Alternative Measurements of Poverty in Latent Curve Models of Maltreated Children's Development

Elvon Christopher Lloyd

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Social Work.

Chapel Hill
2008

Approved by:

Richard P. Barth

Mimi S. Chapman

Shenyang Guo

Paul Smokowski

Robert C. MacCallum

Abstract

E. CHRISTOPHER LLOYD: Alternative Measurements of Poverty in Latent Curve Models of Maltreated Children's Development
(Under the direction of Richard P. Barth)

Differing methods of including poverty in analyses of development may produce differing analytic outcomes. Poverty is modeled in a nationally representative sample of maltreated infants using latent curve models, controls for demographic and maltreatment characteristics, and using infant development as the outcome of interest. Poverty was specified as a dichotomous variable (in poverty or not in poverty), a continuous variable (income-to-needs ratio), as low socioeconomic status (SES), as a moderator of developmental predictors, and as being mediated by development predictors for each of four developmental outcomes. Multiple imputation is used to address the problem of missing data. Findings for the complete sample support the use of income-to-needs ratios as the preferred method of measuring poverty based on component and global fit of the model, though effects were generally only found on the intercept factor. The slope factor had few or no predictors, perhaps as a result of relatively small amounts of developmental change in the infants. Some support was found for the more informative mediated and moderated models of poverty as well, and may be of use in the development of interventions to remediate the effects of poverty. Subsamples were created based on gender, membership in a racial minority group, maltreatment type experienced, and type of child welfare placement. In these models, predictors varied compared to each other and the complete sample. Females in foster care and membership in a racial minority were associated with lower scores on the intercept and

negative slopes, respectively. When latent dependent variables were used in the latent curve models, fit and precision of estimates improved while the shape of the trajectories did not change. This was similar to prior research.

Dedications

To the participants in the National Survey of Child and Adolescent Well-Being and to Carter and Conner.

Acknowledgements

I am deeply grateful for the guidance and support of Richard P. Barth and many faculty members at the University of North Carolina at Chapel Hill as well as the patience of my wife, Carrie.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Abbreviations

CDM	Composite Developmental Measure
CFI	Comparative Fit Index
FC	Foster Care
INH	In-Home
LCM	Latent Curve Model
LDV	Latent Dependent Variable
LONGSCAN	Longitudinal Study of Child Abuse and Neglect
LV	Latent Variable
MAR	Missing at Random
MCAR	Missing Completely at Random
MCMC	Markov Chain Monte Carlo
MDV	Manifest Dependent Variable
MI	Multiple Imputation
MV	Manifest Variable
NSCAW	National Survey of Child and Adolescent Well-Being
PSU	Primary Sampling Unit
RMSEA	Root Mean Squared Error of Approximation
SE	Standard Error
SEM	Structural Equation Modeling
SES	Socioeconomic Status
SRMR	Standardized Root Mean Squared Residual
TLI	Tucker-Lewis Index

List of Symbols

□ Measured (or Manifest) Variable

○ Latent Variable

→ Direction of Influence (between variables) or Influence of Error (pointing into a variable)

Chapter 1

The Project in Brief

Infants are at greater risk of maltreatment compared to older children and are more vulnerable to the effects of poverty, maltreatment, and other hazards because of their rapid development across multiple developmental domains (Wulczyn et al., 2005). Data indicate that at least 150,000 infants are affected by maltreatment in the United States each year (ACF, 2005). Many, if not most, experience a deficit in at least one domain of developmental functioning.

Poverty commonly co-occurs with child maltreatment. Maltreated infants involved with child welfare systems, however, may come to stay in foster care or kinship care and consequently are somewhat less likely to reside in poverty at that time (ACF, 2005). None the less, poverty is an all too common characteristic of their home of origin. Like maltreatment (see Commission on Behavioral and Social Science and Education [CBSSE], 1993), no clear consensus exists regarding the conceptualization and resulting operationalization of poverty in studying infant development (McLoyd, 1998).

Poverty may not have a direct effect on child development (McLoyd, 1998; Cicchetti & Lynch, 1995). Rather, poverty may interact with other influences which do have direct causal effects on children's development. For example, the effects of poverty on young children may be mediated through adult behaviors (Garrett, Ng'andu, & Ferron, 1994) such as parental affection or the provision of developmentally stimulating toys. No clear consensus exists regarding the optimal conceptualization and resulting measurement of poverty in studying infant development (Aber, Bennett, Conley, & Ji, 1997; McLoyd, 1998), and investigators may use whatever data are available rather than adhering to a particular conceptual or operational model.

There are few longitudinal studies of maltreated infants' development (Dubowitz, Pappas, Black, & Starr, 2002) and findings about the relationship between poverty and development have been inconsistent in some areas. That chronic poverty is a negative influence on infants' and children's development is not in question (McLoyd, 1998; Guo, 1998; Costello, Compton, Keeler, & Angold, 2003; McLeod & Nonnemaker, 2000). But some longitudinal research (i.e., research using two or more time points) has found a more negative effect on development for poverty experienced early in life (i.e., during infancy or early childhood) compared to poverty experienced later in childhood (Brooks-Gunn and Duncan, 1997; National Institute of Child Health and Human Development [NICHHD] Early Child Care Network, 2005; Teo, Carson, Mathieu, Egeland, & Sroufe, 1996) using a variety of development outcomes. Other research (Guo, 1998; Duncan, Brooks-Gunn, & Klebanov, 1997) has found no stronger effect for early poverty. Interpretation of the existing research is problematic as each study uses differing outcome times (e.g., elementary school or adolescence) and outcome measures (e.g., language development or academic achievement).

At the same time, theories describing and operationalizing antecedents, pathways, and consequences of child maltreatment on development are becoming more complex, incorporating transactional effects between the developing infant and the environment (English, Graham, Litrownik, Everson, & Bangdiwala, 2005). The transactional perspective (Sameroff, 1995) posits that the infant and environment simultaneously influence each other. A transaction differs from an interaction in that the participants cause change in each other in addition to producing an effect, whereas in an interaction, only an effect is produced. Development is conceptualized as the result of these dynamic, continuous transactions between infant and environment. As a result, the experiences provided by the environment are not independent of the infant's actions. This is in contrast to prior developmental models in which the environment is seen as being independent of the developing infant's influence. Little research has been done to validate these newer perspectives, however.

The goal of this project was a better understanding of the relationship of poverty to the trajectories of multiple domains of development in the context of maltreatment. Specifically, differing models of poverty reflecting differing conceptualizations of poverty were evaluated using a common set of data. These models were then re-evaluated in smaller, more homogenous sub-groups to identify and assess differences in gender, race, maltreatment experience, and child welfare placement experience. Finally, an alternative and possibly more valid measure of development was evaluated against a more commonly used, simpler measure of development. Results address not only these issues but whether early poverty has an effect on the resulting developmental trajectories of maltreated, often poor, infants.

Common Data and Methods

A common set of methods was used for data management and statistical analyses. Using an identical, or very similar, set of methods allows for the most valid comparisons among resulting estimates in the sense of comparing 'apples to apples'. The trade-off is that statistical models often have one or more extraneous variables that would probably have otherwise been removed from the model for the sake of parsimony.

The first step in the common methods employed in this project was the decision to attempt to model missing data using a process called *multiple imputation* (MI). Based on earlier similar research (Lloyd, 2007) and the researcher's prior experience with the data set, the data were judged to be appropriate for MI. That is, the pattern of missing data was judged to be at least missing at random (see Allison, 2002, for a detailed typology of missing data) and to occur at rates that would risk bias in resulting estimates if addressed using mean substitution or a similar single imputation strategy.

Multiple imputation requires the researcher to develop a 'model' of the missing data. The model is ideally composed of all variables to be used in subsequent analytic modeling as well as any other available variables that might be potentially informative (i.e., related to one or more of the variables of interest). The researcher then selects and executes a method of imputation (SAS Institute Inc., 2006b; Little & Rubin, 2002; Allison, 2002). Because of the lack of structure among the missing data (i.e., non-monotone), Markov Chain Monte Carlo (MCMC) was the only option available in SAS. In MCMC-based MI, repeated estimates of the parameter of interest are made based on random numbers developed using Markov Chains, which accurately simulate even very complex probability distributions. More simply, MCMC introduces a degree of randomness to the imputations while still reflecting the characteristics of the distribution of the parameter of interest. In so doing, a pre-specified number of data sets were created, each with a single imputation of each missing variable on which standard analyses were performed. The results obtained from each data set were then combined to obtain a single set of results.

As detailed in subsequent chapters, the time intervals between data collections did not coincide with any conceptually useful units of time in the study of child development.

Consequently, the unit of time in the study was changed to years by treating the task as a problem of missing data using methods described by Bollen and Curran (2006). Using MI, developmental data for years of life 2, 3, 4, and 6 were created, and the metric of time was moved from time between waves of data collection to years of life.

The result of MI was ten data sets, each with a differing set of imputations for each missing variable, which were used for all analytic modeling. Ten was chosen based on the work of Rubin (1987) who showed ten imputations were adequate in virtually all circumstances. Standard analytic methods assuming complete data were then used for each data set and the results combined using formulas developed by Rubin (1987) and implemented in Mplus 4.21 (Muthen & Muthen, 1998-2007a).

A common set of analytic methods was developed in addition to a common set of data. Latent Curve Models (LCMs; Bollen & Curran, 2006) were selected to best model developmental trajectories of the NSCAW infant sample. The dependent variables were scores on standardized instruments assessing cognitive, communication, and adaptive behavior development. The mean of these three scores was used to represent overall development. Five commonly-used conceptualizations were then operationalized in LCM methodology. In all analyses described later, an unconditional model was first estimated to assess whether a trajectory could be modeled and, if so, to identify the shape of the trajectory. Then a set of controls and the models of poverty were added as predictors of the identified trajectory. Controls included gender, racial minority status, type and severity of maltreatment, and type of child welfare placement (i.e., in-home or in foster care). Implicitly, age is controlled for by the scoring method used to produce the scale scores used as dependent variables except as noted in chapter four.

Dependent variables were included in the imputation model and imputed dependent variables were used in analysis. While there is some question about introducing bias into parameter estimates when using imputed dependent variables, Allison (2002) reports that analyses with both real and simulated data have failed to demonstrate that this effect occurs in practice. As a result, the dependent variables were retained to maximize sample size, which is important in MI (McKnight, McKnight, Sidani, & Figueredo, 2007).

The common data and common methods were then used to produce results based on the following topics of interest. Chapter two compares selected models of poverty across the four developmental indicators. Chapter three compares overall development for all infants to sub-groupings based on minority status, gender, maltreatment type, and child welfare placement. Chapter four compares latent curve modeling using a single measured variable as the dependent variable to second-order latent curve models having a latent dependent outcome with multiple indicators. Chapter five summarizes and connects key findings.

Chapter 2

Alternative Measurements of Poverty in Latent Curve Models of Maltreated Children's Development

It is well established that poverty has a negative effect on infants' development in maltreated populations as well as the general population (McLoyd, 1998; Guo, 1998; Costello, Compton, Keeler, & Angold, 2003; McLeod & Nonnemaker, 2000). However, the precise definition of poverty used in studies of infant development varies, sometimes substantially. This leads to the question of what effect common conceptualizations and operationalizations of poverty in statistical models have on developmental or other findings.

Poverty is common among families involved in child welfare services. Fifty percent of all families involved with child welfare fall below the poverty line (Administration for Children and Families [ACF], 2005)—substantially higher than the approximately five percent of the general population for 1999 (United States Census Bureau [USCB], 2001). Occurrence among children ages birth to two is slightly more common. Poverty occurrence among children placed out of their home of origin was less common than those remaining in their homes of origin (52% versus 29%), but both percentages remained markedly larger than the general population (ACF, 2005). Other studies have placed estimates as high as 82% (Connell-Carrick & Scannapieco, 2006), albeit with less representative samples.

No clear consensus exists regarding the optimal conceptualization and resulting measurement of poverty in studying infant development (McLoyd, 1998). Often, investigators measure poverty in an *ad hoc* fashion using whatever data are available, sometimes markedly simplifying those measures. Regardless of specific measures employed, there are several common approaches to the formulation of poverty as a predictor of development.

A simple and common method is to use a dichotomous system: either the subject is in poverty or not in poverty. The classification is often based on measures of household size, composition, and income (United States Census Bureau [USCB], 2006) or some similar set of criteria. Tables published annually by the USCB identify a poverty threshold based on household size and number of adults and children residing in it. Households making less than the threshold amount are identified as being in poverty. The resulting measurement is a two-level categorical (i.e., dichotomous) variable.

An alternative to the dichotomous concept of poverty is to conceptualize poverty as a continuum that specifies a distance above or below the poverty threshold of interest. *Income-to-needs ratios* (ITNRs) specify poverty as a ratio of the total household income compared to the threshold established in USCB tables for the year of interest (USCB, 1999). The resulting variable is continuous with a minimum value of zero. Thus, the ratio for a household earning exactly the amount of the poverty line for its composition would be 1:1 or, more commonly, 1.0. The ‘to one’ part of the ratio is understood and, consequently, not normally stated. In 1999 a household of two adults and two children earning an annual income of $20,000 has a poverty threshold of $16,895 (as identified by the USCB table for 1999) and would have an ITNR of 1.18. That same family with an annual income of $60,000 would have an ITNR of 3.55 (using the same 1999 threshold).

Socio-economic status (SES) has been preferred over the USCB poverty line by some researchers. Income, parental education, and parental occupation and prestige are included as these are believed to reflect wealth, knowledge, and access to other aspects of social capital (Bradley & Corwyn, 2002; McLoyd, 1998), all of which may influence infant development. It has been hypothesized to be a broader and more stable indicator of poverty or affluence than household income (Duncan, 1984), given the lack of correlation between social class and income and the often variable nature of household income over even brief time intervals (Duncan, Brooks-Gunn, and Klebanov, 1994). While household income may vary widely depending on the job status of wage earners, indicators of SES such as parental education and social status vary less markedly on a month-to-month basis and are asserted to provide a more consistent measure of the household's typical poverty status.

Indirect models of poverty hypothesize that poverty changes the effects of direct sources of developmental influence. These models offer more evidence of causality compared with dichotomous conceptualizations, SES, and income-to-needs ratios, because they specify how poverty exerts its influence on direct effects as well as what the magnitude of that influence is (McLoyd, 1998; Cicchetti & Lynch, 1995). The effects of poverty on development are hypothesized to be mediated or moderated by variables which do have a direct—and ideally causal—effect on children's development.

Maltreated Infant Development

There are few longitudinal studies of maltreated infants' development (Dubowitz, Pappas, Black, & Starr, 2002). Despite the stated importance of early childhood experience, especially of infancy, studies of psychosocial development often do not begin until the child is pre-school age or older (e.g., Longitudinal Studies of Child Abuse and Neglect [LONGSCAN] which begins at age four at all but one site). Theories describing antecedents, pathways, and consequences of child maltreatment on development are becoming more complex, incorporating ecological and transactional effects (English, Graham, Litrownik, Everson, & Bangdiwala, 2005), but little research has been done to validate them.

Exposure to maltreatment during infancy, so-called *early poverty*, is believed to be associated with subsequent negative developmental outcomes (National Institute of Child Health and Human Development [NICHHD] Early Child Care Network, 2005). Overall academic achievement has been found to be lower (Teo, Carson, Mathieu, Egeland, & Sroufe, 1996) and neglect is associated with lower cognitive and language development (Gowen, 1993). Prior research using data from the National Survey of Child and Adolescent Well-Being (NSCAW) found declines in all three developmental domains measured (cognitive, language, and adaptive behaviors) between the baseline and 18 month follow-up assessment (NSCAW Research Group, 2005). Brooks-Gunn and Duncan (1997), after aggregating a few large, longitudinal studies which included timing and duration of poverty, conclude that early poverty (birth through elementary school age) may have a different effect than later poverty in that children experiencing poverty early in life seem to perform more poorly on academic outcomes during adolescence. Other research is more equivocal. Neglect, often associated with poverty, experienced prior to age three did not predict developmental outcomes at age three or five in children at risk for health and developmental problems (Dubowitz, Pappas, Black & Starr, 2002). Other research, that focuses more on the maltreatment experience, links early maltreatment and subsequent behavior problems which may negatively affect development (Trickett, 1997; Lansford et al., 2006; Herrenkohl, Herrenkohl, Egolf, & Ping, 1991).

Reviews of the biology of psychosocial development clearly point to the important role this time period plays in overall development of the child. That is, experiences in infancy and early childhood play a role in proximal developmental outcomes as well as more distal ones (De Bellis, 2005; Dawson, Ashman, & Carver, 2000). As research into the effects of maltreatment and poverty has shifted into the biological realm, new consequences of maltreatment are being identified and investigated. For example, a recent study linked maltreatment to the development of problems with inflammatory diseases in adulthood, even after controlling for SES, gender, birth weight, heart disease, and other possible influences (Danese, Pariante, Caspi, Taylor, & Poulton, 2007).

As a result of these findings, there are three objectives:

1. To describe a population of infants involved with child welfare
2. To model developmental change using constructs for cognitive, communication, and adaptive behavior development using latent trajectory analysis (LTA)
    a. To identify predictors of developmental trajectories
    b. To compare how differing conceptualizations of poverty explain developmental change in a large, representative sample of maltreated infants

METHOD

Sample

The sample was obtained from the National Survey of Child and Adolescent Well-Being (NSCAW), a national probability sample of children entering child welfare services (see ACF, 2005, for a complete description of the sampling design). For the infants, four full waves of data collection were completed at baseline and approximately 18, 36, and 66 months post-baseline. An additional, reduced wave of data was collected, primarily over the telephone from caregivers, at 12 months post-baseline (Research Triangle Institute [RTI], 2007).

All infants who were less than 13 months old at the baseline data collection were included in analyses yielding an unweighted sample size of 1,196 infants. The age limit was based on prior work done using LONGSCAN data in which it was argued that children entering child welfare services between zero and 18 months of age represent a common developmental group (English, Graham, Litrownik, Everson & Bangdiwala, 2005). More practically, NSCAW data collection at the 66 month follow-up only included children up to 12 months of age at baseline.

Legal substantiation of maltreatment was not used as a criterion for inclusion in the present study. Not all participants had a legal finding that maltreatment had taken place (i.e., substantiation of maltreatment). Herrenkohl (2005) argues that defining maltreatment only by substantiation probably understates the actual level of maltreatment based on analyses showing no differences on ten developmental measures administered to maltreated children by Hussey et al. (2005). The term ‘maltreated’ is used for the sake of efficiency and brevity, though in some cases the maltreatment was not legally substantiated.

Measures

Four dependent variables were used. Each of the three NSCAW developmental domains, cognitive, adaptive, and communication, was included while the fourth measure was an average developmental score. All scores were standardized to a mean of 100 and a standard deviation of 15 to facilitate estimation and comparison.

Vineland Adaptive Behavior Scales Screener – Daily Living is a brief instrument used to screen children for problems in the domain of adaptive behavior and daily living skills. The Vineland Screener (Sparrow, Carter, & Cicchetti, 1993) is completed by a caregiver or other person knowledgeable about the child. The version for child ages zero to two was used at baseline with the three to five year old version used at subsequent waves as the cohort aged. The Vineland Screener strongly correlates (r = .87 to .98) with the full Vineland instrument.

Pre-School Language Scales (PLS). The PLS-3 was used to assess the developmental domain of language. It produces two sub-scales, expressive communication and auditory comprehension, and a total scale in children younger than six years old (Zimmerman, Steiner, & Pond, 1992). The scores are based on observations of the child. Interrater reliability is .98.

Battelle Developmental Inventory and Screening Test (BDI). The BDI was used to assess the developmental domain of cognitive development in children younger than five years old. It produces scores for five sub-domains and a total score. It is administered by an examiner. Despite the fact that the BDI does not require training for the administrator, it has a test-retest reliability of greater than .90 in most domains and in the total score (Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984).

Kaufman Brief Intelligence Test (K-BIT). The K-BIT was used to assess cognitive development in children older than four years. The K-BIT assess four sub-domains as well as provides a total score. It is a self-administered, paper and pencil instrument. The test-retest reliability of the K-BIT varies by construct considered, but ranges from .74 to .95 (Kaufman & Kaufman, 1990).

Composite Developmental Variable. A variable was constructed to measure overall development of the child. The standardized scores for the cognitive, language, and communication domains were averaged to yield a single variable. It should be clearly understood, however, that development across the three domains of development in NSCAW is not parallel. As will be observed in subsequent analyses, both the scores and trajectories of the developmental outcomes vary across the domains. Given this heterogeneity of data, the meaning of results obtained using the CDM should be carefully interpreted, because it is unlikely to clearly apply to any specific domain of development that was used in its construction.

The following are independent variables used in the imputation model or the analytic model to predict developmental scores.

Demographics. Age, race/ethnicity, and gender variables were constructed. Age is measured in months. Race/ethnicity was initially conceptualized as having four levels, but estimation requirements forced the use of a dichotomous variable, minority or non-minority with non-minority serving as the reference. Gender was also coded as a dichotomous variable with female as the reference level.

HOME-SF Scales. Home Observation for Measurement of the Environment (Short Form) was used to assess emotional nurturing as well as cognitive stimulation in the infant's current caregiving environment (Bradley & Caldwell, 1984). Internal consistency for the total scales is .89 with a median of .74 for the subscales. Stability for the total scales is r=.62.

A newer description of scoring the HOME-SF is available and was used (Bradley, Corwyn, McAdoo, & Coll, 2001).

Maltreatment. Maltreatment data was collected from the child welfare workers and case data. Several variables were created based on the dimensions of maltreatment suggested to be important (English, Bangdiwala, & Runyan, 2005). First, the most serious type of maltreatment was identified. To facilitate estimation, the most serious type of maltreatment was dummy-coded as neglect or 'other' maltreatment with abuse (primarily physical or emotional) serving as the reference level. Second, severity of harm, as judged by the child welfare worker and rated as none, low, moderate, or severe, was coded as 1 to 4, respectively. Third, a dichotomous variable was created to indicate whether one or more than one type of maltreatment had occurred. Some children experienced multiple types of maltreatment, in which case the most serious was coded as the primary type (used to identify type in this analysis). Age of first episode of maltreatment was not included because all children in the study are defined by being in a common age group.

Socioeconomic Status Characteristics. Three variables represented the socioeconomic status of the infant's home of origin or of the foster home in which the child resided at baseline. A key concern is that SES lacks a consistent definition, both conceptually and operationally (Mueller & Parcel, 1981). In the analyses detailed below, SES is defined using three variables to represent the social capital model of SES (Bradley & Corwyn, 2002; Coleman, 1988).

Household Income. Annual household income is a scale of 1 to 11. Each number is a 5,000 dollar increment (e.g., 2 represents 5,001-10,000 dollars) while 11 represents any income over 50,000 dollars.

Index of Social Capital Indicators. Six indicators were used to construct an index indicating SES. Items included a primary caregiver having had at least some education beyond high school, being part of a first generation immigrant family, having a low-skill or unskilled type of job or having held such a job recently if unemployed, being unemployed, receiving one or more types of social assistance in the household, and being a single (i.e., not married or in a stable romantic relationship) caregiver. Organizing these items into a scale, rather than retaining them as individual independent variables, was necessary to avoid model estimation problems as well as the tendency for more complex structural equation models to exhibit a better fit than simpler models (Preacher, 2003). There is no generally accepted means of specifying SES in statistical models (Bradley & Corwyn, 2002).

Neighborhood Safety. Caregivers were asked a series of questions about their perceptions of criminal activity in their neighborhoods. Their responses were aggregated into a brief scale measuring caregiver perception of neighborhood safety.

Poverty. The household having total income below the official poverty line for that year and household composition is included as a dichotomous variable with not in poverty as the reference. An income-to-needs ratio was also created using household income divided by the dollar value of the poverty line for the household composition and year (USCB, 2006). Because household income was indicated as being in a given range of values, the midpoint of each range was used.

Setting. In child welfare, setting is the type of residence the child resides in. It may be in-home with nuclear family of origin, in kinship care with a relative or close family friend, in foster care, or in residential care of some sort such as group home or residential treatment facility. The data used was collected at baseline and most likely reflects where the child was placed early in the case. The placement may change as the case proceeds, but the setting variable identifies where the child was initially placed. Setting was collapsed to in-home (INH) or foster care (FC) to facilitate estimation with in-home placement as the reference. Kinship foster care is considered a foster care placement despite placement with a relative being considered in-home under the TANF program.

Child Health Scale. This variable was created using a current caregiver rating of the child's health from 1 (poor) to 5 (excellent).

Caregiver Health and Mental Health Problem. Caregivers were administered the SF-12, which assess general health and mental health (Ware, Kosinski, & Keller, 1998). Two variables were created—one each for health and mental health—using the standardized results.

Service Receipt. Both child welfare workers and caregivers identified services provided to the child. These data were used to create dichotomous variables indicating a need for developmental services, having an Individual Family Service Plan (IFSP), and receipt of services based on the worker-reported information.

Number of Children in the Household. The number of persons under age 18 residing in the household was used only during imputation modeling.

NSCAW Analysis Weight. This variable is unique for each case and is designed to produce results which are representative of almost all American children involved with child welfare systems. The details of its creation are given by the Data File Users Manual (RTI, 2007).

Stratum. This indicates which of the nine strata in NSCAW the data came from. Each of the first eight strata is a single large population state, while the ninth stratum includes data from several smaller states. It is part of the complex survey sample design data necessary to correctly account for clustering in NSCAW data (RTI, 2007).

NSCAWPSU. This indicates which primary sampling unit (PSU) the data came from. Each PSU was typically a child welfare agency serving a geographical locale (normally a county). It is also part of the complex survey sample data necessary to correctly account for clustering in NSCAW data (RTI, 2007).

RESULTS

Analysis

Two distinct sets of methods were required for completion of analyses. First, multiple imputation was employed to address the problem of missing data and a problematic metric of time. Because expectation-maximization (EM) analytic algorithms are not available for data which include weighting and complex sample design variables (i.e., stratum and PSU), multiple imputation was the only model-based option available. Second, Latent Curve Models (LCMs; Bollen & Curran, 2005) were estimated for trajectories of all four developmental domains as well as the composite developmental indicator.

Developmental data gathered at baseline was not used for two reasons. First, data collectors failed to correctly gather data on many occasions, leaving questions about the validity of the data gathered (RTI, 2007). Because brief screening versions of the instruments were used, even missing responses on a few questions were sufficient to cast the final scores into doubt. Second, assessment of infant development requires specialized expertise. Some have questioned the appropriateness of using inexpert, though not untrained, data collectors in the gathering of infant developmental data (Barth, Scarborough, Lloyd, & Casanueva, 2007). As a result of these two related concerns, infant developmental data gathered at baseline was omitted from both imputation and analysis.

