternal psychological adjustment (PI agree, extroversion, neuroticism), 1m maternal child rearing attitudes, 1m maternal separation anxiety, 1m maternal attitudes about benefits of work, missing value variables.

Table 3: "Change Models": Regression Coefficients And Standard Errors For Various Regression Models Relating Continuous Measures Of Child Care Quality To Change in Children's Cognitive and Achievement Outcomes Between 24 And 54 Months

Independent variables	**Simple change models: cognitive scores**					**Residualized change models: cognitive scores**				
	Model 6	Model 7	Model 8	Model 9	Model 10	Model 6'	Model 7'	Model 8'	Model 9'	Model 10'
N	1032	1032	1032	1032	1032	1032	1032	1032	1032	1032
R^2	.02*	.08***	.10***	.13***	.13***	.46***	.48***	.53***	.56***	.57***
Coefficients	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)
ORCE Quality 36-54 mo. Average	1.17** (.44)	.87* (.44)	.72 (.44)	.58 (.45)	.59 (.45)	1.73*** (.40)	1.55*** (.40)	1.11** (.39)	.88* (.38)	.86* (.38)
(1,0) Mother care 36-54 mo.=1	-2.18 (1.61)	-2.11 (1.58)	-1.83 (1.57)	-1.50 (1.57)	-1.52 (1.58)	-3.20* (1.45)	-3.20* (1.43)	-2.13 (1.37)	-1.45 (1.34)	-1.38 (1.34)
(1,0) Quality 36-54 mo. Missing	-.93 (1.53)	-.93 (1.50)	-1.12 (1.50)	-.19 (1.51)	-.31 (1.51)	-1.76 (1.38)	-1.40 (1.36)	-1.39 (1.30)	-.27 (1.28)	-.38 (1.29)
Hours of care 27-54 mo. Average	-.03 (.04)	-.02 (.04)	-.02 (.04)	-.00 (.04)	-.01 (.04)	-.10** (.04)	-.08* (.04)	-.05 (.04)	-.04 (.04)	-.04 (.04)
Prop. Center care 27-54m	3.41* (1.47)	4.04** (1.46)	3.82** (1.47)	3.26** (1.48)	3.22** (1.49)	4.56*** (1.33)	5.37*** (1.32)	4.70*** (1.28)	3.95** (1.26)	3.84** (1.26)

(Table 3 continued)

	Simple change models: achievement score					Residualized change models: achievement scores				
	Model 6	Model 7	Model 8	Model 9	Model 10	Model 6'	Model 7'	Model 8'	Model 9'	Model 10'
N	1014	1014	1014	1014	1014	1014	1014	1014	1014	1014
R-squared	.02*	.08***	.09***	.10***	.11***	.34***	.36***	.41***	.43***	.44***
Coefficients	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B(se)	B (se)
ORCE Quality 36-54 mo. Average	.38 (.45)	.25 (.45)	.39 (.45)	.45 (.46)	.43 (.46)	1.23*** (.35)	1.28*** (.35)	.95** (.34)	.88** (.34)	.84* (.34)
(1,0) Mother care 36-54 mo.=1	-.38 (1.66)	-.51 (1.63)	-.84 (1.63)	-.95 (1.64)	-1.04 (1.64)	-1.83 (1.27)	-2.12 (1.26)	-1.26 (1.22)	-.87 (1.21)	-.89 (1.21)
(1,0) Quality 36-54 mo. Missing	-.13 (1.56)	-.62 (1.54)	-.90 (1.54)	-.54 (1.56)	-.57 (1.57)	-1.26 (1.19)	-1.47 (1.19)	-1.41 (1.15)	-.79 (1.15)	-.78 (1.15)
Hours of care 27-54 mo. Average	.08 (.04)	.07 (.04)	.06 (.04)	.08 (.04)	.07 (.04)	-.01 (.03)	-.00 (.03)	.02 (.03)	.04 (.03)	.04 (.03)
Prop. Center care 27-54m	1.54 (1.51)	1.73 (1.50)	1.73 (1.50)	1.66 (1.55)	1.82 (1.54)	3.15** (1.16)	3.57** (1.16)	2.87* (1.13)	2.54* (1.13)	2.57* (1.13)

Notes: ORCE quality was rescaled to have a mean of 0 and standard deviation of 1.

