



NATIONAL EARLY INTERVENTION LONGITUDINAL STUDY

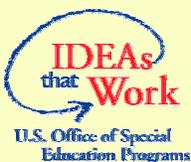
**SAMPLING AND WEIGHTING PROCEDURES:
ENROLLMENT FORM, FAMILY INTERVIEW,
SERVICE RECORDS**

November 2002

NEILS Methodology Report No. 1

Harold Javitz
Donna Spiker
Kathleen Hebbeler
Mary Wagner

SRI Project 07879



SRI International, 333 Ravenswood Avenue, Menlo Park, CA 94025

The National Early Intervention Longitudinal Study (NEILS) is being conducted by SRI International under cooperative agreement (number H329E50001) to SRI International from the Office of Special Education Programs, U.S. Department of Education. NEILS is a collaborative effort of SRI International, the Frank Porter Graham Child Development Institute at the University of North Carolina Chapel Hill, Research Triangle Institute, and the American Institutes for Research.

NEILS is part of a program of longitudinal studies funded by the U.S. Department of Education and being conducted by SRI International. Other studies in the program include the Special Education Longitudinal Study (SEELS) and the National Longitudinal Transition Study-2 (NLTS2). For more information about these studies, see <http://www.seels.net/> or <http://www.nlts2.org/>. The NEILS Web site is <http://www.sri.com/neils/>.

Contents

Effect of Major Research Questions on the Sampling Plan for the National Early Intervention Longitudinal Study	1
Results of the Sample Design Process	2
Determination of the Child Sample Size.....	2
Sampling States.....	3
Allocating the Child Sample among States.....	6
Sampling Local Communities	7
Allocating Each State’s Child Sample among Counties.....	9
Recruiting the Child Sample.....	8
The Enrollment and Study Samples.....	11
Source for Enrollment Information.....	12
Procedures for Weighting the Enrollment Sample.....	12
Weighting the Enrollment Sample to County Totals.....	12
Weighting County Totals to State Totals.....	14
Weighting State Totals to Regional Totals.....	14
Replicate Weights for the Enrollment Sample.....	15
Weighting Enrollment-Eligible Children	15
Weighting Enrollment-Eligible Children with NEILS Participation Permission.....	16
Weighting Enrollment Interviews	17
Weighting the First Interim Interview	17
Weighting the 36-Month Family Interview	17
Weighting the First Service Record	18
Data Analysis Procedures.....	20
Data Limitations.....	21
References	22
Appendix: NEILS Family Information Form – Form A.....	23

Effect of Major Research Questions on the Sampling Plan for the National Early Intervention Longitudinal Study

The National Early Intervention Longitudinal Study (NEELS) has selected a nationally representative sample of 3,338 Part C participants, rather than attempt data collection on the universe of participants, to reduce the burden and cost of data collection.

The major research questions for NEELS had important implications for the sampling approach:

- All questions focus on the national picture of Part C. The sample must support descriptions of participants and services nationally. Geographic dispersion is critical if the sample is to be credible and to generalize to Part C participants and services nationally.
- All questions require estimates of participant or service characteristics. The sample must be sufficiently large to yield estimates with acceptable precision.
- A focus on Part C services raises important issues of the variability from state to state in service organization, eligibility definitions, and other factors. To describe services nationally, the sample must capture key elements of service variability among states. Thus, the sample design must specify criteria for selecting states as one stage in the sampling approach.
- A focus on characteristics and outcomes of children implies, of course, that individual children and their families are the ultimate unit of analysis and, therefore, that selecting children is one of the sampling stages that must be specified in the sample design. In selecting children/families, a specific focus on children receiving services under Part C rules out such child sampling approaches as random digit dialing, through which perhaps hundreds of thousands of families would need to be screened to identify those with children ages birth to 3 and then the even smaller proportion who receive services through Part C. An efficient sampling approach must identify children who are deemed eligible for Part C by appropriate local agencies. Thus, the sampling approach must specify how local agencies are to be identified, as another sampling stage.

Perhaps the most important of the requirements listed above was the need to select a nationally representative sample, which has fundamental implications for the study design and the resulting data obtained. Further, the requirements specified above dictated a three-stage sampling approach: (a) states, (b) local communities and agencies, and (c) individual children/families. The following considerations pertained to these stages of the sample:

- **States** differ in their approach to the implementation of Part C, including the lead agency administering the program. States also are obviously different in geography and demographics, which affect the populations they serve and the challenges they face in doing so.
- **Local communities and agencies** differ in size, economic health, and population density, among other characteristics. Many of these factors can affect the delivery of early intervention services. For example, the service delivery issues in rural areas differ greatly from those in urban ones. Also, local service delivery for Part C is organized in many different ways. A study of Part C needs to capture some of the important ways in which local communities differ.

- *Individual children/families* receiving Part C services also differ in many ways. Part C explicitly addresses the need to provide early intervention services to members of traditionally underrepresented groups. There is some speculation that children and families from minority backgrounds, as well as those living in rural areas, do not receive services commensurate with their proportions in the population. The sample for the proposed study needs to include enough families from traditionally underrepresented groups to be able to examine whether the legislation is addressing one of its specified purposes.

The NEILS sample design balances the requirements for a nationally representative, three-stage sample against a further crucial criterion: a fixed level of resources. The sample design must yield a sample that can be tracked and measured repeatedly throughout the years of the study within the currently available budget.

Results of the Sample Design Process

The first 6 months of the NEILS design phase were spent in considering options for balancing the multiple requirements for the design. In brief, the following sample parameters emerged from the sample design and recruitment process:

- The NEILS study obtained a sample of 3,338 children entering Part C services from September 1997 through November 1998, from a universe of almost 200,000 Part C children served nationally.
- The NEILS sample children were selected from 20 geographically dispersed states (of a universe of 50 states participating in Part C) that represent variation in important features related to Part C and that serve a large proportion of the entire universe of Part C participants.
- The NEILS sample children were selected from 3 to 7 counties within each sampled state, representing both rural and nonrural areas. The total of 93 counties was drawn from a universe of 1,441 counties in the 20 sample states.

