

Data User's Guide for the Nine-Year Follow-Up Wave of the Fragile Families and Child Wellbeing Study

October 2011

***Bendheim-Thoman
Center for Research on Child Wellbeing***
Wallace Hall
Princeton University
Princeton, NJ 08544
<http://crcw.princeton.edu>

Columbia Population Research Center
1255 Amsterdam Avenue, Room 715
Columbia University
New York, NY 10027
<http://cupop.columbia.edu/>

Prepared by the staff at the Bendheim-Thoman Center for Research on Child Wellbeing (CRCW), Princeton University. For more information about Fragile Families, please visit our web site at <http://www.fragilefamilies.princeton.edu/> or email ffdata@princeton.edu.

Table of Contents

Overview of the Nine-Year Follow-Up Data Collection	4
Sample Size.....	5
Nine-Year Completion Rates	6
Changes to Questionnaires.....	6
General Notes on File Contents and Structure.....	6
Nine-Year Weights	9
Interviewing Incarcerated Respondents	9
Non-Parental Caregiver Survey	9
I. Notes on Missing Data.....	10
II. Additional Variable Information.....	10
Primary Caregiver Survey.....	10
I. Notes on Missing Data.....	11
II. Additional Variable Information.....	11
III. Scales Information	11
Biological Mother and Father Surveys	11
I. Notes on Missing Data.....	12
II. Additional Variable Information.....	12
III. Scales Information	15
Child Survey	16
I. Notes on Missing Data.....	16
II. Additional Variable Information.....	16
III. Scales Information	16
Primary Caregiver Self-Administered Questionnaire (PCG SAQ).....	17
I. Notes on Missing Data.....	17
II. Additional Variable Information.....	17
III. Scales Information	18
In-Home Observations	18
I. Notes on Missing Data.....	18
II. Additional Variable Information.....	19
III. Scales Information	19
Height and Weight Measurements.....	19

I. Notes on Missing Data.....	20
II. Additional Variable Information.....	21
Child Assessments	21
I. Notes on Missing Data.....	22
II. Additional Variable Information.....	22
Teacher Survey	22
I. Notes on Missing Data.....	23
II. Additional Variable Information.....	23
III. Scales Information	23

Overview of the Nine-Year Follow-Up Data Collection

The Fragile Families and Child Wellbeing Study's fifth wave of data collection, around focal children's ninth birthdays, was conducted from August 2007 through April 2010. Westat, Inc., a survey research firm in Rockville, MD., hired interviews and oversaw data collection activities for this wave. This wave of data collection was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD).

The Nine-Year wave of data collection integrated interviews with 1) core biological parents, 2) primary caregivers (and in certain circumstances, a non-parental caregiver), 3) "focal" children, and 4) teachers. Home Visits were also conducted and included cognitive tests, in-home observations, a primary caregiver self-administered questionnaire, and saliva sample collection for genetic analysis.¹ Interviewers completed "In-Home Observations" of the home environment following the Home Visit.

This wave of data collection was fielded to allow researchers to answer the following questions: How do children develop over time, and how do family resources influence children's health and development? How do the resources of unmarried parents evolve over time, relative to those of married parents? How do children's genetic endowments interact with their environments to influence their outcomes? How do school environments influence children's social and academic outcomes?

These survey components were typically administered in the following order: In most cases, the primary caregiver (PCG) survey² was completed by Computer-Assisted Telephone Interviewing (CATI) followed by the core biological parent interviews. Home Visits were typically scheduled during the primary caregiver and core biological parent phone interviews. During the Home Visit, a 20-minute interview was administered to the focal child (using Computer-Assisted Personal Interview (CAPI) technology), the primary caregiver completed a self-administered questionnaire, height (focal child only) and weight (focal child and biological mother) measurements were taken, a speech sample was taken from the primary caregiver, and cognitive assessments were conducted with the focal child. Saliva samples were also collected from biological mothers and focal children. Interviewers also collected consent and contact information in order to mail hard-copy interviews to focal children's teachers.

Families were provided with compensation for their participation in the Nine-Year wave of data collection. The compensation schedule for the survey components was: primary caregiver- \$25, biological mother- \$30, biological father- \$75, child survey- \$30 (given to the primary caregiver to administer to the child as they saw appropriate), and Home Visit activities- \$65. Families were typically paid cash during the Home Visit. If they did not participate in the Home Visit, a check was mailed to the respondent.

The "Introduction to the Fragile Families Public Use Data," located at http://www.fragilefamilies.princeton.edu/documentation/core/4waves_ff_public.pdf, provides information for the previous four waves of the Fragile Families and Child Wellbeing Study. This document provides similar information for the fifth wave of data collection. It describes participation levels and provides general information on file contents and structure. The data available for each of the Nine-Year survey components are detailed in the concluding sections. Pages 9 through 15 describe the

¹ Analysis files for the genetic component will be available through CRCW's Restricted Use Contract process beginning in mid-2012. The variables, **cm5saliva** and **ck5saliva**, indicate which mothers and children contributed saliva samples for genetic analysis.

² The primary caregiver questionnaire was conducted by phone with the biological mother in situations where she or she and the biological father had custody of the "focal child" for half or more of the time. If the biological mother did not have primary custody of the child, the primary caregiver interview was conducted with the father, relative, or friend who had custody of the child half or more of the time. An additional set of questions were administered to non-parental caregiver at the beginning of the primary caregiver interview in situations where both biological parents were not the primary caregiver.

telephone surveys (Non-Parental Caregiver, Primary Caregiver, Biological Mother and Biological Father). Pages 15 through 22 describe the Home Visit activities (Child Survey, PCG SAQ, In-Home Observations, Height and Weight Measurements, Child Assessments). The Teacher Survey is described on pages 22 through 23.

While this guide also provides a brief description of some key scales and question sets in each survey, the Nine-Year Scales Documentation contains a more complete listing of the variables that were derived from other surveys or scales.

Sample Size

The following chart describes the number of completed interviews associated with each survey component in these files. A minimal number of “breakoff” (interview ended before it was fully complete) cases were included in this file; for these few cases, there was a significant amount of data collected prior to the “breakoff.”

A series of variables exists on the individual data files that identify whether or not a participant is included “in the wave” for that particular survey. Those variables are as follows: **cm5mint** (biological mother), **cf5fint** (biological father), **ck5kint** (child), **cn5nint** (non-parental caregiver), **o5oint** (in-home observations), **cp5pint** (primary caregiver) and **t5tint** (teacher).

Survey	Variable Name	Completes
Biological Mother	cm5mint	3,515
Biological Father	cf5fint	2,652
Non-Parental Caregiver	cn5nint	132
Primary Caregiver (PCG)	cp5pint	3,630
Child	ck5kint	3,377
In-Home Observations	o5oint	3,392
Teacher Survey	t5tint	2,254

Before administering the survey, interviewers determined which parent is the “primary caregiver” (PCG) for the focal child. This would determine who the PCG survey is administered to. The variable **pcg5idstat** (included on the PCG file) will either be “61 – biological mother”, “62 – biological father” (in cases where the biological mother and biological father do not live together), “63 – non-parental caregiver”, or “-9 – not in wave/not ascertained”. Mothers will be coded as the “primary caregiver” (and **pcg5idstat** will be “61”) in cases where mothers and fathers live together with the child; the survey used the mother as the default primary caregiver.