*p<.05 ** p<.01 ***p<.001
Model 6: *Covariates:* none *Child care variables*: mean of ORCE child care quality at 6-24m; mean of hours of care/week 3-24m; proportion of time in center care at least 10 hours/week 3-24m; exclusive mother care 6-24m (yes or no); missing quality of care data 6-24m; mean of ORCE child care quality at 36-54m; mean of hours of care/week 27-54m; proportion of time in center care at least 10 hours/week 27-54m; exclusive mother care 27-54 (yes or no); missing quality of care data 27-54m.
Model 7: adds to Model 6 *Covariates*: site only.
Model 8 adds to Model 7 *Covariates:* gender, ethnicity, and maternal education, missing value variables.
Model 9 adds to Model 8 *Covariates:* income/poverty thresholds averaged between 6-24 months & 36-54 months, 6m maternal sensitivity, 6m HOME total score, maternal depression averaged between 6-24 months & 36-54 months, maternal PPVT, % time partner in household for 6-24 months & 36-54 months, missing value variables.

Model 10 adds to Model 9 *Covariates*: 6m child temperament, 6m maternal psychological adjustment (PI agree, extroversion, neuroticism), 1m maternal child rearing attitudes, 1m maternal separation anxiety, 1m maternal attitudes about benefits of work, missing value variables.
Models 6'-Model 10' include all listed child care variables and covariates and add the MDI 24m as a right hand variable.

Table 4: Regression Coefficients and Standard Errors For Various Regression Models Relating Categorical Child care Quality To Children's Cognitive Outcomes At 24 And 54 Months

		24m Cognitive score: Level Model	54m Cognitive score: Level Model	54m Cognitive score: Change Model	54m Cognitive score: Residualized Change Model	54m Achievement score: Level Model	54m Achievement score: Change Model	54m Achievement score: Residualized, Change Model
N		1062	1078	1032	1032	1056	1014	1014
R^2		.35***	.43***	.12***	.56***	.36***	.10***	.44***
6-24m Quality Group Comparison		F(4,1119)= 4.04 p=003.	F(4,1023)= 1.82, p=.12			F(4,1001)= 2.76, P=.027		
Linear Quality		F(1,1119) = 7.12, p=.009	F(1,1023)= 3.13, p=.08			F(1,1001)= 2.20, p=.14		
6-24m Child Care Categories		B (se)	B (se)	B (se)	B (se)	B (se)	B (se)	B (se)
ORCE Quality 1.0-2.0(n=32)		Reference	Reference			Reference		
2.0-2.5 (n=107)		.44 (2.43)	.25 (2.41)			.72 (2.02)		
2.5-3.0 (n=246)		.55 (2.31)	.65 (2.27)			-.79 (1.90)		
3.0-3.5 (n=262)		3.40 (2.30)	2.85 (2.27)			1.49 (1.90)		
3.5-4.0 (n=87)		5.68* (2.55)	3.46 (2.54)			2.94 (2.12)		
Exclusive Mother care (n=242)		1.96 (2.42)	.14 (2.42)			.08 (2.02)		
Formal care- not observed (n=47)		.81 (2.76)	-3.23 (2.76)			-2.41 (2.31)		
Informal care- not observed (n=37)		2.51 (2.98)	-1.51 (2.92)			-.45 (2.46)		