Imputation Modeling

Prior to attempting multiple imputation (MI), missing-ness and ignore-ability of missing data must be determined by the researcher (Allison, 2002; Little & Rubin, 2002). There are no mathematical or other formal tests which may be applied (Allison, 2002). Missing data for each variable was plotted and compared to other variables. Rates varied between less than one percent and 74 percent. Missing rates were highest for the developmental measures and lowest among the independent variables. No mechanism causal of missing-ness was identified nor does the researcher know of any causal mechanism in NSCAW. As a result of these facts and research experience, missing data were judged to be at least missing-at-random (MAR) and ignorable.

The metric of time to be used in the analyses had to be created. In NSCAW, data were collected at baseline as well as 18 months, 36 months, and 66 months post-baseline. In the field of child development, a more useful metric of time is the child's age in years. Bollen and Curran (2006) describe how to change the unit of time from intervals of data collection to years of chronological age using a direct maximum likelihood estimator, but they state that MI can be used with identical results.

Using MI brings with it several potential limitations. First, the data may not have been MAR despite the researcher's best efforts. Second, it is not possible to fully account for the clustering in the sample design, though recommendations from Allison (2002) were followed. Finally, the imputation model may lack some degree of validity. Guidance from key texts (Little & Rubin, 2002; Rubin, 1987; Allison, 2002) and more recent papers (Croy & Novins, 2005) was minimal when theoretically important variables show low covariances.

Multiple imputation was completed in SAS 9.0.1 using Proc MI (SAS Institute, 2006a) using Markov Chain Monte Carlo (MCMC). In keeping with the recommendations of experts (SAS Institute, 2006b; Little & Rubin, 2002), data from the dependent variables as well as the independent variables to be used in the analytic models were included in the imputation model and variables dropped from analytic modeling were retained in imputation modeling. A complete set of developmental values was possible only for years of life 2, 3, 4, and 6 years old, so final, imputed data sets contain developmental values for those years as well as the independent variables of interest.

Several variables used in analytic modeling were simple combinations of other variables and were calculated after MI was completed. The composite developmental variable was the mean of the three standardized developmental indicators. Variables moderated by poverty were created by multiplying the variable of interest by the income-to-needs ratio.

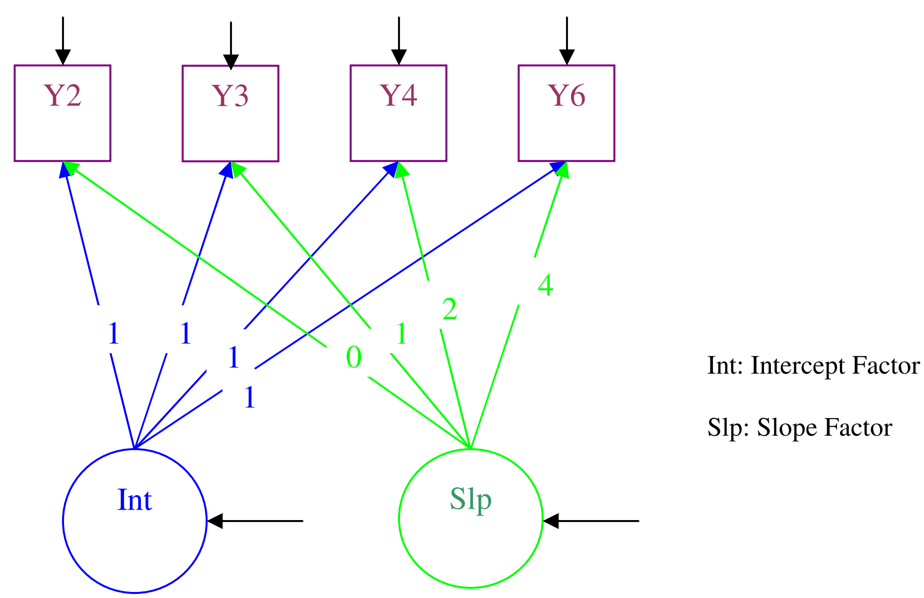
SAS Proc MI executed the imputation model without errors or warnings, and convergences were achieved in fewer than 30 iterations of the Proc MI algorithms. Ten imputed data sets were created. The computed variables were added using SPSS 12.0.0 (SPSS Inc., 2003), and the data sets were separated into individual files and converted to a format usable by Mplus.

Analytic Modeling

All analyses were executed using Latent Curve Modeling (LCM: Bollen & Curran, 2006). This methodology is based in structural equation modeling (SEM: Bollen, 1989) and makes calculation of fit indices possible. Models were estimated using Mplus version 4.21 (Muthen & Muthen, 1998-2007a), which allows combination of the results for each of the ten data sets created during the imputation step. A robust maximum likelihood estimator was used (see Muthen & Muthen, 1998-2007b) with weighting, stratification, and clustering data included.

Modeling was completed in a stepwise fashion. Initially, an unconditional model was fit to the data (see Figure 2.1). The goal of this step was to determine whether a latent curve model was appropriate for the data in question. The dependent variables were developmental scale scores at each time point. If necessary, a non-linear model was attempted. After a model was identified as fitting the data, a set of controls were added and each of the five substantive models of poverty being tested were fit to the data (see Figure 2.2, which has the measurement model omitted for the sake of brevity).

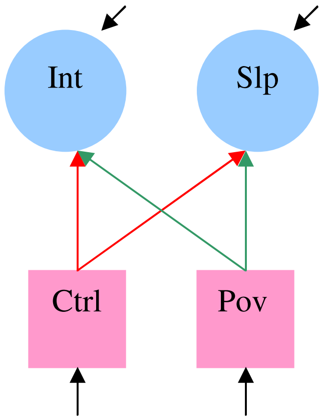
Figure 2.1: Unconditional Latent Curve Model[1]



[1] Circles represent latent variables and squares represent manifest variables in path models.

Figure 2.2: Modeling Effects of Poverty on the Slope and Intercept Factors in LCMs

Simple and Ratio Poverty



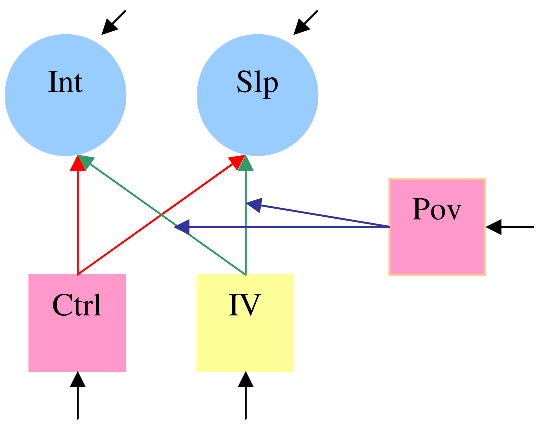
IV: Independent Variables

Ctrl: Control Variables

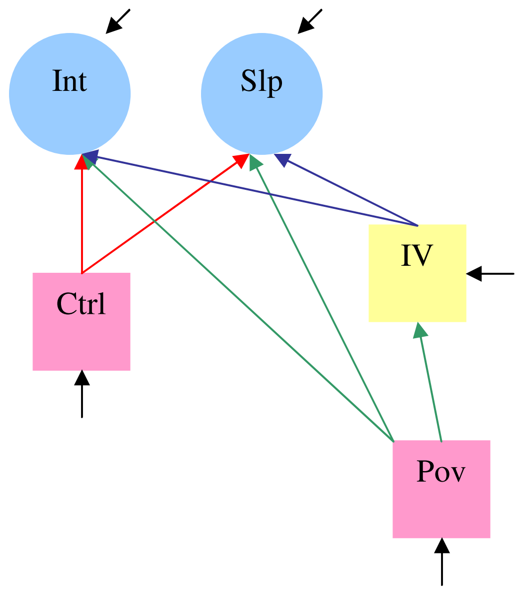
Pov: Poverty Variable

SES: Socio-Economic Status Variable

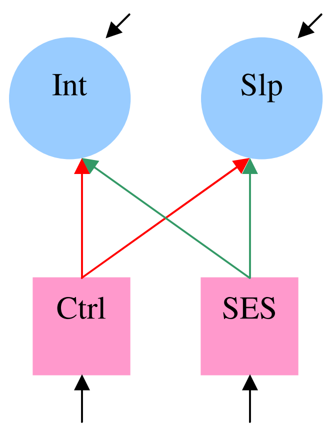
Moderated Poverty



Mediated Poverty



Socioeconomic Status



All models included controls for case and demographic characteristics. Maltreatment was modeled by including type, severity, and experiencing more than a single identified type. The reference for type of maltreatment was abuse. The reference for gender was female. The reference for race was non-minority. Finally, the reference for child welfare placement type was in-home (INH). Kinship care was considered a foster care (FC) placement in most circumstances.

Evaluation of models is based on several criteria. First, the model had to produce acceptable parameter estimates. Ideally, no improper solutions—negative variances for example— should be reached, though they may be produced, even when the model is a good fit to the data (Bollen & Curran, 2006). In a few instances, noted in the results, a small, non-significant (from zero) negative residual variance was produced and tolerated.

A second important indicator of fit is the presence of reasonable estimates. That is, the magnitude and sign of the estimates should be appropriate to the data in use (Bollen & Curran, 2006; Schumaker & Lomax, 2004). In planning, this was to be evaluated in part using confidence intervals. Due to the use of MI, however, this was not possible. Statistical significance was used instead, though it provides less information.

Finally, as with most SEM-based results, fit indices are produced. Mplus 4.21 (Muthen & Muthen, 1998-2007a) produces four fit indices; Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), Standardized Root Mean Square Residual (SRMR; Joreskog & Sorbom, 1981; Bentler, 1995), and Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993; Steiger & Lind, 1980). Using the CFI and TLI, greater than .94 indicates a good fit while a result greater then .90 indicates acceptable fit. When using the SRMR and RMSEA, less than .06 indicates a good fit while less than .09 or .11, varying slightly by author, indicates an acceptable fit (Bollen & Curran, 2006; Garson, 2007; MacCallum, 2005; Hu & Bentler; 1999; Schumacker & Lomax, 2004; Bollen, 1989). Because of the use of MI, results reported below are given as means and standard errors of the fit indices.

Fit indices may offer differing assessments of fit. Each fit index uses a differing conceptual and mathematical definition of what 'good fit' is. Preference is given to the RMSEA and the SRMR, which do not depend on substantively meaningless null models for comparison. As has been shown elsewhere (MacCallum, 2005; Garson, 2007; Rigdon, 1996), their meaning is substantively clear and readily interpretable compared with those obtained using null model comparisons. The TLI and CFI are reported to fully disclose findings and permit alternative interpretations.

Fit indices may also understate the fit of a model to the data in the context of latent curve modeling. The form (e.g., linear, curvilinear) chosen for the latent curves is not intended to be a perfect fit to all data. It is intended only to be a good approximation of that data. This lack of perfect global fit (as opposed to component fit), despite being intentional or at least acknowledged, often results in understated fit indices (Coffman & Millsap, 2006). Global fit may be considered with *component fit* (Bollen, 1989), which is how well the individual substantive parameters fit as assessed by an appropriate sign and a substantive magnitude of the independent variables.

Findings

Each developmental measure had six unique models completed using ten common imputed data sets. Results are reported in combined form. Parameter estimates were combined by the software using appropriate formulae as described previously. Fit indices are reported as means and, when necessary, standard errors.

Key demographic and case characteristic variables were described using frequency counts (Table 2.1). These counts are based on the un-weighted, un-imputed data. A substantial majority of infants (68%) were neglected. Despite the generally low developmental scores on standardized instruments, few infants (N=247) were judged to be in need of developmental services by their child welfare workers. Approximately one-third were in poverty, regardless of whether poverty status or an income-to-needs ratio was used.

Table 2.1: Distribution of Infant Characteristics at Baseline

Age (Months)	*N*
0-3	162
4-6	471
7-9	317
10-12	246
Gender	
Male	619
Female	577
Race	
Minority	780
Majority	410
Child Welfare Placement	
In-Home	760
Foster care	436
Poverty Status	
In Poverty	470
Not in Poverty	620
Income-to-Needs Ratios	
<1	414
1-1.99	291
2-2.99	210
3+	175
Primary Type of Maltreatment	
Abuse	266
Neglect	730
Other	89
Number of Types Listed	
One	779
Two or More	320
Severity of Primary Maltreatment	
None/Minor	297
Mild	252
Moderate	279
Severe	260

Developmental scores are described by mean, median, and standard error. Results were obtained using un-weighted, un-imputed data and are presented in Table 2.2.

Table 2.2: Characteristics of Infants' Developmental Scores

	Year 2	Year 3	Year 4	Year 6
Adaptive				
Mean (SE)	86.8 (16.1)	88.9 (17.5)	95.4 (17.9)	83.1 (18.3)
Median	86	89	97	82
Cognitive				
Mean (SE)	88.5 (22.6)	85.8 (19.1)	85.2 (16.1)	92.1 (13.3)
Median	87	82	82	92
Communication				
Mean (SE)	84.8 (19.1)	87.1 (20.3)	86.2 (18.4)	93.6 (19.1)
Median	82	85	85	95

Sample sizes vary from 625-757.

Composite Development Measure

Table 2.3: Mean and Variance for the Slope and Intercept Factors of the Unconditional CDM Model

	Int Mean	**Int Var.**	**Slp Mean**	**Slp Var.**
Estimate (SE)	87.24 (0.76)	91.96 (14.27)	0.60 (0.29)	4.80 (1.99)

Table 2.4: Sample Estimates for the Dependent Variables in the Unconditional CDM Model

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	86.7 (14.1)	87.3 (14.1)	89.0 (12.9)	89.4 (12.7)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 2.5: Variance Estimates and R-Squared Values for the Dependent Variables in the Unconditional CDM Model

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var.	107	101	62	39
R-Squares	.463	.471	.613	.786

Table 2.6: Fit Indices for the Unconditional CDM Model

	CFI	**TLI**	**SRMR**	**RMSEA**

Estimate (SE)	.993 (.006)	.992 (.008)	.064 (.026)	.022 (.014)

The composite developmental measure (CDM) was successfully estimated using an unconditional linear LCM (Tables 2.3, 2.4, and 2.5). The mean (standard error) of the intercept is 87.24 (.76) and the slope is .59 (.29). Both parameters are statistically significant, indicating maltreated children in their second year of life are nearly a standard deviation below the normative mean, but they are making small developmental gains as well. In addition, some variability exists in the data as well. The variance of the intercept factor was 91.96 (14.27) while the variance of the slope factor was 4.80 (1.99). Residual variances and r-square values for the CDM at each time point (Table 2.5) indicate the model accounts for an increasingly large amount of the total variance.

The model fit appears acceptable (Table 2.6). The CFI and TLI are .993 (.006) and .992 (.008), respectively. The RMSEA is .022 (.014) and the SRMR is .064 (.026). All values are within the suggested values for good or acceptable fit. Parameter estimates have expectable signs. Expected and observed proportions of values for all four fit indices are similar, further suggesting an acceptable fit of the model to the data.

Poverty and Development

Assessing the fit of a model to the data is based on several criteria including fit indices, parameter estimates, and the presence of improper solutions (Bollen & Curran, 2006; Bollen, 1989). Models of poverty in maltreated children were first assessed using fit indices (see Table 2.7). Results suggest the mediated poverty model was a poor fit, especially when compared with the other four models, all of which produced near-acceptable or acceptable results.

Table 2.7: Fit of Poverty Models

	CFI	TLI	RMSEA	SRMR
Simple	.955 (.021)	.919 (.038)	.034 (.009)	.039 (.007)
Ratio	.947 (.020)	.905 (.036)	.037 (.008)	.040 (.007)
Mediated	.794 (.010)	.707 (.014)	.045 (.002)	.065 (.003)
Moderated	.933 (.017)	.875 (.031)	.038 (005)	.034 (.006)
SES	.944 (.016)	.898 (.030)	.035 (.007)	.036 (.006)

A second test of model fit is the presence of statistically and substantively significant parameter estimates. All statistically significant estimates of predictors of either the slope or the intercept factors are presented in Table 2.8. Scores marked with a ^ are significant at the $p<.11$, sometimes termed a trend or trending towards significance.

All models produced at least several significant predictors of the intercept. Two predictors of intercept appear in all five models. Males were between four and five points lower compared to females in four of the five models. Infants who experienced multiple types of maltreatment were two to three points lower compared to those who experienced only a single type of maltreatment in all five models. Gender had smaller standard errors relative to its estimate compared to multiple types of maltreatment. Poverty, as either a simple dichotomous variable or an income-to-needs ratio variable, has a substantively and statistically significant effect on development and a relatively small standard error when included. Income also produced a statistically significant result, although the parameter estimate is small compared with its standard error.

In the mediated poverty model, income-to-needs ratio was a predictor of parent-related variables, but only child-related variables directly affected developmental scores. In addition, income-to-needs ratio continued to have a substantive and statistically meaningful effect directly on the developmental scores. The parent and child variables moderated by poverty did produce two health-related effects, but these were not substantively meaningful.

No model produced poverty-related predictors of the model's slope. Placement into a foster care (FC) living situation produced statistically significant effects, but the parameter size is not particularly large, ranging from 0.73 to 0.91. The model of SES also produced a positive effect on slope for minority status.

Table 2.8: Predictor Estimates of Poverty Models

Simple (Int)	Estimate (SE)
2 Types	-2.62 (1.32)
Male	-4.35 (1.49)
Poverty	-4.02 (1.54)
Simple (Slp)	
FC	0.73 (0.45)^

Ratio (Int)	Estimate (SE)
2 Types	-2.29 (1.37)^
Male	-4.45 (1.48)
FC	-2.99 (1.77)^
Ratio	1.71 (0.56)
Ratio (Slp)	
FC	0.73 (0.45)^

Mediated (Int)	Estimate (SE)
2 Types	-2.43 (1.38) ^
Male	-4.80 (1.51)
Ratio	1.49 (0.57)
Child Health	1.16 (.061) ^
Cog. Stim.	4.68 (2.64) ^
Mat. Warmth	6.02 (3.26) ^
Mediated (Slp)	
FC	0.91 (0.44)
Mediated(Rto)	
Parent Health	0.93 (0.67)
Parent MH	1.34 (0.49)

Moderated(Int)	Estimate (SE)
2 Types	-2.38 (1.34)
Male	-4.68 (1.50)
Mod Ch. Health	0.58 (0.31)
Mod Prt Health	-0.06 (0.04)
Moderated(Slp)	
FC	0.81 (0.46)

SES (Int)	Estimate (SE)
FC	-3.16 (1.90)
2 Types	-2.52 (1.35)
Male	-4.32 (1.36)
Income	0.45 (0.24)
SES (Slp)	
FC	-0.81 (0.46)
Minority	0.98 (0.52)

The amount of variance for each model accounts for varies by model (see Table 2.9). The simple and ratio models of poverty produced nearly identical results for both intercept and slope, .114. As might be expected, large or more complex models produced larger r-square values because they have more pathways, .163 and .206 for mediated poverty and .161 and .146 for moderated poverty models. Socio-economic status produced mixed results having a small intercept r-square, .102, but a relatively moderately sized slope r-square, .147.

Table 2.9: R-Squared Values for Poverty Models

	Int R-Sqr	**Slp R-Sqr**
Simple	.114	.113
Ratio	.114	.115
Mediated	.163	.206
Moderated	.161	.146
SES	.102	.147

Modeling Specific Domains of Development

Development may not have proceeded along a common trajectory for all developmental domains. As seen in the unconditional model, significant heterogeneity may exist in the data. As a result, modeling specific domains of development may produce differing results compared to the CDM.

Cognitive Development

Cognitive development proceeded along a non-linear trajectory. While the paths from the intercept remain fixed at one, the paths from the slope are freely estimated after fixing two—either the first two paths or the first and the last paths—as a reference (see Bollen & Curran, 2006). The first and last time points were fixed to zero and one, respectively. The second and third time points were freely estimated, and the results should be interpreted as change relative to the net change between Year 2 and Year 6 (Bollen & Curran, 2006). Fit was assessed using the criteria outlined above.

The specific trajectory appears to be curved. The infants' cognitive development falls from Year 2 to Year 4 then rises sharply at Year 6. This is indicated by the data presented in Table 2.11. While the mean scores at the initial and final time points are similar, the intervening mean scores are lower. Further evidence is found in the estimated slope paths which indicated a negative or downward slope. Large standard errors of the path estimates suggest marked variation exists and explain why the estimates do not test as significantly different from zero.

Figure 2.10: Mean and Variance for the Intercept and Slope Factors of the Unconditional Model of Cognitive Development

	Int Mean	**Int Var.**	**Slp Mean**	**Slp Var.**
Estimate (std err)	88.71 (1.87)	60.76 (20.85)	2.58 (1.98)	Heywood

Table 2.11: Sample Estimates of the Dependent Variables in the Unconditional Model of Cognitive Development

	Year 2	**Year3**	**Year4**	**Year 6**
Mean Score	88.5 (22.6)	85.8 (19.1)	85.2 (16.1)	92.1 (13.3)
Path Estimate	0 (fixed)	-2.63 (8.53)	-2.99 (9.86)	1 (fixed)

Fit indicators are not consistent (Tables 2.13). While fit indices are acceptable, the standard errors are large and that indicates the model probably does not fit all of the data sets well. Also, the model accounts for much less of the variance in the data compared to the models of other developmental indicators. In addition, this model, as well as the subsequent conditional model, produced an improper solution for either the estimate of the slope's variance or residual variance. Despite this information that suggests a somewhat less than good fit of the model, other modeling attempts proved more problematic. Either a model could not be estimated for all ten data sets or indicators of fit were well beyond generally accepted standards for good fit.

Table 2.12: Residual Variances and R-Square Values of the Unconditional Model of Cognitive Development

	Year 2	Year 3	Year 4	Year 6
Residual Var.	435	276	197	179
R-Squares	.123	.217	.274	.069

Table 2.13: Fit Indices of the Unconditional Model of Cognitive Development

	CFI	TLI	SRMR	RMSEA
Estimate (std err)	.979 (.024)	.963 (.054)	.026 (.022)	.050 (.032)

The income-to-needs ratio model of poverty (Tables 2.14 and 2.15) produced the most meaningful results. Children with a greater income-to-needs ratio started out with somewhat larger developmental scores. In addition, infants in a racial minority started out over three points lower than their non-minority peers, even after accounting for the effects of poverty, maltreatment, and the other variables outlined previously. While no variables predicted the slope factor, two influences—minority status and having multiple types of maltreatment—trended (i.e., close to the $p<.05$ threshold) towards being significant ($p=.12$). If they were indeed true differences, minority children lost 2.9 points per time point compared to their non-minority peers while those with two types of maltreatment lost 2.1 points per time point.

Table 2.14: Fit Indices for the Income-to-Needs Ratio Poverty Model for Cognitive Development

	CFI	TLI	RMSEA	SRMR
Ratio	.912 (.033)	.823 (.066)	.032 (.006)	.036 (.006)

Table 2.15: Predictor Estimates for the Income-to-Needs Ratio Poverty Model for Cognitive Development

	Estimate (SE)

Ratio (Int)	1.64 (0.75)
Minority (Int)	-3.17^ (1.92)
Int R-Square	.156
Slp R-Square	Heywood

Other models of poverty in cognitive development were rejected. The simple model of poverty produced slightly better fit indices but failed to produce significant parameter estimates indicating poorer component fit of the model (Bollen, 1989). Mediated and moderated poverty as well as SES models produced unacceptably poor fit and few or no significant parameter estimates.

Communication

It proved difficult to fit a latent curve model to explain the measure of child communication. A linear and several non-linear models were fit to the data. Non-linear models failed to produce estimates for all ten data sets and were discarded. The linear model proved to be the best model but was not a good fit to the data. As is suggested by the data in Tables 2.16-2.19, developmental scores plateau during the third and fourth years of life—they are not significantly different means—and this may be the source of the less than good fit observed.

Figure 2.16: Mean and Variance for the Intercept and Slope Factors of the Unconditional Model of Communication Development

	Int Mean	**Int Var.**	**Slp Mean**	**Slp Var.**
Estimate (std err)	85.07 (1.38)	181.52 (38.31)	2.05 (0.53)	12.11 (4.47)

Table 2.17: Sample Estimates of the Dependent Variables in the Unconditional Model of Communication Development

	Year 2	**Year3**	**Year4**	**Year 6**
Mean Score	84.8 (19.1)	87.1 (20.3)	86.2 (18.4)	93.6 (19.1)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 2.18: Residual Variances and R-Square Values of the Unconditional Model of Communication Development

	Year 2	Year 3	Year 4	Year 6
Residual Var.	180	249	174	90
R-Squares	.504	.410	.521	.765

Table 2.19: Fit Indices for the Unconditional Model of Communication Development

	CFI	TLI	RMSEA	SRMR
Estimate (std err)	.834 (.068)	.800 (.082)	.105 (.018)	.078 (.024)

No conditional model of communication development emerged as definitively superior. Both the simple and ratio models demonstrated very similar characteristics and none of the other three models produced better fit, especially component fit. Because the ratio model provides more information than the simple model of poverty, it is the preferred model of poverty in communication development among maltreated infants (Table 2.21).

While the CFI and TLI indicate poor fit, the RMSEA and SRMR suggest a modestly good fit (Table 2.20). The differences are probably due to the method by which each indicator is computed. The RMSEA and SRMR do not compare the model in question to a null model. The SRMR reflects the average deviation of the observed data from the model-implied variance/covariance structure. The RMSEA is the chi-square discrepancy per degree of freedom. The CFI and TLI rely on comparing the model to a null model, which often has no substantive meaning and simply provides a reference against which to consider the model chi-square (Garson, 2007; Bollen & Curran, 2006; Bollen, 1989). As a result, greater weight is given to the RMSEA and SRMR when overall model fit is considered.

As in the model of cognitive development, component fit is not particularly strong. Only two parameters demonstrate significance for the intercept (ratio and minority status) and no slope predictors are significant. The standard errors for the ratio variable are relatively small and the effect is substantive. Maltreated infants with an income-to-needs ratio of 3:1

start nearly seven points or approximately half a standard deviation higher than their peers below the poverty line (less than 1:1). Minority infants who are maltreated had a deficit of 3.63 points compared to their non-minority peers on the intercept factor.