Pcg5idstat may have been captured even in cases where surveys were not completed. In these cases, a primary caregiver was identified but the survey was not completed. The distribution of **pcg5idstat** in this round of data collection is:

- PCG = Biological Mother (n=3,828); **pcg5idstat** of “61”
- PCG = Biological Father (n=180); **pcg5idstat** of “62”
- PCG = Non-parental Caregiver (n=194); **pcg5idstat** of “63”

Additionally, the variables **pcg5stat** (primary caregiver), **cm5_bmomstat** (biological mother), and **cf5_bdadstat** (biological father) provide a description of the mode of data collection (phone or in-person). These variables will have codes in the 60's when the interview has been fully completed.

For those parents who did not participate in the biological mother or biological father interviews, the variables are **cm5samp** (biological mother) and **cf5samp** (biological father) describe the reasons for their non-response.

Interviewing began in the first two pilot cities and progressively rolled-out over the following three years. Most cities had a 9-month to 1-year interviewing period; however, a few cities remained open longer. Data collection for this wave of interviewing concluded in April 2010. The following chart provides information on the number of completed interviews in each sample city.

Nine-Year Completion Rates

The following chart provides information on the response rates for the primary caregiver, biological mother, biological father, and Home Visit survey components.

	Survey Component			
	Primary Caregiver	Biological Mother	Biological Father	Home Visit
Sample Eligible³ for Interview	4688	4654	4464	4688
Completed Interviews	3630	3515	2652	3391
Complete Rate (Completes/Eligibles)	77%	76%	59%	72%

Changes to Questionnaires

A series of changes were made to the survey instruments between the first two pilot cities and the remaining 18 cities. In many cases, these were formatting changes that did not affect the substantive content of the survey questions. Values in variables where the content of the question changed between survey cities are denoted “-5 not asked” in individual data files. A binary variable indicating if the family was in the first two pilot cities, **cm2twoc** and **cf2twoc**, can be used to differentiate these cases, as well.

General Notes on File Contents and Structure

Variable coding, specifically for missing information, follows the same format as earlier waves – valid skips are coded as -6, “Don’t Know” as -2, “Refused” as -1, general missing as -3, and -9 as “Not in Wave.” To note, the “-9 Not in Wave” coding is used for each individual survey and does not refer to the entire case. For example, variables within a case can be coded as “-9” for the Home Visit and biological father interview; however, the primary caregiver and biological mother interviews may be complete, containing valid values for these surveys’ variables and codes in the 60’s for the **cm5_bmomstat** and **pcg5stat**.

³ Families are no longer considered eligible for interviewing if the child is deceased or formally/legally adopted. A portion of non-randomly selected cases were also ineligible for interviewing in the Nine-Year wave. For the parent interviews, deceased parents are not included in the eligible sample. Additionally, fathers with conflicting information about their paternity were not eligible for follow-up.

Open-Ended Response Codes - As with previous core data files, free response questions (open-ended questions) were coded by CRCW staff. When appropriate, open-ended responses were recoded into the main response categories of the questions. Open-ended responses that did not fit into the existing response categories are recoded into new categories in the 100 range (101, 102, etc). Cases that indicate an “other” but were vague or unique remain coded simply as “Other (specify).” In cases where new response categories were constructed from “Code All That Apply” questions, the binary will only have a value of “1” for those who volunteered this information.

Constructed Variables: The file contains a series of constructed variables similar to those available at previous waves and a few new constructs. Researchers may find these variables useful, but are free to construct them in other ways. Please note that in the files raw yes/no questions are coded as 1=Yes and 2=No. Constructed yes/no variables are coded as 1=Yes and 0=No.

	<i>Sample and Interview Completion Flags</i>
c*5*int	Binary indicators for completed biological mother (cm5mint), biological father (cf5fint), primary caregiver (cp5pint), non-parental caregiver (cn5nint), child (ck5kint), in-home observations (o5oint), and teacher (t5tint) interviews
c*5_bmomstat c*5_bdadstat	Type of complete for biological mother (cm5_bmomstat) or biological father (cm5_bdadstat) interview (“64”/”67” – Telephone; “65”/”66” – In-Person, “71” – Breakoff)
c*5samp	Reason for incomplete biological mother (cm5samp) or biological father (cf5samp) interview
c*5natsm	Binary indicator if parents are in the national sample
c*5citsm	Binary indicator if parents are in the city sample.
c*5twoc	Binary indicator if biological parents are in the first two cities sample
pcg5idstat	Primary caregiver type (“61- Biological Mother”, “62-Biological Father, “63-Non-Parental/Other”)
pcg5stat	Type of complete for primary caregiver interview (“64”– Telephone; “65- In-Person, “71” – Breakoff, “80- Final Refusal”, “89- Final Non-Complete Other”)
c*5saliva	Binary indicator for collection of saliva sample from biological mothers (cm5saliva) and children (ck5saliva)
	<i>Additional Description of Interviews</i>
c*5tele	Binary indicator for completing biological parent interview on the telephone
c*5span	Binary indicator for completing biological parent interview in Spanish
o5g1	Indicator for completing the Primary Caregiver Interview in Spanish
p5qlang	Indicator for whether the Primary Caregiver Self-Administered Questionnaire was in English or Spanish
cf5new108	Binary indicator for a father newly interviewed in Nine-Year wave
cm5tdiff	Time difference between biological mother and biological father interviews; only constructed when both biological parent interviews were completed
c*5intyr	Interview year for biological parent interview (“2007” through “2010”)
c*5intmon	Interview month for biological parent interview
*5_intmon	Interview month for the child assessments (hv5_intmon), interviewer observations (o5_intmon) and primary caregiver interview (p5_intmon)
*5_intyr	Interview year for the child assessments (hv5_intyr), interviewer observations (o5_intyr) and primary caregiver interview (p5_intyr)
	<i>Relationship Status Variables</i>
cm5relf	Mother’s report of biological parents’ relationship status; coded from m5a4 .
c*5mar*	Binary indicator that biological parents were married at time of that parent’s interview
c*5marp	Binary indicator that parents were married to a new partner at time of interview (only corresponds to parents asked m5d2b or f5d2b)
c*5coh*	Binary indicator that biological parents were living together (and unmarried) at time of that parent’s interview
c*5cohpc	Binary indicator that parents were living with a new partner at time of interview (only corresponds to parents asked m5d2d or f5d2d)