Determination of the Child Sample Size

It was important that the NEILS sample have enough statistical power to detect differences between groups and other kinds of relationships that are relevant to the research questions. The statistical power in a given study is determined by the interplay between the desired significance level, the characteristics of the measures being used, the expected size of differences likely to be observed, and the sample size. A large sample size is a key component to identifying statistically significant differences between groups or relationships between variables.

Two statistical power analyses with general linear modeling analysis techniques established that a sample size of approximately 1,300 children would be sufficient to capture all but quite small group differences in NEILS analyses. However, the initial sample needed to be quite a bit larger if the sample of children remaining in Year 5 was to include a minimum of 1,300 children.

The initial sample size needed to be large enough to account for attrition from the sample over time. Attrition can result from families' relocating outside of the sampled communities, moving and becoming untraceable within the communities (although this is less likely if they stay in early intervention), or voluntarily withdrawing from the study. Mobility was expected to be the major source of attrition, although this is a less significant problem when the main data collection mechanism is a telephone survey. Attrition of 10% per year from the sample was assumed in identifying the original sample size, based on following the children through kindergarten.

The initial sample size needed to account for the fact that not every child is likely to have all data elements in the extremely complex data set that will be constructed in and NEELS. For example, not all children will have a complete set of service records or service records and family interviews. Others may have individual items missing from some instruments. Many complex multivariate analyses will eliminate a sample member with incomplete data for the variables included in the analyses. Hence, to ensure that a sample of 1,300 is available for many of the analyses, data were sought for many more children and families at the outset.

This array of issues was considered carefully in determining a reasonable NEELS study sample size. The size of the child sample that maximized the precision of estimates for the study within the available resources was 3,300 children and families. Using the attrition assumptions noted above, 2,400 children were expected to remain in the sample at the end of the 3 years of data collection, more than required for many analyses and a solid basis to justify continued data collection until kindergarten entry. In actuality, more than 3,000 children remained after 3 years. With an initial sample of 3,338, each sampled child represents approximately 55 Part C children nationally.

The sample children needed to be selected so that their experiences represent the wide range of Part C services and delivery characteristics nationwide. The following sections describe how the child sample was allocated among states and local areas.

Sampling States

A sample of 20 states was deemed adequate to represent the key dimensions of Part C variation at the state level. The final state sample reflects the following considerations.

Number of children served. Selecting states that serve large numbers of Part C children enhances the representativeness and generalizability of the final sample. States were assigned probabilities of being selected for the sample on the basis of the percentage of the Part C-eligible population served. Data reported to OSEP for December 1, 1994, indicated that the total number of children served in Part C was 165,072, with the numbers for individual states ranging from 210 to 19,471 (U.S. Department of Education, 1996).¹ An estimate of the percentage of the birth-to-age-3 population that are served under Part C

¹ Part C child count data are from the U.S. Department of Education (1996). Population data are from the U.S. Census Bureau (1995). These were the most current data available when the state sample was selected in March 1996.

also varied considerably among states, ranging from .34% to 6.78%. Although the mechanisms contributing to this variation are not well understood, there may be something fundamentally different about the provision of Part C services in states serving such different percentages of young children. Those differences are potentially important to a national study of the program and must be accounted for in the sample by including enough states to represent important dimensions of variation.

Eligibility. States also vary in the eligible populations they serve, particularly with regard to the inclusion of children designated as “at risk.” Part C leaves it to states’ discretion whether to include children at risk in their eligible population. Not only do states differ with regard to whether they are serving children at risk, they also differ in how they have defined the at-risk population (Shackelford, 1998). Some states that have opted to serve children at risk would need to be included in the state sample.

Geographic dispersion and population size. Selecting states that represent different geographic regions of the country and population sizes will increase the credibility of the sample. The administrative challenges associated with state population size and geography have permeated many discussions about Part C. Again, these appear to be critical variations for a national study, and the sample should be balanced with regard to geographic region of the country and population.

Administrative variations. Important program variations may result from differences in states’ designated lead agencies for Part C. The impact of differences in lead agencies, especially at the local level, is not well understood, but lead agency is such a salient characteristic of Part C implementation that it could not be overlooked in state selection.

Traditionally underrepresented groups. These groups, identified in the Part C legislation (minority, low income, inner city, rural), are not distributed evenly among the states in the nation nor among local areas. The state and county sample must include sufficient numbers of states and counties with reasonably large populations from underrepresented groups to ensure their representation in the family sample, particularly African-American and Hispanic families.

Preliminary work indicated that nearly all combinations of 20 out of 50 states would produce a range of lead agencies, some states serving at-risk children, and some states with underrepresented groups, so these considerations did not need to be included directly as sampling criteria.

Given these considerations, states were selected for inclusion in the Part C study in the following way. An examination of the available state-reported counts of children served (December 1, 1994) showed that the nine states serving the largest numbers of children under Part C accounted for approximately 60% of the children/families participating in Part C nationally. These nine states were sampled with certainty: California, Florida, Illinois, Massachusetts, New York, North Carolina, Ohio, Pennsylvania, and Texas.

To select the remaining states in the sample, the country was divided into regions as follows:

1. The South: Alabama, Arkansas, Georgia, Florida, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia, West Virginia.
2. The West: all states west of the Mississippi River, except Texas, which is included in the South.
3. The East: all remaining states.

An additional 11 states were selected randomly from these regions with probability proportional to the size of the state's birth-to-age-3 population. The number of states selected per region took into account the number of states from each region already selected as part of the nine states selected with certainty. The supplemental states were as follows:

1. The South: Arkansas, Maryland, South Carolina.
2. The West: Colorado, Hawaii, Idaho, Kansas, New Mexico, South Dakota.
3. The East: Michigan, Minnesota.

Table 1 shows the 1995 U.S. Department of Education data for the number of Part C children served, by state, which became available shortly after the state sample was selected and were used for the next step in the sampling process. The 20 sampled states served 70% of Part C children nationally at the time of selection, indicating strong generalizability from the sample to the nation. The table also indicates the percentage of the birth-to-age-3 population served for the 20 states selected, which varied from 1.06 to 6.73, almost as wide a variation as among the states nationally. Overall, the percentage of the birth-to-age-3 population served in the 20 sampled states was 1.62, slightly above the percentage served in all 50 states, 1.46.