Table 2.20: Fit Indices for the Income-to-Needs Ratio Poverty Model for Communication Development

	CFI	**TLI**	**RMSEA**	**SRMR**
Ratio	.821 (.048)	.676 (.087)	.060 (.008)	.037 (.009)

Table 2.21: Predictor Estimates for the Income-to-Needs Ratio Poverty Model for Communication Development

	Estimate (SE)
Ratio (Int)	2.31 (0.82)
Minority (Int)	-3.63^ (2.05)
Int R-Square	.089
Slp R-Square	.055

Adaptive Behavior

Adaptive behavior proceeded along a non-linear trajectory. In the unconditional LCM, the final model fixed the first two time points as a reference, then freely estimated the final two time points (Table 2.23). Fit was assessed as discussed previously (Tables 2.24 and 2.25). The resulting trajectory is curved in that scores increase steadily during the second, third, and fourth years of life, but drop off sharply at the sixth year of life. This sharp drop coincides with a change in version of the VABS instrument as children move from the 3 – 5 year-old version and into the 6 – 10 year old version. It may have been that the two instruments are not tapping the same construct (i.e., validity problems). Alternatively, it may be that a substantial portion of the NSCAW infant group is not ready to move to the 6 – 10 year old version despite their age. Perhaps this is because they are lagging behind their normative peers in this developmental domain. In either case, when the year six data is excluded from analysis, a linear model fits the data very well, adding further evidence to the idea that the year six data is somehow different than what had been gathered at previous data collections.

Table 2.22: Mean and Variance for the Intercept and Slope Factors of the Unconditional Model of Adaptive Behavior Development

	Int Mean	**Int Var.**	**Slp Mean**	**Slp Var.**
Estimate (std err)	87.62 (1.44)	112.25 (25.19)	3.29 (1.70)	0.66 (9.05)

Table 2.23: Sample Estimates of the Dependent Variables in the Unconditional Model of Adaptive Behavior Development

	Year 2	**Year3**	**Year4**	**Year 6**
Mean Score	86.8 (16.1)	88.9 (17.5)	95.4 (17.9)	83.1 (18.3)
Path Estimate	0 (fixed)	1 (fixed)	2.94 (1.43)	-1.69 (1.70)

Table 2.24: Residual Variances and R-Square Values of the Unconditional Model of Adaptive Behavior Development

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var.	151	188	137	224
R-Squares	.425	.400	.523	.304

Table 2.25: Fit Indices of the Unconditional Model of Adaptive Behavior

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (std err)	.948 (.048)	.896 (.097)	.059 (.018)	.096 (.029)

In contrast to other developmental domains and the CDM, neither simple poverty nor the income-to-needs ratio was statistically significant. Rather, there was evidence of interactions in the moderated poverty model. Parental health moderated by poverty had a statistically significant but very small effect size. More meaningfully, cognitive stimulation moderated by poverty accounted for a 2.13 point increase in developmental scores on the intercept factor, though the standard errors indicate a lack of precision. Mediated poverty proved a very poor fit to the data and neither SES nor the mediated poverty model produced significant parameter estimates.

In all models of poverty, several variables significantly affected the intercept factor. First, male infants lagged behind females by 5.36 points (1.55). Second, when compared with abused infants, those whose maltreatment type was classified as 'neglect' or 'other' fared better—4.06 (1.66) and 4.32 (2.42), respectively. No variables influenced the slope factor. Models excluding the questionable year six data also showed a strong (6.01) negative of FC placement on the intercept factor while the effect of FC was positive (1.98) on the slope. Regardless, the amount of the intercept's variance explained by the moderated poverty model was high compared to other models at 0.19. The r-squared of the slope was not estimated due to a non-significant negative estimate of the slope's residual variance.

Table 2.26: Fit Indices for the Moderated Poverty Model of Adaptive Behavior Development

	CFI	**TLI**	**RMSEA**	**SRMR**
Moderated Poverty	.882 (.045)	.764 (.090)	.041 (.007)	.038 (.007)

Table 2.27: Predictor Estimates for the Moderated Poverty Model for Adaptive Behavior Development

	Estimate (SE)
Mod Parent Health (Int)	-0.07 (0.04)
Mod Cognitive Stim (Int)	2.13^ (1.56)
Male (Int)	-5.36 (1.55)
Outhome (Int)	-3.80 (1.61)
Neglected (Int)	4.06 (1.66)
Other Maltx (Int)	4.32 (2.42)
Int R-Square	0.19
Slp R-Square	Heywood

DISCUSSION

Results presented here tend to confirm the finding that poverty has a negative influence on the psychological development of maltreated infants and children (McLoyd, 1998; Guo, 1998; Costello, Compton, Keeler, & Angold; McLeod & Nonnemaker, 2000; NSCAW Research Group, 2005). In their second year of life, infants enrolled in NSCAW were significantly delayed in all three domains of development measured and poverty was associated with those delays after demographics and maltreatment were controlled for.

In contrast to prior NSCAW research on development that only covered the first 18-months (NSCAW Research Group, 2005) and other research on development in the context of neglect (Dubowitz, Pappas, Black & Starr, 2002), during the approximately four years of study, infants enrolled in NSCAW appeared to be slowly narrowing the gap between normative and achieved development, both overall and in the three specific domains. But as the study cohort approaches school-age they do not appear to be school-ready based on their developmental scale scores. The domains of cognitive, communication, and adaptive behavior development were regarded by NSCAW planners and consultants as factors important to subsequent school success (RTI, 2007), so the lack of developmental achievement in those domains is an ominous predictor of eventual academic achievement.

Little evidence was found to support a lingering effect for early poverty on infants' developmental trajectories as they approach school-age. That is, the developmental trajectories (i.e., the slopes) of the NSCAW infant cohorts were not affected by their poverty status during their first year of life. Poverty was clearly associated with their initial developmental scores (i.e., intercepts). This result was consistent with prior research indicating early poverty is not more detrimental to development than poverty experienced later in life (Duncan, Brooks-Gunn, & Klebanov, 1994; Guo, 1998).

The income-to-needs ratio model demonstrated the best results of the five models of poverty evaluated. The income-to-needs model of poverty met the criteria outlined for evaluation; substantively and statistically meaningful parameter estimates, fit indices within generally accepted limits, and only a single improper solution. Estimates of fit indices and parameter estimates were nearly identical to those obtained using the simple poverty model and definitively superior to those obtained using SES or interaction-based models of poverty.

Given their similarity and the fact that ordinal data provide less information and require stronger assumptions than interval data (Allison, 1999; Frankfort-Nachmias & Nachmias, 2000), the clear preference is for the income-to-needs model. When the results obtained here are taken together with the fact that no information beyond what is used in the official poverty line (Dearing, McCartney, & Taylor, 2001; McLoyd, 1998; USCB, 1998) is required to compute an income-to-needs ratio and the interval nature of income-to-needs data, there seem to be few reasons to continue measuring poverty as a dichotomous, ordinal variable.

Mediated and moderated poverty, while not producing acceptable results using the criteria outlined above, did produce results that support further research. First, as will be discussed, both mediated and moderated models of poverty did produce some significant estimates of mediators of poverty or variables moderated by poverty on year two developmental measures. Second, fit indices may overstate the degree to which these interaction-based models demonstrated poor fit. Model fidelity was given priority over parsimony in modeling trials to avoid specification searches and potentially non-generalizable model modifications (MacCallum, Roznowski, & Necowitz, 1992). In so doing, model parsimony was reduced, and irrelevant or otherwise poor-fitting components were retained.

There is little prior research into poverty as a moderator of parent- and child-level predictors of children's development. Human capital of mothers has been found to moderate adolescent well-being among low-income families (Coley, Bachman, Votruba-Drzal, Lohman, & Li-Grining, 2007). Also, maternal depression (often associated with being in poverty) has been found to moderate infant development in early life (Cornish, MacMahon, Ungerer, Barnett, Kowalenko, & Tennant, 2005). Both the cognitive stimulation and maternal responsiveness scales of the HOME instrument were moderated by poverty in the model of adaptive behavior (Tables 2.22 - 2.27). However, results reported above, and results completed but not reported for the sake of brevity, also indicate differing domains of development may have differing predictors, each of which may be moderated by poverty. Much further research is necessary to clarify when, if ever, poverty is a moderator of other developmental predictors.

The result of the mediated poverty model, the most complex model evaluated, also appeared promising. This model specifies both a direct effect and an indirect effect of poverty. The indirect effect is mediated by five variables suggested by other researchers (McLoyd, 1998; Yeung, Linver, & Brooks-Gunn, 2002; Gershoff, Aber, Raver, & Lennon, 2007) thought to be influenced by poverty while also influencing development (see above). Results from all four developmental domains indicated that poverty does have an effect on parental health and mental health, but neither of those had an effect on developmental scores in the second year of life. The income-to-needs ratio used as the direct measure of poverty in the mediated model did have an effect on development at year two. Depending on the domain of development being modeled, maternal responsiveness, cognitive stimulation, or child health also had substantively and statistically significant results. Effects were smallest in adaptive behavior development and larger in the other two domains. They also extend prior research (Yeung, Linver, & Brooks-Gunn, 2002; Gershoff, Aber, Raver, & Lennon, 2007) into the domain of language development.

Mediated and moderating models of poverty are especially important for policy development. While the income-to-needs ratio model of poverty is currently the best tool for modeling the effects of poverty, it does not suggest interventions beyond anti-poverty programs such as income supporting programs. This research, together with prior work (Gershoff, Aber, Raver, & Lennon, 2007), suggests that developmental intervention methods might be directed towards programs addressing other topics such as parenting skills development, provision of stimulating toys, and improving the mental and physical health of parents. Further research using mediated models of poverty is necessary to clearly identify mediators of poverty for each domain of development.

From a programmatic perspective, these findings have implications as well. First, the findings highlights the fact that defining poverty as being below a certain amount of income for a given household type is arbitrary. Fewer than half of the infants in the current study are below the USCB poverty line (Table 2.1), yet mean scores for the developmental outcomes were approximately one standard deviation below the norm for all four measures at the outset and improved relatively little over time. Taken in the context of findings by others (e.g., Dearing, McCartney, & Taylor, 2001) that developmental outcomes are sensitive to changes in income among lower-income children, the implication is that even children in near poverty are prone to experience developmental problems. Consequently, programs addressing children's development in the context of poverty may wish to use a more liberal standard of poverty than that of the USCB, because achieving normative developmental outcomes seems to require more than simply the financial ability to meet basic necessity needs—the standard of the USCB.

A key question is the degree to which the findings presented here can be generalized to include all infants in, or nearly in, poverty. Put another way the question is; how similar are poor infants and maltreated infants? This is a difficult question because, by their defining characteristic—poverty or maltreatment—they are different.

From a developmental perspective, the two groups overlap in exposure to risk factors. While poverty researchers often control for risks such as reduced cognitive stimulation at home, health status of the infant or child, parental characteristics (e.g., education, marital status, and substance abuse), and service receipt (McLoyd, 1998), they rarely ask about or consider maltreatment. However, maltreated infants and children are exposed to very similar risk factors (ACF, 2005). Moreover, while not all maltreated infants are in poverty, many are or are close to the poverty threshold (Table 2.1).

The experience of deprivation, or lack thereof, is important in determining the eventual development of infants whether they are in poverty (Yeung, Linver, & Brooks-Gunn, 2002) or have experienced maltreatment. It is precisely the experience of deprivation that is thought to mediate the link between poverty and development, though the process is not well studied or understood (Yeung, Linver, & Brooks-Gunn, 2002; McLoyd, 1998). Most of the infants in NSCAW—and the national population it is intended to represent—experienced some form of neglect which is, by definition, an experience of deprivation of emotional or material resources that is considered injurious because it, in part, harms the child's development.

Children in poverty and maltreated children are probably similar, but it is not possible to be more definitive with the data at hand and in the literature. They share many developmental risk factors and a common experience of deprivation. It is not known to what extent infant maltreatment occurs in households in or near poverty because some, if not many, cases of maltreatment go undetected or at least unreported. As seen here, many maltreated infants live at or near poverty, suggesting a degree of commonality.

In summary, poverty has an effect presently best measured by an income-to-needs ratio regardless of developmental domain studied. There appears to be support for the concept of mediated poverty or poverty as a moderator of other developmental predictors. Findings did confirm some elements of previously hypothesized mediated and moderated models of poverty's effects, but these varied by developmental domain being? modeled. It may be that poverty is mediated by or moderates differing variables by differing magnitudes according to domain of development being analyzed. Further research is necessary to confirm this phenomenon and, if validated, evaluate which specific developmental predictors are moderated by or mediate poverty for each domain.

No support was found for a differential or persisting effect of early poverty on any of the four developmental trajectories of maltreated infants estimated and analyzed. Initial (i.e., early) poverty measures had no impact on the slope of the infants' developmental trajectories as the infants aged. Further research using time-varying covariates or another longitudinal method of investigation is indicated to further investigate causes of change, or the lack of change, in the development of infants and young children. A particular focus should be on understanding how living in poverty interacts with other developmental predictors and under what circumstances poverty might have enduring effects on children's development, whether experienced early in life or later in life.

Chapter 3

Modeling Infant Developmental Trajectories by Gender, Minority Status, Poverty Status, Maltreatment Type, and Placement Type in the Context of Child Welfare

Maltreated infants may not develop uniformly or follow similar developmental trajectories. As with non-maltreated infants, some infants will have typical or better development, achieving most developmental milestones at or before the norm. Others will do less well. Prior research has suggested the following groupings may demonstrate a common developmental trajectory. Evidence related to the likelihood of different outcomes for important sub-groups is reviewed below.

Gender

Expectations for growth and physical development are stratified by gender beginning at birth (National Center for Health Statistics, 2007). Domains of growth include length, weight, and head circumference. Physical development, such as bone development in the fingers, is also related to gender (Cox & Jordan, 2006; National Center for Health Statistics, 2007).

Domains of psychosocial development, such as language, also differ according to gender. Females develop many language and cognitive skills earlier than their male peers. Among a sample of middle-class children, at pre-school age (4 ½ years) females scored significantly higher on the Test of Language Development, particularly in the area of grammar use and comprehension. Females also scored significantly higher on the General

Cognitive Index of the McCarthy Scales of Children's Abilities, including all sub-domains[2] except the quantitative domain (Bornstein, Hahn, Gist, & Haynes, 2006).

In addition to findings of gender differences among samples of children with few or no developmental risks, males and females may have differing developmental trajectories when experiencing developmental risk factors. Among children with pre-natal exposure to cocaine[3], exposed females did significantly more poorly on an assessment of expressive language than their unexposed female peers. While males scored lower than females when not exposed, there was no significant difference between exposed and unexposed males. Males also showed less variability than their female peers. This difference was consistent with the literature reviewed including animal studies (Beeghly et al., 2006).

In addition to being found among children at developmental risk, gender risks have also been found among children born prematurely. At 6 ½ years of age males and females differed both in language developmental achievement and in the effect premature birth had on language development. Females demonstrated greater overall language development using a variety of assessments, but, as in the Beeghly et al. (2006) study, males were less affected by being born pre-term (Jennische & Sedin, 2003).

At the biological level, expectable developmental trajectories of maltreated children are neither well studied nor understood, particularly in the context of gender differences (De Bellis & Keshavan, 2003). It seems likely that the brains of maltreated children—presumably the biological seat of psychosocial development—develop differently over time (De Bellis &

---

[2] Verbal, Perceptual/Performance, Memory, and Motor

[3] Exposure was determined by assay of the fetal meconium in addition to maternal self-report yielding a much higher reliability than maternal self-report alone.

Keshavan, 2003). However, sample sizes and characteristics are small and probably unrepresentative so more definitive understandings are not possible.

The degree to which developmental differences between genders are products of nature or nurture is not understood. Parental and societal behaviors towards children differ by gender from birth, but there is evidence that biology has a role to play as well (Eisenburg, Martin, & Fabes, 1996). Transactional developmental theory (Sameroff, 1995) suggests it is an interaction between nature and nurture that causes differential outcomes. Differing environmental experiences will cause differing expressions of genetic predispositions. The resulting behaviors of the developing child may further reinforce a tendency for gender-stratified outcomes by cuing caregivers to provide certain kinds of stimuli (e.g., certain kinds of toys or activities).

Minority Status

Membership in a racial or ethnic (hereafter shortened to racial) minority group may carry some risk of developmental problems that persist over time. But it is not clear precisely what the underlying cause of differences between racial groups might be. That is, precisely what it is about membership in such a group that is associated with, or causal of, developmental problems is much less clear.

There are differences in developmental achievement over time among children as described by racial categories. Using baseline data, developmental risk as assessed using the Bayley Infant Neurodevelopmental Screener (BINS) was found to be greater among children birth to two years old who are Black compared to their White counterparts. White children ages birth to five scored higher on developmental assessments of language and cognitive ability as well. Differences were apparent in the overall sample as well as among children placed in-home, but were less, or not, apparent in children placed out-of-home (Administration for Children and Families [ACF], 2005). Other research also has found some evidence for racial differences in developmental achievement. Cognitive development and stimulation has been found to be lower in children who are Black compared to their White peers in a study of non-maltreated children (Linver, Brooks-Gunn, & Kohen, 2002).

Reasons given for these differences vary. McLoyd (1998) makes the case that poverty is more common among Black children and that this so-called *urban poverty* is worse because it is concentrated into small geographic urban areas also associated with few job opportunities, single parenthood, and high crime. Most simply, she is asserting that urban poverty, an aggravated form of poverty, is the root cause.

Others have argued (with some empirical support) that differential effects of race on development can be found in parental and intergenerational experiences. On average, black parents have had less education and that education has often been of poorer quality compared to their White peers. Moreover, their parents (i.e., the grandparents) may have resided in a segregated environment and resided in poverty with minimal education that was also of poor quality (Linver, Brooks-Gunn, & Kohen, 2002).

The effects of classification as a racial minority on developmental trajectories in the context of child welfare are not well-studied. That is, there seems to be little research on the developmental achievements of racial minorities over time (i.e., as a trajectory). Most of the studies forming the basis for the previous discussion are of children in poverty, but do not account for maltreatment. Nevertheless, to the extent children in poverty resemble children who have been maltreated (see subsequent discussion of neglect and poverty), the findings and inferences may apply or at least offer a starting point for generating hypotheses.

Maltreatment Type

The type of maltreatment experienced appears to play a role in the subsequent development of the child (National Clearinghouse on Child Abuse and Neglect Information, 2005; English, 1998). While maltreatment may be sub-typed into numerous categorical structures, the simplest breakdown yields two categories; abuse and neglect. Abuse is usually an act of commission (e.g., striking the child with an object) whereas neglect is usually an act of omission (e.g., failing to provide proper supervision of a child). By definition, child abuse and neglect can generally only be perpetrated by a parent or caregiver of the child; such actions by others constitute other crimes such as assault or rape (English, 1998). Few studies have compared the developmental effects of different types of maltreatment with each other (Maikovich & Jaffe, 2006), despite acknowledgement of its significance (Hagele, 2005; Cicchetti, 1994).

Abuse is often found to cause, or be strongly associated with, behavior problems, which are often associated with developmental problems (Hoffman-Plotkin & Twentyman, 1984). Children who were abused have been found to display more behavior problems than children who were neglected (Hoffman-Plotkin & Twentyman, 1984; Landsford et al., 2006). These include externalizing problems such as aggression towards others and internalizing problems such as anxiety and depression (Lansford et al., 2006). Sometimes the effects of abuse are not seen until much later in life (English, 1998), suggesting a more subtle or insidious developmental effect.

The result is that the effects of abuse seem to fall largely in the domain of behavior and associated developmental achievements such as developing trust in others and tolerance for frustration (English, 1998). In prior analyses using the National Survey of Child and

Adolescent Well-Being (NSCAW), there has been little evidence for differences in verbal and non-verbal cognitive development among the various sub-types of abuse and neglect in a sample composed of school-aged children (Maikovich & Jaffe, 2006).

Neglect has generally been found to have both direct and indirect negative effects on development, and its effects are somewhat different from those associated with abuse. In a brief review of the literature, it was found that neglect (regardless of sub-type) was associated with poor cognitive development in infants. By school age, neglected children were the poorest cognitive performers. Neglect was also linked to problems in language development in infants (Hildyard & Wolfe, 2002). When neglect is linked to aggression (Knutson, DeGarmo, Koeppl, & Reid, 2005), it has been proposed that a key mediating link is problematic emotional regulation, as opposed to a willingness (or even preference) to use aggression in pursuit of goals (Lee & Hoaken, 2007).

In contrast, results from Longitudinal Studies of Child Abuse and Neglect (LONGSCAN) analyses found an effect for only a single sub-type of neglect (psychological neglect) on development. However, the LONGSCAN sample was composed of a group of children already at risk for developmental problems (Dubowitz, Pappas, Black, & Starr, 2002). So a history of early deprivation may not have made their development any worse.

A key point in consideration of this data is that definitions of maltreatment, abuse, neglect, aggression, antisocial behavior, and other key concepts vary widely. To at least some degree, this reality confounds attempts to compare and integrate findings from research (Lee & Hoaken, 2007; Commission on Behavioral and Social Science and Education, 1993).

Child Welfare Placement Setting

Children placed into foster care (FC) settings[4] may have a different developmental trajectory from their peers remaining in their home of origin. This may be because they have differing characteristics or a greater number of risk factors, or it may be because of the foster care experience itself. At the least, they are very likely to have had different experiences compared with children who remain at home. Infants and young children represent the largest group of children enter fostering care (Vig, Chinitz, & Shulman, 2005). Because infancy and early childhood are key times for development, identifying and understanding differences in developmental risk and achievement are important (American Academy of Pediatrics [AAP], 2000).

Children placed into FC settings have more health problems than similar groups of children such as those receiving Medicaid (Hansen, Mawjee, Barton, Metcalf, & Joye, 2004) or the general population (AAP, 2000; Rosenfeld et al., 1997). Physically, children placed into FC have been found to have lower height, smaller head circumference, and poorer sensorimotor functions compared to a community control group (Pears & Fisher, 2005). Children in foster care were specifically found to have higher rates of abnormal physical exams (in a variety of medical specialties such as dermatology and ear, nose, and throat), dental problems, and delayed or absent immunizations (Hansen, Mawjee, Barton, Metcalf, & Joye, 2004). Merely having a serious or chronic medical problem places a child at greater risk for maltreatment and the developmental sequelae associated with it (AAP, 2000), and young children in foster care are among the medically needy groupings of children (Vig, Chinitz, & Shulman, 2005).

Children in foster care also had high rates of developmental delay or risk for delay. One study found 62% of maltreated infants and toddlers (3 to 36 months old) removed from

[4] Defined as both kinship and non-kinship foster care placements

their biological parents potentially had a developmental problem. When further tested, 73% of this group scored at least two standard deviations below the normative mean on a standardized developmental assessment instrument (Leslie, Gordon, Ganger, & Gist, 2002). Prior NSCAW research, including one of the few studies to compare children who are maltreated and remain in-home and maltreated children placed into foster care, supports this finding. Children in FC were found to be more often at risk of developmental problems than their in-home peers, 59.8% high risk and 51.1% high risk, respectively (ACF, 2005). Reviewing developmental literature, Vig, Chinitz, and Shulman (2006) find that problems in all developmental domains commonly occur in infants and young children in foster care—particularly among children who have been neglected.

Children placed into a FC situation, even in kinship care arrangements, will experience a significant disruption in their lives. Younger children are more susceptible to the consequences of such a disruption because they have fewer psychological resources to use in coping and may lack the cognitive faculties necessary to grasp that the placement is temporary (AAP, 2000). As a result of placement into a FC situation, children—especially younger children—may have problems with attachment (AAP, 2000) and behavior problems which may be quite severe and persist into at least adolescence (Lawrence, Carlson, & Egeland, 2006; AAP, 2000). Some research has indicated that children who remain in FC have better outcomes—developmental and otherwise—than their peers who were returned to their home of origin (Taussig, Clyman, & Landsverk, 2001), so results are not uniform.

Explanations for the high level of problems for children who remain in foster care abound. A caveat given by some (e.g., Hansen, Mawjee, Barton, Metcalf, & Joye, 2004) is that children placed into foster care gain access to more professionals, such as pediatricians and social workers, who are likely to notice symptoms of developmental problems or do a formal assessment of development. In addition, services are often more readily available to children removed from their families of origin. In contrast, a recent study using NSCAW data indicated that foster parents were better at identifying developmental problems than biological parents (Berkoff, Leslie, & Stahmer, 2006), so it may be that substitute caregivers are bringing the children's developmental problems to the attention of professionals, rather than increased contact with professionals who are subsequently identifying developmental problems. Alternatively, it may be that it is the number of placement changes a child endures, rather than the type of placement, that brings about behavior problems. Research indicates that the children who experience even a few placement changes are apt to also display behavior problems (James, Landsverk, Slyman, & Leslie, 2004; Harden, 2004; Newton, Litrownik, & Landsverk, 2000).

Research Question

Do infants' developmental trajectories or developmental predictors vary according to their membership in smaller, potentially more homogeneous groups? Prior efforts to identify common developmental trajectories suggest that there may be a substantively significant amount of variability in developmental trajectories. Some results—large standard errors of some estimates—suggested further modeling, using smaller, possibly more homogeneous groups may better identify significant parameter estimates despite the loss of sample size. This may have limited the number of parameters which were statistically significant, as well as prevented identification of effects which occur in only some of the infants among those placed in a foster care setting. Complete findings from earlier research using the Composite

Developmental Measure (CDM) may be found in chapter two, but results using the complete sample and the composite developmental measure (see below) are found in Tables 3.1 - 3.6.

Table 3.1: Means and Variances Using the CDM and the Complete NSCAW Infant Sample – Unconditional Linear Model of CDM

	Int Mean	**Int Var.**	**Slp Mean**	**Slp Var.**
Estimate (std err)	87.24 (0.76)	91.96 (14.27)	0.60 (0.29)	4.80 (1.99)

Table 3.2: Mean Estimates of the Dependent Variables in the Unconditional Model of the CDM and the Complete NSCAW Infant Sample

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score	86.7 (14.1)	87.3 (14.1)	89.0 (12.9)	89.4 (12.7)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.3: Variance Estimate and R-Squared Estimates for the Dependent Variables in the Unconditional CDM Model Using the Complete NSCAW Infant Sample

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var.	107	101	62	39
R-Squares	.463	.471	.613	.786

Table 3.4: Fit Indices for the Unconditional CDM Model and the Complete NSCAW Infant Sample

	CFI	**TLI**	**SRMR**	**RMSEA**
Estimate (std err)	.993 (.006)	.992 (.008)	.064 (.026)	.022 (.014)

Table 3.5: Results of Latent Curve Modeling Using the CDM and the Complete NSCAW Infant Sample - Ratio Model of Poverty

	CFI	**TLI**	**RMSEA**	**SRMR**
Ratio	.947	.905	.037	.040

Table 3.6: Predictor Estimates for the Ratio Poverty Model in the CDM and Using the Complete NSCAW Infant Sample

Ratio	**Estimate** (SE)
Ratio (Int)	1.71 (0.56)
Male (Int)	-4.45^ (1.48)
FC (Int)	-2.99^ (1.77)
2 Types (Int)	-2.29^ (1.37)
FC (Slp)	0.73^ (.045)
Int R-Square	.114
Slp R-Square	.115

METHODS

Sample

The sample was obtained from the National Survey of Child and Adolescent Well-Being (NSCAW), a national probability sample of children entering child welfare services (see ACF, 2005, for a complete description of the sampling design). Four full waves of data collection were completed at baseline and approximately 18, 36, and 66 months post-baseline. An additional, reduced wave of data was collected, primarily over the telephone from caregivers, at 12 months post-baseline (Research Triangle Institute [RTI], 2007).