<i>Household and Demographic Information</i>	
cf5hhgridflag	Flag –Father was not asked about household grid because he reported being married to or cohabiting with the biological mother. Biological mother’s household grid information can be substituted for all household grid related variables.
c*5age	Parents’ age in years at time of interview
c*5b_age	Focal child’s age in months at time of biological parent interview
c*5b_ageyrs	Focal child’s age in years at time of biological parent interview
c*5adult, c*5kids, c*5gmom, c*5gdad	Number of adults (c*5adult) and children (c*5kids) living in biological parents’ household; binary indicators for whether the focal child’s grandmother (c*5gmom) or grandfather (c*5gdad) is living in their biological parents’ household. Note fathers who reported being married to or cohabiting with the biological mother (f5a4=1 or f5a4=4) skipped the house hold grid. They are coded as “-5 not asked” for these variables. cf5hhgridflag flags these fathers.
c*5edu	Categorical variable for biological parents’ education level at time of interview; coded as “Less than high school,” “High school or equivalent,” “Some college or technical school” or “College or graduate school”
c*5hhinc, cf5hhincb	Income reported during the interview including imputed values for participants indicating a range of income or no income. cf5hhincb uses the mother’s report of household income for married and cohabiting couples. See file containing notes on poverty and household income at year 9 for more information.
c*5hhimp. cf5hhimpb	Binary indicator that imputation was used to generate income values
c*5povco	Ratio of total household income (as defined in the variables described above) to the official poverty thresholds established by the U.S. Census Bureau. The thresholds vary by family composition and year. Thresholds for the year preceding the interview were used. We calculated separate thresholds based on mother and father reports of household size and composition. However, calculations for married/cohabiting mothers and fathers rely on mother reports of household size and composition. A small number of missing values (don’t know, refused) were treated as 0 in household membership counts.
c*5povca, cf5povcab	Categorical variable created from c*5povco . The cf5povcab version of this variable for fathers is based on the “b” version of his household income variables. The imputation flags created for the household income variables also refer to the poverty variables.
<i>Occupational Recoding</i>	
*5i12a_code	Occupation variables for biological parents were constructed based on the 3-digit codes from the U.S. Bureau of Labor Statistics (BLS) <i>Occupational Classification System</i> by Major Occupational Groups. These categories are summarized below: “101 - Professional, Technical, and Related Occupations (Group A)” “102 - Executive, Administrative, and Managerial Occupations (Group B)” “103 - Sales Occupations (Group C)” “104 - Administrative Support Occupations, including Clerical (Group D)” “105 - Precision Production, Craft, and Repair Occupations (Group E)” “106 - Machine Operators, Assemblers, and Inspectors (Group F)” “107 - Transportation and Material Moving Occupations (Group G)” “108 - Handlers, Equipment Cleaners, Helpers, and Laborers (Group H)” “109 - Service Occupations, except Private Household (Group K)” “110 – Unspecified” “112 – Military” The variables describe main (*5i12a_code) employment.
<i>Depression</i>	
c*5md_case_con c*5md_case_lib	Binary indicators for biological parents’ and non-parental caregiver’s experience of a major depressive episode in proceeding year; see Nine-Year Scales Documentation for further description.
<i>Incarceration History</i>	
c*5fevjail c*5finjail cmf5finjail	Binary indicator based on parents’ report of whether the biological father has ever been incarcerated (c*5fevjail, cmf5fevjail) or if the biological father is currently in jail or prison (c*5finjail, cmf5finjail) and were constructed using information collected by interviewers

cmf5fevjail	<p>during fielding and data reported directly from parents during the interview. Parents who were not interviewed are coded as missing in these variables.</p> <ul style="list-style-type: none"> • cm5finjail constructed from m5b29 (value “5”) and m5b30 (value “3”) and additional information collected by interviewers during fielding • cm5fevjail constructed from m5b29 (value “5”), m5b30 (value “1” or “3”), m5a4k (value “4”), m5b17d_7 (value “1”), cm1finjail (value “1”), cm2finjail (value “1”), cm3finjail (value “1”), cm4finjail (value “1”), cm2fevjail (value “1”), cm3fevjail (value “1”), cm4fevjail (value “1”). Note mothers who were not interviewed are coded as missing for this variable. • cf5finjail constructed from father’s final incarcerated complete (value “66” and “67”) and incomplete (value “81”) status codes, f5a5 (value “-10”), f5f2 (value “9”) and additional information collected by interviewers during fielding. • cf5fevjail constructed from f5a5 (value “-10”), f5b17d_7 (value “1”), f5f2 (value “9”), f5f28b (value “1”), f5f28e (value “1”), f5a4k_4 (value “1”), cf1finjail (value “1”), cf2finjail (value “1”), cf3finjail (value “1”), cf4finjail (value “1”), cf5finjail (value “1”), cf2fevjail (value “1”), cf3fevjail (value “1”), and cf4fevjail (value “1”) • cmf5fevjail constructed from cm5fevjail (value “1”) or cf5fevjail (value “1”) • cmf5finjail constructed from cm5finjail (value “1”) or cf5finjail (value “1”)
--------------------	---

Nine-Year Weights

Statisticians at Columbia University are currently working on constructing weights for the Nine-Year data. We anticipate that a set of weights, comparable to earlier waves, will be available in late-December 2011. The data files currently contain 2 sets of variables indicating which cases are in the national (**cm5natism** and **cf5natism**) or city (**cm5citsm** and **cf5citsm**) samples. Corresponding national and city weights will be appended to the files as they become available.

Interviewing Incarcerated Respondents

In the first two pilot cities and early interviewing in the next five cities, we attempted to conduct interviews (either by phone or in-person) with parents that were incarcerated. However, shortly after the five cities interviewing began, the project discontinued interviews with currently incarcerated parents. Therefore, only six interviews were conducted on the phone (**cm5_bmomstat** of “67”) and in-person (**cm5_bmomstat** of “66”) with incarcerated mothers. Twenty-five interviews were conducted on the phone (**cf5_bdadstat** of “67”) with incarcerated fathers; no interviews with incarcerated biological fathers were conducted in person.

Non-Parental Caregiver Survey

The Primary Caregiver Questionnaire is asked of the person who is identified as the primary caregiver – that is, the person with whom the child lives at least half time or more. The default respondent is the mother in cases where the mother and father live together with the child. The questionnaire consists of two major parts: one part is asked only if the primary caregiver is someone other than the mother or father (this person is referred to as the non-parental caregiver) called the “Non-Parental Caregiver Survey”; the second part is asked of all persons identified as the primary caregiver.

The Non-Parental Caregiver survey contains a subset of questions asked in the Mother and Father questionnaires and contains seven sections. The sections are as follows: A. Introduction; B. Mother-child relationship; C. Father-child relationship; D. Demographics; E. Income, Education, and Employment; F. Health and Wellbeing; and G. Environment. Below is a crosswalk between questions contained in the

non-parental caregiver survey and the biological parent surveys. The source variables for this survey begin with the prefix “n5.”

Variable/ Concept	Survey Component		
	Non-Parental Caregiver	Biological Mother	Biological Father
Split time	n5a5-a6b	m5a3f-h	f5a3f-h
Why not w/parent	n5a9/n5a11	m5a3a1/m5a3i	f5a3a1/f5a3i
Kind of mother	n5b1	m5k1	f5b3x
Mother seen child	n5b2	m5a3e	f5b2x/f5b4x
Kind of father	n5c1	m5b2e	f5k1
Father seen child	n5c2	m5b2	f5a2a/f5a3e
Father buys things	n5c3	m5b22	f5b22
Number in HH	n5d2a	m5a5	f5a5
HH grid	n5d3	m5a5	f5a5
Citizenship	n5d5c	m5e1d	f5e1d
Employment	n5e2	m5i4-10	f5i4-10
HH income	n5e3	m5j1	f5j1
Life satisfaction	n5f0	m5g0	f5g0
Overall health	n5f1	m5g1	f5g1
Serious health issue	n5f2	m5g2	f5g2
Prescribed medicine	n5f2b	m5g2b	f5g2b
Depression	n5f3-15	m5g3-15	f5g3-15
Control/contentment	n5f16	m5g16	f5g16
Smoking	n5f17-18	m5g17-18	f5g17-18
Hardship	n5g1	m5f23	f5f23

I. Notes on Missing Data

Those coded with “-7” throughout completed the primary caregiver survey (n=3,489), but the non-parental caregiver survey portion was not applicable. The non-parental caregiver survey was not applicable in these cases because either the biological mother or the biological father is the primary caregiver.