Table 1
NEILS STATE SAMPLE WITH NUMBER AND PERCENTAGE OF CHILDREN
SERVED IN PART C AS OF DEC. 1, 1995

State	Number of Children Served	Percentage of Birth-to-3 Population Served
Arkansas	1,328	2.14
California	18,119	1.10
Colorado	3,917	2.47
Florida	10,771	1.89
Hawaii	3,874	6.73
Idaho	845	1.60
Illinois	8,029	1.46
Kansas	1,429	1.32
Maryland	3,695	1.71
Massachusetts	8,484	3.49
Michigan	4,384	1.10
Minnesota	2,622	1.39
New Mexico	1,747	2.14
New York	13,317	1.66
North Carolina	4,336	1.43
Ohio	15,205	3.39
Pennsylvania	6,845	1.49
South Carolina	1,897	1.23
South Dakota	376	1.22
Texas	10,078	1.06
TOTAL	121,298	

Source: U.S. Department of Education (1997).

Allocating the Child Sample among States

To identify the number of children/families to be enrolled per state, the total sample of 3,300 children first was allocated among the regions in proportion to the number of Part C children in each region, based on OSEP data for December 1, 1995. For example, the South included about 28% of Part C children served nationally; states sampled from that region received an allocation of 932 sample members, or 28.2% of the total sample. Similarly, the East, which had 45% of the national Part C population, received an allocation of 1,492 sample participants. Regional samples were further allocated among states within regions in proportion to the number of Part C children in each state relative to the other sample states. For example, New York had 13,317 participants out of a total of 53,500 Part C children in all states sampled in the East, or 25%; hence, the sample assigned to New York was 25% of the 1,492 Eastern sample participants, or 371 children.

It was important that the allocation approach assign a large enough number of children to each state to adequately represent that state; a sample of 40 children was deemed to be a reasonable minimum per state to address this concern. With the allocation approach described above, all states received 40 or more participants, except Idaho (with 33) and South Dakota (with 14). To represent these states more adequately, their sample sizes were increased to 50 and 40, respectively. These sample sizes would make the expected weights

for the sample members in these states comparable to those of other states in the same region.

Table 2 shows the final allocation of the NEILS sample to each state.

Table 2
ALLOCATION OF CHILD SAMPLE BY STATE

State	Child Allocation
Arkansas	61
California	347
Colorado	147
Florida	308
Hawaii	166
Idaho	50
Illinois	246
Kansas	62
Maryland	104
Massachusetts	236
Michigan	122
Minnesota	73
New Mexico	66
New York	371
North Carolina	122
Ohio	243
Pennsylvania	200
South Carolina	53
South Dakota	40
Texas	283
TOTAL	3,299

Sampling Local Communities

The concept of “community” within the context of Part C implementation is difficult to describe because of the great within-state variation in how Part C is being implemented. States are divided into local jurisdictions for the provision of services, but the nature and size of the jurisdictions vary from state to state. For example, in Maryland, early intervention is organized by county, but California organizes Part C services around regional centers, which may span more than one county or there may be several regional centers within a single county.

For sampling purposes, a single type of local unit had to be identified for which the population of children as known (so that the appropriate weights could be calculated). Counties were selected as the local sampling unit because the boundaries of counties are clear (unlike cities or towns), and counties organize many public service-providing agencies.

Three to five counties per state were considered sufficient to recruit the required sample of children.² Counties were selected with the probability of selection proportional to the size of the birth-to-age-3 population in the county, with the additional criterion that there be at least 10 children projected to be served under Part C in a county.³ This latter provision was necessary for efficiency and cost-effectiveness in sample recruitment efforts.

If one or more of the selected counties declined to participate in the study, an alternate county was selected as a replacement. The alternate counties had been preselected and represented the counties with probabilities of selection closest to those of the original counties. Three of the counties that were selected for the sample were not willing or able to participate, so three replacement counties were selected.

Because the urbanicity or ruralness of a community is important to the delivery of early intervention services, an adequate sample of children residing in rural areas was needed. Many of the selected counties include rural areas. On the basis of U.S. Census counts of the percentage of each sampled county's population that reside in rural versus nonrural areas,⁴ in conjunction with the 1995 data on the number of children served under Part C for the sample states, it was estimated that the final sample would include 400 (12%) children from rural areas.

Further, it was essential to the purposes of the study that the sample of families include enough children and families from traditionally underrepresented groups, particularly low-income residents and ethnic minorities, to allow for a valid examination of their outcomes in the analyses. According to 1990 Census data, the selected counties would have more than adequate representation of minority families. The 1990 Census data indicate that the total percentage of nonwhite persons in the United States was 19.7%, and the percentage of nonwhite persons in the selected counties was 28.9%. The total nonwhite population in the selected counties represents 35.3% of the total U.S. nonwhite population.

A total of 83 counties initially were selected for participation in the study. Near the end of the recruitment period (described below), it became clear that some of the originally selected counties would not be able to recruit the number of children that had been assigned to them. To ensure that the overall total number of children and families recruited would be adequate, 11 supplemental counties were added to the study, for a total of 94 counties.

² Study designers sought to minimize the number of counties per state because of the costs and logistical complexities of spreading the state sample among many counties. The liaison effort required to recruit counties and the multiple service agencies within them was substantial. These costs were weighed against the desire to select enough counties to recruit from both rural and nonrural areas and a range of community types. Three to five counties, usually four per state, was deemed an acceptable number from the perspectives of both cost and representativeness.

³ The projected number of children in Part C in a county in a year was calculated by multiplying the proportion of the state's birth-to-3 population that lived in the county by the total Part C state population.

⁴ The Census definition of *rural* refers to areas with fewer than 2,500 persons living in places incorporated as cities, villages, boroughs (except in Alaska and New York), and towns (except in the six New England states, New York, and Wisconsin), but excluding the rural portions of "extended cities."

