

**THE CONTINUING SURVEY OF FOOD INTAKES
BY INDIVIDUALS AND THE DIET AND
HEALTH KNOWLEDGE SURVEY**

- 1989 -

SURVEY OPERATIONS REPORT

Prepared for:

**HUMAN NUTRITION INFORMATION SERVICE
UNITED STATES DEPARTMENT OF AGRICULTURE**

Prepared by:

NATIONAL ANALYSTS
A Division of Booz•Allen & Hamilton Inc.

OCTOBER 1991

TABLE OF CONTENTS

	<u>Page #</u>
I. INTRODUCTION	I-1
II. SAMPLE DESIGN, SELECTION AND WEIGHTING	II-1
III. SURVEY PREPARATION	III-1
IV. DATA COLLECTION	IV-1
V. DATA REDUCTION AND FILE PREPARATION	V-1

I. INTRODUCTION

The United States Department of Agriculture (USDA) monitors food consumption patterns of the American public. USDA contracted with National Analysts, a division of Booz, Allen & Hamilton, to conduct the Continuing Survey of Food Intake by Individuals (CSFII) 1989. The procedures and protocol used to design the multi-year CSFII survey system and the effort involved in the conduct of CSFII 1989 are the subject of this report.

II. SAMPLE DESIGN, SELECTION AND WEIGHTING

This chapter presents a discussion of design, sample selection and sample weighting methods associated with all years of the Continuing Survey and with the implementation of the basic and low-income sample designs for Year 1, CSFII 1989. The discussion opens with an overview of National Analysts' Master Sample which is the framework for this and the prior CSFII surveys of 1985 and 1986. It then focuses on the specific elements of the current sampling plan and closes with a detailed description of the weighting of the CSFII and DHKS interviews.

A. The Sample Design for the CSFII Is a Three-Stage National Probability Sample of Households That Utilized National Analysts' Master Sample

National Analysts' Master Sample, the fourth in its history, is a stratified, clustered sample of 240 Census-defined areas, using population counts projected to 1985. It was designed to serve as the first stage for all of National Analysts' multi-stage national probability samples of households. Half of the Master Sample -- 120 areas -- were used for the CSFII survey. Within each of these 120 first-stage units, two additional stages of sampling -- area segments and households -- were conducted.

The first step in selecting the Master Sample was to partition the 48 conterminous states and D.C. into over 3,000 land areas referred to as *primary sampling units* (PSUs). There are three basic types of PSUs that were defined for the Master Sample. The definition of each of these types involves **metropolitan statistical areas (MSAs)** that are defined by the Office of Management and Budget (OMB) based on information and recommendations provided by the Census Bureau. MSAs are large population centers which are defined as one or more adjacent counties or county equivalents (i.e., independent cities and parishes), except in some New England states where MSAs are defined in terms of subcounty units, Minor Civil Divisions (MCDs). Based on the 1980 Census, OMB defined 316 MSAs.

The three types of PSUs that were defined for the Master Sample correspond to three levels of urbanization, as follows:

- **Central Cities** -- All OMB-designated central cities, as defined by their corporate city limits, located in 1980 MSAs. These are primarily the urban cores of the MSAs. Though some MSAs contain no central city, most MSAs contain one or more. In total there were over 400 central city PSUs defined among the 316 MSAs.
- **Suburban Counties** -- All counties, or county equivalents, located in MSAs, excluding any central cities. In those states where MCDs are used to define

MSAs, a suburban PSU was defined as the MSA part of a county (excluding any central cities). There were nearly 700 PSUs of this type defined for the Master Sample.

- **Non-metropolitan counties** -- All counties, or county equivalents, that were located outside of 1980 MSAs. In defining PSUs, counties that had a projected 1985 population of less than 10,000 persons were combined with one or more near by non-metro counties in the same state. (The 1985 projected populations were based on extrapolations of Census information for 1970, 1975, and 1980.) Also, in those states for which MSAs are defined in terms of MCDs, PSUs were defined in terms of the non-metropolitan parts of the counties. Roughly 2,000 PSUs of this type were defined for the Master Sample.
1. PSUs defined for the Master Sample were stratified in two steps