II. Additional Variable Information

Questions involving the household grid begin with the prefix “n5d3.” Information on the focal child is contained in the first slot of the household grid. In a few cases, information about the focal child was manually added to the grid where it should have been preloaded.

Primary Caregiver Survey

The main purpose of the Primary Caregiver interview is to gather information about the focal children and their environment including their health, family life, family routines, and educational experiences. Information about primary caregivers’ attitudes towards raising children is also collected. This interview was administered as a 45-minute CATI survey. There are six main sections in the Primary Caregiver

survey: H. Child Health, I. Family Routines and Home Life, J. Nutrition, K. Parental Stress and Mastery, L. Child's Education, and M. Child's Neighborhood. The source variables for this survey begin with the prefix "p5." Variables with the prefix "p5q" were generated from the Primary Caregiver Self-Administered Questionnaire (PCG SAQ) which was conducted during the Home Visit.

I. Notes on Missing Data

In the primary caregiver file, there are 9 cases with a **pcg5stat** of "71 Breakoff". These cases will have some valid information but then "-3" missings after the point where the break-off occurred.

There are also two cases with a code other than "71" or "64," "65," "66," or "67" for **pcg5stat** that contain some non-missing data. In these cases, the status code indicates a final refusal or other final incomplete code. The data have been left in to the point the surveys were completed, and afterwards missing answers are indicated by "-3."

II. Additional Variable Information

Due to an error in the CAPI programming, seven primary caregiver mothers are missing for variables **p5h15**, **p5j6**, and **p5i30c – p5i47c**.

There are 15 cases in which a child's gender was not coded correctly or was missing in the preload file. These are identified with the variable **p5childgen_wrong**. This error resulted in missing information for 12 boys for **p5h19** & **p5h19a** and 3 girls for **p5h22** through **p5h22b**.

III. Scales Information

Pubertal Development Scale These items (**p5h17**, **p5h17a**, **p5h17b**, **p5h19**, **p5h19a**, **p5h21**, **p5h22**, **p5h22a**, **p5h22b**, **p5h23**) are taken from the Mother questionnaire of the Phase IV NICHD Study of Early Child Care and Youth Development Survey. The Pubertal Development Scale was created by Anne Peterson to provide an instrument for self-assessment of pubertal development by adolescents, which could be used in school. It consists of a series of questions about physical development that ask the respondent to evaluate the degree to which a specific physical change (such as pimply skin, growth spurt, breast development, and facial hair) has occurred. Development is related on a scale with the following values 1 (*No*), 2 (*Yes, barely*), 3 (*Yes, Definitely*), and 4 (*Development completed*). This measure has been widely used for assessment of pubertal development by parents and other observers. There are separate questions for girls and boys.

Biological Mother and Father Surveys

The Biological Mother Survey is an 11-section CAPI survey lasting approximately 50 minutes. The subsections include information on the following: A. Family Characteristics, B. Father Contributions and Resources, C. Mother's Relationship with Father, D. Current Partner, E. Family Background and Support, F. Environment and Programs, G. Health and Health Behavior, H. Religion, I. Education and employment, J. Income, and K. Secondary Caregiver. The biological mother only completed Section K when she was not the primary caregiver for the child (ie. the child lived with the biological mother less than half of the time) or neither biological parent had primary custody of the child. The source variables for this survey begin with the prefix "m5."

The Biological Father Survey is also an 11-section CAPI survey lasting approximately 50 minutes. Similar to the Biological Mother Survey, the subsections include information on the following: A. Family Characteristics, B. Father/Mother Contributions and Resources, C. Father's Relationship with Mother, D.

Current Partner, E. Family Background and Support, F. Environment and Programs, G. Health and Health Behavior, H. Religion, I. Education and Employment, J. Income, and K. Secondary Caregiver. The biological father completed Section K in most situations, except when the father was the primary caregiver or neither parent had primary custody of the child. The source variables for this survey begin with the prefix “f5.”

I. Notes on Missing Data

In the biological mother file there are 5 cases with non-complete codes (**cm5_bmomstat** codes other than “64”, “65”, “66”, “67”) that were not coded as “-9 not in wave.” These cases, with a **cm5_bmomstat** of “71 Breakoff” will have some valid information but then “-3” missing after the point where the break-off occurred.

Similarly, in the biological father file there are 2 cases with non-complete codes (**cf5_bdadstat** codes other than “64”, “65”, “66”, “67”) that were not coded as “-9 not in wave.” These cases, with a **cf5_bdadstat** of “71 Breakoff,” will have some valid information but then “-3” missing after the point where the break-off occurred.

II. Additional Variable Information

1. Relative or friend serves as the primary caregiver

- When the primary caregiver of the focal child is a friend or other relative, the biological mother should answer questions **m5a3b1a** (n=25) and **m5a3c** (n=8). A programming error resulted in data not being collected on these variables in a minimal number of cases.

2. Household Grid information in both surveys

The aims for this series of questions are to: 1) capture the household size and the characteristics of household members, including the biological children living with the parent (**m5a5*** and **f5a5*** questions), 2) capture information about biological children who reside outside parents’ household (**m5a6*** and **f5a6*** questions), and 3) record parents’ relationship status with each of their children’s other biological parent at the time of the child’s birth and distinguish when the other biological parent is similar or different across their children (**m5a8*** and **f5a8*** questions).

- *Biological Mother Survey*

If the biological mother is coded as living with the child half-time or more (**m5a2=1** or **2**), the focal child is included in the first series of questions (**m5a5b01/c01/c01a/d01/e01**) of the household grid. The child is not counted in **m5a51** (number of people living with mother excluding self and focal child).

- When the focal child does not live with the biological mother at least half of the time (**m5a2=3, 4** or **7**), the focal child is preloaded to **m5a6c01** and is counted in **m5a6a** (number of biological children, including child, who do not live with biological mother).

- *Biological Father Survey*

If the biological father reports being married to or cohabiting with the biological mother (**f5a4=1** or **4**), he does not answer the traditional household grid (**f5a5** through **f5a5j**). Instead, married or cohabiting fathers report the total number of biological children (other than focal child, where applicable) living in their household (**f5a5l**).

- As in the biological mother survey, if the biological father is coded as living with the child half-time or more (**f5a2**=1 or 2), the focal child is included in the first series of questions (**f5a5b01/c01/c01a/d01/e01**) of the household grid, even though the **f5a5** and **f5a51** questions explicitly specify not to include the focal child in responses. However, for married or cohabiting fathers (who also do not complete the **f5a5b*** – **f5a5e*** questions), the child is not counted in **f5a51** (number of people living with father excluding self and focal child).
- When the focal child does not live with the biological father at least half of the time (**f5a2**=3, 4 or 7), the focal child is preloaded to **f5a6c01** and is counted in **f5a6a** (number of biological children, including child, who do not live with biological father)
- *Additional Notes on the Total Number of Children Reported in the Grids*
Because the CATI programming did not contain important internal checks, a minimal amount of data cleaning was required for both biological mother and father surveys in order to reconcile the total number of biological children living with the parent in the household and residing outside of their household. For example, household grids for the biological mother and father were checked for duplicate entries. Entries were considered duplicates when the name, age, and relationship entered in two slots in the grid matched exactly. When duplicate entries occurred they were removed from the grid.