Allocating Each State's Child Sample among Counties

Once the counties were selected, sample designers selected Part C recipients in counties with subsampling rates proportional to the inverse of the probability of selecting the county. This two-step sampling process yielded child weights that were approximately equal within states, thus approximating the ideal of random sampling of children with equal probabilities of selection. Therefore, this two-step procedure assured that children from counties of different sizes had approximately the same chance of being selected for inclusion. Minimizing the differences between the sampling weights improved the sampling efficiency (defined as the inverse of the ratio of the actual sample size to the size of a completely random sample that would yield equivalent standard errors of estimation). Typically, for very small counties, researchers specified that Part C recipients would be selected for an entire year; for moderate-size counties, Part C recipients would be selected for 5 to 9 months; for very large counties, Part C recipients would be selected for only 1 to 3 months. The exact duration of Part C recruitment for each county was dependent on the probability of selection for that county and each other county in the sample, as well as the total sample size requirement for the state. Using this approach, sampling efficiency for most states (not counting potential clustering effects) was 95% or greater, and the lowest sampling efficiency was a respectable 70%. The size of the child sample allocated to each county ranged from 4 to 134; 96% of the counties were projected to have 10 or more children; 71% were projected to have 20 or more children.

Recruiting the Child Sample

Children were recruited for the sample at the point they were found eligible for Part C services (i.e., at the time of signing the initial IFSP). During the recruitment period, every family with a child who was found eligible for Part C services and had not received Part C services previously and who met the study criteria was invited to participate in the study. The eligibility criteria for the study were:

- Being less than 31 months of age at the time the IFSP was signed.
- Having an English- or Spanish-speaking adult in the household who could answer questions about the child and family.
- Being the only child in the family recruited for the study (siblings and other children of multiple-birth sets were excluded).

Children 31 months of age or older would not be in early intervention long enough to reasonably expect much change by the children's 3rd birthdays. Thus, the sample was between the ages of birth and 31 months at enrollment. The study's limited resources could only support a family interview in English or Spanish, necessitating the second criterion. Only one child was recruited per family because the family would have needed to be interviewed for each child in the study, and the study team considered that level of burden excessive. If a family had more than one child entering early intervention (e.g., twins who both were eligible), data were collected on only one of the eligible children, selected by chance.

In recruiting children and families, it was necessary for researchers to work with local staff in selected counties to determine all possible points of entry into early intervention services. There often were multiple such programs or agencies in counties that served infants and toddlers who were eligible for Part C. All such programs or agencies within each selected county (or a sample of programs/agencies to represent the county) were invited to participate. For programs that served wider areas than the selected counties, only residents of the selected counties were recruited for the child sample.

Recruitment of the entire national sample of children and their families occurred from September 1997 until November 1998. The time needed to enroll the required sample varied from 1 to 14 months among the counties (a few counties needed an additional 2 months beyond the originally planned 12 months), depending on the sample size assigned to the county and the number of children routinely enrolled in early intervention in a year from the county.

The expected number of children who were eligible for service does not equal the number who were enrolled in the sample, however, because some families refused to participate. Families' participation in the study was entirely voluntary, and written consent was obtained at the time of recruitment for the study.

The total number of months that each county was projected to take to recruit the assigned child sample was calculated by estimating the actual number of children who were likely to be found newly eligible in a year, discounting that number by a factor of 15%, and dividing the total by 12 to obtain an expected monthly sample yield. Using this monthly projected yield, the number of months each county was expected to take to recruit the assigned sample was calculated.

To spread recruitment out over the recruitment period and avoid any potential seasonal bias in enrollment, each county was assigned a starting recruitment month on the basis of the expected length of its recruitment period plus a 1-month "fudge factor" in case extra time was needed to generate the sample. For example, all counties estimated to need an 11- or 12-month recruitment period began in month 1 of the enrollment period and recruited for the subsequent 11 months (with a 1-month addition if required) or 12 months (no "fudge factor"). All counties estimated to have a 4-month recruitment period were randomly assigned a starting month between month 1 and month 8, to allow for 4 months of recruitment and a possible additional month if needed. Counties recruited all children found eligible for service in the subsequent months until the total assigned sample was reached. Study enrollment originally was scheduled to be completed in 12 months but was extended by 2 months to achieve an adequate sample size in all locales.

The final number of children recruited in each state is shown in Table 3, along with the projected sample size shown previously in Table 2. Some states recruited more than the originally projected sample because of the supplemental counties that were added during the latter half of sample recruitment.

**Table 3
CHILD SAMPLES BY STATE: FINAL ENROLLMENT**

State	Projected Child Sample Size	Enrolled Child Sample Size
Arkansas	61	66
California	347	363
Colorado	147	94
Florida	308	289
Hawaii	166	92
Idaho	50	61
Illinois	246	253
Kansas	62	68
Maryland	104	135
Massachusetts	236	252
Michigan	122	126
Minnesota	73	81
New Mexico	66	47
New York	371	287
North Carolina	122	108
Ohio	243	277
Pennsylvania	200	191
South Carolina	53	61
South Dakota	40	47
Texas	283	320
TOTAL	3,300	3,338

The Enrollment and Study Samples

Two samples have been recruited as part of NEILS. The first sample is referred to as the enrollment sample. It includes all children who lived in the sample counties⁵ and were being enrolled in Part C early intervention for the first time⁶ during their county's recruitment period. This sample includes 5,668 children.

A second sample, which is a subset of the first, is used in the majority of NEILS analyses; it is referred to as the study sample. These are the 3,338 children who met the criteria for study participation and for whom an adult provided written consent for the child's and family's participation in the study. As noted previously, eligibility for the study sample required that the child and family meet the criteria for the enrollment sample, as well as:

⁵ Some enrolling programs served a population that spanned more than one county. In this case, only children who lived in the sample county were included in the enrollment data collection.

⁶ Children who were enrolled during the recruitment period but who had been in Part C early intervention previously were excluded from the enrollment data collection. However, children were not excluded if they were enrolled for the first time but were part of a family that had received early intervention previously for a different child. Multiple children in a family were included in the enrollment sample, although not in the study sample, as long as all children were being enrolled for the first time during the recruitment period.

- Being less than 31 months of age at the time the IFSP was signed.
- Having an English- or Spanish-speaking adult in the household who could answer questions about the child and family.
- Being the only child in the family recruited for the study (siblings and other children of multiple-birth sets were excluded).

Source for Enrollment Information

The enrollment data were provided by early intervention personnel in sample counties who enrolled families into early intervention. They were asked to complete a NEILS Family Information Form A (see Appendix) for all children who met the criteria for the enrollment sample during their county's study recruitment period. Local personnel were asked to complete an enrollment form within 30 days of the development of the family's IFSP. Completed forms were mailed to the study team. There was no personally identifying information on these forms. An additional form was completed for children who met the criteria for the study sample and whose parents agreed to participate in NEILS.