All children who were less than 13 months old at the baseline data collection were included in analyses yielding an unweighted sample size of 1,196 infants. The age limit was based on prior work done using LONGSCAN data in which it was argued that children entering child welfare services between zero and 18 months of age represent a common developmental group (English, Graham, Litrownik, Everson & Bangdiwala, 2005). More practically, NSCAW data collection at the 60 month follow-up only included children up to 12 months of age at baseline (i.e., infants). Legal substantiation of maltreatment was not used as a criterion for inclusion in the present study. Not all participants had a legal finding that maltreatment had taken place (i.e., substantiation of maltreatment). Herrenkohl (2005) argues that defining maltreatment only by substantiation probably understates the actual level of maltreatment based on analyses showing no differences on ten developmental measures administered to maltreated children by Hussey et al. (2005). The term ‘maltreated’ is used for the sake of efficiency and brevity, though in some cases the maltreatment was not legally substantiated.

Measures

Four dependent variables were used. One each for three NSCAW developmental domains—cognitive, adaptive, and communication—were included while the fourth measure was a composite developmental score. All scores were standardized to a mean of 100 and a standard deviation of 15 to facilitate estimation and comparison.

Composite Developmental Measure (CDM). A variable was constructed to measure overall development of the child. The standardized scores for the cognitive, language, and communication domains were averaged to yield a single variable. This is one method of measuring overall or mean developmental gains. It should be clearly understood, however, that development across the three domains of development in NSCAW does not occur in parallel. As will be observed in subsequent analyses, both the scores and trajectories of the developmental outcomes vary across the domains. Given this heterogeneity of data, the meaning of results obtained using the CDM should be carefully interpreted because it is unlikely to clearly apply to any of the specific domains of development that were used in its construction.

Vineland Adaptive Behavior Scales Screener – Daily Living. A brief instrument used to screen children for problems in the domain of adaptive behavior and daily living skills. The Vineland Screener (Sparrow, Carter, & Cicchetti, 1993) is completed by a caregiver or other person knowledgeable about the child. The version for child ages zero to two was used at baseline with a version for children ages three through five used at subsequent waves as the cohort aged. The Vineland Screener strongly correlates (r = .87 to .98) with the full Vineland instrument.

Pre-School Language Scales. The PLS-3 was used to assess the developmental domain of language. It produces two sub-scales, expressive communication and auditory comprehension, and a total scale in children younger than six years old (Zimmerman, Steiner, & Pond, 1992). The scores are based on observations of the child. Interrater reliability is .98.

Battelle Developmental Inventory and Screening Test. The BDI was used to assess the developmental domain of cognitive development in children younger than five years old. It produces scores for five sub-domains and a total score. It is administered by an examiner. Despite the fact that the BDI does not require training for the administrator, it has a test-retest reliability of greater than .90 in most domains and in the total score (Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984).

Kaufman Brief Intelligence Test. The K-BIT was used to assess cognitive development in children older than four years. The K-BIT assesses four sub-domains as well as provides a total score. It is a self-administered, paper and pencil instrument. The test-retest reliability of the K-BIT varies by construct considered, but ranges from .74 to .95 (Kaufman & Kaufman, 1990).

The following are independent variables used in either the imputation model or the analytic model to predict developmental scores.

Demographics. Age, race/ethnicity, and gender variables were constructed. Age is measured in months. Race/ethnicity was initially conceptualized as having four levels, but estimation requirements forced the use of a dichotomous variable, minority or non-minority with non-minority serving as the reference. Gender was also coded as a dichotomous variable with female as the reference level.

HOME-SF Scales. Home Observation for Measurement of the Environment (Short Form) was used to assess emotional nurturing as well as cognitive stimulation in the infant's current caregiving environment (Bradley & Caldwell, 1984). Internal consistency for the total scales is .89 with a median of .74 for the subscales. Stability for the total scales is r=.62. A newer description of scoring the HOME-SF is available and was used (Bradley, Corwyn, McAdoo, & Coll, 2001).

Maltreatment. Maltreatment data was collected from the child welfare workers and case data. Several variables were created based on the dimensions of maltreatment suggested to be important (English, Bangdiwala, & Runyan, 2005). First, the most serious type of maltreatment was identified. To facilitate estimation, most serious type of maltreatment was dummy-coded as neglect or 'other' maltreatment with abuse (primarily physical and emotional) serving as the reference level. Second, severity of harm, as judged by the child welfare worker and rated as none, low, moderate, or severe, was coded as 1 to 4, respectively. Third, a dichotomous variable was created to indicate whether one or more than one type of maltreatment had occurred. Some children experienced multiple types of maltreatment, in which case the most serious was coded as the primary type (used to identify type in this analysis). Age of first episode of maltreatment was not included because all children in the study are defined by being in a common age group.

Socioeconomic Status Characteristics. Three variables represented the socioeconomic status of the infant's home of origin or of the foster home in which the child resided at baseline. A key concern is that SES lacks a consistent definition, both conceptually and operationally (Mueller & Parcel, 1981). In the analyses detailed below, SES is defined using three variables to represent the social capital model of SES (Bradley & Corwyn, 2002; Coleman, 1988).

Household Income. Annual household income is a scale of 1 to 11. Each number is a 5,000 dollar increment (e.g., 2 represents 5,001-10,000 dollars) while 11 represents any income over 50,000 dollars.

Index of Social Capital Indicators. Six indicators were used to construct an index indicating SES. Items included a primary caregiver having had at least some education beyond high school, being part of a first generation immigrant family, having a low-skill or unskilled type of job or having held such a job recently if unemployed, being unemployed, receiving one or more types of social assistance in the household, and being a single (i.e., not married or in a stable romantic relationship) caregiver. Organizing these items into a scale rather than retaining them as individual independent variables was necessary to avoid model estimation problems as well as the tendency for more complex structural equation models to exhibit a better fit than simpler models (Preacher, 2003). There is no generally accepted means of specifying SES in statistical models (Bradley & Corwyn, 2002).

Neighborhood Safety. Caregivers were asked a series of questions about their perceptions of criminal activity in their neighborhoods. Their responses were aggregated into a brief scale measuring caregiver perception of neighborhood safety.

Poverty. The household having total income below the official poverty line for that year and household composition is included as a dichotomous variable with not in poverty as the reference. An income-to-needs ratio was also created using household income divided by the dollar value of the poverty line for the household composition and year (USCB, 2006). Because household income was indicated as being in a given range of values, the midpoint of each range was used.

Setting. In child welfare, setting is the type of residence the child resides in. It may be in-home with nuclear family of origin, in kinship care with a relative or close family friend, in foster care, or in residential care of some sort such as group home or residential treatment facility. The data used was collected at baseline and most likely reflects where the child was placed early in the case. The placement may change as the case proceeds, but the setting variable identifies where the child was initially placed. Setting was collapsed to in-home (INH) or foster care (FC) to facilitate estimation with in-home placement as the reference. Kinship foster care is considered a foster care placement despite placement with a relative being considered in-home under the TANF program.

Child Health Scale. This variable was created using a current caregiver rating of the child's health from 1 (poor) to 5 (excellent).

Caregiver Health and Mental Health Problem. Caregivers were administered the SF-12, which assesses general health and mental health (Ware, Kosinski, & Keller, 1998). Two variables were created—one each for health and mental health—using the standardized results.

Service Receipt. Both child welfare workers and caregivers identified services provided to the child. These data were used to create dichotomous variables indicating a need for developmental services, having an Individual Family Service Plan (IFSP), and receipt of services based on the worker-reported information.

Number of Children in the Household. The number of persons under age 18 residing in the household was used only during imputation modeling.

NSCAW Analysis Weight. This variable is unique for each case and is designed to produce results which are representative of almost all American children involved with child welfare systems. The details of its creation are given by the Data File Users Manual (RTI, 2007).

Stratum. This indicates which of the nine strata in NSCAW the data came from. Each of the first eight strata is a single, large population state while the ninth stratum includes data from several smaller states. It is part of the complex survey sample data necessary to correctly account for clustering in NSCAW data (RTI, 2007).

NSCAWPSU. This indicates which primary sampling unit (PSU) the data came from. Each PSU was typically a child welfare agency serving a geographical locale (normally a county). It is also part of the complex survey sample data necessary to correctly account for clustering in NSCAW data (RTI, 2007).

RESULTS

Analysis

Two distinct sets of methods were required for completion of analyses. First, multiple imputation was employed to address the problem of missing data and a problematic metric of time. Second, Latent Curve Models (LCM; Bollen & Curran, 2005) were estimated for trajectories of the composite development measure (CDM) using the following variables to create sub-groupings: Gender [male/female]; Minority [white/minority]; Maltreatment Type [abuse/neglect]; Child Welfare Setting [In-Home/Foster Care].

Prior to imputation and analysis, the data were cleaned and formatted for analysis. This involved creating a single designation for missing data as missing data is coded by reason for missing-ness (if known) in the NSCAW data set. In some cases, a designation of missing was equivalent to a meaningful response. For example, a question about the amount of parenting training received might be skipped if the respondent indicated on a prior question that no services were received. While coded as missing, the actual amount of parenting training received was zero hours for that time period. In addition, all dichotomous responses were dummy-coded. In some cases, new variables were created or derived from one or more existing variables (e.g., number of children in the household).

Developmental data gathered at baseline was not used for two reasons. First, data collectors failed to correctly gather data on many occasions, leaving questions about the validity of the data gathered (RTI, 2007). Because brief screening versions of the instruments were used, even missing responses on a few questions were sufficient to cast the final scores into doubt. Second, assessment of infant development requires specialized expertise. Some have questioned the appropriateness of using inexpert, though not untrained, data collectors in the gathering of infant developmental data (Barth, Scarborough, Lloyd, & Casanueva, 2007). As a result of these two related concerns, infant developmental data gathered at baseline was omitted from both imputation and analysis.

Imputation Modeling

Multiple imputation was employed for two reasons. First, the problem of missing data needed to be addressed since LCM assumes complete cases are used. Second, time between waves of data collection was not a meaningful unit of time in the context of infant development. As demonstrated by Bollen and Curran (2006), changing the unit of time in longitudinal research may be conceptualized as a problem of missing data and addressed as such. Complete details of the MI process may be found in chapter two.

SAS Proc MI (SAS Institute, 2006) executed the imputation model without errors or warnings, and convergences were achieved in fewer than 30 iterations of the Proc MI algorithms. Ten imputed data sets were created. The computed variables were added using

SPSS 12.0.0 (SPSS Inc., 2003), and the data sets were separated into individual files and converted to a format usable by Mplus.

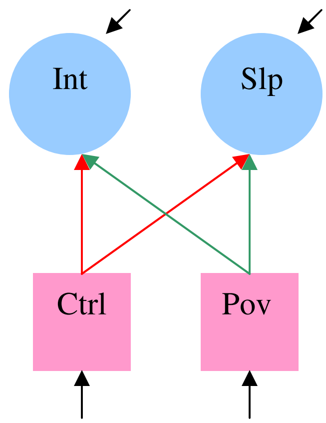
Analytic Modeling

The initial plan of data analysis called for multi-sample analysis (also termed Multiple Group Analysis). While possible, even when using clustered and stratified data, these models were not successfully estimated in this study. The most likely cause is a combination of using clustered and stratified data, multiple data sets with imputed data, and analytically complex models. As a result, what comparisons are made are without the benefit of statistical testing.

The process of analytic modeling was similar for each group analyzed. Initially, five conceptualizations of poverty were operationalized in the LCM framework. They may be seen in Figure 3.1 with the dependent variables omitted for the sake of clarity and space. The simple and ratio models of poverty are identical except that the independent variable is dummy-coded or continuous, respectively. In the moderated poverty model, poverty moderates or interacts with other variables which may predict development. The mediated poverty model allows poverty to have a direct effect, as in the ratio model, but also allows poverty to influence independent variables which, in turn, influence the slope and intercept factor. The socioeconomic status (SES) model utilized a social capital approach based on the work of Coleman (1988).

Figure 3.1: Operationalizations of Poverty in LCM

Simple and Ratio Poverty



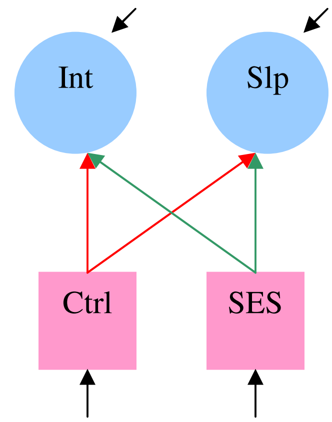
IV: Independent Variables

Ctrl: Control Variables

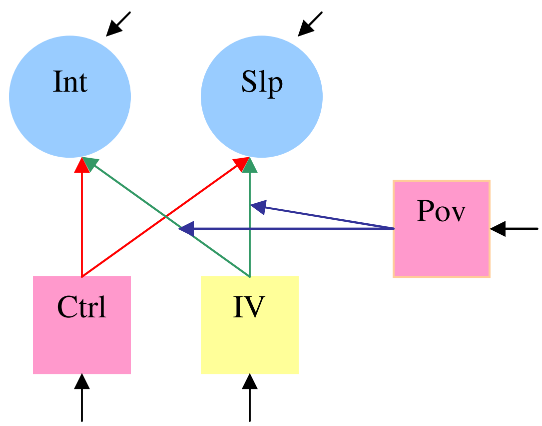
Pov: Poverty Variable

SES: Socio-Economic Status Variable

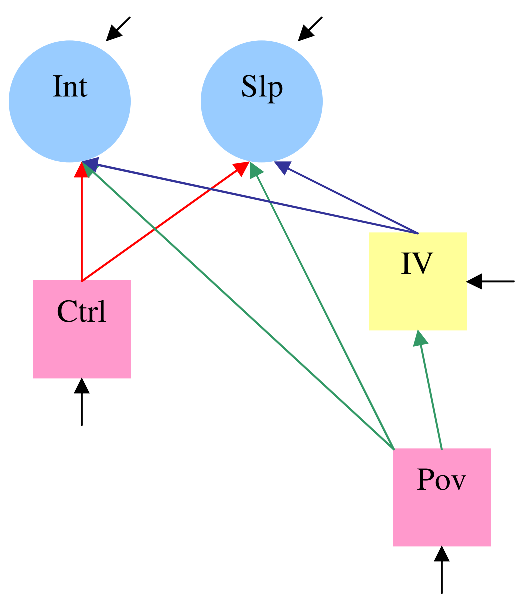
Socioeconomic Status



Moderated Poverty



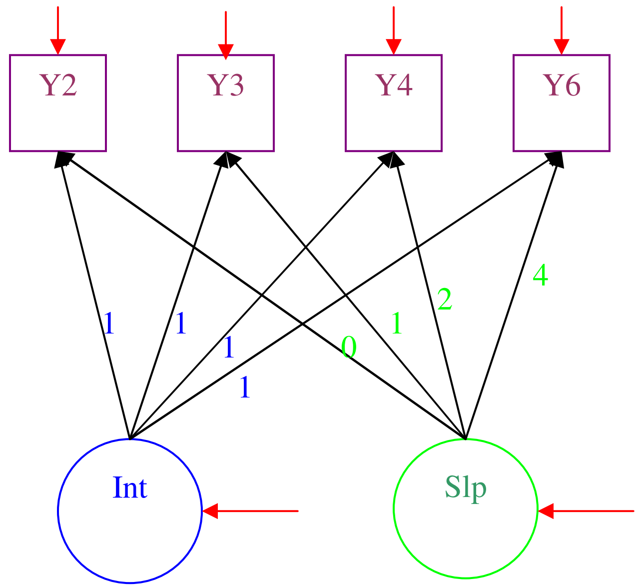
Mediated Poverty



All analyses were executed using Latent Curve Modeling (LCM; Bollen & Curran, 2006). This methodology is based in structural equation modeling (SEM; Bollen, 1989) and calculation of fit indices is possible. Models were estimated using Mplus version 4.21 (Muthen & Muthen, 1998-2007a), which allows combination of the results for each of the 10 data sets created during the imputation step. A robust maximum likelihood estimator was used (see Muthen & Muthen, 1998-2007b) with weighting, stratification, and clustering data included. The desired group was selected using the subpopulation command.

Models were fit to the data in a stepwise fashion. Initially, a basic unconditional model was fit to the data (see Figure 3.2). During this step it was determined whether the LCM was viable for the data in question. Dependent variables were the developmental data for each domain (composite, cognitive, etc.). If a linear model did not fit, a non-linear model was attempted. Typically, this involved using fixed paths from the slope factor to the initial time point and to either the second time point or the final time point. If the unconditional model was found to fit the data, a set of controls was added and each of the five substantive models of poverty being tested was fit to the data (Figure 3.1). This process was completed for each conceptualization of poverty for each group of interest. When the sub-grouping was by a variable included in the model (e.g., gender), then that variable was dropped in all models for that sub-group.

Figure 3.2: Unconditional Latent Curve Model



All models included controls for case and demographic characteristics. Maltreatment was modeled by including type, severity, and experiencing more than a single identified type. The reference for type of maltreatment was abuse. The reference for gender was female. The reference for race was non-minority. Finally, the reference for child welfare placement type was in-home (INH). Kinship care was considered a foster care (FC) placement in most circumstances.

Evaluation of models is based on several criteria. First, the model had to produce acceptable parameter estimates. Ideally, no improper solutions—negative variances, for example— should be reached by the EM, though they may be produced even when the model is a good fit to the data (Bollen & Curran, 2006). In a few instances noted in the results, a small, non-significant (from zero) negative residual variance was produced and tolerated.

A second important indicator of fit is the presence of reasonable estimates. That is, the magnitude and sign of the estimates should be appropriate to the data in use (Bollen &

Curran, 2006; Schumaker & Lomax, 2004). In planning, this was to be evaluated in part using confidence intervals. Due to the use of MI, however, this was not possible. Statistical significance was used instead, though it provides less information.

Finally, as with most SEM-based results, fit indices are produced. Mplus 4.21 produces four fit indices; Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), Standardized Root Mean Square Residual (SRMR; Joreskog & Sorbom, 1981; Bentler, 1995), and Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993; Steiger & Lind, 1980). Using the CFI and TLI, greater than .94 indicates a good fit while a result greater then .90 indicates acceptable fit. When using the SRMR and RMSEA, less than .06 indicates a good fit while less than .09 or .11, varying slightly by author, indicates an acceptable fit (Bollen & Curran, 2006; Garson, 2007; MacCallum, 2005; Hu & Bentler; 1999; Schumacker & Lomax, 2004; Bollen, 1989). Because of the use of MI, results reported below are means and their standard errors, of the fit indices.

Fit indices may offer differing assessments of fit. Each fit index uses a differing conceptual and mathematical definition of what ‘good fit’ is. Preference is given to the RMSEA and the SRMR, which do not depend on substantively meaningless null models for comparison. As has been shown elsewhere (MacCallum, 2005; Garson, 2007; Rigdon, 1996), their meaning is substantively clear and readily interpretable compared with those obtained using null model comparisons. The TLI and CFI are reported to fully disclose findings and permit alternative interpretations.

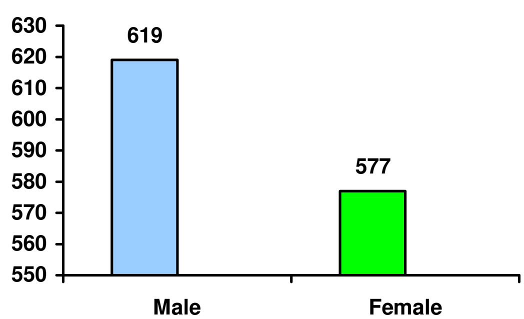
Fit indices may also understate the fit of a model to the data in the context of latent curve modeling. The form (e.g., linear, curvilinear) chosen for the latent curves is not intended to be a perfect fit to all data, rather it is chosen because it is a good approximation of those data. This lack of perfect global fit, despite being intentional or at least acknowledged, often results in understated fit indices (Coffman & Millsap, 2006). Global fit may be considered with *component fit* (Bollen, 1989), which is how well the individual substantive parameters fit as assessed by an appropriate sign and a substantive magnitude of the independent variables.

Findings

Infants were divided into groups based on the variables discussed previously. Analyses were completed for both groups, and the results compared with those obtained in prior analyses of the complete group of infants. The descriptive statistics were completed using un-imputed, un-weighted data because it provides the clearest information about the data that were used in the imputation model. Put another way, these data are the 'starting point' from which subsequent estimates are derived. Further descriptive statistics are available in chapter 2.

Gender

Figure 3.3: Distribution by Gender



Latent curve models are compared by gender (Figure 3.3). Analysis of all 1,196 NSCAW infants indicated males scored three to four points lower on developmental instruments during the second year of life (i.e., on the intercept factor), but no differences appeared in the slope factor (see chapter two).

Table 3.7: Sample Estimates for the Ratio Poverty Model for Males Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	84.7 (13.7)	84.6 (12.8)	86.7 (12.7)	88.5 (12.9)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.8: Residual Variance and R-Squared Estimates for the Ratio Poverty Model for Males Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var. (SE)	104 (20.5)	98 (15.2)	63 (13.0)	25 (32.8)
R-Squares	.449	.466	.606	.863

Table 3.9: Fit Indices for the Ratio Poverty Model for Males Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.915 (.020)	.848 (.035)	.050 (.007)	.054 (.008)

Table 3.10: Predictor Estimates for the Ratio Poverty Model for Males Using the CDM

	Estimate (SE)
Ratio (Int)	1.67 (0.69)
2 Types (Int)	-3.56^ (1.91)
Int R-Square	.103
Slp R-Square	.197

Table 3.11: Sample Estimates for the Ratio Poverty Model for Females Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	88.8 (14.3)	90.1 (14.7)	91.5 (12.7)	90.4 (12.4)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.12: Residual Variance and R-Squared Estimates for the Dependent Variables in the Ratio Poverty Model for Females Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var. (SE)	100 (24.3)	107 (16.5)	54 (12.6)	37 (12.7)
R-Squares	.517	.478	.655	.803

Table 3.13: Fit Indices for the Ratio Poverty Model for Females Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.945 (.033)	.902 (.059)	.040 (.014)	.046 (.014)

Table 3.14: Parameter Estimates for the Ratio Poverty Model for Females Using the CDM

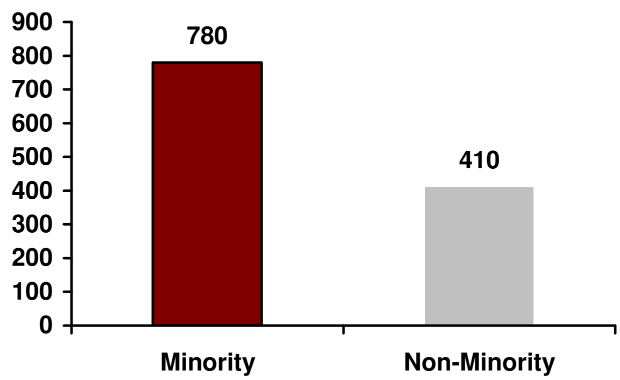
	Estimate (SE)
Ratio (Int)	1.98 (0.91)
FC (Int)	-6.77 (2.70)
Int R-Square	.097
Slp R-Square	.137

The ratio model proved to be the best of the five models for both males and females (Tables 3.7 – 3.14) but also differed in several respects. As expected, males had lower mean scores at each year of development (Tables 3.7 and 3.11). Fit indices were better for the female group (Tables 3.9 and 3.13). Their standard errors were larger suggesting somewhat greater variability among females (cf. Beeghly et al., 2006). The key difference between the groups is with regard to which of the predictors demonstrated substantive and statistical significance. Among males, having experienced two types of maltreatment was associated with a 3.56 point lower score at intercept. Females placed into an out-of-home setting had a score that was 6.77 points lower than males—almost half of one standard deviation—in scores at intercept (Tables 3.10 and 3.14).

No parameter was predictive of the slope, but each group had a different variable which trended (i.e., close to the $p<.05$ threshold) towards being significant. In the male group, being minority trended ($p<.20$) towards a 0.8 point drop per time unit. In the female group, being placed into an FC setting trended towards association with a 0.9 point drop per time unit (p<.20), further increasing the size of the discrepancy with males that was identified at the 2 year assessment.

Minority

Figure 3.4: Distribution by Racial Status



Moderated poverty proved the best model of poverty for the racial minority grouping (Figure 3.4). As was typically the case, scores did not change much over time—only two points over the five years of life studied. This is supported by the relatively modest associated standard errors (normative standard deviation is 15). As a result, there may be little change in the scores (i.e., slope) to attempt to predict. Fit indices were within acceptable limits and several predictors were substantively and statistically significant for the intercept (Tables 3.15 - 3.18).

In addition to the effects listed, three predictors also trended towards significance on the intercept factor. Moderated parental health and mental health had small effects (-.06, -.05, respectively; p <.15) and children who experienced the 'other' type of maltreatment did 3.54 points better than their abused or neglected peers. For the slope factor, moderated cognitive stimulation trended towards a beneficial effect of 0.66 (p<.20).

The mediated poverty model allows poverty to have had an effect on predictors of development as well as a direct effect. Comparison-based fit indicators were poor (CFI and TLI were .811 and .729, respectively). But many predictor effects were significant and the RMSEA and SRMR were within acceptable limits. Poverty influenced parental health and mental health while maternal responsiveness (among several variables) had a very strong (10.58) effect on the intercept factor. Poverty continued to exert a direct effect on intercept. Ultimately, the mediated model was not chosen because it did not fit the data well, and the pattern of significance among the independent variables did not validate the hypothesized mediated link between poverty and development (though results were suggestive).