(table continued)		24m Cognitive score: Level Model	54m Cognitive score: Level Model	54m Cognitive score: Change Model	54m Cognitive score: Residualized Change Model	54m Achievement score: Level Model	54m Achievement score: Change Model	54m Achievement score: Residualized, Change Model
36-54m Child Care Group Comparison			F(1,1023)= 2.43, p=.046	F(4,989)= 1.42, p =.23	F(4,988)= 1.77 P=.12	F(1,1001)= 3.90, p=.005	F(4,971)= .99, p =.41	F(4,971)= 2.69 P=.03
Linear Quality Trend			F(1,1023)= 7.89, p=.005	F(1,989)= 3.00, p=.08	F(1,988)= 6.04, p=.014	F(1,1001)= 4.00, p=.046	F(1,971)= .32, p=.57	F(1,971)= 1.91, p=.17
36-54m Child Care Categories		B (se)	B (se)	B (se)	B (se)	B (se)	B (se)	B (se)
ORCE Quality 1.0-2.0 (n=26)		Reference	Reference	Reference	Reference	Reference	Reference	Reference
2.0-2.5 (n=125)			3.24 (2.54)	5.17 (2.62)	3.99 (2.24)	-1.44 (2.13)	.61 (2.75)	-1.26 (2.01)
2.5-3.0 (n=308)			4.16 (2.41)	3.55 (2.48)	3.64 (2.11)	-.40 (2.02)	-.81 (2.61)	-1.09 (1.91)
3.0-3.5 (n=341)			5.60* (2.42)	4.02 (2.49)	4.62* (2.12)	1.38 (2.03)	-.15 (2.61)	.38 (1.91)
3.5-4.0 (n=103)			6.55* (2.54)	5.47* (2.72)	5.60* (2.31)	3.18 (2.21)	2.05 (2.85)	2.17 (2.08)
Exclusive Mother care (n=79)			4.33 (2.72)	2.24 (2.81)	2.45 (2.39)	.23 (2.28)	-1.23 (2.95)	-1.22 (2.15)
Formal care-not observed (n=26)			6.07 (3.26)	2.21 (3.38)	3.29 (2.88)	1.10 (2.76)	-1.76 (3.56)	-1.01 (2.61)
Informal care-not observed (n=52)			4.21 (2.92)	4.00 (3.00)	3.67 (2.55)	-.33 (2.44)	-.63 (3.13)	-1.33 (2.29)

Notes: *p<.05 ** p<.01 ***p<.001
The 24-month cognitive score regression includes the same covariates as Table 2, Model 1.
The 54-month cognitive score regressions includes the same covariates as Table 2, Model 5.

The 24 to 54-month change regression includes the same covariates as Table 3, Model 10
The 54-month cognitive score regression model with 24-month cognitive scores as right-hand variables includes the same covariates as Table 3, Model 10'

Table 5: Child care quality coefficients for children with very low vs. higher scores on the 15m MDI cognitive assessment Predicting cognitive skills at 54m

Independent variables	54m Cognitive Score			54m Achievement Score		
	Level Model	Change Model	Resid. Change Model	Level Model	Change Model	Resid. Change Model
	B (se)	B (se)	B (se)	B (se)	B (se)	B (se)
Models with all covariates[a]						
Quality 6-24 months	1.36** (.49)			.90* (.41)		
Quality 36-54 months	1.17** (.43)	.59 (.45)	.86* (.38)	1.12** (.36)	.43 (.46)	.84* (.34)
Models that include interactions for low MDI and all child care covariates						
15m MDI Low (n=265)						
Quality 6-24 months	2.33* (1.02)	.44 (1.08)	1.23 (.91)	2.39** (.83)	.54 (1.08)	1.68* (.81)
Quality 36-54 months	1.12 (.77)	2.02* (.85)	1.50* (.75)	1.60* (.65)	2.08* (.86)	1.54* (.64)
15m MDI Normal (n=757)						
Quality 6-24 months	1.26* (.54)	-.76 (.57)	.24 (.49)	.63 (.46)	-1.27 (.59)	-.02 (.44)
Quality 36-54 months	1.11* (.50)	-.00 (.53)	.66 (.45)	.78 (.42)	-.40 (.54)	.48 (.40)

Notes: ORCE quality was rescaled to have a mean of 0 and standard deviation of 1.
[a-] these coefficients are also reported in Tables 2, 3, and 4. All covariates included in model.
*p<.05 ** p<.01 ***p<.001