Procedures for Weighting the Enrollment Sample

Three weights were calculated for each child enrolled in early intervention in sampled counties during the enrollment period for whom enrollment data were provided:

- Weight 1 (denoted W1) was the value necessary to project the enrollees in each sampled county to the total number of early intervention enrollees in that county during an entire year.
- Weight 2 (denoted W2) was the value necessary to project the total number of enrollees during a year in the sampled counties to the total number of enrollees in the state on December 1, 1997⁷
- Weight 3 (denoted W3) was the value necessary to project the total number of enrollees on December 1, 1997, in the sampled states to the number of early intervention enrollees on that date in the regions containing those states.

The final weight for each sample enrollee was the product of the three weights (that is, $W1 \times W2 \times W3$). The procedure for determining each weight is described below.

Weighting the Enrollment Sample to County Totals

The first step in weighting the sample enrollees was to weight the enrollees in each sampled county to the total number of early intervention enrollees in that county during a 1-year interval. For example, if there were 20 sample enrollees in a county and it had been estimated that during a year the county would enroll 200 early intervention participants, then the first weight (W1) for each sample enrollee in the county would be 10. The challenge in performing this calculation was that the yearly number of early intervention enrollees in a county was not known and needed to be estimated.

⁷ December 1 is the date on which the state child count is taken for federal reporting purposes.

Two estimates of the number of early intervention enrollees in the sampled counties were available. The first estimate was derived via the following two-step process:

- The number of early intervention enrollees in the state on December 1, 1997, was divided by 3 to account for the fact that the participants on this date were enrolled over the preceding 3-year period and the study needed a yearly enrollment value.
- The resulting number of early intervention enrollees in the state were allocated to the counties on the basis of U.S. Census Bureau information concerning the number of residents in the county.

This estimate was relatively easy to calculate but had the following limitations:

- The number of early intervention participants on December 1, 1997, underestimates the total number of enrollees over a 3-year period, because some who enrolled in the prior 3 years would have left the program before this date.
- The estimation procedure does not adjust for county wealth. Research suggests that poorer counties should have proportionally more early intervention enrollees.⁸

The second estimate of the number of early intervention enrollees over a 1-year period was derived from the IFSP dates of the early intervention enrollees during the study enrollment period. That is, the number of enrollees per day during the enrollment period was calculated and multiplied by the number of days per year. This estimate had the following limitations:

- Some counties appeared to start the enrollment process, record the IFSP dates of some early intervention participants, then stop enrollment for an extended period of time. Presumably, these were counties with a “false start.” An outlier estimation procedure was used to identify these false starts, the outliers were not counted in the enrollment estimation process.
- Some counties appeared to start the enrollment process, eventually stopped enrolling, and then many days later enrolled one or two more children. Presumably, these were counties that stopped enrolling thinking that they had reached their study enrollment target, realized that they were still marginally short of the target, and then enrolled the last few children that were needed. An outlier estimation procedure was used to identify these “restarts,” and the outliers were not counted in the enrollment estimation process.
- Some counties enrolled only a portion of their early intervention participants. Typically, this occurred when there were multiple sites or persons involved in enrollment and only a portion of those sites or persons participated in the study.

Our final estimate of the number of yearly early intervention enrollees in each county was the larger of the first and second estimates, subject to the limitation that the number estimated could not exceed three times the first estimate (derived from the county population projection procedure).

⁸ Special education students are disproportionately poor, relative to the general student population (Marder & Cox, 1991). The relationship between poverty and disability could also apply to infants and toddlers with disabilities.

One variation on this procedure involved substitute or supplemental counties. If the originally selected county did not participate, enrollees from the substitute county were weighted as if the substitute county had been selected from the start. If the originally selected county was able to provide some enrollees, the additional enrollees from the supplemental counties were weighted as if they had been enrolled in the originally selected county. However, in the latter case, they were not counted when determining the number of enrollees in a year from the originally selected county.

Weighting County Totals to State Totals

The estimated numbers of enrollees in sampled counties over a 1-year period were weighted to state totals of the numbers of early intervention participants on December 1, 1997, via a four-step process:

- The inverse of the probability of county selection was calculated for each selected county, denoted as Y . For example, if the probability of selecting a particular county was 20%, then the inverse would be $Y = 5$. (Counties were originally selected from a state with probabilities that were proportional to the population-based estimate of the number of early intervention enrollees. Consequently, large counties tended to have larger probabilities of selection and smaller Y values, and vice-versa for small counties. However, counties with fewer than 10 estimated early intervention enrollees were not candidates for selection for budgetary reasons.)
- The quantity $W1 \times Y$ was calculated for each sample enrollee and summed over all enrollees in the state, denoted as Z . This is an estimate of the number of early intervention enrollees per year.
- Let Q denote the total number of early intervention participants in the state on December 1, 1997. The quantity (Q/Z) was calculated. This is the ratio of the estimated number of early intervention participants in the state on December 1, 1997, to the estimated number of yearly enrollees in the state.
- The second weight was calculated as $W2 = (Q/Z) \times Y$. This weight projects the number of yearly enrollees in the sampled counties during a year to the number of early intervention participants on December 1, 1997, in the state.

Weighting State Totals to Regional Totals

The numbers of early intervention enrollees on December 1, 1997, in the sampled states were weighted to the numbers of early intervention enrollees on December 1, 1997, in the regions of the country containing those states. For example, the Southern region contained 10 states, of which 6 were sampled. The total early intervention enrollment on December 1, 1997, in those 10 states was 52,895, and the total early intervention enrollment in the sampled states on that date was 36,283. The third weight for all enrollees in the sampled Southern states was calculated as $W3 = 52,895 / 36,283 = 1.458$.