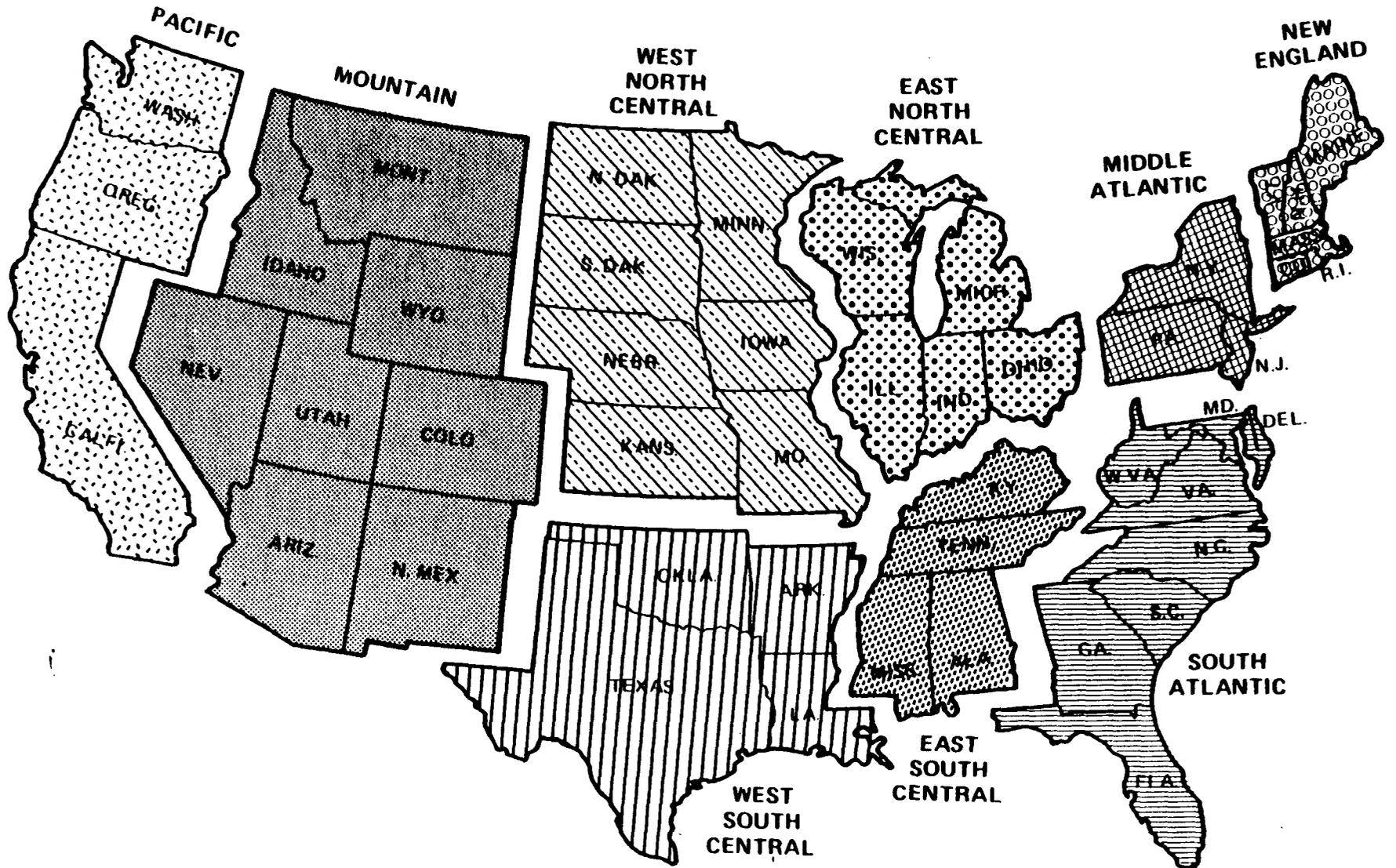
As the first step in the stratification of PSUs, the PSUs were placed into 27 super strata, defined by the nine Census divisions and three levels of urbanization (central cities, suburbia, and non-metro). The nine Census divisions are shown on the next page and are listed below:

- New England
- Middle Atlantic
- East North Central
- West North Central
- South Atlantic
- East South Central
- West South Central
- Mountain
- Pacific

The second step in the stratification of PSUs was to define 60 (final) strata by subdividing the PSUs in each of the super strata into one or more strata. The number of strata defined in a super stratum was determined by the projected 1985 population for the super stratum. The goal was to create strata that were comprised of geographically proximal areas and that were approximately equal in terms of their 1985 projected populations -- each stratum containing roughly 4,000,000 persons (i.e., 240,000,000/60).