Table 6: Unstandardized coefficients from analyses relating child outcomes to caregiver education and observed ratios and group sizes

	Level Model			Change Model			Residualized Change		
	1	2	3	1	2	3	1	2	3
Caregiver Education									
24m Cognitive Development	1.56***	.72**	.61*						
(n=541)	(.28)	(.27)	(.30)						
54m Cognitive Development	.64*	.10	.00	.46*	.46*	.42^{+}	.48*	.24	.17
(n=716)	(.26)	(.23)	(.22)	(.22)	(.22)	(.23)	(.20)	(.20)	(.20)
54m Academic Achievement	.61**	.22	.19	.50*	.64**	.65**	.52**	.34*	.32^{+}
(n=716)	(.20)	(.18)	(.18)	(.23)	(.23)	(.23)	(.17)	(.17)	(.17)
Observed Group Size									
24m Cognitive Development	.05	.11	.07						
(n=642)	(.18)	(.16)	(.16)						
54m Cognitive Development	-.02	-.04	-.01	-.27***	-.27***	-.27**	-.19**	-.17**	-.15*
(n=833)	(.09)	(.08)	(.08)	(.08)	(.08)	(.08)	(.07)	(.07)	(.07)
54m Academic Achievement	.03	.01	.03	-.21**	-.20*	-.22**	-.07	-.06	-.05
(n=833)	(.07)	(.06)	(.07)	(.08)	(.08)	(.08)	(.06)	(.06)	(.06)
Observed Child-Adult Ratio									
24m Cognitive Development	-.11	.03	.03						
(n=642)	(.30)	(.27)	(.27)						
54m Cognitive Development	.22	.07	.08	-.27^{+}	-.26^{+}	-.24^{+}	-.10	-.11	-.08
(n=832)	(.17)	(.15)	(.14)	(.14)	(.14)	(.14)	(.13)	(.13)	(.13)
54m Academic Achievement	.25*	.16	.17	-.22	-.18	-.16	.04	.05	.05
(n=832)	(.13)	(.12)	(.11)	(.14)	(.14)	(.15)	(.11)	(.11)	(.11)

Note: $^{+}$ .1 < p < .05; * p<.05; ** p<.01; *** p<.001

Model 1: covariates are type of care (center-yes or no) and hours of care/week

Model 2: covariates are gender, ethnicity, maternal education, missing value variables, type of care (center-yes or no) and hours of care/week
Model 3: covariates are income/poverty thresholds averaged between 6 and assessment age, 6m maternal sensitivity, 6m HOME total score, maternal depression averaged 6 and assessment age, maternal PPVT, % time partner in household for 6m and assessment age 6m child temperament, 6m maternal psychological adjustment (PI agree, extroversion, neuroticism), 1m maternal child rearing attitudes, 1m maternal separation anxiety, 1m maternal attitudes about benefits of work, missing value variables.
gender, ethnicity, maternal education, type of care (center-yes or no), hours of care/week

Appendix Table 1: Complete Regression Results for All-Covariate 54-month Cognitive Level and Change Models

	Dependent variables			Descriptive Statistics		
Independent variables	54m cognitive score	24m to 54m change	54m with 24m as control	Mean	SD	N
ORCE Quality 6-24 mo. Average	1.36** (.49)	-.33 (.51)	.48 (.43)	2.91	.49	734
ORCE Quality 36-54 mo. Average	1.17** (.43)	.59 (.45)	.86* (.38)	2.93	.45	903
(1,0) Mother care 6-24 mo.	-1.51 (1.19)	-.31 (1.23)	-.95 (1.05)	.23		1060
(1,0) Mother care 36-54 mo.	-.33 (1.49)	-1.52 (1.58)	-1.38 (1.34)	.07		1060
(1,0) Quality 6-24 mo. Missing	-4.19** (1.40)	-1.84 (1.49)	-2.69* (1.27)	.08		1060
(1,0) Quality 36-54 mo. Missing	-.02 (1.45)	-.31 (1.51)	-1.38 (1.34)	.07		1060
Hours of care 3-24 mo. Average	.06 (.04)	-.01 (.04)	.03 (.03)	23.24	16.89	1060
Hours of care 27-54 mo. Average	-.06 (.04)	-.01 (.04)	-.04 (.04)	28.90	15.85	1060
Prop. Center care 3-24m	-1.09 (1.94)	-6.07** (2.00)	-3.76* (1.70)	.12	.25	1060
Prop. Center care 27-54m	4.10** (1.44)	3.22* (1.49)	3.84** (1.26)	.33	.35	1060
Covariates						
Intercept	97.76*** (1.27)	2.73 (1.36)	51.43*** (2.83)			
Site	***	***	***			
Male	-2.33** (.72)	2.41** (.75)	-.14 (.65)	.50		1060
Black	-6.15*** (1.45)	.96 (1.51)	-2.75* (1.29)	.12		1060