In calculating $W3$, no adjustment was made for the relative probabilities of state selection. States with the largest early intervention enrollments in a region were selected with certainty; a few others in each region were selected with probability proportional to early intervention enrollments. Use of traditional weighting formulas would have resulted in the states selected with certainty having $W3 = 1.0$ and states selected with less than certainty

having larger weights. (For example, in the Southern region, the three states selected with less than certainty would have had weights averaging about 3.0.) Allowing the few states not selected with certainty to represent the remaining nonselected states would result in substantial variance in the estimates. Available information indicated that the set of selected states (randomly selected and purposefully selected) was similar to the nonselected states and that there was justification for treating all respondents in the selected states as if they were a random selection of all NEILS-eligible children in all states. To the extent this is not true, some bias would be introduced into the estimates. Sample designers considered it reasonable to trade off the possibility of some small bias for the resulting reduction in sample estimate variance.

Replicate Weights for the Enrollment Sample

There were 5,668 children for whom the NEILS Form A was completed. These children were all assigned Form A weights. The total of the Form A weights is 193,322, which represents the universe of all children enrolled to receive services during a 1-year period.

The NEILS sample is both stratified and clustered, so that calculating standard errors by formula is not straightforward. Standard errors for means and proportions can also be estimated using pseudo-replication, a procedure that is widely used by the U.S. Census Bureau and other federal agencies involved in fielding complex surveys. To that end, a set of weights was developed for each of 20 balanced half-replicate subsamples. Each half-replicate involved selecting half of the total set of counties that agreed to participate using a partially balanced factorial design (resulting in about half of the counties being selected within each stratum) and then weighting that half to represent the entire universe. (Since the primary sampling units were states, if the state was selected to be in a half-replicate, all of the counties within that state also were selected for that half-replicate. The exception was California, which comprised its own region. For California, half of the counties were assigned to each half-replicate.) The half-replicates were used to estimate the variance of a sample mean by: (1) calculating the mean of the variable of interest on the full sample and each half-sample using the appropriate weights; (2) calculating the squares of the deviations of the half-sample estimate from the full sample estimate; and (3) adding the squared deviations and dividing by $(n-1)$, where n is the number of half-replicates. Sixteen replicate Form A weights were derived.

Weighting Enrollment-Eligible Children

Of the children with Form A weights, some children were not eligible to participate in NEILS, typically because they were 31 months of age or older, they were a second twin in the family, or their families did not have an English- nor Spanish-speaking respondent. Of the 5,668 children with Form A weights, only 4,867 were eligible for enrollment in NEILS. (See "Recruiting the Child Sample" above.) When weighted by their Form A weights, these children represent the portion of the universe that was eligible for NEILS. This portion of the universe consists of an estimated 168,218 children.

Sixteen replicate enrollment-eligible weights were derived. The process for a replicate involved: (1) taking the Form A weights for that replicate and (2) setting any weights for individuals who were ineligible to zero. The replicate enrollment rates do not each total 168,218, although the totals are close to this amount.

Weighting Enrollment-Eligible Children with NEILS Participation Permission

Of the 4,867 children who were eligible for NEILS, permission to enroll in the study was received for 3,338. Weights (called the "NEILS participation weights") were calculated for these children so that they also represent the 168,218 children who were eligible for NEILS. These weights were calculated by modifying the enrollment eligibility weights for the 3,338 children using Deming's algorithm (Deming & Stephan, 1940), so that the marginal totals for the 3,338 children with NEILS participation permission equaled the marginal totals for the 4,867 NEILS-eligible children, on the following balancing variables:

1. Child's age in months at referral to early intervention (with categories of birth to less than 6 months, 6 months to less than 12 months, 12 months to less than 18 months, 18 months to less than 24 months, 24 months to less than 30 months, 30 months and above, and unknown).
2. Child's age in months at IFSP signing (with categories of birth to less than 6 months, 6 months to less than 12 months, 12 months to less than 18 months, 18 months to less than 24 months, 24 months to less than 30 months, 30 months and above).
3. Child's gender (male, female, or unknown).
4. Child's race (white, black, Hispanic, Asian/Pacific Islander, other, and unknown).
5. Whether the family has a functioning phone at home (yes, no, and unknown).
6. Whether the family receives public assistance (yes, no, and unknown).
7. Whether the child is in foster care (yes, no, and unknown).
8. The reason for eligibility (developmental delay, diagnosed condition with no mention of developmental delay, at risk of delay, and unknown).
9. Mutually exclusive categories that describe the child's delays and disabilities⁹:
 - Delayed development (except physical growth abnormality) and/or intellectual/cognitive delay, with no other impairment.
 - Delayed development (except physical growth abnormality) and/or intellectual/cognitive impairment/delay, with other impairment.
 - Speech/communication impairment or delay with no other impairment.
 - Speech/communication impairment or delay with other impairment (except hearing).
 - Motor delays, abnormal muscle tone, and/or cerebral palsy with no other impairment.
 - Motor delays, abnormal muscle tone, and/or cerebral palsy with other impairment.

⁹Children were assigned to one of these categories on the basis of information provided on the enrollment form. The classification system was developed for the study and is described in Spiker et al. (2001). Although many children had multiple disability descriptors, a set of decision rules was developed to allow each child to be placed in one category for purposes of this analysis.

- Hearing impairments.
- Visual impairments.
- Social/behavioral impairments/delays.
- Congenital disorders.
- Prenatal/perinatal disorders.
- Central nervous system disorders.
- Social/environmental risk factors.
- All other impairments.
- Unknown.

Sixteen replicate participation permission weights were derived. The process for a replicate involved: (1) taking the enrollment-eligible weights for that replicate, (2) setting to zero, any weights for individuals who did not receive permission to enroll and (3) balancing the resulting set of weights to the marginal totals from the entire set of enrollment-eligible weights for that replicate.

Weighting Enrollment Interviews

Of the 3,338 children who were enrollment eligible and for whom permission for granted for NEILS participation, partial or complete enrollment interview data were received for 3,234 children (96.9%). These respondents were weighted up to the 168,218 children who were eligible for NEILS. The weighting process involved assigning the 3,234 children their NEILS participation weights and then adjusting those weights so that they had the same marginal distribution as the 3,338 children who were NEILS participants weighted by their NEILS participation weights.

Sixteen replicate enrollment interview weights were derived. The process for a replicate involved: (1) taking the participation weights for that replicate, (2) setting to zero any weights for individuals who were not interviewed and (3) balancing the resulting set of weights to the marginal totals from the NEILS participation weights for that replicate.