Please note, however, that there is a small (less than 15 in each survey) subset of cases where the number of children described in the fertility grid (a8 questions) does not directly sum to the number of biological children in the household grid (a5 questions) and the number of children reported living away from the parent (a6 questions).

- *Additional Notes on the Fertility Grid*
 - Information pertaining to the focal child will be listed as the final populated slot in the a8b* and a8f* series of questions (for example, if the mother has three biological children, information for the focal child will be stored in m5a8b03 and m5a8f03)
 - Apart from the focal children, information about other biological children will be listed in the order in which appeared in the household grid (a5b/c/d/e questions), first, and, then in the grid of children living outside of the household (a6c/f/g).
 - No distinguishing name information is listed in the grids. For the fertility grid, the first instance in which an *a8b* variable refers to a partner other than the focal child will be denoted with a “91.” Each new, unique partner that the parent had a child with will be coded as a “91” and subsequent references to the new, added partner will be denoted with a “2.” The number of unique partners that parents had a child with can be obtained by adding the number of *a8b* variables with “91” codes (plus 1 for the biological father/mother).

3. *Reasons romantic relationship ended (a4k)*

- The variable for why the biological parents’ romantic relationship ended (**m5a4k** and **f5a4k**) is asked in different formats between the two surveys. In the biological mother

survey, only one response was permitted; however, in the biological father survey, multiple responses were permitted.

4. *Sources of Support (e8 and e9)*

- Information from other questions in the survey were used to populate the response categories for these variables. The following CAPI instructions were used for the displays:
 - DISPLAY ALL ANSWER CATEGORIES UNLESS OTHERWISE INDICATED BY THE FOLLOWING: IF A4 = -13, DO NOT DISPLAY ANSWER CHOICES 2 OR 3. ELSE, IF A4 = -14, DO NOT DISPLAY ANSWER CHOICE 2. ELSE, IF E1G=4, DO NOT DISPLAY CHOICE 1 "Your parents". ELSE, IF E1M=4, DO NOT DISPLAY "parents or" IN ANSWER CHOICE 3. IF A4 = 1 OR D2D = 1, DISPLAY ANSWER CHOICES 6 AND 7. IN ANSWER CHOICES 6 AND 7, DISPLAY "spouse" IF A4=1, AND DISPLAY "partner" IF D2D=1.

5. *Incarceration skip*

- Due to a programming error, all cases with **m5b29=5 (f5b23x** in the father interview) erroneously skip the questions on beginning and projected ending dates of the current incarceration (items **m5b30d1** through **m5b30e2** in the mother interview and **f5b29x1** through **f5b30x2**). Westat took extra steps to retrieve these data and a separate file may be available with any recovered data.
- Please also note that **m5b30** uses either the baseline, one-year or three-year interview as the reference period, rather than the time since the year five follow-up as in other questions.

6. *Biological Father's report of same partner from previous interview (f5d2h)*

- Due to an error in the CATI program, most fathers that had a partner at their previous interview (n=81) do not answer this question.

7. *Current Partners' non-resident children reported in Biological Mother interview*

- An error in the CATI program caused information about current partner's non-resident biological children (for those who reported having any in **m5d13b**) to be uncollected for a subset of cases (n=115) in the variables **m5d13c**, **m5d13d**, and **m5d13e** and 113 cases in **m5d13f**.

8. *Biological Mother's CIDI Scale for depression*

- A small number of moms (n=45) are missing information on three variables in the CIDI scale for depression due to an error in the CATI program. Moms affected are coded as “-3, missing” for the variables listed below. When coding the binary caseness variables for conservative (cm5md_case_con) and liberal estimates (cm5md_case_lib) of depression, moms affected by this error that already met criteria for depression were left as depressed cases (coded as 1). Moms affected by this error that did not yet meet the criteria for depression were coded to missing on the caseness variables, as they might have met the criteria for depression if these data were collected (coded to -3). Five of the

mothers affected by this error meet the criteria for a conservative case of depression and 40 were coded to missing. Thirty-three moms affected by this error meet the criteria for a liberal case of depression and 12 were coded to missing.

- i. **m5g13**: During those two weeks, did you have a lot more trouble concentrating than usual?
- ii. **m5g14**: People sometimes feel down on themselves, no good, or worthless. During that two week period did you feel this way?
- iii. **m5g15**: Did you think a lot about death –either your own, someone else’s or death in general during those two weeks?

9. *Biological Father’s questions on work*

- There are 47 fathers missing on the variables **f5i16a** through **f5i23**. These fathers indicated they did not work in the past week (**f5i4=2**), but had worked since the birth of the child. They reported not knowing the last month in which they worked, which erroneously caused them to skip the variables mentioned above.

III. Scales Information

The following scales were embedded in the Biological Mother and Biological Father Surveys. Brief information about these scales is listed below. Complete information about scoring and basic statistics on these scales may be found in the Nine-Year Scales Documentation.

Alcohol Use Scale A subset of three questions on frequency of alcohol use is derived from the Composite International Diagnostic Interview Short-Form (CIDI-SF).

Drug Use Scale Questions on drug are derived from the Composite International Diagnostic Interview Short Form Alcohol and Drug Dependence scales (CIDI-SF). The survey includes questions regarding the use of the nine drugs and one of the seven symptoms (role interference as a result of use).

Mental Health Scale for Depression The Major Depressive Episode (MDE) nine-year questions are derived from the Composite International Diagnostic Interview - Short Form (CIDI-SF), Section A. The short form of the CIDI interview takes a portion of the full set of CIDI questions and generates from the responses the probability that the respondent would be a “case,” (i.e., a positively diagnosed respondent), if given a full CIDI interview. The CIDI questions are consistent with the Diagnostic and Statistical Manual of Mental Disorders – Fourth Edition.

Couple Relationship Quality Three questions about whether the relationship might be in trouble and consideration of breaking up (d6g-i) are modifications of NSFH questions, which were originally developed by Booth, Johnson, and Edwards as part of the Marital Instability Index . These items assess the couple’s relationship with respect to their commitment, satisfaction with the sexual relationship, and trust. It also questions parents on whether their relationship is in trouble and the frequency with which they discuss breaking up.

Economic Hardship These items are taken from the “Basic Needs – Ability to Meet Expenses” section of the Survey on Income and Program Participation (SIPP) 1996 Panel Wave 8 Adult Well-Being Topical Module Questionnaire (*Survey on Income and Program Participation*, 1998), the 1997 & 1999 New York

City Social Indicators Survey (SIS) (Social Indicators Survey Center, 1997 & 1999), and the 1999 Study of Work, Welfare, and Family Well-Being of Iowa families on FIP (Iowa's assistance program) (Mathematica Policy Research).

Child Survey

The focal child is administered an 8-section questionnaire during the Home Visit that typically took 20-25 minutes to complete. It was administered by the interviewer on CAPI. The beginning section of the survey instrument contains a series of confirmation questions asked to the child to determine the skip patterns for the questionnaire (**k5conf2**, **k5conf3**, **k5conf6**, **k5conf9**, **k5conf10**). The remaining substantive sections of the Child Survey include: A. Parental Supervision and Relationships, B. Parental Discipline, C. Sibling Relationships, D. Routines, E. School, F. Early Delinquency, G. Task Completion and Behavior, and H. Health and Safety. The source variables for this survey begin with the prefix "k5."