Exhibit II-1

CENSUS DIVISION STRATA



In grouping PSUs within a super stratum to form the (final) strata, an attempt was made to put PSUs together that were homogeneous with respect to geography, industry, political structure, and social and demographic characteristics.

The choice of 60 (final) strata for the Master Sample was based on the desire to select a total of 240 PSUs in four independent replications of 60 PSUs each (i.e., one PSU from each stratum).

2. Two of the four replicates selected for National Analysts' Master Sample were used for the CSFII

National Analysts' Master Sample was selected in four independent replications. Each replication consisted of the selection of one PSU from each of the 60 strata, with probability proportional to the PSU's projected 1985 population. In order to provide 120 PSUs for CSFII, the first two replicates of the Master Sample were used for the CSFII.

Since the PSUs were selected independently for each replicate, several of the PSUs were selected for more than one replicate. For the first two replicates, seven of the PSUs were selected for both, leaving 113 unique PSU selections for CSFII.

The selection of PSUs in independent replicates allows for straightforward variance estimation. For the CSFII, since two independent PSU selections are made from each of the 60 strata, 60 degrees of freedom (one from each stratum PSU pair) are available for variance estimation.

Within each of the 120 PSUs selected for the CSFII, two additional stages of sampling were conducted -- one for area segments and one for households within segments.

3. The second stage sampling units -- area segments -- for four years of the CSFII were identified and selected prior to the start of Year 1 data collection

In each PSU, specific areas which define groups of housing units were identified for contact in the CSFII study. These areas, known as *area segments*, are small land masses defined in terms of Census geographic units [i.e., blocks, block numbering areas (BNAs) and enumeration districts (EDs)] and contain known numbers of housing units, based on 1980 Census data.

Segments were defined to contain a minimum of 75 households in terms of the 1980 Census. In areas of the country where Census blocks were used to define the area segment -- primarily heavily populated areas -- two or more adjacent blocks often were combined in constructing segments.

On the other hand, in areas for which EDs were the smallest Census unit defined -- primarily sparsely populated areas -- EDs rarely had to be combined. More typically, EDs contained too many housing units for a segment and covered large land areas. In order to reduce the field burden in such areas, EDs were often subdivided in the office into "chunks" that contained an average of about 100 households. One of these chunks was selected at random for the survey from each ED of this type. For example, suppose that an ED selected for the survey as a segment contained 516 households as enumerated in the 1980 Census. Using Census ED maps, an attempt would be made to define five ED chunks that would average 103 households. One of these chunks would be randomly selected for the survey.

In defining the boundaries of ED chunks, roads, railroad tracks, streams, and other identifiable landmarks were used. The goal was to define chunks that could easily be located in the field and that contained about the same number of households. It was difficult to equalize the number of households in a chunk since the Census ED maps did not indicate the location of households. Consequently, the number of households in some selected chunks varied considerably from the ED average. (In a few cases selected chunks contained no housing units.) Though the variation in chunk sizes adds to the variance of survey estimates, no bias is introduced by it.

The total number of area segments to be selected for the four-year draw was dictated by the desire to maintain continuity with past CSFII and other food surveys performed for USDA. Historically, National Analysts has targeted completion of an average of not more than six interviews per segment in order to achieve a healthy dispersion of surveyed households throughout the designated areas. With this goal in mind, a total of 1040 basic segments (260 per year) and 2000 low-income sample segments were selected (500 per year).

In order to draw the CSFII sample efficiently and to avoid selection of the same areas into more than one year's sample, all area segments for use in the multi-year CSFII monitoring effort were selected prior to the onset of Year 1 interviewing. The total number of segments to be selected was allocated to each PSU in proportion to the PSU's 1985 projected stratum population. The allocated number of segments was selected with probabilities proportional to the 1980 Census count of housing units. Selections were made using a systematic selection with a random start (systematic PPS) among PSU block and ED listings.

Two separate draws were programmed, one for area segments to be used in the basic sample and for area segments for consideration in the low-income sample. Although the possibility of overlap in area segments between the basic and low-income samples existed, no areas were identified as falling into both samples by chance in Year 1.