Weighting the First Interim Interview

The first interim interview was conducted 1 year after the IFSP date (± 2 months). The interview was attempted with all families whose child was less than 33 months of age at the time of the interview scheduling. Of the 2,047 families who met that criterion, interim interviews were completed with 1,629, or 79.6%. For purposes of calculating the first interim interview weights, the 3,338 NEILS participating children were divided into the following categories:

- A. Children who were too old to be eligible for the first interim interview (i.e., 644 days old or older at Part C enrollment). There were 1,259 children in this category. This category also includes children who had died previously but would have satisfied the conditions of the previous statement had they not died.
- B. Children who had died but otherwise would not have satisfied the conditions for belonging to category A. There were 32 children in this category.

- C. Children who had acceptable interim interview data (i.e., interview data completed through at least the end of Section A of the interview). There were 1,629 children in this category.
- D. Children who had unacceptable interim interview data (including those whose families were not interviewed). There were 418 children in this category.

Weighting was accomplished in the following steps:

1. Children in category C were weighted to represent children in categories C and D. Symbolically, $C \rightarrow C + D$. That is, the NEILS participation weights for children in C were modified so that they summed to the same total as the NEILS participation weights for the children in categories C and D, and so that the marginal distributions on the balancing variables were approximately equal.
2. For children in categories A and B, the first interim interview weight is their NEILS participation weight; for children in category C, the first interim interview weight is that derived in step 1; for children in category D, the first interim interview weight is zero, since the portion of the universe that they represent has been assigned to children in category C. However, since only children in category C have adequate data for analysis, analyses on the first interim interview data are restricted to category C children.

Sixteen replicate first interim interview weights were derived. The process for a replicate involved: (1) taking the full-sample first interim interview weight for that replicate (including weights for category A, B, and C children) and (2) balancing the resulting set of weights to the marginal totals from the participation permission weights for that replicate.

Weighting the 36-Month Family Interview

All of the 3,338 children with enrollment permission potentially were eligible for the interview when the child turned 36 months of age. Some children (59) had died by the time of the interview reducing the number of potential interviews to 3,279. Interview data were collected for 2,756 (84%). For purposes of weighting, children were divided into three categories: (A) 59 children who were deceased by the time of the interview (i.e., before 40 months of age), (B) 2,756 children with acceptable transition interviews, and (C) 523 children who were not deceased and who did not have acceptable transition interviews. Category A is self-representing. Category B has been weighted to represent categories B and C.

Weighting the First Service Record

Service records were sent to the service provider identified at study enrollment as the one most knowledgeable about the services received by the child and family. Service records were sent out every 6 months to cover the previous 6-month period as long as the child stayed in early intervention. Service data for the first 6 months in early intervention were provided for 2,662 children (79.7%).

For purposes of calculating the first service record weights, the 3,338 NEILS participating children were divided into the following categories:

- A. Children who were enrolled in Part C for the entire 6-month period and had acceptable service data. A child was known to be enrolled in Part C services if the response to question 1a of the service record (“Is this child still enrolled in early intervention in your community?”) was “yes.” A child was considered to have acceptable service information if any response (including “none”) to question 2 (“Which of the following early intervention services were provided to this child or family in the past 6 months?”) was provided. There were 2,255 children in this category.
- B. Children who were enrolled in Part C for the entire 6 months and had unacceptable service data. There were no children in this category.
- C. Children who were enrolled in Part C for less than the entire 6-month period and had acceptable service data. (This group includes those who died during the 6 months and had acceptable service data.) There were 407 children in this category.
- D. Children who were enrolled in Part C for less than the entire 6 months and had unacceptable service data. (This group includes those who died during the 6 months and had unacceptable service data.) There were 36 children in this category.
- E. Children who were missing data as to whether they were enrolled for the full 6 months and whether or not they had received services during the 6-month period covered by the first service record. These were children missing a first service record. There were 640 children in this category.

The first step in the weighting was to weight the children in categories A, B, C, and D so that they represented the children in categories A, B, C, D, and E. Symbolically, $A + B + C + D \rightarrow A + B + C + D + E$. That is, the NEILS participation weights for children in categories A through D were modified so that they summed to the same total as the NEILS participation weights for the children in categories A through E, and so that the marginal distributions on the balancing variables were approximately equal.

The second step in weighting was to weight the children in category A so that they represented the children in categories A and B, and to weight the children in category C so that they represented the children in categories C and D. Symbolically, $A \rightarrow A + B$ and $C \rightarrow C + D$. That is, the weights from step 1 for children in category A were modified so that they summed to the same total as the step 1 weights for the children in categories A and B, and so that the marginal distributions on the balancing variables were approximately equal.

The resulting first service record weights for the children in categories A and C are those derived in step 2; the weights for categories B, D, and E are zero (since the portion of the universe represented by those children has been assigned to children in categories A and C). However, since only children in categories A and C have adequate data for analysis, analyses on the first service record were restricted to category A and C children.

Sixteen replicate first service record weights were derived. The process for a replicate involved: (1) taking the full-sample first service record weight for that replicate (including weights for category A and C children) and (2) balancing the resulting set of weights to the marginal totals on the balancing variables using the participation weights for that replicate.

Data Analysis Procedures

All data were analyzed with SUDAAN (Shah, Barnwell, & Bieler, 1997), which is a software package designed for analyses of correlated data resulting from complex sampling designs. The data tables included in NEILS reports present data for the full sample of infants and toddlers with disabilities and for important subgroups (e.g., children who were eligible for early intervention for different reasons). Most of the variables presented in tables are reported as weighted percentages of the children in the full sample or in the subgroups. In some cases, rather than percentages, the figures refer to means, such as the mean age at which children entered early intervention. Percentages and means are weighted to represent the national population of children entering early intervention. However, the percentages and means are only estimates of the actual percentages and means that would be obtained if all children entering early intervention were included in the study. These estimates vary in how closely they approximate the true measures that would be derived from a study of all children entering early intervention. To aid the reader in determining the precision of the estimates, for each percentage and mean, NEILS reports present the approximate standard error and the unweighted number of cases on which the statistic is based.