I. Notes on Missing Data

Data collected during the Home Visit are coded to have the same number of "-9 Not in Wave", 1,498. These people did not participate in any portion of the Home Visit including the child assessments, child survey, primary care giver self-administered questionnaire, interviewer observations, and physical measurements. Those who participated in some portion of these activities are coded as "-3 Missing" for any incomplete parts of the Home Visit. In the child survey, 23 children who participated in the Home Visit did not start the survey, and 9 more failed to complete it.

II. Additional Variable Information

In a few instances, answering patterns were inconsistent with pcgs' status codes and answers given in the beginning confirmation section of the survey. The variable **ck5pcg_mismatch** flags cases where the primary care giver is listed as one individual and questions are answered as if the primary care giver is a different person. There are 6 specific cases flagged in **ck5pcg_mismatch**.

III. Scales Information

Child Delinquency Measures In questions **k5f1a** through **k5f1q**, the survey questions the child on whether he or she had ever engaged in a series of seventeen delinquent behaviors. These items are modeled after the Things That You Have Done scale. Similar items were included in the National Longitudinal Survey of Youth (1997). Please refer to the Nine-Year Scales Documentation for information on subscales and scoring for this measure.

Connectedness at School These items (**k5e1a**, **k5e1b**, **k5e1c**, **k5e1d**) were compiled by Jacquelynn Eccles for the PSID-CDS-III to measure the degree of inclusiveness, closeness, happiness, and safety the child experiences at school. Each question is rated on a 5-point Likert Scale ranging from 0 (*not once in past month*) to 4 (*everyday*).

Peer Bullying These items (**k5e2a**, **k5e2b**, **k5e2c**, **k5e2d**) are modeled after the peer bullying assessment from the PSID-CDS-III. The questions were confirmed for use in the PSID confirmatory factor analysis, which loaded onto one factor.

Behavior and Emotions Data about child’s behaviors and emotions (**k5g2b, k5g2d, k5g2f, k5g2h, k5g2m, k5g2n, k5g2a, k5g2c, k5g2e, k5g2g, k5g2i, k5g2j, k5g2k, k5g2l**) were collected via self-interview using items from the internalizing and externalizing subscales of the Self-Description Questionnaire (SDQ). The Nine-Year child interview contains 14 of the SDQ items, on which the child is asked to rate their frequency of emotions and behaviors from 0 (Not at all true) to 3 (Very true).

Primary Caregiver Self-Administered Questionnaire (PCG SAQ)

During the Home Visit, primary caregivers are asked to complete a hard-copy booklet with questions regarding discipline style and care of the child, the focal child’s behavior and the family’s involvement with Child Protective Services. The variables for this survey begin with the prefix “p5q.” The PCG SAQ questions are:

p5q1a through **p5q1n**: The frequency with which the caregiver has used selected disciplinary practices on the focal child in the past year (e.g., time out; shouted or yelled; spanked); 14 items from the Parent Child Conflict Tactics Scales (CTSPC).

p5q2a through **p5q2e**: The frequency with which problems interfered with respondent’s ability to care for the child in the last year (e.g., had to leave child alone, not able to make sure child got to a doctor or hospital when he or she needed it); 5 items from the CTSPC.

p5q3a through **p5q3do**: A series of questions about child behavior problems and whether they are not true, somewhat true, or very true of subject child (e.g., argues a lot; acts too young for age; drinks alcohol without parent approval); 111 items from the Achenbach Child Behavior Checklist (CBCL) and 13 items from Adaptive Social Behavior Index (ASBI). While the origin of questions **p5q3dc** through **p5q3do** is the ASBI, no scoring information for the ASBI is provided in the scales documentation. The PCGSAQ does not contain the complete ASBI or an entire subscale from the instrument.

p5q4 through **p5q14**: Biological parents are asked this series of 11 questions about their family’s contact with the Child Protective Services system.

I. Notes on Missing Data

Data collected during the Home Visit are coded to have the same number of “-9 Not in Wave”, 1,498. These people did not participate in any portion of the Home Visit including the child assessments, child survey, primary care giver self-administered questionnaire, interviewer observations, and physical measurements. Those who participated in some portion of these activities are coded as “-3 Missing” for any incomplete parts of the Home Visit. In the PCG SAQ, 61 people who participated in the Home Visit did not even start the survey, and twomore failed to complete it.

II. Additional Variable Information

The PCG SAQ administered in either English or Spanish. The variable **p5qlang** will indicate whether it was completed in English or Spanish.

III. Scales Information

The following scales were embedded in the PCG SAQ: the Child Behavior Checklist (CBCL) and the Conflict Tactics Scale. Brief information about these scales is listed below. Complete information about scoring and basic statistics on these scales may be found in the Nine-Year Scales Documentation.

Child Behavior from the CBCL Data about child's behavior were collected using questions taken from the behavioral, emotional, and social problems scales of the CBCL/6-18 (See Child Behavior Checklist/6-18). The PCG SAQ contains 111 items and scales of the CBCL/6-18 on which a parent, or surrogate parent, is asked to rate their child's behavior from 0 (*Not true*) to 2 (*Very true or often true*). Several items from the other problems scale were excluded. These items include the following questions: child has bowel movements outside of the toilet, child does not eat well, child bites fingernails, child sleeps more than most children, child sucks thumb, child wets self during the day, child wets the bet, and child wishes to be opposite sex. Selected items in the CBCL comprise the following eight constructs or syndromes: aggressive behavior, withdrawn/ depressed, anxious/ depressed, attention problems, social problems, rule-breaking behavior, somatic complaints, and thought problems.

Conflict Tactics The PCG SAQ uses 14 of the 22 items on the Parent Child Conflict Tactics Scales (CTSPC). The original Conflict Tactic Scales (1979) was designed for use with partner in a marital, cohabiting, or dating relationship. The CTSPC was created in 1996 in response to limitations of the original scale as a measure of child maltreatment (Straus, et al., 1998). The main CTSPC subscales are Physical Assault, Non-Violent Discipline, and Psychological Aggression. Our survey eliminates eight questions from the CTSPC that ask about severe physical maltreatment. However, we include the CTSPC's supplemental scale on Neglect (5 additional questions).

In-Home Observations

Immediately following the completion of the Home Visit, the interviewer completed a series of questions based on their observations of the respondent's home and neighborhood and their experiences with the family and focal child during the Home Visit. Information collected includes observations of the neighborhood, areas immediately outside the home, interior common areas (for apartments only), interior of the home/apartment, and child's appearance; HOME Scale measures regarding PCG's affect and demeanor during the Home Visit; and ratings of the behavior and demeanor of the child. The source variables for this survey begin with the prefix "o5."

I. Notes on Missing Data

Data collected during the Home Visit are coded to have the same number of "-9 Not in Wave", 1,498. These people did not participate in any portion of the Home Visit including the child assessments, child survey, primary care giver self-administered questionnaire, interviewer observations, and physical measurements. Those who participated in some portion of these activities are coded as "-3 Missing" for any incomplete parts of the Home Visit. For the in-home observations, 8 interviewers who conducted some portion of the Home Visit failed to complete it.

This survey component was completed by the interviewer following the home visit. As a result, there are no questions for which answers are coded as "-1 Refused." All missing information is coded as "-2 Don't Know."