Hispanic	-3.45 (1.66)	-.87 (1.73)	-2.10 (1.47)	.06		1060
Other	1.70 (1.82)	-.71 (1.97)	-1.27 (1.67)	.04		1060
M education	.54** (.21)	-.22 (.22)	.15 (.19)	14.39	2.47	1060
M sensitivity at 6 months	.27 (.43)	-.41 (.44)	-.04 (.38)	.03	.99	1035
M sensitivity missing	.54 (.83)	-1.66 (5.39)	-.31 (4.57)	.03		1060
HOME total at 6 months	1.99*** (.50)	.20 (.52)	1.02* (.44)	.05	.98	1040
HOME total missing	-21.53 (11.99)	2.22 (12.20)	-10.38 (10.38)	.02		1060
M PPVT at 36 months	.19*** (.03)	.12*** (.03)	.16*** (.02)	99.62	18.32	1032
PPVT missing	-1.77 (2.24)	-2.64 (2.66)	-2.04 (2.26)	.03		1060
M depression 6-24m	-.08 (.08)	.01 (.08)	-.07 (.07)	9.17	6.99	1060
M depression 36-54m	-.20** (.07)	-.10 (.07)	-.14* (.06)	9.54	7.48	1060
Income/poverty threshold 6-24m	.20 (.27)	.39 (.27)	.28 (.23)	3.69	2.85	1060
Income/poverty threshold 36-54m	.32 (.25)	-.23 (.26)	.08 (.22)	3.60	2.93	1060
Partnered 6-24m	-2.73 (1.85)	-3.13 (1.97)	-3.19 (1.67)	.86	.32	1060
Partnered 36-54m	-.91 (1.64)	1.53 (1.72)	.60 (1.46)	.83	.34	1060
Separation anxiety at 1 months	-.03 (.03)	-.03 (.03)	-.03 (.03)	69.83	13.13	1048
Separation anxiety missing	.35 (3.51)	-3.17 (3.57)	-1.35 (3.03)	.01		1060
Benefits of M employment at 1 months	-.10 (.13)	.09 (.14)	-.03 (.12)	19.09	3.12	1060
Child temperament at 6 months	.16 (.97)	.38 (1.01)	.29 (.85)	3.17	.40	1041
Child temperament missing	13.54 (13.93)	.01 (14.41)	8.38 (12.24)	.02		1060
M agreeableness at 6 months	.14 (.08)	.09 (.09)	.10 (.07)	46.46	5.21	1034

M extraversion at 6 months	-.04 (.07)	.02 (.07)	-.01 (.06)	42.41	5.76	1034
M neuroticism at 6 months	.20** (.07)	.14* (.07)	.18** (.06)	29.76	7.03	1034
M agree.missing	3.23 (4.55)	-4.05 (4.63)	-.55 (3.93)	.02		1060
Progressive idea for raising kids at 1m	.22* (.11)	.11 (.12)	.18 (.10)	32.84	3.54	1060
R-squared	.43***	.13***	.57***			
Mean (SD) case count	100.0 (11.6) n=1060	7.51 (12.4)	n=1030			

Notes:

*p<.05 ** p<.01 ***p<.001