To determine the precision of a particular percentage or mean, the reader can construct a confidence interval for the estimate by multiplying the standard error by 1.96. The result is the range around the estimate within which the true measure would be found 95 out of 100 times. For example, NEILS estimates that 60.91% of the children entering early intervention are male. The standard error of that estimate, 1.09, is multiplied by 1.96, letting us assume with 95% confidence that the true percentage of males falls within a range of ± 2.14 percentage points of 60.91, or between 58.77% and 63.05%.

Readers also may want to compare percentages or means for different subgroups to determine, for example, whether the difference in the average age at enrollment into early intervention between children who are eligible because of a developmental delay and those who are eligible because of a diagnosed condition is statistically significant. To calculate whether the difference between two values is statistically significant with 95% confidence (denoted as $p < .05$), the squared difference between the two values of interest is divided by the sum of the two squared standard errors. If the result is larger than 3.94, the difference is statistically significant. Presented as a formula, a difference between two values is statistically significant at the .05 level if:

$$\frac{(V_1 - V_2)^2}{SE_1^2 + SE_2^2} > 1.96^2$$

where V_1 and SE_1 are the first value and its standard error and V_2 and SE_2 are the second value and its standard error.

Data Limitations

The most notable limitation on the data involves the potential for bias resulting from incomplete data. The study team made every effort to enlist all counties and to obtain enrollment information for every family who lived in the sample counties and enrolled in early intervention for the first time during the study recruitment period. However, there were three reasons that information was not received for all such families.

The first reason involves the dropping out of three counties that never enrolled any children. These three counties were replaced by three other counties that resembled the originals closely. However, it is not possible without a nonresponse study to know the extent or amount of bias that may have resulted from the unaccounted-for differences between the original and replacement counties. Study resources were not sufficient to conduct a nonresponse study.

The second reason involves the refusal to participate in the study by some early intervention programs in a few counties. In this case, the target numbers of families needed for recruitment into the study were allocated to the remaining programs in the county. However, it is possible that the families who would have been enrolled by the refusing programs would be different from those enrolled by the participating programs in the county. This could occur, for example, if a refusing program served a geographic area in the county with a population that differed markedly from that of the geographic area served by the other programs (e.g., an inner city vs. a suburban area) or if the refusing program enrolled children with a specific kind of disability and those children generally were not enrolled by the participating programs. If refusing programs did enroll a different clientele than participating programs, bias would be introduced into the enrolled sample. Again, without a nonresponse study, it is impossible to know the extent to which bias has been introduced into the study or the kind or amount of bias that may have resulted from the refusal of some programs to participate in the study.

The third reason that information was not collected on all enrolling families is uneven adherence to study recruitment procedures in some enrolling programs. Local enrollment personnel were told in the recruiting instructions that "It is critical that we get the information on the NEILS Family Information Form A for all families who completed their initial IFSP during the county's enrollment period. This information will allow us to describe the population of children and families entering early intervention for the first time and to determine the extent to which families who choose to participate differ from those who choose not to participate in NEILS." No family identification information was included on Form A, to remove any disincentive to completing the form that might have resulted from concerns about anonymity.

Despite the importance placed on completing Form A for all enrolling families, there is reason to believe that not all early intervention personnel completed forms for all families. Sometimes this failure resulted from a misunderstanding of enrollment procedures that was corrected by the study site liaisons when it was discovered. In other sites, program staff reported that excessive workloads led them to complete forms during some time periods only for families who also agreed to participate in the study. When this circumstance was

discovered, sites were asked to complete enrollment information for the missed families retroactively, using data available in their program files. Data were not recovered in all cases, however.

There is no reason to believe that the families for whom Form A data were provided differ in any systematic way from families for whom data were not provided, so there is no known bias introduced by the incompleteness of information that results from this cause.

References

- Deming, W. E., & Stephan, F. F. (1940). On least square adjustment of a sampled frequency table when the expected marginal totals are known. *Annals of Mathematical Statistics, 11*, 427-444.
- Shackelford, J. (1998). *State and jurisdictional eligibility definitions for infants and toddlers with disabilities under IDEA*. Chapel Hill, NC: National Early Childhood Technical Assistance Team, Frank Porter Graham Child Development Center, University of North Carolina at Chapel Hill.
- Shah, B. V., Barnwell, B. G., & Bieler, G. S. (1997). *SUDAAN: Software for the statistical analysis of correlated data*. Research Triangle Park, NC: Research Triangle Institute.
- Spiker, D., Hebbeler, K., Wagner, M., Scarborough, A., Simeonsson, R., & Collier, M. (2001). *A first look at the characteristics of children and families entering early intervention services*. Menlo Park, CA: SRI International.
- Marder, C., & Cox, R. (1991). More than a label: Characteristics of youth with disabilities. In M. Wagner, L. Newman, R. D'Amico, et al. (Eds.), *Youth with disabilities: How are they doing? The first comprehensive report from the National Longitudinal Transition Study of Special Education Students*. Menlo Park, CA: SRI International.
- U.S. Department of Education. (1996). *Eighteenth annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: Author.
- U.S. Department of Education. (1997). *Nineteenth annual report to Congress on the implementation of the Individuals with Disabilities Education Act*. Washington, DC: Author.
- U.S. Census Bureau. (1995). *Population profile of the United States: 1995*. Washington DC: U.S. Government Printing Office.

NEILS INVITATION CRITERIA

1. Was the child younger than 31 months old on the date the IFSP was signed? *(Circle one number)*

1---Yes

2---No

2. Is there an English- or Spanish-speaking adult in the household to respond to a phone interview?

(Circle one number)

1---Yes

2---No

3. **ONLY ONE CHILD PER FAMILY IS ELIGIBLE FOR NEILS.** Is this **the only child** from this family starting early intervention **OR** if there are multiple children from the same family in early intervention, is this the **one child** eligible for NEILS (per instructions on pages 5-6 of the NEILS Enrollment Procedures Manual)?

1---Yes

2---No

- o If **YES** to **ALL** of the criteria above, invite the family to participate in NEILS.

Go to the NEILS Enrollment Form–Part B.

- o If **NO** to **ANY** of the criteria above, the family is not eligible for NEILS.

STOP HERE. Return forms to the NEILS Enrollment Coordinator for your site for mailing to SRI.

Questions? Call the NEILS Hotline toll free: 1-800-682-9319