II. Additional Variable Information

- Four constructed flag variables were created from **o5g7**, **o5g8**, and interviewer comments to provide additional description of the administration of the Home Visit activities:
 - **o5notinhouse**: Binary variable indicating cases in which the interviewer conducted the home visit in a location other than the child’s primary residence, such as a grandparent’s house (if grandparent is not primary care giver), the residence of the non-custodial parent, the library, or the child’s school. In some cases the interviewer marked “do not know” for information about the characteristics of the child’s home and neighborhood, but in other cases he/she may have responded based on this alternative location.
 - Binary variable indicating cases in which the interviewer noted that family members were present during the child assessment (**o5FamAtCA**), child’s interview (**o5FamAtCI**), and during other parts of the home visit and interview with the primary caregiver (**o5FamAtInt**).
- **o5c2**: The wording in this question for the pilot cities asked if wires in the house were covered; after the pilot cities, this question was altered to ask if wiring was exposed. The responses in the data reflect the latter wording, and yes/no responses have been switched for respondents from the first two cities.

III. Scales Information

The origin of the observation questions is the Home Observation for Measurement of the Environment (HOME) Inventory developed by M. Caldwell and R.H. Bradley. We do not provide scoring or normative information for the HOME scale in the scales documentation because questions were pulled from three versions of the HOME scale: HOME for early childhood, HOME for middle childhood, and HOME for early adolescents.

Height and Weight Measurements

Height measurements, in centimeters (using a large plastic standing ruler called a “stadiometer”), of focal children and weight measurements, in pounds, of mother and focal child were taken during the Home Visit. In addition to the basic height and weight measures, the file contains a series of variables for Body Mass Index (BMI), as well as height-for-age and weight-for-age percentiles, constructed from these measurements. The variables for these measurements begin with the prefix “hv5.”

<i>Child Weight and Height Source Variables</i>			
hv5_9	Status of child weight measurements	hv5_11	Status of child height measurements
hv5_9a1	Child weight measurement #1	hv5_11a1	Child height measurement #1
hv5_9a2	Child weight measurement #2	hv5_11a2	Child height measurement #2
hv5_9a3	Child weight measurement #3	hv5_11a3	Child height measurement #3
hv5_9b	Child self-reported weight	hv5_11b	Child self-reported height
<i>Child Weight and Height Constructed Variables</i>			
hv5_cwtlb	Child weight (lb.) from mult. measurements	hv5_cbmi	Child’s Body Mass Index (BMI)
hv5_cwtkg	Child weight (kg.) from mult. measurements	hv5_waz	Child’s Weight-for-Age Z-Score
hv5_chtcm	Child height (cm.) from mult. measurements	hv5_wap	Child’s Weight-for-Age Percentile
hv5_flag_cm	Flag: child’s height may have been improperly measured (recoded)	hv5_bmiz	Child’s BMI Z-Score
hv5_flag	Flag of possible issue with anthropometric or BMI measures	hv5_bmip	Child’s BMI Percentile

hv5_agem	Child's age in months at time of measurement	hv5_haz	Child's Height-for-Age Z-Score
		hv5_hap	Child's Height-for-Age Percentile
<i>Mother Weight and Height Source and Constructed Variables</i>			
hv5_8	Status of mother weight measurements		
hv5_8a1	Biological Mother weight measurement #1	hv5_mhtcm	Biological Mother's height (cm)
hv5_8a2	Biological Mother weight measurement #2	hv5_mmis_ht	Biological Mother's height is missing
hv5_8a3	Biological Mother weight measurement #3	hv5_selfht	Biological Mother self-reported height
hv5_8b	Biological Mother self-reported weight		
hv5_ovscale	Biological Mother's weight too large for scale to process	hv5_mflag	Flag of possible issue with Biological Mother's height/weight
hv5_selfwt	Biological Mother's weight was self-reported	hv5_mbmi	Biological Mother's BMI
hv5_mmis_wt	Missing weight for Biological Mother	hv5_mbmiz	Biological Mother's BMI Z-Score
hv5_mwtlb	Biological Mother's weight (lb.) from multiple measurements		
hv5_mwtkg	Biological Mother's weight (kg.) from multiple measurements		

Please note that the biological mothers' height was not collected at during the Nine-Year Home Visit. In order to construct mothers' BMI, height measurements taken at the Home Visit at Years 3 or 5 were used in the calculation. Additionally, self-reported height from Year 3 was used when measurement of height was not available at both Years 3 and 5.

Due to a programming error, -1 and -2 values in **hv5_cbmiz**, **hv5_mbmiz**, **hv5_haz**, and **hv5_waz** are incorrectly labeled as missing but are actually valid z-scores.

The Nine-Year Scales Documentation further describes these source variables, constructed measures, and flag variables.

I. Notes on Missing Data

Data collected during the Home Visit are coded to have the same number of "-9 Not in Wave", 1,498. These people did not participate in any portion of the Home Visit including the child assessments, child survey, primary care giver self-administered questionnaire, interviewer observations, and physical measurements. Those who participated in some portion of these activities are coded as "-3 Missing" for any incomplete parts of the Home Visit.

There are currently 260 cases where the Home Visit was conducted, at least in part with the family, but height and/or weight data are missing for the mother and 51 cases where the height and/or weight data are missing for the child, so **hv5_cbmi** is unavailable. Although some mothers were present and measured at the home visit despite not being the primary caregivers, some mothers were not present, even if they were primary caregivers, and thus there are 265 additional cases in which the mother's information is not available. In total, there are 525 cases where the Home Visit was completed but **hv5_mbmi** is unavailable.

II. Additional Variable Information

The file contains a constructed variable, **hv5_flag_cm**, which flags cases where, due to considerably higher and lower values, we believe a value for inches (not cms as instructed) for the height may have erroneously been recorded by the interviewer or that the pieces of the stadiometer may have been inserted incorrectly when the height measurement was taken. This would have yielded a value 50 cms shorter or taller than the true height value. For the 71 cases where **hv5_flag_cm** equals 1, we have recoded the height based on these assumptions. Heights of fewer than 60 cms were presumed to be in inches and have been multiplied by 2.54 to yield cms. Heights between 60 cms and 110 cms were assumed to result from incorrect stadiometer construction and 50 cms were added to the height measurement. Heights equal to or greater than 174 cms were assumed to also result from incorrect stadiometer construction but in the opposite manner, and 50 cms were subtracted from the height. The resulting range of heights in cms for these 71 cases was 120-158 cms. *Data users should be aware of the issues and recoding if these cases are included in analyses.*

For weight measurements, the interviewer should have taken a third measurement of the mother or child if the difference between the two was 2 lbs or greater. In 7 cases, the child should have been measured a third time and was not. In 8 cases, the mother should have been measured a third time and was not. For each of the 7 child and 8 mother cases, the second measurement was used.

The file contains another variable, **hv5_cflag** (similar in format to the Year 5 In-Home variable), that also describes inconsistencies with the measurement data; values of 1 through 7 denote specific, potential problems with the measurement for this record. There is a parallel variable, **hv5_mflag**, that notes potential problems with records for the mothers.

When taking height measurements of the child, the interviewer took two measurements and should have taken a third if the two were not the same. In 57 cases, this third measurement should have been taken but was not, although the difference between the two measurements was 3 cm or fewer in 55 cases. The remaining cases had differences of 9 and 24 cms. In all 57 cases, the second measurement was used for the composite height measure.