Table 3.15: Sample Estimates for the Moderated Poverty Model for the Minority Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	86.1 (14.3)	87.3 (13.8)	88.6 (13.0)	88.2 (12.7)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.16: Variance and R-Squared Estimates for the Moderated Poverty Model for the Minority Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var. (SE)	117 (24.1)	111 (18.0)	58 (12.6)	30 (25.0)
R-Squares	.456	.481	.635	.835

Table 3.17: Fit Indices for the Moderated Poverty Model for the Minority Group Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.949 (.019)	.905 (.036)	.034 (.008)	.037 (.009)

Table 3.18: Predictor Estimates for the Moderated Poverty Model for the Minority Group Using the CDM

	Estimate (SE)
Mod. Resp.(Int)	4.53 (2.03)
Mod. Ch. Health	0.64

(Int)	(0.32)
Male (Int)	-4.07 (1.69)
2 Types (Int)	-2.73^ (1.50)
Int R-Square	.224
Slp R-Square	.149

Table 3.19: Sample Estimates for the Ratio Poverty Model for the Minority Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	88.0 (13.9)	87.3 (14.6)	89.7 (12.7)	92.1 (12.2)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.20: Residual Variance and R-Squared Estimates for the Ratio Poverty Model for the Minority Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var. (SE)	86 (24.8)	106 (15.1)	56 (13.2)	39 (36.5)
R-Squares	.548	.484	.656	.798

Table 3.21: Fit Indices for the Ratio Poverty Model for the Minority Group Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.935 (.024)	.883 (.043)	.051 (.011)	.053 (.013)

Table 3.22: Predictor Estimates for the Ratio Poverty Model for the Minority Group Using the CDM

	Estimate (SE)
Ratio (Int)	1.58^ (0.94)
Male (Int)	-4.76 (2.32)
Int R-Square	.100
Slp R-Square	Heywood

For the non-minority group of infants, the ratio model had the best fit among the poverty models (Tables 3.19 - 3.22). The ratio poverty variable had 1.58 per unit influence on the slope factor while males continued to score over four points (4.76) less than females on the intercept factor. As above, the change in scores over time was small, perhaps reflecting little developmental change. There also seemed to be slightly less variability among the scores compared to the minority group, based on the smaller standard errors of the developmental scores as well as the somewhat smaller residual variances despite no better fit of the model. In addition to the effects shown in Table 3.22, placement into an out-of-home child welfare setting (1.13) trended towards significance in predicting the slope factor for non-minority children ($p<.20$).

No model of poverty fit the data well for the non-minority group. The ratio poverty model had the most numerous and largest effects among the five poverty models, though the TLI is poor and the model produced a Heywood case. However, the other models of poverty did not produce an effect for poverty or mediators of poverty at all.

Maltreatment Type

Figure 3.5: Distribution of Primary Maltreatment Type

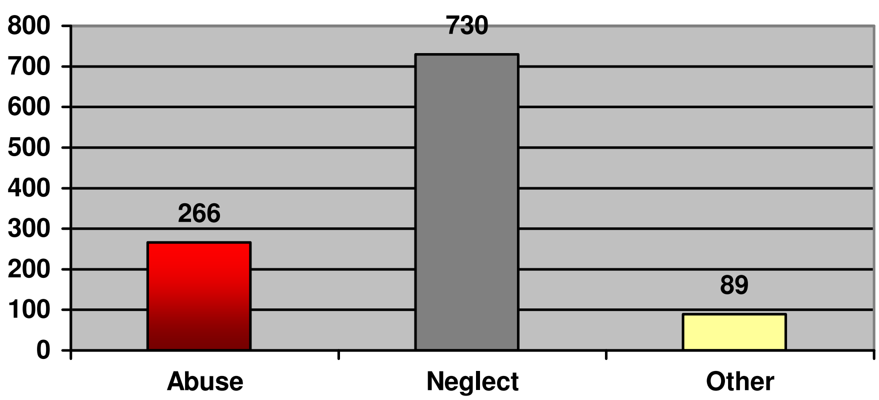


Table 3.23: Sample Estimates for the Moderated Poverty Model for the Abused Group Using the CDM

	Year 2	Year 3	Year 4	Year 6
Mean Score (SE)	85.2 (13.9)	86.3 (13.7)	88.8 (13.6)	89.2 (12.6)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.24: Variance and R-Square Estimates for the Moderated Poverty Model for the Abused Group Using the CDM

	Year 2	Year 3	Year 4	Year 6
Residual Var.	102	125	57	43

(SE)	(31.2)	(30.9)	(15.8)	(32.0)
R-Squares	.474	.428	.643	.766

Table 3.25: Fit Indices for the Moderated Poverty Model for the Abused Group Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.909 (.036)	.832 (.066)	.053 (.012)	.049 (.010)

Table 3.26: Predictor Estimates for the Moderated Poverty Model for the Abused Group Using the CDM

	Estimate (SE)
Mod. Resp.(Int)	7.35^ (4.04)
Mod. Par. Health (Int)	-0.12 (0.06)
Male (Int)	-3.81^ (2.23)
Int R-Square	.279
Slp R-Square	Heywood

Moderated poverty was the best model of poverty for abused infants (Tables 3.23 – 3.26). Effects for moderated maternal responsiveness and gender were substantive (7.35 and 3.81, respectively) and statistically significant. R-Square for the intercept factor was .279, but a slope r-square value could not be computed. In addition to the effects of predictors shown in Table 3.26, having experienced two types of maltreatment trended towards an effect of 1.40 ($p<.15$) on the slope factor. Other models of poverty were simply a poor fit to the data.

Table 3.27: Sample Estimates for the Ratio Poverty Model for Neglect Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	87.3 (14.1)	87.6 (14.0)	89.1 (12.6)	89.0 (12.8)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.28: Variance and R-Squared Estimates for the Ratio Poverty Model for Neglect Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var. (SE)	95 (22.7)	91 (12.3)	52 (11.3)	33 (29.0)

R-Squares	.537	.532	.673	.826

Table 3.29: Fit Indices for the Ratio Poverty Model for Neglect Group Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.948 (.022)	.908 (.038)	.045 (.013)	.051 (.008)

Table 3.30: Predictor Estimates for the Ratio Poverty Model for Neglect Group Using the CDM

	Estimate (SE)
Ratio (Int)	2.04 (0.72)
Male (Int)	-4.20 (2.01)
2 Types (Int)	-3.57 (1.71)
Int R-Square	.116
Slp R-Square	.097

Neglected infant data was best modeled by the ratio model (Tables 3.27 – 3.30). Substantive and significant effects were found for the ratio poverty, gender, and having two types of maltreatment variables (2.04, -4.20, and -3.57, respectively). R-Square values were smaller, however, at .116 for the intercept factor and .097 for the slope factor. Placement into an out-of-home child welfare setting trended towards an association with the intercept (-3.51) and the slope (0.82) factors (Table 3.30). Both were significant at the $p<.15$ level. Estimates of the mean developmental scores suggest less variability in the data for neglected infants (Table 3.28). While the simple model of poverty had an identical pattern of loadings, its estimates were smaller and none of the other models of poverty was a good fit to the data.

Analyses of the remaining group of infants, whose maltreatment type was ‘other’, were omitted for two reasons. First, there are less than 100 subjects in the group. Estimates based on small samples are particularly vulnerable to distortion. Second, there is no readily identifiable reason to believe the group may be more homogeneous than the complete sample.

Child Welfare Placement

Figure 3.6: Distribution of Child Welfare Placement Types

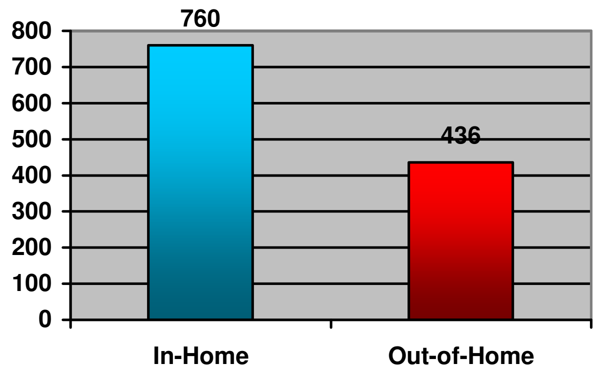


Table 3.31: Sample Estimates for the Ratio Poverty Model for In-Home Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	87.3 (14.4)	87.2 (13.6)	89.4 (12.6)	89.4 (12.7)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.32: Variance and R-Squared Estimates for the Ratio Poverty Model for In-Home Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var. (SE)	102 (21.8)	104 (14.1)	56 (10.9)	37 (28.4)
R-Squares	.504	.484	.648	.792

Table 3.33: Fit Indices for the Ratio Poverty Model for In-Home Group Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.936 (.026)	.885 (.046)	.045 (.011)	.044 (.007)

Table 3.34: Predictor Estimates for the Ratio Poverty Model for In-Home Group Using the CDM

	Estimate (SE)
Ratio (Int)	1.86 (0.78)
Male (Int)	-5.84 (2.11)
Int R-Square	.136
Slp R-Square	.204

The ratio model of poverty was selected as the best model of poverty for the in-home (INH) group (Tables 3.31 – 3.34). There seemed to be significant variability among the infants in INH placement. Large standard errors relative to the estimates were common and limited the number of estimates which were statistically significant (Tables 3.31, 3.32, and 3.34). This occurred in the ratio, mediated, and moderated poverty models in particular.

Table 3.35: Sample Estimates for the Ratio Model of Poverty for the Foster Care Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Mean Score (SE)	85.7 (13.6)	87.5 (14.8)	88.1 (13.6)	89.5 (12.7)
Path Estimate	0 (fixed)	1 (fixed)	2 (fixed)	4 (fixed)

Table 3.36: Variance and R-Squared Estimates for the Ratio Poverty Model for the Foster Care Group Using the CDM

	Year 2	**Year 3**	**Year 4**	**Year 6**
Residual Var. (SE)	105 (23.4)	99 (19.3)	65 (15.6)	16 (27.9)
R-Squares	.466	.487	.628	.921

Table 3.37: Fit Indices for the Ratio Poverty Model for the Foster Care Group Using the CDM

	CFI	**TLI**	**RMSEA**	**SRMR**
Estimate (SE)	.967 (.029)	.943 (.055)	.029 (.019)	.056 (.021

Table 3.38: Predictor Estimates for the Ratio Poverty Model for In-Home Group Using the CDM

	Estimate (SE)
Ratio (Int)	1.68 (0.68)
Minority (Slp)	-1.60 (0.62)
Int R-Square	.093
Slp R-Square	.145

For infants placed into FC settings at year 2, the distribution of substantive and statistically significant parameter estimates was different compared to other groups of infants, including those placed INH (Tables 3.35 – 3.38). Gender was not a significant predictor of either the slope or the intercept factors. The ratio measure of poverty (1.68) was a significant predictor of the intercept, indicating that children who were placed into very poor households had lower developmental standing at year 2. Minority status (-1.60) was a significant predictor of the slope (Table 3.38). That is, infants in foster care who were in a minority group were doing more poorly compared to their non-minority peers in foster care, as time passed. By contrast, racial category was not found to be a significant predictor of developmental trajectory among the infants remaining in-home (Table 3.34).

DISCUSSION

The research question asked whether infants' developmental trajectories or predictors of those trajectories are different when sub-grouped by case characteristics. In practice, this is really two questions. First, are the developmental trajectories substantively different for the sub-groups compared to the full sample? Second, are the predictors of those trajectories different compared with the full sample?

Developmental Trajectories

The developmental trajectories of all sub-groups varied little from the trajectory of the complete sample when using the composite developmental measure (CDM) as the outcome of interest. All models successfully fit to the data of the various subgroups were linear LCMs. The means of the developmental scores at each time point for each sub-group also do not vary substantively from the complete sample.

Predictors of Developmental Trajectories

Sub-grouping NSCAW infants produced some marked changes in predictor estimates compared to the complete sample. In particular the female and foster care sub-groupings revealed important contrasts compared with the complete sample. Moreover, in the minority and abused sub-groupings a moderated model of poverty was preferred over the ratio model preferred in the complete sample and many sub-groupings.

Among female maltreated infants, being placed into foster care initially was associated with a developmental score that was almost a half a standard deviation (6.77 points) lower compared with those remaining in-home at year 2. In addition, foster care placement trended toward being associated with a 0.9 point drop at each time point. By contrast, in the complete sample, males generally scored approximately five points lower than females at year 2 and gender was not a predictor of the slope at all. Among male maltreated infants, the only material difference was the r-squared value for the slope factor was almost .2 in the male-only model whereas it was slightly over .1 in the complete sample model. This suggests that among males, the model better accounted for the changes in the slope factor than it did in the complete sample model.

It is difficult to compare results presented here to the available literature. The size and composition of the samples are quite different, though both the Beeghly et al. (2006) and Jennische and Sedin (2003) samples shared some developmental risk with the NSCAW sample (e.g., parental substance abuse). In the Beeghly et al. (2006) and Jennische and Sedin (2003) analyses, developmental hardship seemed to affect female infants more strongly than males resulting in approximately equal developmental achievements among both genders.

In the analyses here, both gender groups performed well below their normative peers—probably reflecting their experiences of maltreatment and other risk factors—but females retained a slight advantage of two to four points over males in mean scores at each time point. However, the negative effects of foster care placement on both the starting point and slope of the female group's trajectory indicated that those females remaining in-home are doing better than males, whereas females placed into foster care were doing less well.

The net result seems to be that female infants are more developmentally challenged by their foster care experiences or females being placed into foster care have experienced more developmentally injurious events prior to their placement in a foster setting (i.e., it's an effect of which girls are selected for foster care). If the former case holds, these findings reflect a gender difference in the effect of foster care and suggest there is not only a difference between the genders in how their overall development is affected by hardships such as maltreatment and poverty but also a difference in how specific hardships affect development. This may be so, because female infants develop more quickly than males and so are more vulnerable to early developmental hardships. A related explanation is that since girls are expected to develop more quickly, deficits are observed more quickly since children placed into foster care are typically given a physical examination by a physician or other primary care provider.

That females who were placed into foster care experienced more serious maltreatment is not supported by subsequent analyses. Male and female infants experienced approximately similar likelihoods of experiencing multiple types of maltreatment, perceived harm from maltreatment, occurrence of poverty, likelihoods of being in a racial minority, cognitive stimulation, and parental responsiveness. Males were more likely to have experienced physical abuse and females were more likely to have experienced sexual abuse, but the numbers involved in either maltreatment were small and very unlikely to have impacted analyses to the degree observed. Hence, it seems unlikely that females had a systematically different maltreatment or environmental experience prior to entering the child welfare systems.

The result is that the cause of the gender difference for foster care placement observed among NSCAW infants is unknown. It is unlikely to be because they had a different developmental experience than males prior to entering child welfare. A reasonable explanation is that females are more vulnerable to early hardship than males because they develop sooner. This finding may also reflect a statistical artifact that will not be replicated in subsequent national samples.

Modeling by sub-groups based on initial child welfare placement—either remaining in-home or placement into foster care—produced results that help clarify results found using the complete sample. The effect of male gender at year 2 was not observed in the foster care group while a strong effect for minority group membership on the slope was observed. The effects of poverty were similar for both the INH and FC groups and the complete sample as were mean scores at each time point. There are sharp contrasts between the INH and FC groups while both sets of predictors appear less clearly in the model using the complete sample.

Also, the INH and complete samples have similar good fit indices, but the FC group shows much better fit on three of the four indices (.02 higher on the CFI, .04 higher on the TLI, and .008 lower on the RMSEA). The FC sub-group shows the highest r-square values for the dependent variables, suggesting that the experience of foster care is one of the most clearly-defined experiences among the sub-groups evaluated.

A final substantive finding is that among infants placed into foster care initially, those classified as being in a racial minority had lower scores than those where were not as time passed. Infants in a racial minority group had slopes 1.60 points lower than the non-minority group while there was not difference at year 2. This means that the racial minority groups' score fell by 1.60 points at each subsequent time point compared to their non-minority peers. More simply, children placed into foster care as infants have lower developmental achievement scores and, subsequently, are probably less school-ready then their non-minority peers.

Prior research, which shows that children who are placed come from families with the most risk factors and have the poorest development achievement (ACF, 2005; NSCAW Research Group, 2006; Vig, Chinitz, & Schulman, 2006; Leslie, Gordon, Granger, & Gist, 2002), suggested the infants placed into foster care should have had the poorest developmental scores and trajectories. Rather, what was found using NSCAW data is more nuanced. It seems that FC placement is associated with lower developmental scores in only some infants—those who are female or racial minorities in particular. The prior research outlined previously is based on samples that include older children who would have had greater opportunity to experience prolonged (and hence, presumably more serious) maltreatment than the group analyzed here, and this may account for at least some of the difference in findings. Moreover, few of the infants in the NSCAW sample analyzed here performed at the normative, or better, level. Rather, the question is the degree to which the infant group in question performed below the norm.

Further comparison to the existing literature is difficult. Most literature focuses on the effect of being placed into foster care on infants (see above) rather than which infants seem to achieve a better developmental outcome once in a FC setting. Moreover, it is not clear from these data whether it is that foster care itself is causing developmental problems or if it is that infants placed into foster care have experienced greater hardship prior to placement and, consequently, are more likely to have developmental problems.

Conclusions

What improved component fit provided was greater clarity of predictor effects on the slope and intercept factors. Several models had predictors that were trending towards significance in the complete sample (e.g., gender and race) which became statistically significant or were more clearly not significant. Also, new effects emerged not seen in the complete sample in some instances. This leads to new understanding that minority status among males as a predictor of worse developmental trajectories for those involved with child welfare services.

The most substantive impact of this last finding is to suggest that differing effects of poverty and some other predictors of development have been, and will continue to be, seen in the literature because poverty affects differing groups in differing ways. These results suggest that the specific methodologies and variables chosen for a study of poverty and development should be chosen with the characteristics of the expected sampling frame, and eventual sample, in mind. In particular, the findings presented here underline the importance of interpreting study results with clear respect to the characteristics of the sample.

In contrast to expectations, no marked gains in model fit were observed by modeling subgroups. Only the FC-placed infants had better fit indices than those obtained for the complete sample. Residual variance and r-square values for the dependent variables were also similar in the complete sample and sub-groups. Likewise, standard errors for the mean CDM scores at each time point were not smaller in the sub-groups than in the complete sample. The only parameters to show almost uniform improvement over the complete sample estimates were r-square values for the intercept and slope, which ranged from slightly below the complete sample in a few sub-groups to a high of .279 for the intercept for the abused infants sub-group. Mean scores at the final time point were generally similar in each sub-group regardless of what was used to segregate them. Specific predictors of developmental starting points and eventual trajectories varied, though the preferred model of poverty did not change among the sub-groups markedly either.

The most significant difference found was in the pattern of significant relationships observed in each model, suggesting that membership in a particular sub-group does matter. Clearly poverty has been shown to have a negative effect on development overall and should be considered by policymakers and practitioners when addressing the needs of maltreated infants. It also suggests interventions may require tailoring to the specific characteristics of the child to best meet their needs. For example, minority children placed into foster care seem to need special attention to avoid a negative developmental track. Understanding what psychological or ecological variables cause this phenomenon requires additional research.

The preferred model of poverty continued to be the ratio model, which offers more information than the traditional dichotomous measure of poverty. The sporadic success of the moderated model of poverty and the mixed results from the mediated poverty model suggest poverty may act through different predictors in different groups. Based on findings presented here and in chapter two, it continues to be recommended that researchers use an income-to-needs ratio when possible to improve the quality of findings about poverty.

To the extent practical and technically feasible, further research should continue into how maltreated infants vary in developmental predictors, trajectories, and outcomes according to group membership. Some argue that it is the accumulation of risk, rather than any specific risk that triggers negative developmental outcomes (Sameroff, 1998). The findings presented here suggest it is possible to identify specific risks when the characteristics of the sample are available. What is much less clear is what groups of maltreated infants they become relevant in. Overall the only differences identified between the complete sample and sub-groupings modeled was among the predictors of slope and intercept factors. The question remains as to what, if anything, maltreated infants with similar slopes and intercepts actually do have in common.

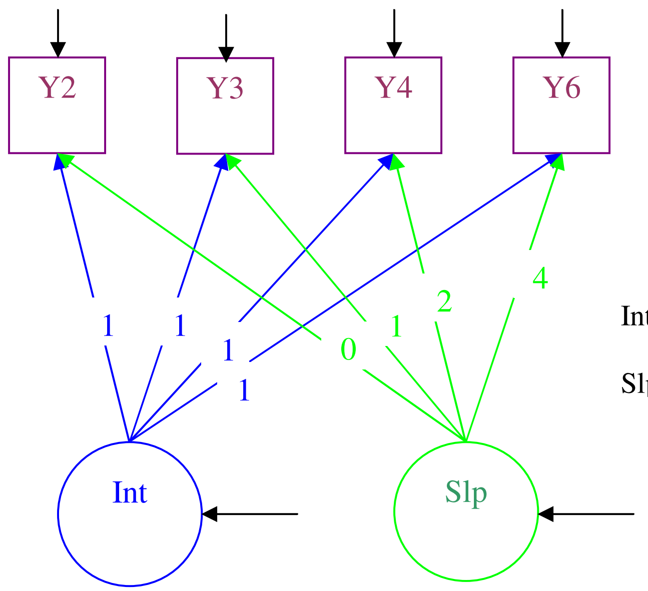
Chapter 4

Comparison of Results Obtained Using Latent Dependent Variables as Compared to Measured Dependent Variables

Latent curve models (LCM; Bollen & Curran, 2006) are a means of modeling change in a dependent variable of interest measured at different time points. Latent slope and intercept factors predict the values of a dependent variable measured repeatedly at three or more time points. The paths from the latent factors are fixed to create a specific trajectory. Fit indices and others measures of fit are assessed to understand how well the models fit the data. Manifest or latent predictors of the slope and intercept factors may be added as well, creating a conditional model, to assess what variables predict the slope and intercept factors (Bollen & Curran, 2006).

Latent curve models (Bollen & Curran, 2006) typically have a single manifest dependent variable (MV), such as a composite scale score, as a dependent variable (see Figure 3.1). However, because the LCM is specified and estimated using structural equation modeling (SEM; Bollen, 1989), the dependent variable need not be a MV. Some have suggested that a latent dependent variable (LV) may produce superior results (Coffman & MacCallum, 2005). Additional research is needed to help inform the choice between a latent dependent variable or a manifest variable.

Figure 4.1: Basic Latent Curve Model Path Diagram



Int: Intercept Factor

Slp: Slope Factor

There are two major methodological issues that must be addressed when estimating LCMs. Latent variables are, by definition, constructs and cannot be measured directly by a researcher. Rather, they are indirectly measured by other MVs or indicators that are thought to be caused by the latent construct (Bollen, 1989). Care must be taken to insure that the indicators used are caused by the latent construct as opposed to being properties of a larger concept when using normal methods of estimating structural equation models. For example, the latent construct of intelligence might be thought to cause cognitive abilities (e.g., logical reasoning and memory) so measured variables would include tests of such cognitive abilities. The measures of the latent intelligence variable would be expected to vary together to some degree. By contrast, stress—when conceptualized as a latent variable—might be measured by poverty, substance abuse, and family conflict. But these variables need not vary together nor are they necessarily caused by stress—they may be symptomatic of it. This latter type of latent construct may be termed *emergent* and must be treated differently than the former case or risk biased parameter estimates because the assumptions of SEM are being violated. So the researcher must insure, using theory or prior research, that a latent construct could plausibly cause changes in the MVs (Cohen, Cohen, Teresi, Marchi, & Velez, 1990).

A second issue is the assumption that manifest, or measured, variables are assumed to be measuring the construct of interest without error. This is an often unsupported assumption in research (Coffman & MacCallum, 2005; Sayer & Cumsille, 2001), and the situation is aggravated by the difficultly in predicting how resulting parameter estimates will be biased (Bollen, 1989). Coffman and MacCallum (2005) and Sayer and Cumsille (2001) state the unmodeled measurement error will tend to cause resulting estimates to understate effects among the variables, raising the probability of a Type II error in significance testing.

Using a latent variable in place of a single manifest variable, such as a composite scale, may have some advantage. The assumption in LCM is that a MV perfectly measures the construct of interest. However, a LV allows for a discrepancy between the observed MVs and the latent construct of interest they indicate (Bollen, 1989). This is termed unique variance and is estimated for each latent variable, becoming part of the model. The result of this change may be less biased estimates (Coffman & MacCallum, 2005). Because LCM methodology as described by Bollen and Curran (2006) typically calls for the use of measured dependent variables, this is a significant concern.

Latent curve models do allow for the use of latent dependent variables (Bollen & Curran, 2006; Sayer & Cumsille, 2001). Specifically, a composite scale measuring a latent construct (e.g., cognitive development) might be converted to a latent variable. Individual items, groups of items, or sub-scales are then used as indicators of the new latent construct. However, there is little research comparing the use of composite scales with latent variables where the components of the scale, either individually or in groups, are used as indicators (recent exceptions include Sayer & Cumsille, 2001 and Coffman & MacCallum, 2005).

While making the dependent variables in LCM latent has an advantage in modeling error variance, it also raises the issue of how to best measure the latent construct of interest when the dependent variable in LCM is a scale score. A scale score is typically derived from combining the results of the individual scale items. If there are too many items to use as individual measures, they will need to be combined into groups which are then used as indicators of the latent construct. These groups are called *parcels.* A parcel is defined as, "…an aggregate-level indicator comprised of the sum (or average) of two or more items, responses, or behaviors," (Little, Cunningham, Shahar, & Widaman, 2002, p. 152).

Parceling is often necessary because an excess of indicators in the measurement models of a structural equation model may cause bias in the fit indices. Prior research has suggested that as the number of indicators per latent variable increases, fit indices improve. Also, individual items tend to be less reliable and less normally distributed than parcels (Hall, Snell, & Foust, 1999) resulting in violation of the assumptions of SEM and LCM methodologies.

Parcels may be organized along several characteristics (Bagozzi & Edwards, 1998). First is *dimensionality*. Items may be assigned to parcels so as to make the parcels either unidimensional or multidimensional. In a unidimensional parcel, all items in a parcel measure a single common concept. That is, if a factor analysis were completed on them, all items would load on a common single factor with no meaningful loadings on any additional factors. In a multidimensional parcel, the constituent items do not load on a single common factor. Each parcel has one or more items from each dimension or aspect of the underlying construct (Kishton & Widaman, 1994; Bagozzi & Edwards, 1998; Hall, Snell, & Foust, 1999; Little, Cunningham, Shahar, & Widaman, 2002). Multidimensional parcels may also be termed domain representative parcels (Kishton & Widaman, 1994).

The requirements for the parcels are different according to their dimensionality. In unidimensional parceling, the parcels must be at least minimally reliable (~ 0.8) to obtain the best resulting estimates possible. They must also unambiguously represent only one domain or topic. In multidimensional parceling, no such requirements are necessary. Each parcel is a representative indicator of the underlying construct (Kishton & Widaman, 1994).

A second characteristic of parcels is their *level of aggregation* (Bagozzi & Edwards, 1998; Coffman & MacCallum, 2005). Parcels may be constructed by using raw items composing a scale or other research instrument or they may be composed of raw items that have been combined in some fashion. Parcels may be classified as follows:

Total Disaggregation: All individual items are used to directly represent the latent construct.

Partial Disaggregation: Individual items are combined into parcels—by summing or averaging, for example—and these groupings are then used to represent the underlying latent construct.

Partial Aggregation: This assumes the latent construct has at least two distinct aspects. Individual items are grouped such that each parcel represents one such aspect.

Total Aggregation: This assumes the latent construct has at least two highly correlated aspects or that the latent construct causes all variation in the individual items other than measurement error. Items are grouped such that each parcel represents one such aspect.

If there are not aspects, as in the latter assumption, then the items could be randomly assigned to parcels (Bagozzi & Edwards, 1998; Coffman & MacCallum, 2005).

Misspecified parcels may bias estimates even though still producing acceptable values on fit indices (Hall, Snell, & Foust, 1999; Kim & Hagtvet, 2003). Because decisions about how to parcel may be mandated by the characteristics of the data, it is important to be aware of the relationships among the items to be parceled and, to the extent possible, to choose a parceling strategy that reflects these relationships. The essential advantage of aggregating item-level indicators is improvement of the overall psychometric properties of the construct of interest (Kishton & Widaman, 1994).

There is little prior research on the modeling effects of using a latent dependent variable in the context of LCM. Prior research, using simulated and real data to compare scale scores and partially disaggregated parcel models, found SEM path coefficients increased and residual variances decreased. The effect was most pronounced when parcels were unidimensional. Nonetheless, multidimensional parcels still performed better than a single scale score (Coffman & MacCallum, 2005). Other research found similar results to those of Coffman and MacCallum (2005) in that fit indices were very good and parameter estimates were substantive and fit with expectations. No direct comparison with a single score dependent variable was made, but it was thought that the advantage gained from the use of latent outcomes and resulting change in modeling measurement error allows for more subtle analyses such as multi-sample comparisons (Hancock, Kuo, & Lawrence, 2001).

Based on this information, the following research question is asked: How do estimates of independent variables change when dependent variables are specified as latent as compared to when they are specified as measured variables?

METHODS

The sample was obtained from the National Survey of Child and Adolescent Well-Being (NSCAW), a national probability sample of children entering child welfare services. The NSCAW sample was created using stratified cluster design to represent the target population as precisely as possible. The data are gathered from child welfare workers, caregivers, and children in nine strata and 92 primary sampling units, typically a child welfare catchment area. Baseline data collection took place between October, 1999, and December, 2000. For children less than 13 months old at baseline, three additional full waves of data collection were completed at approximately 18, 36, and 66 months post-baseline. An additional, reduced wave of data was collected, principally from telephone interviews, at 12 months post-baseline (ACF, 2005).

All children who were less than 13 months old at the baseline data collection were included in the analyses yielding a sample of 1,196 infants. The age limit was based on prior work done using the Longitudinal Study of Child Abuse and Neglect (LONGSCAN) data in which it was argued that children entering child welfare services between zero and 18 months of age represent a common developmental group (English, Graham, Litrownik, Everson & Bangdiwala, 2005). More practically, NSCAW data collection at the 66 month follow-up only included children up to 12 months of age at baseline (i.e., infants).

Having maltreatment allegations substantiated was not used as a criterion for inclusion. While all participants in NSCAW had at least one allegation of maltreatment, not all had a finding that maltreatment had taken place (i.e., substantiation of maltreatment). Herrenkohl (2005) argues that defining maltreatment only by substantiation probably understates the actual level of maltreatment. Moreover, LONGSCAN researchers compared the development of children who had substantiated cases of maltreatment with those where allegations of maltreatment were not founded. On 10 unique developmental measures, no differences between the groups were found. Based on these findings and a review of the literature, they argue that, from a developmental perspective, these two groups should not be distinguished from each other (Hussey et al., 2005). The term 'maltreated' is used for the sake of efficiency and brevity, though in some cases the maltreatment was not legally substantiated.

Measures

Four dependent variables were used. Each of the three NSCAW developmental domains—cognitive, adaptive, and communication—were included while the fourth measure was an average developmental score. All scores were standardized to a mean of 100 and a standard deviation of 15 to facilitate estimation and comparison.

Vineland Adaptive Behavior Scales Screener – Daily Living. A brief instrument used to screen children for problems in the domain of adaptive behavior and daily living skills. The Vineland Screener (Sparrow, Carter, & Cicchetti, 1993) is completed by a caregiver or other person knowledgeable about the child. The version for child ages zero to two were used at baseline with the three to five year old version used at subsequent waves as the cohort aged. The Vineland Screener strongly (.87 to .98) correlates with the full Vineland instrument. It has a Cronbach's alpha of .88 for the Daily Living screener component items only (Sparrow, Carter, & Cicchetti, 1993).

Parcels for the Vineland Screener were constructed by randomly assigning responses to the fifteen items to one of three groups of five items each. Because all items are assessing a single, common construct, no other parceling strategy is available. The number of parcels was chosen based on the observation that having several indicators per latent variable reduced the risk of problems with convergence (Bollen, 1989). Three groups of five were chosen because three evenly divides into fifteen and parsimoniously fulfills the requirement for multiple indicators. The process yields completely disaggregated, domain-representative parcels using the naming system described previously. The Vineland instrument's reliability meets the criteria established by Kishton and Widaman (1994).

Pre-School Language Scales. The PLS-3 was used to assess the developmental domain of language. It produces two sub-scales, expressive communication and auditory comprehension, and a total scale in children younger than six years old. The scores are based on observations of the child. Reported Cronbach's alpha for the expressive and receptive subscales have means of .81 and .76, respectively with mean values for the complete scale of .87 (Zimmerman, Steiner, & Pond, 1992).

The PLS-3 is administered by having the respondent answer questions until four in a row are responded to incorrectly. As a result, each respondent may answer a different number of items. This makes a parceling impossible because each case (respondent) does not have the same number of items to parcel. Scoring the PLS-3 yields two sub-scales expressive communication and receptive communication, and these were used in place of item parcels. As a result, the parcels may be considered to be partially aggregated, but unidimensional using the nomenclature described previously. The reliability estimates for this scale are acceptable using Kishton & Widamans (1994) criteria.

Battelle Developmental Inventory and Screening Test. The BDI was used to assess the developmental domain of cognitive development in children younger than five years old. It produces scores for 4 sub-domains and a total score. It is administered by an examiner.

Despite the fact that the BDI does not require training for the administrator, it has a test-retest reliability of greater than .90 in most domains and in the total score (Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984).

The BDI was parceled by its four sub-domains; Perceptual Discrimination, Memory, Reasoning, and Conceptual Development, which have been previously established for the BDI (Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984). The items were summed to produce a score for each parcel. This yields parcels that are partially disaggregated and unidimensional (i.e., homogeneous) using the nomenclature described above.

Kaufman Brief Intelligence Test. The K-BIT was used to assess cognitive development in children older than four years. The K-BIT assess 4 sub-domains as well as provides a total score. It is a self-administered, paper and pencil instrument. The test-retest reliability of the K-BIT varies by construct considered, but ranges from .74 to .95 (Kaufman & Kaufman, 1990). Even though the K-BIT is believed to assess the same developmental domain, cognitive development, as the BDI, it was not included in the analyses reported later (other than imputation) because it is not organized along the same sub-domains and has differing items. Consequently, creating a parallel parceling system is not possible.

Composite Developmental Variable. A variable was constructed to measure overall development of the child. The standardized scores for the cognitive, language, and communication domains were averaged to yield a single variable. This composite variable is also scaled to a mean of 100 and a standard deviation of 15. It should be clearly understood, however, that development across the three domains of development in NSCAW does not occur in parallel. As will be observed in subsequent analyses both the scores and trajectories of the developmental outcomes vary across the domains. Given this heterogeneity of data, the meaning of results obtained using the CDM should be carefully interpreted because it is unlikely to clearly apply to any specific domain of development that were used in it's construction. When specified as a latent variable, the scale scores of each domain served as indicators.

The following were independent variables used in either the imputation model or the analytic model to predict developmental scores.

Demographics. Race/ethnicity and gender variables were constructed. Race/ethnicity was a binary variable, minority or non-minority, or a categorical variable consisting of the following levels: black, Hispanic, white, other. Gender was coded as a binary variable.

HOME-SF Scales. Home Observation for Measurement of the Environment (Short Form) was used to assess emotional nurturing as well as cognitive stimulation (Bradley & Caldwell, 1984). Internal consistency for the total scales is .89 with a median of .74 for the subscales. Stability for the total scales is r=.62.

Maltreatment. Maltreatment data was collected from the child welfare workers and case data. Several variables were created based on the dimensions of maltreatment suggested to be important (English, Bangdiwala, & Runyan, 2005). First, the most serious type of maltreatment was identified. To facilitate estimation, most serious type of maltreatment was dummy-coded as neglect or 'other' maltreatment with abuse (primarily physical and emotional) serving as the reference level. Second, severity of harm, as judged by the child welfare worker and rated as none, low, moderate, or severe, was coded as 1 to 4, respectively. Third, a dichotomous variable was created to indicate whether one or more than one type of maltreatment had occurred. Some children experienced multiple types of maltreatment in which case the most serious was coded as the primary type (used to identify type in this analysis). Age of first episode of maltreatment was not included because all children in the study are defined by being in a common age group.

Socioeconomic Status Characteristics. Three variables represented the socioeconomic status of the infant's home of origin or of the foster home in which the child resided at baseline. A key concern is that SES lacks a consistent definition, both conceptually and operationally (Mueller & Parcel, 1981). In the analyses detailed below, SES is defined using three variables to represent the social capital model of SES (Bradley & Corwyn, 2002; Coleman, 1988).

Household Income. Annual household income is a scale of 1 to 11. Each number is a 5,000 dollar increment (e.g., 2 represents 5,001-10,000 dollars) while 11 represents any income over 50,000 dollars.

Index of Social Capital Indicators. Six indicators were used to construct an index indicating SES. Items included a primary caregiver having has at least some education beyond high school, being part of a first generation immigrant family, having a low-skill or unskilled type of job or having held such a job recently if unemployed, being unemployed, receiving one or more types of social assistance in the household, and being a single (i.e., not married or in a stable romantic relationship) caregiver. Organizing these items into a scale rather than retaining them as individual independent variables was necessary to avoid model estimation problems as well as the tendency for more complex structural equation models to exhibit a better fit than simpler models (Preacher, 2003). There is no generally accepted means of specifying SES in statistical models (Bradley & Corwyn, 2002).

Poverty. The household having total income below the official poverty line for that year and household composition is included as a dichotomous variable with not in poverty as the reference. An income-to-needs ratio was also created using household income divided by the dollar value of the poverty line for the household composition and year (USCB, 2006). Because household income was indicated as being in a given range of values, the midpoint of each range was used.

Setting. In child welfare, setting is the type of residence the child resides in. It may be in-home with nuclear family of origin, in kinship care with a relative or close family friend, in foster care, or in residential care of some sort such as group home or residential treatment facility. The data used collected at baseline and most likely reflects where the child was placed early in the case. The placement may change as the case proceeds, but the setting variable identifies where the child was initially placed. Setting was collapsed to in-home (INH) or foster care (FC) to facilitate estimation with in-home placement as the reference. Kinship foster care is considered a foster care placement despite placement with a relative being considered in-home under the TANF program.

Child Health Scale. This variable was created using the current caregiver instrument that asked the caregiver to rate the child's health from 1 (poor) to 5 (excellent).

Caregiver Health and Mental Health Problem. Caregivers were administered the SF-12, which assess general health and mental health (Ware, Kosinski, & Keller, 1998). Two variables were created—one each for health and mental health—using the standardized results.

Service Receipt. Both child welfare workers and caregivers identified services provided to the child. These data were used to create dichotomous variables indicating a need for developmental services, having an Individual Family Service Plan (IFSP), and receipt of services based on the worker-reported information.

Number of Children in the Household. The number of persons under age 18 residing in the household was used only during imputation modeling.

NSCAW Analysis Weight. This variable is unique for each case and is designed to produce results which are representative of almost all American children involved with child welfare systems. The details of its creation are given by the Data File Users Manual (Research Triangle Institute [RTI], 2007).

Stratum. This indicates which of the nine strata in NSCAW the data came from. Each of the first eight strata is a single, large population state while the ninth stratum includes data from several smaller states. It is part of the complex survey sample data necessary to correctly account for clustering in NSCAW data (RTI, 2007).

NSCAWPSU. This indicates which primary sampling unit (PSU) the data came from. Each PSU was typically a child welfare agency serving a geographical locale (normally a county). It is also part of the complex survey sample design information necessary to correctly account for clustering in NSCAW data (RTI, 2007).

RESULTS

Two distinct sets of methods were required for completion of analyses. First, multiple imputation was completed to address the problem of missing data and a problematic metric of time. Because expectation-maximization (EM) analytic algorithms are not available for data which include weighting and complex sample design variables (i.e., stratum and PSU), multiple imputation was the only model-based option available. Second, Latent Curve Models (LCM; Bollen & Curran, 2005) using latent dependent variables were estimated for trajectories of all three developmental domains as well as the composite developmental indicator.

Developmental data collected during the baseline data collection was not used for two reasons. First, NSCAW researchers have acknowledged problems in the collection of developmental data at baseline. Data collectors in the field failed to correctly gather data on many occasions, leaving questions about the validity of the data gathered (RTI, 2007). Because short, screening versions of the instruments were used, even problems with a few questions are sufficient to cast the final scores into doubt. Second, assessment of infants' (i.e., children one year or less in age) development is a challenging task requiring a modicum of skill. Some have questioned the appropriateness of using inexpert, though not untrained, data collectors in the gathering of infant developmental data (Barth, Scarborough, Lloyd, & Casanueva, 2007). As a result, infant developmental data gathered at baseline was not used.

Imputation Modeling

Prior to attempting multiple imputation (MI), missing-ness and ignore-ability of missing data must be determined by the researcher (Allison, 2002; Little & Rubin, 2002). There are no mathematical tests to detect patterns of missing-ness which may be applied (Allison, 2002). Missing data for each variable was plotted and compared to other variables. Rates varied between less than one percent and 74 percent. Missing rates were highest for the developmental measures and lowest among the independent variables. No mechanism causal of missing-ness was identified nor does the researcher know of any causal mechanism in NSCAW. As a result of these facts and research experience, missing data were judged to be at least missing-at-random (MAR) and ignorable.

The metric of time to be used in the analyses had to be created. In NSCAW, data were collected at baseline as well as 18 months, 36 months, and 66 months post-baseline. In the field of child development, a more useful metric of time is the child's age in years. Bollen and Curran (2006) describe how to change the unit of time from intervals of data collection to years of chronological age using a direct maximum likelihood estimator, but they state that MI can be used with identical results.

In MI, missing values are replaced with a distribution of m possible values based on values of other variables in the data set as well as a degree of randomness. The distribution of possible values need not be large; $4 < m < 10$ generally (Rubin, 1987). Each data set is then analyzed using complete data procedures for the modeling strategy, and results were then combined to produce a single result (Little & Rubin, 2002; Rubin, 1987).

Using MI brings with it several potential limitations. First, the data may not have been MAR despite the researcher's best efforts. Second, it is not possible to fully account for the clustering in the sample design, though recommendations from Allison (2002) were followed. Finally, the imputation model may lack some degree of validity. Guidance from key texts (Little & Rubin, 2002; Rubin, 1987; Allison, 2002) and more recent papers (Croy & Novins, 2005) was minimal when theoretically important variables demonstrate low covariances. Because EM analytic algorithms are not available for data which include weighting and complex sample design variables (i.e., stratum and PSU), multiple imputation was the only model-based option available.

Multiple imputation was completed in SAS 9.0.1 using Proc MI (SAS Institute, 2006a) using Markov Chain Monte Carlo (MCMC). In keeping with the recommendations of experts (SAS Institute, 2006b; Little & Rubin, 2002), data from the dependent variables, as well as the independent variables to be used in the analytic models, were included in the imputation model, and variables dropped from analytic modeling were retained in imputation modeling. A complete set of developmental values was possible only for years of life 2, 3, 4, and 6 years old, so final imputed data sets contain developmental values for those years as well as the independent variables of interest.

Several variables used in analytic modeling were simple combinations of other variables and were calculated after MI was completed. The composite developmental variable was the mean of the three standardized developmental indicators. Variables moderated by poverty were created by multiplying the variable of interest by the income-to-needs ratio.

SAS Proc MI using MCMC executed the imputation model without errors or warnings, and convergences were achieved in fewer than 30 iterations of the Proc MI algorithms. Ten imputed data sets were created. The computed variables were added using SPSS 12.0.0 (SPSS Inc., 2003), and the data sets were separated into individual files and converted to a format usable by Mplus.

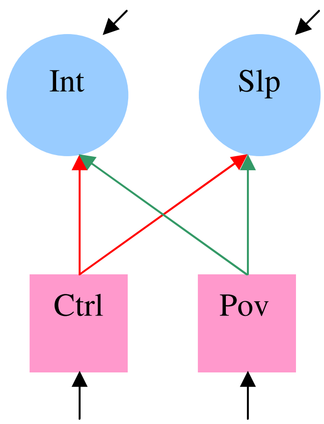
Analytic Modeling

All analyses were executed using Latent Curve Modeling (LCM; Bollen & Curran, 2006). This methodology is based on structural equation modeling (SEM; Bollen, 1989) and makes calculation of fit indices possible. Models were estimated using Mplus version 4.21 (Muthen & Muthen, 1998-2007a), which allows combination of the results for each of the ten data sets created during the imputation step. A robust maximum likelihood estimator was used (see Muthen & Muthen, 1998-2007b) with weighting, stratification, and clustering data included.

Modeling was completed in a stepwise fashion. Initially, an unconditional model was fit to the data (see Figure 4.1). During this step it was determined whether the LCM was viable for the data in question. The dependent variables were developmental scale scores at each time point. If necessary, a non-linear model was attempted. After a model was identified as fitting the data, a set of controls were added and each of the five substantive models of poverty being tested were fit to the data (see Figure 4.2).

Figure 4.2: Modeling Effects of Poverty on the Slope and Intercept Factors in LCMs[5]

Simple and Ratio Poverty



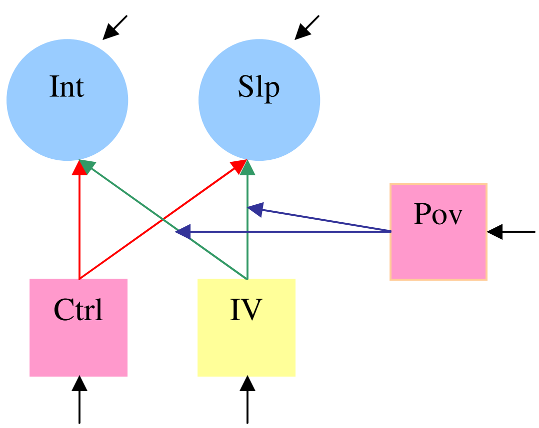
IV: Independent Variables

Ctrl: Control Variables

Pov: Poverty Variable

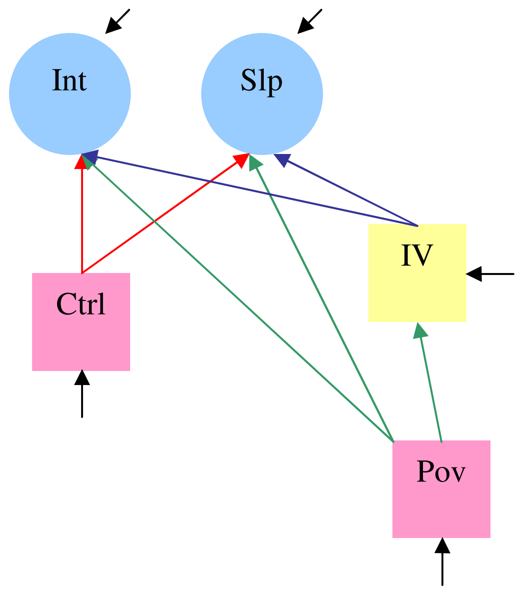
SES: Socio-Economic Status Variable

Moderated Poverty

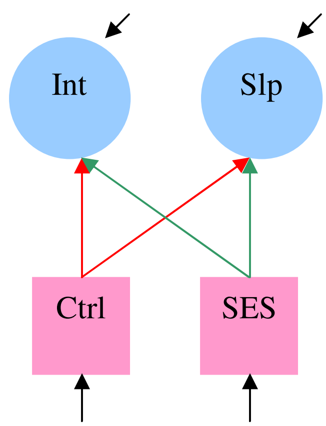


Mediated Poverty

[5] Dependent variables and their paths are omitted for the sake of clarity and space.

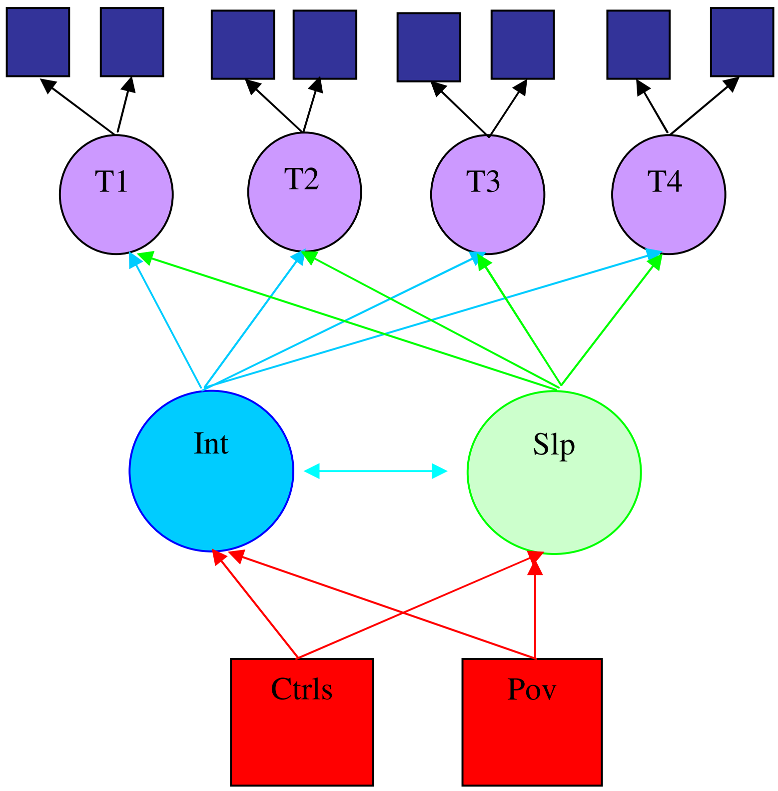


Socioeconomic Status



Modeling was completed using latent variables with multiple indicators as dependent variables in place of manifest scale scores. The SEM path model may be seen in Figure 4.3. Squares represent parcels being used as indicators and circles are the latent constructs at each time point. Each model had a different specification of poverty and this is represented in Figure 4.3 by a single square labeled ‘pov’. As described by Bollen and Curran (2006, pp. 245-254), these models are similar to second order factor models. Intercepts were fixed to zero for the indicator serving as a reference (Sayer & Cumsille, 2001).

Figure 4.3: LCM Using Latent Dependent Variables[6]



All models included controls for case and demographic characteristics. Maltreatment was modeled by including type, severity, and experiencing more than a single identified type. The reference for type of maltreatment was abuse. The reference for gender was female. The reference for race was non-minority. Finally, the reference for child welfare placement type was in-home (INH). Kinship care was considered a foster care (FC) placement in most circumstances.

Evaluation of models is based on several criteria. First, the model had to produce acceptable parameter estimates. Ideally, no improper solutions are reached by the EM, though they may be produced even when the model is a good fit to the data (Bollen & Curran, 2006). In some instances a small, non-significant negative residual was tolerated.

[6] Error terms have been omitted.

A second important indicator of fit is the presence of reasonable estimates. That is, the magnitude and sign of the estimates should be appropriate to the data in use. In planning, this was to be evaluated in part using confidence intervals. Due to the use of MI, however, this was not possible. Statistical significance was used instead, though it provides different information. Expectable sign, substantive magnitude, and statistical significance do not necessarily indicate good fit, but are useful to consider with other available data (Bollen & Curran, 2006; Schumaker & Lomax, 2004).

Finally, as with most SEM-based results, fit indices are produced. Mplus 4.21 produced four fit indices; Comparative Fit Index (CFI; Bentler, 1990), Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), Standardized Root Mean Square Residual (SRMR; Joreskog & Sorbom, 1981; Bentler, 1995), and Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993; Steiger & Lind, 1980). Using the CFI and TLI, greater than .94 indicates a good fit while a result greater then .90 indicates acceptable fit. When using the SRMR and RMSEA, less than .06 indicates a good fit while less than .09 or .11, varying slightly by author, indicates an acceptable fit (Bollen & Curran, 2006; Garson, 2007; MacCallum, 2005; Hu & Bentler; 1999; Schumacker & Lomax, 2004; Bollen, 1989). Because of the use of MI, results reported below are means and their standard errors of the fit indices.

Fit indices may offer differing assessments of fit. Each fit index uses a differing conceptual and mathematical definition of what ‘good fit’ is. The TLI and CFI define good fit by comparing model-implied values to a null model in which the covariances are zero. In practice, this system has several problems. First, the null model has no substantive meaning. It simply serves as a reference against which to compare the model in question (Garson, 2007;

Schumaker & Lomax, 2004; Rigdon, 1996). This model has no substantive meaning, and the immediate consequence of this is that fit indices based on the null model are of somewhat questionable substantive meaning. A further concern is that the more the sample covariance matrix has elements close to zero, the less meaning the TLI and CFI have because there is simply very little or no relationship to explain (Garson, 2007). Because of the need for fidelity to theory-implied relationships in the modeling reported and the need to retain all controls so that the models are comparable, there are low covariances in the sample matrix on which the analyses of poverty are based.

The practical consequence is that preference is given to the RMSEA and the SRMR, which do not depend on null models for comparison. The RMSEA is essentially a measure of discrepancy per degree of freedom in the model. Ideally, this ratio should be low and result in a value less than .10 (Garson, 2007; Rigdon, 1996; MacCallum, 2005). Similarly, the SRMR is computed by taking the square root of the mean of the squared residuals when the implied covariances are subtracted from the observed covariances (MacCallum, 2005; Garson, 2007). As has been shown, their meaning is substantively clear and readily interpretable compared with those obtained using a null model comparison system (Garson, 2007; Rigdon, 1996; MacCallum, 2005). The TLI and CFI are reported to fully disclose findings and permit alternative interpretations.