Child Assessments

Woodcock Johnson Passage Comprehension and Applied Problems The initial Passage Comprehension (WJ Subtest 9) items involve symbolic learning, or the ability to match a rebus (pictograph representation of a word) with an actual picture of the object. The next items are presented in a multiple-choice format and require the individual to point to the picture represented by a phrase. The remaining items require the person to read a short passage and identify a missing key word that makes sense in the context of that passage. The items become increasingly difficult by removing pictorial stimuli and by increasing passage length, level of vocabulary, and complexity of syntactic and semantic cues.

Applied Problems (WJ Subtest 10) requires the focal child to analyze and solve math problems. To solve the problems, the focal child must listen to the problem, recognize the procedure to be followed, and then perform relatively simple calculations. Because many of the problems include extraneous information,

the focal child must decide not only the appropriate mathematical operations to use but also which numbers to include in the calculation. Item difficulty increases with complex calculations.

PPVT-III A The PPVT-III test measures receptive vocabulary and screens for verbal ability. Like the Woodcock Johnson III tests, it is administered with an “easel” or activity book. The interviewer reads a word and asks the child to identify the picture in the easel (among a set of four pictures) that corresponds to that word.

Digit Span Data about child’s auditory short term memory, sequencing skills, attention, and concentration were measured using the Wechsler Intelligence Scale for Children, Digit span subtest (WISC- IV Digit Span: (Wechsler, D., 2003). The In-Home Assessment contains 16 items in two sections from the WISC- IV Digit Span forward and backward tests. Each item contains two trials, or chances for a child to repeat the span correctly. Each trial is different, but trials for each individual item are equivalent. Interviewers read a number and ask the child to repeat the number, forward or backward, dependent on the section. Interviewers score 1 (*correct*) or 0 (*incorrect*). Children who do not respond receive a rating of 0: incorrect.

I. Notes on Missing Data

Some children completing other Home Visit activities may not have completed the full battery of assessments (PPVT, Woodcock Johnson and Digit Span). Additionally, amongst those completing a particular assessment, a minimal number of cases do not have the full set (including the standard score, percentile rank, and age equivalent) of constructed variables.

II. Additional Variable Information

Focal children’s specific responses that were used to score the assessments are not included in the file. Westat programmers used the focal child’s responses, in conjunction with scoring software, to construct raw score, standard score, percentile rank, and age equivalent variables for each of these assessments. The final, scored Child Assessment variables contained on this file are:

	Raw Score	Standard Score	Percentile Rank	Age Equivalent
PPVT III-A (n=3346)	hv5_ppvtraw	hv5_ppvtss	hv5_ppvtpr	hv5_ppvtae
Digit Span (n =3366 to 3369)	hv5_dsraw	hv5_ds	hv5_dspr	hv5_dsae
WJ-Applied Problems Test 10 (n=3342 to 3343)	hv5_wj10raw	hv5_wj10ss	hv5_wj10pr	hv5_wj10ae
WJ-Passage Comp. Test 9 (n=3332 to 3333)	hv5_wj9raw	hv5_wj9ss	hv5_wj9pr	hv5_wj9ae

More detailed information about these assessments is available in the Nine-Year Scales Documentation.

Teacher Survey

The Nine-Year Teacher Survey collected a) general information on the child’s teacher and school, b) classroom behavior and social skills specific to the participating child, c) information specific to the participating child, d) parent/guardian involvement, e) classroom characteristics, f) school climate and

environment, and g) general information about the teacher. The source variables for this survey begin with the prefix “t5.”

During the Nine-Year Home Visit, interviewers collected written consent from the child’s primary caregiver and verbal assent from the child to contact the teacher, as well as school and address information for the appropriate teacher. If the primary caregiver reported that the child has more than one teacher, we attempted to contact the child’s Language Arts teacher. If the interview was conducted from June through October, the teacher from the previous school year was contacted. Otherwise, the child’s current teacher was contacted to participate. Prior to contacting the teacher, the principal of the child’s school, superintendent of the child’s school district, and state education officials were advised of the study.

Teachers were mailed paper copies of the consent forms and survey instrument to complete and mail back to our survey subcontractor, Westat, Inc.

I. Notes on Missing Data

Answers left blank by the respondent were coded as “-1, Refused.” In cases where teachers made written indication of not knowing particular answers, responses were coded as “-2 , Don’t Know.” “Don’t Know” was also included as an option for 15 variables.

II. Additional Variable Information

Please note the coding of the following three variables below:

t5a2b indicates the grade of the child when taught by the teacher, with “-7” indicating that the teacher answered “not enrolled in grade.”

t5g9_age is a constructed variable calculated by the teacher’s date of birth and the date of the interview, if both were provided. The original day of the interview and full date of birth have been removed to protect teacher’s confidentiality.

III. Scales Information

Below is a brief description of scales included in the Teacher’s Survey. Please see the Nine-Year Scales Documentation for further information on these measures.

Conner’s Teacher Rating Scale Data about the child’s behavior were collected (**t5b4a - t5b4ab**) via report by the child’s teacher using items from the oppositional, cognitive problems/ inattention, hyperactivity, and Attention Deficit Hyperactivity Disorder (ADHD) subscales of the Conner’s Teacher Rating Scale—Revised Short form (CTRS-R:S). The Teacher survey contains 28 items and scales of the CTRS-R:S on which a teacher is asked to rate the child’s behavior from 0 (*Not true at all: never, seldom*) to 3 (*Very much true, very often, very frequently*).

Social Skills Data about child’s social skills were collected (**t5b1a, t5b1b, t5b1c, t5b1d, t5b1e, t5b1f, t5b1g, t5b1h, t5b1i, t5b1j, t5b1k, t5b1l, t5b1m, t5b1n, t5b1o t5b1p, t5b1q, t5b1r, t5b1s, t5b1t, t5b1u, t5b1v, t5b1w, t5b1x t5b1y, t5b3a, t5b3b, t5b3c, t5b3d, t5b3e, t5b3f, t5b3g, t5b3h, t5b3i, t5b3j, t5b3k, t5b3l**) via report by the child’s teacher using items from the cooperation, assertion, self-control, and social problems subscales of the Social Skills Rating System (SSRS). The Nine-Year

Teacher survey contains 37 items. The SSRS uses a 3-point Likert scale in which never is equal to zero, sometimes is equal to 1, and very often is equal to 2. Following the question format in the Early Childhood Longitudinal Study-Kindergarten Class (ECLS-K), we included an additional point on the scale so that never is equal to 1, sometimes is equal to 2, often is equal to 3, and very often is equal to 4. The modified scale was provided by the US Department of Education's National Center for Education Statistics with permission from the copyright owner.

Several items from the instrument were excluded. These items include the following questions: introduces herself or himself to new people without being told, appropriately questions rules that may be unfair, initiates conversations with peers, appropriately tells you when he or she thinks you have treated him or her unfairly, and volunteers to help peers with classroom tasks. It should be noted that scale scores are only calculated for participants with responses to each item in the scale. When a participant responds with don't know, refuse, or missing to any item on a given scale, their scale score will be missing. Several items from the hyperactivity subscale were also excluded. These items include the following questions: is easily distracted, interrupts conversations of others, disturbs ongoing activities, doesn't listen to what others say, acts impulsively, and fidgets or moves excessively.