Fit indices may also understate the fit of a model to the data in the context of latent curve modeling. The form (e.g., linear, curvilinear) chosen for the latent curves is not intended to be a perfect fit to all data. Instead it is chosen because it is a good approximation of that data. This lack of perfect global fit (as opposed to component fit), despite being intentional or at least acknowledged, often results in understated fit indices (Coffman &

Millsap, 2006). Global fit may be considered with *component fit* (Bollen, 1989), which is how well the individual substantive parameters fit as assessed by an appropriate sign and a substantive magnitude of the independent variables.

Results reported in chapter two were used for comparison and employed nearly identical procedures. The essential differences are use of a single imputation model with all three sets of developmental scale scores and manifest scale scores were used as dependent variables in the latent curve models. Details of the procedures may be found in chapter two. Using the terminology outlined above, these scale scores previously obtained are fully aggregated.

Findings

SAS Proc MI executed the final model without errors or warnings, and convergences were achieved in less than 30 iterations of the Proc MI algorithms. A total of ten imputed data sets were created using the variables described previously. Imputations were not done for the dichotomous poverty variable. Rather, cases with missing data on this variable were manually completed using data from the income-to-needs ratio for that case. This avoided the possibility of conflicting data in the two measures. As a result, differences in estimates between the simple and ratio models of poverty are more likely to result from substantive differences rather than a consequence of conflicting data.

Analytic modeling was completed for three of the four dependent variables. The unconditional composite LCM with latent dependent variables did not fit the data. Both linear and non-linear models were attempted. Consequently, no further modeling was completed using the CDM.

Results of LCM with latent dependent variables for each of the three domains of development; adaptive behavior, cognitive development, and communication, are presented separately. They are accompanied by results using MVs as dependent variables obtained in chapter two. Descriptive statistics of the independent variables may be found in chapter two as well.

Adaptive Behavior

Using a latent dependent variable (LDV) model, the following results were obtained for the moderated poverty model, which was preferred in prior work (see chapter two). Adaptive behavior development modeled using a measured scale (MDV) as the dependent variable is reported to facilitate comparison. Predictors have been fully standardized using the latent and measured variables' variances to make them comparable (Table 4.1).

Table 4.1: Fit Indices for the Moderated Poverty Model for Adaptive Behavior Development Using Latent and Manifest Dependent Variables

Model Type	**CFI**	**TLI**	**RMSEA**	**SRMR**
Moderated Poverty (LDV)	.759 (.026)	.704 (.032)	.039 (.003)	.051 (.002)
Moderated Poverty (MDV)	.882 (.045)	.764 (.090)	.041 (.007)	.038 (.007)

Table 4.2: Standardized Estimates of Predictors for the Adaptive Behavior Development Model

Moderated Poverty (LDV)	**Estimate**	**Moderated Poverty (MDV)**	**Estimate**
Mod Parent Health (Int)	-.359	**Mod Parent Health** (Int)	-.421
Mod Child Health (Int)	.353	**Mod Child Health** (Int)	NS
Mod Cog Stim (Int)	NS	**Mod Cog Stim** (Int)	.198^
Age (Int)	.359	**Age** (Int)	N/A
Male (Int)	-.213	**Male** (Int)	-.258
Minority (Int)	.116	**Minority** (Int)	NS
Outhome (Int)	-.136	**Outhome** (Int)	-.162
Neglected (Int)	.159	**Neglected** (Int)	.183
Other Maltx (Int)	NS	**Other Maltx** (Int)	.117^
Intercept R-Square	0.34	**Intercept R-Square**	0.19
Age (slp)	-.650	**Age** (slp)	NS

Slope R-Square	0.71	**Slope R-Square**	Heywood

^: p<.11; NS=Not Significant; N/A: Not Applicable; Age was not in the MDV models.

The moderated poverty model was the preferred model of poverty for adaptive behavior development using LDVs. Other models of poverty failed to produce significant results for the indicators of poverty. There was evidence that parental mental health mediated the effect of poverty on the slope factor; otherwise the previously seen pattern of poverty affected parent-level variables while only child-level variables affected intercept. In the MDV models, other models failed to produce substantive parameter estimates for the indicators of poverty and in the case of mediated poverty, fit more poorly and did not demonstrate the theoretically predicted relationships among the slope, intercept, and predictor variables.

In contrast to the findings presented by Coffman and MacCallum (2005), the LDV modeling did not produce better fit indices or larger predictor estimates. Estimates of the fit indices were more precise as evidenced by their smaller standard errors. The pattern of significant predictor estimates has changed as well. A greater number of significant loadings were found, and those that were approaching statistical significance ($p<.11$) in the MDV model were not significant in the LDV models. In addition, two additional variables appeared to be significant in the MDV model that were not in the LDV model—minority status and moderated child health. The r-squared values of the LDV model were both successfully estimated (i.e., improper solutions were not obtained) and the values were markedly higher than any of those produced by the MDV models. A final point is that in all models of poverty when latent dependent variables were used, all ten imputed data sets were successfully estimated. By contrast, in the MDV models, one or sometimes two data sets were not successfully estimated.

Cognitive Development

Cognitive development proved difficult to model. In the LDV models, only three time points are available. At year 6, the K-BIT replaced the BDI as the instrument used to assess cognitive development (Research Triangle Institute [RTI], 2007). As a result, parcels could be constructed for only three time points—years 2, 3, and 4—during which the BDI was used. Little change occurred in cognitive development scores during this time (see chapter two).

Models using latent dependent variables did not demonstrate any more success in modeling the cognitive development data compared to the manifest outcome models. In fact, in the LDV models there were no significant predictors of either the slope or intercept factors. In addition, the LDV model produced a non-significant slope factor and an improper solution for the r-squared values of both the slope and intercept factors. Both of these further indicate poor fit. Using the LDV model, fit improved using the TLI and (slightly) the CFI, but worsened in the RMSEA and SRMR, though both still indicated good fit. As seen previously, precision of the fit indices improved in the LDV models..

Table 4.3: Fit Indices for the Income-to-Needs Poverty Model for Cognitive Development Fit Using Manifest and Latent Dependent Variables

	CFI	**TLI**	**RMSEA**	**SRMR**
Ratio Model (MDV)	.912 (.033)	.823 (.066)	.032 (.006)	.036 (.006)
Ratio Model (LDV)	.913 (.009)	.893 (.011)	.041 (.002)	.051 (.003)

Estimate (SE)

The lack of good fit in the LDV models is most likely attributable to a lack of a clear trajectory. The sample correlation[7] matrix of the parcels at each time point is suggestive (Table 4.4). While each parcel strongly correlates with other parcels at a given time point (as highlighted using bold, italicized, and underlined text), they do not correlate with each other

[7]The correlation matrix rather than the covariance matrix is used for clarity.

across time points. For example, a sub-domain score at Year 2 does not correlate with that same sub-domain score at Year 3 or 4.

The most plausible conclusion is that the sub-domains are closely related but that the infants do not have a common trajectory. If this is the case, then latent dependent variables will do no better in modeling the data than a single measured variable. The key difference is that by parceling in a unidimensional fashion, enough information is provided to form a specific hypothesis as to why the modeling was not successful.

Table 4.4: Correlation Matrix for the Four Unidimensional (Homogeneous), Partially Disaggregated Parcels of Latent Cognitive Development

	PARCEL 2A	*PARCEL 3A*	PARCEL 4A
PARCEL 2A	**1.000**		
PARCEL 3A	0.148	*1.000*	
PARCEL 4A	0.165	0.207	1.000
PARCEL 2B	**0.690**	0.123	0.165
PARCEL 3B	0.068	*0.786*	0.202
PARCEL 4B	0.023	0.258	0.781
PARCEL 2C	**0.672**	0.097	0.062
PARCEL 3C	0.084	*0.665*	0.092
PARCEL 4C	0.147	0.103	0.649
PARCEL 2D	**0.604**	0.041	0.102
PARCEL 3D	0.082	*0.699*	0.207
PARCEL 4D	0.232	0.157	0.696

	PARCEL 2B	*PARCEL 3B*	PARCEL 4B
PARCEL 2B	**1.000**		
PARCEL 3B	0.097	*1.000*	
PARCEL 4B	0.254	0.335	1.000
PARCEL 2C	**0.731**	0.055	0.126
PARCEL 3C	0.124	*0.716*	0.156
PARCEL 4C	0.205	0.107	0.653
PARCEL 2D	**0.619**	-0.005	0.142
PARCEL 3D	0.119	*0.705*	0.236
PARCEL 4D	0.257	0.150	0.698

	PARCEL 2C	*PARCEL 3C*	PARCEL 4C
PARCEL 2C	**1.000**		
PARCEL 3C	0.133	*1.000*	
PARCEL 4C	0.114	0.040	1.000

PARCEL 2D	**0.601**	0.035	0.200
PARCEL 3D	0.079	*0.613*	0.102
PARCEL 4D	0.168	0.101	0.708

	PARCEL 2D	*PARCEL 3D*	PARCEL 4D
PARCEL 2D	**1.000**		
PARCEL 3D	0.041	*1.000*	
PARCEL 4D	0.195	0.177	1.000

Bold is used for parcels at Year 2, *Italics* is for Year 3, and Underline is for Year 4. Each number is the Year and each letter is a different sub-domain.

An alternative explanation is that there are many distinct developmental trajectories in the BDI data. Each trajectory represents a small group of infants with a common set of characteristics and predictors. Taken in aggregate form—as is done here—they appear to follow no readily identifiable trajectory but further analyses might reveal these patterns (e.g., latent class analysis).

Communication Development

The preferred model of poverty using manifest dependent variables when modeling the communication domain of development was the income-to-needs ratio model. The data were diverse, and the resulting trajectory appears linear but for the scores at Year 4 in both the latent and manifest dependent variable models. The preferred model of poverty, the ratio model, did not produce many significant estimates for the predictors. The fit of the model was acceptable, however, according the RMSEA and SRMR (Table 4.5).

Modeling using latent dependent variables also resulted in the ratio model of poverty being preferred. Fit indices generally improved as did the precision of the estimates, though not necessarily by a marked amount. The LDV model produced a non-significant negative estimate of a residual, which resulted in the r-square of the slope not being calculated. The r-square of the intercept factor is larger, as expected.

Estimates of predictors (Table 4.6) became more numerous and more clearly statistically significant. The effect of being minority became definitively statistically significant and the effect of being male became significant as well. The effect of the ratio variable became smaller in magnitude, but other effects were larger.

Table 4.5: Fit Indices for the Ratio Poverty Model for Communication Development Using Latent and Manifest Dependent Variables

	CFI	**TLI**	**RMSEA**	**SRMR**
Ratio (MDV)	.821 (.048)	.676 (.087)	.060 (.008)	.037 (.009)
Ratio (LDV)	.917 (.023)	.883 (.032)	.035 (.005)	.041 (.008)

Table 4.6: Standardized Estimates of Predictors for the Adaptive Behavior Development Model

Ratio Model (MDV)	**Estimate**	**Ratio Model (LDV)**	**Estimate**
Ratio (Int)	.200	**Ratio** (Int)	.145
Minority (Int)	-.135^	**Minority** (Int)	-.161
Male (Int)	NS	**Male** (Int)	-.149
Age (Int)	N/A	**Age** (Int)	.141
Intercept R-Square	.089	**Intercept R-Square**	.114
Slope R-Square	.055	**Slope R-Square**	Heywood

^: p<.11; NS=Not Significant; N/A: Not Applicable; Age was not in the MDV models.

DISCUSSION

The analyses completed showed that the use of latent dependent variables yielded improved model fit as assessed using fit indices and their standard errors (with the exception of adaptive behavior models that had mixed results). Component fit—as assessed by the magnitude and statistical significance of the independent variables—was varied in that some estimates became larger or statistically significant while others became smaller or were no longer significant. Results were then more mixed than those of Coffman and MacCallum (2005) who were using simulated data and a simple random sample. They found both model fit and component fit improved when using a LDV model compared with results obtained using manifest variables.

Whereas Coffman and MacCallum (2005) used only a single fit index (the RMSEA), the results above indicated that the improvements in the RMSEA value are likely to be observed in other fit indices as well. These improvements were not shown by the SRMR. This may be a consequence of how the SRMR is calculated in that if one or more of the additional paths in the LDV model is not a good fit to the model, it may cause the mean of the values derived from the fitted residuals—on which the SRMR is based—to rise (Garson, 2007). The standard errors associated with the fit indices also improved suggesting greater precision of estimate. This is consistent with prior findings by Sayer & Cumsille (2001).

These findings also help illuminate the effect of the shift to LDV modeling on the estimates of the predictors, although this remains complicated. Prior research (Coffman & MacCallum, 2005) has indicated parameter estimates should be larger in magnitude compared with estimates obtained using MDV models. This proved to be only partially true. Parameters that were found to be statistically significant in the MDV models at times became smaller in magnitude in the LDV models. For example, the income-to-needs ratio variable in Table 4.6 dropped from a fully standardized value of .200 in the MDV model to a value of .145 in the LDV model, a significant drop.

Predictors that were close to significance ($p < .11$) in the MDV models were not consistent in how they changed. In the LDV model of adaptive behavior, they were clearly not significant whereas in the LDV model of communication, they became clearly significant. Based on the findings of Coffman and MacCallum (2005) and others who have investigated the effects of differing parceling strategies (e.g., Bandalos, 2002), it was expected that estimates would become greater in magnitude and become statistically significant.

The reasons for these somewhat conflicting results are not clear. It may be that the near-significant estimates obtained in the adaptive behavior model using MDVs were simply chance-produced, and the improved precision of estimates obtained using LDVs (as identified by Sayer & Cumsille, 2001) better demonstrated their true, non-significant status. Similarly, removing some chance-produced variability removed enough 'noise' to enable significant effects to be more readily identified as observed in models of both communication and adaptive behavior.

What did not change is also important. Using LDVs did not cause the form of the LCM to change. Models that were best fit by a linear form in MDV-based models were also best fit by a linear form in the LDV-based models. Similarly, when the data are heterogeneous, inconsistent, or otherwise unruly, using LDV-based modeling will not necessarily improve the robustness of the modeling procedure against producing improper solutions. The basic characteristics of the data were not altered.

More generally, latent curve modeling strategies using LDVs will not fix problems with the data. The improvement in the fit indices based on a comparison to a null model highlights this. Because the LDV models are more complex in that they have more parameters to freely estimate and the fact that the parcels used were strongly related, the model appears to fit better. In reality, this is a consequence of adding a new measurement model, where fit relative to a null model is excellent, to a structural model where fit is less precise. The result is a gain in the mean fit of the model compared to a null model as captured in the CFI and TLI, even after accounting for expected improvement because of the increase in model complexity (Bollen, 1989).

This research adds to the growing body of knowledge about latent curve models in general and the use of latent dependent variables in latent curve models in particular. Findings are generally consistent with prior research—with the notable exception of magnitude of the estimates—in that latent curve models using latent dependent variables seemed to fit better and produce a more precise set of estimates compared to models using manifest dependent variables. Fit indices were generally better as well.

This research tests latent dependent variable modeling using complex data and analyses. As seen in work by Coffman and MacCallum (2005), fit of the model improved. Fit of the component independent variables, especially those found to be near .05, was mixed and no pattern of change could be discerned. There is not sufficient information in the literature to clarify if the changes in component fit observed are idiosyncratic to MI-augmented data, latent curve models, complex sample designs, or some combination of the three. Both Sayer and Cumsille (2001) and Coffman and MacCallum (2005) used simple random samples and complete cases. Further modeling using simulated and real data is needed to continue to fill in the knowledge base and refine expectations. In the interim, models using complex data (i.e., because of sample design, use of MI, heterogeneous data, or some similar situation) and a manifest dependent variable in latent curve models should be interpreted with caution.

The results obtained above do not contradict the recommendation by Coffman and MacCallum (2005) to use latent dependent variables in latent curve models whenever possible. The findings above support using LDV models, even when a model failed to estimate, diagnosis of the problem was facilitated. Rather, they suggest there are limitations to the improvements that may be obtained using LDV modeling in some analytic contexts.

Chapter 5

Project Findings

Statistical modeling of differing conceptualizations of poverty using latent curve models and a large sample of maltreated infants produced results of interest from both substantive and methodological perspectives. Also important, results serve to highlight the links between the two perspectives. Substantive findings included initial developmental scores at approximately one standard deviation at the outset of the study period followed by varying patterns of small changes as the infants grew, as well as predictors that vary by developmental measure and population modeled. Methodological findings indicated estimates may vary meaningfully because of methodological decisions.

To facilitate comparison of differing methodologies, a common data set and some common methodologies were employed. Managing the problem of missing data was important. To avoid the reduction in sample size and risk of biased parameter estimates associated with case deletion and single imputation methods for managing missing data, multiple imputation (MI) was employed. An imputation model was developed and ten data sets, each with a differing set of values imputed, were created. These ten data sets, with all missing values being replaced by MI, were used for all analytic modeling.

For the analytic procedure, five commonly employed methods of operationalizing poverty were selected based on a review of the literature of poverty and children's development. Poverty was measured as a dichotomous variable; either the infant was in or was not in poverty. Poverty was measured as a continuous variable by constructing an income-to-needs ratio (ITNR) by dividing household income by the poverty threshold income (as determined by the Census Bureau) for that household composition. Mediated and moderated poverty models were developed using variables identified in the literature as being related to both poverty and development. Finally, a model of socio-economic status (SES) was operationalized based on the social capital model of SES (Coleman, 1986). These representations of poverty—or lack thereof—were then entered into latent curve models (LCMs) as predictors of the infants' developmental trajectories.

Key Findings

Maltreated infants' development lagged behind their normative peers. The scores of the infants at their second, third, fourth, and sixth years of life were well below the normative means across all four developmental measures. This pattern was true for the entire group of infants and the various sub-samples analyzed. It was also observed in the latent dependent variable-based models.

Developmental trajectories varied by domain of interest. Each of the three domains of development—cognitive, communication, and adaptive skills—followed a somewhat differing trajectory. Cognitive development followed a downward trajectory until the sixth year of life when the trajectory turned sharply upwards. Communication development followed a shallow linear trajectory upward. Adaptive behavior also followed a shallow upward trajectory until the sixth year of life when it dropped sharply. It appears development is not necessarily uniform or parallel across domains. Moreover, the shape of the trajectory determined in initial modeling in chapter two did not vary by sub-group or in the LCMs using latent dependent variables in the subsequent chapters.

Poverty exerted a consistent effect on the intercept factor of the trajectory, but the best fitting model of poverty was not the same for all three domains of development or the composite measure. In three of the four developmental outcomes, the best fitting model of poverty was the income-to-needs ratio. The exception was adaptive behavior development in which the best available model involved using poverty as a moderator of other developmental predictors. These findings did not change across sub-groups and in second order models.

Other specific predictors of the infants' developmental trajectories varied widely as well. Models involving the full sample of infants had one or more predictors trending (i.e., close to the $p<.05$ threshold) towards significance (see Tables 2.8, 2.14, 2.20, and 2.26). Once infants were organized into sub-groups, predictors identified in the full sample models were different from those found to be significant in the sub-samples. For example, while the ITNR, multiple types of maltreatment, placement into foster care, and gender were significant in the complete sample of the composite developmental measure, when the group was divided by gender only the ITNR and multiple types of maltreatment were significant in males and the ITNR and foster care in the female group. This phenomenon was repeated when the complete sample models were re-estimated using latent dependent variables.

These results indicate that, in addition to varying by domain, the predictors of the infants' developmental trajectories often vary by the specific make up of the sample involved. Significant predictors identified in the full sample may or may not appear in the sub-samples. As a result, interpretations using the complete sample are strongest when the resulting estimates are well below the $p<.05$ threshold and the magnitude of the estimate is large. Estimates that are at or near the $p<.05$ level or substantively small in magnitude may or may not appear in a more homogeneous sub-sample. Specifically, poverty—when measured by the ITNR—appears in nearly all the complete and sub-sample groups as a significant predictor of at least the intercept factor of the trajectories and a substantive negative effect for being male on the intercept factor.

Using a latent dependent variable with parceled indicators in place of a single measured variable as the dependent variable also caused resulting estimates of the predictors to change. In models of both adaptive behavior and communication, in the LDV models, there were no predictors trending towards significant (that is, near the p=.05 threshold). Rather, the indicators either became unambiguously not significant (as in the adaptive behavior modeling) or significant (as in the communication modeling). In addition, in the adaptive behavior model, both child health moderated by poverty and minority status were not significant in initial models using a manifest outcome but were significant in the LDV modeling in chapter four. Gender followed this pattern in the communication model. These findings are interpreted as evidence of improved precision of estimates as suggested by prior research (e.g., Coffman & MacCallum, 2005).

The parceling strategies that were necessitated by the data make more definitive statements about second-order modeling difficult. Whereas the PLS-3 had two well-defined sub-scales that had to be left partially aggregated and unidimensional because of the system for administering and scoring it, the VABS had no sub-scales (i.e., all individual items assess a single latent construct) and resulting parcels were disaggregated, though also unidimensional. The result was that there were two unidimensional parcels with each assessing a different dimension of the PLS-3 compared to three undimensional parcels with all three assessing a common dimension of the VABS. Had they used an identical system, it is not clear what the results would have been; however they both produced similar effects in the second order LCM modeling completed here.

When analyzing the complete sample of NSCAW infants—as was done in chapter two—the use of second order latent curve models seems well-advised. As was discussed in chapter four, several findings from chapter two did not hold when replicated using LDV models in chapter four. Interpretation of effects that are significant at the $p>.05$ and $p<.10$ is risky for the complete sample of infants. However, when smaller sub-samples are analyzed, few marginal effects (i.e., trends or variables trending towards significance) were observed (see chapter 3 tables). A plausible conclusion is that second order models are most useful when the data are heterogeneous because such models do a better job of identifying effects common to the complete sample rather than those found only in a sub-sample that appear as nearly significant effects in the complete sample. That is, the greater precision of estimates offered by second order models is of greater value when the data are heterogeneous.

Further evaluation of second order modeling is needed to validate the results obtained in this project. Whereas prior researchers (Sayer & Cumsille, 2001; Coffman & MacCallum, 2005) used complete simple random samples, analyses completed in this project employed complex sample information and multiple imputation in addition to limited opportunities to trim models of non-significant variables due to the need to maintain a common set of data. The result is a demanding analytic strategy that is different from prior research. The finding of similar, though not identical, conclusions suggests that second order latent curve modeling is both possible and desirable in a wide array of circumstances.

The results produced in this project do allow for several general conclusions to be made. First, how poverty is measured and entered into statistical models has a potentially significant effect on the magnitude and significance of resulting estimates. This effect was clearly identified in chapter two. While the ‘best’ model of poverty was not always the same one, it appears that poverty is most ideally measured as a continuous variable using an income-to needs ratio, though data from the moderated and mediated models of poverty suggested how poverty exerts its effects and, hence, provide more information to the researcher when such models are successfully implemented.

Researchers generally operationalize variables according to their theoretical or personal preference. The comparisons in the preceding chapters are made without reference to a theoretical perspective on how poverty is best conceptualized. Despite offering greater substantive effect and better fit indices, the ITNR may not be used by researchers who find fault with its implicit assumptions about poverty. But regardless of theoretical assertions, this research highlights that how poverty is operationalized may have substantial effects on what the eventual findings are. All approaches to modeling poverty are not equal in methodological or substantive terms.

Second, infants in poverty are not a homogeneous group, so predictors of their developmental trajectories may vary significantly among various sub-groups. As was observed in chapter three, infants in a racial minority in foster care had trajectories with more negative slopes than their white peers in foster care. Similarly, among girls, those in foster care also had lower scores on the intercept factor than girls remaining in their homes of origin. Some effects, such as that of poverty, were more uniform but in these examples several substantively important effects were only detected when the infants were organized into sub-groups.

In addition to the substantive implications of the findings, this research has methodological significance as well. These findings suggest that a large, heterogeneous sample considered more representative of the population—usually a desirable trait in research—may overlook important effects. Depending on the specific interest of the researcher, a smaller, more homogeneous sub-sample may be the most desirable sample in at least some instances.

Finally, results from chapter four add further evidence for the assertion that when possible, latent dependent outcomes should be used in latent curve models. When operationalizing the outcome as latent with multiple indicators, the resulting estimates of the predictors of the slope and intercept factors seemed less ambiguous. Without knowing the 'true' values of the predictors, however, the differences cannot be shown to be more accurate. But when these findings are combined with prior research, the LDV models do appear to be an improvement in terms of identifying statistically and substantively significant effects.

The overarching implication for research into poverty and development among infants and children is that the methods used in statistical models matter. Resulting predictors of development trajectories will vary according to what decisions the researcher makes in conceptualizing and operationalizing variables used in their investigations. Further research, using both real and simulated data, is important to better identify and understand what effects methodological decisions have on resulting estimates and substantive findings.

A final consideration is that the findings reported in preceding chapters and just reviewed here can probably be generalized to all infants in poverty based on three factors. First, many infants who are maltreated also live in poverty or near poverty. Second, children living in poverty and maltreated infants share many developmental risk factors. Third, both groups of infants are likely to have experienced substantial deprivation of affective or material resources important to their development. As a result, it seems likely that there is significant overlap in their developmental experiences.

Limitations

This research has several limitations that were carefully considered when the results were interpreted. Some have been discussed in preceding chapters when they seemed especially germane. The remaining concerns are laid out below.

The sample of maltreated infants in NSCAW may not be representative of all children living in, or close to, poverty. As discussed in chapter two, younger children who reside in poverty are more likely to be maltreated, but not all poor infants and toddlers will be maltreated by any means. The NSCAW sample represents an especially developmentally challenged group of infants because of their maltreatment experiences and the presence of other developmental risks (e.g., poverty, family dysfunction).

The composite developmental measure (CDM) should be interpreted cautiously. As was observed elsewhere in this project, development does not occur in parallel across the three developmental domains that were averaged to produce the CDM. As a result, the mean tends to conceal the variation in the measures that comprise it. While it is a straightforward, conceptually simple measure of overall mental and, to a lesser extent, physical development of the infant, this simplicity was achieved using a reductionist strategy that fails to provide information about the variability of the developmental data and each domain’s unique developmental trajectories.

Initial child welfare placement was chosen to avoid the necessity for time-varying covariates in already complex and difficult to estimate models. Despite federal standards, placement changes frequently occur in many cases (Administration for Children and Families, 2006). Ideally, placement would have been represented as a time-varying covariate to better represent the on-going effect of placement type. The use of initial placement type does not reflect the effects of either the number of placement changes or other types of child welfare placements—including returning home from foster or kinship placement.

The imputation model had several limitations. First, the model itself may have been suspect. On one hand, MI methodology requires the variables included in the imputation model be influential on each other. That is, that a systematic relationship exists among them. On the other hand, if relevant variables do not show meaningful covariance when prior research or theory has indicated they should—as happened in the model used—the question arises as to whether they should be included. The consequences of excluding them are not well studied. Moreover, at a minimum, all variables to be used in the analytic models must be included regardless of their utility during imputation modeling. Guidance on decision-making from key texts (Little & Rubin, 2002; Rubin, 1987; Allison, 2002) and more recent papers (e.g., Croy & Novins, 2005) remains limited.

Second, clustering inherent in the NSCAW data was not fully accounted for in the imputation model. While clustering data (NSCAWPSU and STRATUM variables) were included in the imputation model in keeping with the recommendations of Allison (2002), this does not fully account for the effects of clustering. One series of software modules for R (The R Foundation, 2007) allows clustering in an imputation model, but as it was not clear how reliable and valid these user-created modules are, they were not used.

Third, it is possible the data was not at least MAR. The NSCAW policy of coding the reason for missing data makes this less likely. Nonetheless, it is possible some systematic mechanism not identified existed, which would confound the missing data model to some degree. Some recent research (see Allison, 2002) has indicated MI is somewhat robust to violations of the MAR and ignorable assumptions, so this may not be a substantive limitation.

The lack of component fit of some models with respect to the slope factor is also a potential limitation. Few variables predicted the slope factor despite numerous models being completed and evaluated. So, while overall fit of the models varied, fit of the slope component was not good by this criterion. The cause is most likely the result of one of several possibilities. First was variability in the data. It has been shown that if a relatively small minority of cases does not fit the model employed, fit indices will tend to indicate a poor fit (Coffman & Millsap, 2006). Alternatively, despite evidence to the contrary (e.g., Guo, 1998) it may be that early poverty does not have an effect on maltreated infants' subsequent development trajectories. Finally, the relatively small developmental changes may simply have few predictors.

Data quality at the sixth year of life for adaptive behavior is a concern. Derived scores based on the published scoring system were used in Chapters 2 and 3, and these scores suggested a sharp drop in scores between the fourth and sixth year of life and an attendant change in trajectory from upward to downward. The scoring system computes a scale score then modifies that score based on the specific age of the child in months in order to allow examiners to use the same version of the VABS for a range of ages. However, the parcels for the second-order LCMs in Chapter 4 are based on domain representative parcels of the raw items. Based on the means of the parcels, this phenomenon was not observed and scores appeared to continue to rise. This contrast may indicate an error in scoring the VABS.

While not a limitation *per se*, it should be emphasized that predictors identified and discussed here are predictors of the developmental trajectory of the maltreated infants composed of the intercept and slope factors in the latent curve models. This is not equivalent to discussing what predictors might influence developmental scores during year six or any point in time of the study beyond the intercept. Predictors of the slope of the model might overlap with predictors of developmental scores at a specific point in time, but this is not required nor is it required that the predictors have similar magnitudes in both models.

Similarly, causality is not necessarily established. Because the predictors of the trajectories were chosen in part because they occurred prior to the second year of the infants' lives (the first time point), this is not sufficient to establish causality. Covariation and reasonable alternative explanations must also be established. The models presented here focus on development as the outcome while evaluating poverty and controlling for demographic and maltreatment characteristics. These are certainly important potential causes of development as the reviewed literature demonstrates, but hardly represent all potential predictors of development. Additional risk factors for poor developmental outcomes exist and are discussed elsewhere (e.g., Barth, Scarborough, Lloyd, & Casanueva, 2007). Because reasonable alternatives are not ruled-out, causality is suggested but not firmly established by these models.

Future Directions

Substantively, the development of maltreated infants remains an under-investigated subject. Given the increasing importance being placed on the first years of life, surprisingly little is known about development in children experiencing maltreatment, poverty, and other risks to development. Research presented previously strongly suggests variation by domain of development and specific characteristics of the infants. Further research is necessary to clarify when and how risk factors like poverty and maltreatment exert their influence on development.

From a methodological perspective, how variables are operationalized and what analytic methods are employed have been shown to have an effect on many parameters being estimated. This may seem a statement of the obvious, but the results presented in preceding chapters clearly indicate that these decisions may have significant implications for the resulting estimates and findings and, when applicable, model fit as well. This is most clearly shown in chapter four comparing models estimated using latent and manifest outcomes. Researchers might consider the effects of their methodological decisions when making substantive interpretations of their findings because, at least in the case of poverty, how the concept of interest is measured appears to influence what effects are found as a result of analysis. Researchers have observed a lack of uniformity in operationalizing poverty (McLoyd, 1998), maltreatment (Commission on Behavioral and Social Sciences and Education, 1993), and other variables of interest. Further research is key to fostering a better understanding of methodological decisions made in prior research, as well as to making recommendations for how to best operationalize these important concepts in quantitative research.

Finally, as a result of investigating how methodology affects outcomes in maltreated infants, a number of substantively significant findings were identified. In particular, foster care seemed a detrimental experience for girls when compared to boys. Infants classified as a racial minority in foster care had more negative developmental trajectories than their non-minority counterparts. Both effects were substantively significant and should be further investigated to better understand why these effects are occurring since it is likely there is an unidentified mediator at work.

An unequivocal effect identified is the detrimental effect of poverty on the development of maltreated infants even after their maltreatment and demographic characteristics are controlled for. While the ‘mechanism of action’ (to borrow a medical concept) of poverty’s effect on development is not yet well understood, this study adds more evidence for the assertion that alleviating poverty among the youngest and most at-risk children should be a key component of any evidence-based model of social intervention to improve both short-term and distal developmental outcomes.

References

Aber, J. L., Bennett, N. G., Conley, D. C., & Li, J. (1997). The effects of poverty on child health and development. *Annual Review of Public Health, 18,* 463-483.

Administration for Children and Families (2005). *NSCAW CPS sample component wave 1 data analysis and report.* Washington, DC: United States Government.

Administration for Children and Families (2006). *Background paper: Child and family services reviews national standards*. Retrieved March 28, 2008 from http://www.acf.hhs.gov/programs/cb/cwmonitoring/legislation/background.htm.

Administration on Children, Youth and Families (2007). *Child maltreatment 2005*. Washington, D. C.: U. S. Government Printing Office.

Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: SAGE Publications, Inc.

American Academy of Pediatrics (2000). Developmental issues for young children in foster care. *Pediatrics, 106,* 1145-1150.

Bagozzi, R. P. & Edwards, J. R. (1998). A general approach for representing constructs in organizational research. *Organizational Research Methods, 1,* 45-87.

Bandolos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling, 9,* 78-102.

Barth, R. P., Scarborough, A., Lloyd, E. C., Losby, J., Casanueva, C., & Mann, T. (2007). *Developmental Status and Early Intervention Service Needs of Maltreated Children.* Washington, DC: U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation.

Beeghly, M., Martin, B., Rose-Jacobs, R., Cabral, H., Heeren, T., Augustyn, M., Bellinger, D., & Frank, D. A. (2006). Prenatal cocaine exposure and children's language functioning at 6 and 9.5 years: Moderating effects of child's age, birthweight, and gender. *Journal of Pediatric Psychology, 31,* 98-115.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107,* 238-246.

Bentler, P. M. (1995). EQS structural equations program manual. Encino, CA: Multivariate Software.

Berkoff, M. C., Leslie, L. K., & Stahmer, A. C. (2006). Accuracy of caregiver identification of developmental delays among young children involved with child welfare. *Development and Behavioral Pediatrics, 27,* 310-318.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: John Wiley & Sons, Inc.

Bollen, K. A. & Curran, P. J. (2006). *Latent curve models: A structural equations perspective.* Hoboken, NJ: John Wiley & Sons, Inc.

Bornstein, M. H., Hahn, C., Gist, N. F., & Haynes, O. M. (2006). Long-term cumulative effects of childcare on children's mental development and socioemotional adjustment in a non-risk sample: The moderating effects of gender. *Early Child Development and Care, 2,* 129-156.

Bradley, R. H. & Caldwell, B. M. (1984).The relation of the infants' home environments to achievement test performance in first grade: A follow-up study. *Child Development, 55*, 803-809.

Bradley, R. H. & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology, 53,* 371-399.

Bradley, R. H., Corwyn, R. F., McAdoo, H. P., & Coll, C. G. (2001). The home environments of children in the United States part I: Variations by age, ethnicity, and poverty status. *Child Development, 72,* 1844-1867.

Brooks-Gunn, J. & Duncan, G. J. (1997). The effects of poverty on children. *Future of Children, 7*(2), 55-71.

Brown, M. W. & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.) *Testing structural equation models* (pp. 136-162). Newbury Park, CA: SAGE.

Cicchetti, D. (1994). Advances and challenges in the study of the sequelae of child maltreatment. *Development and Psychopathology, 6*, 1-3.

Cicchetti, D. & Lynch, M. (1995). Failures in the expectable environment and their impact on individual development: The case of child maltreatment. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 2 risk, disorder, and adaptation* (pp. 32-71). Hoboken, NY: John Wiley & Sons, Inc.

Coffman, D. L. & MacCallum, R. C . (2005). Using parcels to convert path analysis models into latent variable models. *Multivariate Behavioral Research, 40,* 235-259.

Coffman, D. L. & Millsap, R. E. (2006). Evaluating latent growth curve models using individual fit statistics. *Structural equation modeling, 13,* 1-27.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement, 14,* 183-196.

Coleman, J. S. (1988). Social capital in the creation of human capital. *The American Journal of Sociology, 94* (Supplement), S95-S120.

Coley, R. L., Bachman, H. J., Vontruba-Drzal, E., Lohman, B. J., & Li-Grining, C. P. (2007). Maternal welfare and employment experiences and adolescent well-being: Do mothers' human capital characteristics matter? *Children and Youth Services Review, 29,* 193-215.

Commission on Behavioral and Social Sciences and Education (1993). *Understanding child abuse and neglect.* Washington, DC: National Academies Press. Available on-line by chapter at http://fermat.nap.edu/books/0309048893/html/index.html.

Cornish, A. M., MacMahon, C. A., Ungerer, J. A., Barnett, B., Kowalenko, N., & Tennant, C. (2005). Postnatal depression and infant cognitive and motor development in the second postnatal year: The impact of depression chronicity and infant gender. *Infant Behavior and Development, 28,* 407-417.

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *JAMA: Journal of the American Medical Association, 290,* 2023-2029.

Cox, J. & Jordan, L. C. (2006). Nutrition and growth. In V. L. Gunn & C. Nechyba (Eds.), *Harriet lane handbook: A manual for pediatric house officers* (16th ed., pp. 433-503). Philadelphia, PA: Mosby.

Croy, C. D. & Novins, D. K. (2005). Methods for addressing missing data in psychiatric and developmental research. *Journal of the American Academy of Child and Adolescent Psychiatry, 44,* 1230-1240.

Danese, A., Pariante, C. M., Avshalom, C., Taylor, A., Poulton, R. (2007). Childhood maltreatment predicts adult inflammation in a life-course study. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 1319-1324.

Dawson, G., Ashman, S. B., & Carver, L. J. (2000). The role of early experience in shaping behavioral and brain development and its implication for social policy. *Development and Psychopathology, 12,* 695-712.

Dearing, E., McCartney, K., & Taylor, B. A. (2001). Change in family income-to-needs matters more for children with less. *Child Development, 72,* 1779-1793.

De Bellis, M. D. (2005). The psychobiology of neglect. *Child Maltreatment, 10,* 150-172.

De Bellis, M. D. & Keshavan, M. S. (2003). Sex differences in brain maturation in maltreatment-related pediatric posttraumatic stress disorder. *Neuroscience & Behavioral Reviews, 27*, 103-117.

Dubowitz, H., Papas, M. A., Black, M. M., & Starr, R. H. (2002). Child neglect: Outcomes in high-risk urban preschoolers. *Pediatrics, 109*, 1100-1107.

Duncan, G. (1984). *Years of poverty, years of plenty*. Ann Arbor, MI: University of Michigan, Institute for Social Research.

Duncan, G. J., Brooks-Gunn, J., & Klebanov, P. K. (1994). Economic deprivation and early childhood development. *Child Development, 65,* 296-318.

Egeland, B. & Brunnquell, D. (1979). An at-risk approach to the study of child abuse: Some preliminary findings. *Journal of the American Academy of Child Psychiatry, 8,* 219-235.

Eisenberg, N., Martin, C. L., & Fabes, R. A. (1996). Gender development and gender effects. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology.* New York: Prentice Hall International.

English, D. J. (1998). The extend and consequences of child maltreatment. *Future of Children, 8*(1), 39-53.

English, D. J., Bangdiwala, S. I., & Runyan, D. K. (2005). The dimensions of maltreatment: Introduction. *Child Abuse & Neglect, 29,* 441-460.

English, D. J., Graham, J. C., Litrownik, A. J., Everson, M., & Bangdiwala, S. I, (2005). Defining maltreatment chronicity: Are there differences in child outcomes? *Child Abuse and Neglect, 29*, 575-595.

Frankfort-Nachmias, C. & Nachmias, D (2000). *Research methods in the social sciences* (6th ed.). New York: Worth Publishers.

Garrett, P., Ng'andu, N., & Ferron, J. (1994). Poverty experiences of young children and quality of their home environments. *Child Development, 65,* 331-345.

Garson, G. D. (2007). *Structural equation modeling*. Retrieved October 1st, 2007 from http://www2.chass.ncsu.edu/garson/pa765/structur.html.

Gershoff, E. T., Aber, J. L., Raver, C. C., & Lennon, M. C. (2007). Income is not enough: Incorporating material hardship into models of income associations with parenting and child development. *Child Development, 78,* 70-95.

Gowen, J. W. (1993). *Effects of neglect on the early development of children: Final report*. Washington, DC: Department of Health and Human Services.

Grantmakers in Health (2001). *Early childhood development: Putting knowledge into action.* Retrieved October 25th, 2006 from http://www.gih.org/usr_doc/earlychildhood.pdf.

Guo, G. (1998). The timing of the influences of cumulative poverty on children's cognitive ability and achievement. *Social Forces, 77*, 257-288.

Guo, G. & Harris, K. M. (2000). The mechanisms mediating the effects of poverty on children's intellectual development. *Demography, 37,* 431-447.

Hagale, D. M. (2005). The impact of maltreatment on the developing child. *North Carolina Medical Journal, 66,* 356-359.

Hall, R. J., Snell, A. F., & Foust, M. S. (1999). Item parceling strategies in SEM: Investigating subtle effects of unmodeled secondary constructs. *Organizational Research Methods, 2,* 233-256.

Hancock, G. R., Kuo, W., & Lawrence, F. R. (2001). An illustration of second-order latent growth models. *Structural Equation Modeling, 8,* 470-489.

Hansen, R. L., Mawjee, F. L., Barton, K., Metcalf, M. B., & Joye, N. R. (2004). Comparing the health status of low-income children in and out of foster care. *Child Welfare, 83,* 367-380.

Harden, B. J. (2004). Safety and stability in foster care: A developmental perspective. *Future of Children, 14(1),* 31-48.

Herrenkohl, R. C. (2005). The definition of child maltreatment: From case study to construct. *Child Abuse & Neglect, 29*, 413-424.

Herrenkohl, R. C., Herrenkohl, E. C., Egolf, B. P., & Ping, W. (1991). The developmental consequences of child abuse: The Lehigh longitudinal study. In R. H. Starr, Jr. & D. A. Wolfe (Eds.), *The effects of child abuse and neglect: Issues and research* (pp. 33-56). New York: The Guildford Press.

Herrenkohl, R. C. & Russo, M. J. (2001). Abusive early child rearing and early childhood aggression. *Child Maltreatment, 6,* 3-16.

Hildyard, K. L. & Wolfe, D. A. (2002). Child neglect: Developmental issues and outcomes. *Child Abuse & Neglect, 26,* 679-695.

Hoffman-Plotkin, D. & Twentyman, C. T. (1984). A multimodal assessment of behavioral and cognitive deficits in abused and neglected preschoolers. *Child Development, 55,* 794-802.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structural analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55.

Hussey, J. M., Marshall, J. M., English, D. J., Knights, E. D., Lau, A. S., Dubowitz, H., & Kotch, J. B. (2005). Defining maltreatment according to substantiation: Distinction without a difference. *Child Abuse & Neglect, 29,* 479-492.

James, S., Landsverk, J., Slymen, D. J., & Leslie, L. K. (2004). Predictors of outpatient mental health service use—The role of foster care placement change. *Mental Health Services Research, 6*(3), 127-141.

Jennische, M. & Sedin, B. (2003). Gender differences in outcome after neonatal intensive care: Speech and language skills are less influenced in boys than in girls at 6.5 years. *Acta Paediatrica, 92,* 364-378.

Joreskog, K. G. & Sorbom, D. (1981). *LISREL V: Analysis of linear structural relationships by the method of maximum likelihood.* Chicago: National Educational Resources.

Knutson, J. F., DeGarmo, D., Koeppl, G., & Reid, J. B. (2005). Care neglect, supervisory neglect, and harsh parenting in the development of children's aggression: A replication and extension*.* *Child Maltreatment, 10*, 92-107.

Kaufman, A. & Kaufman, N. (1990). *Kaufman brief intelligence test*. American Guidance Services, Inc.

Kim, S. & Hagtvet K. A. (2003). The impact of misspecified item parceling on representing latent variables in covariance structure modeling: A simulation study. *Structural Equation Modeling, 10,* 101-127.

Kishton, J. M. & Widaman, K. F. (1994). Unidimensional versus domain representative parceling of questionnaire items: An empirical example. *Educational and Psychological Measurement, 54,* 757-765.

Lansford, J. E., Malone, P. S., Stevens, K. I., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2006). Developmental trajectories of internalizing and externalizing behavior: Factors underlying resilience in physically abused children. *Development and Psychopathology, 18,* 35-55.

Lawrence, C. R., Carson, E. A., & Egeland, B. (2006). The impact of foster care on development. *Development and Psychopathology, 18,* 57-76.

Lee, V. & Hoaken, P. N. S. (2007). Cognitive, emotional, and neurobiological development: Mediating the relation between maltreatment and aggression. *Child Maltreatment, 12*, 281-298.

Leslie, L. K., Gordon, J. N., Granger, W., & Gist, K. (2002). Developmental delay in young children in foster care by placement type. *Infant Mental Health Journal, 23,* 496-516.

Linver, M. R., Brooks-Gunn, J., & Kohen, D. E. (2002). Family processes as pathways from income to young children's development. *Developmental Psychology, 38,* 719-734.

Little, R. J. A. & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9,* 151-173.

Lloyd, E. C. (2007, January). Modeling the Development of Infants Involved in Child Welfare Services Using Latent Trajectory Analysis. In R. B. Barth (Moderator), *New findings from the National Survey of Child and Adolescent Well-Being (NSCAW): Applying innovative methods to assess child safety and well-being outcomes.* Symposium conducted at the meeting of the Society for Social Work and Research, San Francisco, CA.

MacCallum, R. C. (2005). *Lecture notes.* Unpublished manuscript, University of North Carolina at Chapel Hill.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin, 111,* 490-504.

Maikovich, A. K. & Jaffe, S. R. (2006). *Breaking down "maltreatment": The effects of different types of maltreatment on children's functioning in multiple domains.* Unpublished manuscript, University of Pennsylvania.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figureredo, A. J. (2007). *Missing data: A gentle introduction*. New York: The Guilford Press.

McLeod, J. D. & Nonnemaker, J. M. (2000). Poverty and child emotional and behavioral problems: Racial/ethnic differences in process and effects. *Journal of Health and Social Behavior, 41,* 137-161.

McLoyd, V. C. (1998). Socioeconomic disadvantage and child development. *American Psychologist, 53,* 185-204.

McLoyd, V. C. (1990). The impact of economic hardship on black families and children: Psychological distress, parenting, and socioemotional development. *Child Development, 61,* 311-346.

Mueller, C. W. & Parcel, T. L. (1981). Measures of socioeconomic status: Alternatives and recommendations. *Child Development, 52*, 13-30.

Muthen, B. O. & Muthen, L. K. (1998-2007a). Mplus (4.21). Available at http://www.statmodel.com/

Muthen, L. K. & Muthen, B. O. (1998-2007b). *Mplus user's guide* (4th ed.). Los Angeles: Muthen & Muthen.

National Academy of Sciences (1997). *New findings on welfare and children's development.* Retrieved June 12th, 2006 from http://www.nap.edu/catalog/5479.html.

National Center for Health Statistics (2007). *Clinical growth charts*. Retrieved September 1st, 2007 from http://www.cdc.gov/nchs/about/major/nhanes/growthcharts/clinical_charts.htm.

National Clearinghouse on Child Abuse and Neglect Information (2005). *Long-term consequences of child abuse and neglect.* Retrieved April 10th, 2006 from http://nccanch.acf.hhs.gov/pubs/factsheets/long_term_consequences.cfm.

National Institute of Child Health and Human Development Early Child Care Network (2005). Duration and developmental timing of poverty and children's cognitive and social development from birth through third grade. *Child Development, 76*, 795-810.

National Survey of Child and Adolescent Well-Being Research Group (2005). *NSCAW CPS sample component wave 3 data analysis and report.* Unpublished manuscript, University of North Carolina at Chapel Hill.

Newborg, J., Stock, J. R., Wnek, L., Guidubaldi, J., & Svinicki, J. (1984). *Battelle developmental inventory: With recalibrated technical data and norms: Examiner's manual*. Riverside Publishing.

Newton, R. R., Litrownik, A. J., & Landsverk, J. A. (2000). Children and youth in foster care: Disentangling the relationship between problem behaviors and number of placements. *Child Abuse & Neglect, 24,* 1363-1374.

Pears, K. C. & Fisher, P. A. (2005). Emotion understanding and theory of mind among maltreated children in foster care: Evidence of deficits. *Development and Psychopathology, 17,* 47-65.

Petterson, S. M. & Albers, A. B. (2001). Effects of poverty and maternal depression on early child development. *Child Development, 72,* 1794-1813.

Preacher, K. J. (2003). The role of model complexity in the evaluation of structural equation models (Doctoral Dissertation, The Ohio State University). *Dissertation Abstracts International, 64,* 2979.

R Foundation (2007). R (2.51). Available at: http://www.r-project.org.

Research Triangle Institute (2003). *Data file user's manual.* Cary, NC: Research Triangle Institute.

Research Triangle Institute (2007). *Data file user's manual.* Cary, NC: Research Triangle Institute.

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling, 3,* 369-379.

Rosenfeld, A. A., Pilowsky, D. J., Fine, P., Thorpe, M., Fein, E., Simms, M. D., Halfon, N., Irwin, M., Alfaro, J., Saletsky, R., & Nickman, S. (1997). Foster care: An update. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 448-457.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. Hoboken, NJ: John Wiley & Sons, Inc.

Sameroff, A. J. (1995). General systems theory and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 1. Theory and methods* (pp. 659-695). New York: J. Wiley.

SAS Institute Inc. (2006a). SAS (9.1). Cary, NC: SAS Institute.

SAS Institute Inc. (2006b). *The MI procedure*. Retrieved July 10th, 2006 from http://www.support.sas.com/rnd/app/papers/miv802.pdf.

Sameroff, A. J. (1995). General systems theory and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 1. Theory and methods* (pp. 659-695).

Sameroff, A. J. (1998). Management of clinical problems and emotional care: Environmental risk factors in infancy. *Pediatrics, 102*, 1287-1292.

Sayer, A. G. & Cumsille, P. E. (2001). Second-order latent growth models. In L. M. Collins & A. G. Sayer (Eds.), *New methods for the analysis of change* (pp. 177-200). Washington, DC: American Psychological Association.

Schumacker, R. E. & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Slack, K. S., Holl, J. L., McDaniel, M., Yoo, J., & Bolger, K. (2004). Understanding the risks of child neglect: An exploration of poverty and parenting characteristics. *Child Maltreatment, 9,* 395-408.

Sparrow, S. S., Carter, A. S., & Cicchetti, D. V. (1993). *Vineland screener: Overview, reliability, validity, administration, and scoring*. New Haven, CT: Yale University Child Study Center.

SPSS Inc. (2003). SPSS (12.0.0). Chicago, IL: SPSS Inc.

Stahmer, A. C., Leslie, L. K., Landsverk, J. A., Zhang, J., & Rolls, J. A. (2006). Developmental services for young children in foster care: Assessment and service delivery. *Journal of Social Service Research, 33(2),* 27-38.

Stapleton, L. (2006). An assessment of practical solutions for structural equation modeling with complex sample data. *Structural Equation Modeling, 13,* 28-58.

Steiger, J. H., & Lind, J. (1980, June). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Tausigg, H. N., Clyman, R. B., & Landsverk, J. (2005). Children who return home from foster care: A 6-year prospective study of behavioral health outcomes in adolescents. *Pediatrics, 108,* 10-16.

Teo, A., Carlson E., Mathieu P. J., Egeland, B., & Sroufe, L. A. (1996). A prospective longitudinal study of psychosocial predictors of achievement. *Journal of School Psychology, 34,* 285-306.

Trickett, P. K. (1997). Sexual and physical abuse and the development of social competence. In S. S. Luthar, J. A. Burack, D. Cicchetti, & J. R. Weisz (Eds.) *Developmental psychopathology: Perspectives on adjustment, risk, and disorder*. New York: Cambridge University Press.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38,* 1–10.

United States Census Bureau (1998). *Poverty in the United States: 1997* (Current Population Reports, Series P60-201). Washington, DC: United States Government Printing Office.

U.S. Census Bureau (2001). *Census 2000 supplemental survey profile: United States.* Retrieved November 27, 2007, from http://www.census.gov/acs/www/Products/Profiles/Single/2000/C2SS/Tabular/010/01000US3.htm.

United States Census Bureau (2006). *Poverty thresholds 1999*. Retrieved July 31, 2007 from: http://www.census.gov/hhes/www/poverty/threshld/thresh99.html.

Vig, S., Chinitz, S., & Shulman, L. (2005). Young children in foster care: Multiple vulnerabilities and complex service needs. *Infants & Young Children, 18*(2), 147-160.

Ware, J., Kosinski, M., & Keller, S. (1998). *SF-12: How to score the SF-12 physical and mental health summary scales* (3rd ed.). Lincoln, RI: Quality Metric Incorporated.

Wulczan, Barth, Yuan, Harden, & Landsverk (2005). *Beyond common sense*. New Brunswick, NJ.

Yeung, W. J., Linver, M. R., & Brooks-Gunn, J. (2002). How money matters for young children's development: Parental investment and family processes. *Child Development, 73,* 1861-1879.

Zimmer, M. H. & Panko, L. M. (2006). Developmental status and service use among children in the child welfare system. *Archives of Pediatric and Adolescent Medicine, 160,* 183-188.

Zimmerman, I. L., Steiner, V. G., & Pond, E. P. (1992). *Preschool language scale-3*. The Psychological Corporation.