- a. Virtually all selected area segments were included in the basic CSFII sample, with half being fielded each wave

All 260 area segments selected for CSFII Year 1 were sketched and these sketch maps of the areas, along with the corresponding portions of Census maps, were sent to the field so that a complete enumeration of the housing units contained therein could be made (see Appendix A for Listing Manual). Only one segment in the Miami area could not be listed because, at the time of listing, it was designated a riot area and police were not permitting access to nonresidents. In addition, six areas, mostly ED chunks, were found to have no housing units or so very few that they were subsequently dropped from the sample.

There were two exceptions to the rule of listing all households in a segment. First, for segments that were thought contained more than 200 housing units, based on the 1980 Census -- typically large blocks -- listers were instructed to list only the first 100 households in the segment and to provide an estimate of the number of remaining housing units. In such cases, listers were given a specific location in the segment -- generally a block corner -- to begin listing and a direction to follow so that they could not have any choice of which 100 households to list. The location of these starting points was varied from segment to segment.

The other exception was those segments for which the lister *found* more than 200 housing units in the segment even though there were fewer than 200 enumerated by Census in 1980. In these instances, the lister was instructed to list the first 200 households and provide an estimate of the number of remaining housing units in the segment.

Once all the listing of housing units in all segments had been completed and reviewed by our sampling statisticians, a subsample of units was identified for contact for each wave of Year 1 fieldwork. Because the targeted number of completed interviews would be extremely low on a quarterly basis if all of the segments were activated each wave (6 interviews per segment = 1.5 completed interviews, on average, per wave x 4 waves), we fielded half of the areas in each three-month period. Hence, all of the basic area segments were randomly assigned, on a systematic basis, to one of two subsets. One subset of segments was assigned to Waves 1 and 3 and the other to Waves 2 and 4.

The actual distribution of basic area segments by wave, after adjustment for the riot and "zero" segments, is:

Waves 1 and 3	125
Waves 2 and 4	128

- b. To increase field efficiency, area segments drawn into the low-income sample were subsampled based on degree of near-poverty in the area

To qualify for interview in the low-income survey, households were required to meet an income-and-size criterion. The determination of interview eligibility could only be established at the time of contact. Therefore, sample housing units had to be screened in person and could not be pre-identified for interview. To increase the chances of contacting eligible households in the field, low-income area segments were differentially sampled. Segments in areas with higher rates of near poverty were oversampled while those with lower rates were undersampled.

First, 500 low-income sample area segments were randomly assigned to each of the four sample years. Then, all of the 2000 low-income area segments were classified into one of three income strata based on 1980 Census information for the tract or ED in which they were located. Segments in tracts or EDs with less than 10% of the households at or below the 125% poverty threshold were subsampled at a rate of one in four. Segments with between 10% and 24.9% of the households at or below this threshold were subsampled at a rate of one in 2.5. All segments located in tracts or EDs with 25% or more of the households at or below the threshold were retained for the sample (i.e., subsampled with probability equal 1.0).

The result of this differential sampling of the Year 1 area segments yielded 230 segments which were distributed across the three income strata as shown in Table II-1. As with the basic sample, half of these segments were fielded in Waves 1 and 3 while the other half were put into the field in Waves 2 and 4. In these segments, the differential eligibility hoped for, was achieved (see Table II-2). That is, as many as one in four households screened in the high-poverty segments were eligible for the low-income survey and dropped to only one in twenty in the low-poverty segments.

5. Four times a year, sample housing units were selected for contact in the field to yield a targeted total of 2250 completed CSFII interviews

The number of CSFII interviews to be completed in the basic survey each year was 1500, with approximately 375 interviews to be taken in each of four waves. The total number of completed low-income survey interviews was 750 per year, of which approximately 188 were to be completed each wave. To complete these 2250 interviews, a larger number of *sample housing units* had to be identified for contact to account for occupancy, eligibility and response rates.

Table II-1

**DISTRIBUTION OF LOW-INCOME AREA SEGMENTS
IN CSFII YEAR 1**

	<u>CSFII 1989 Segments</u>
<u>Low Poverty</u> -- Less than 10% of the households in the Census tract or Municipal Civic Division are at or below the 125% poverty threshold	44
<u>Medium Poverty</u> -- Between 10% and 24% of the households are at or below the 125% poverty threshold	100
<u>High Poverty</u> -- Twenty-five percent or more of the households are at or below the 125% poverty threshold	86
(n) =	(230)

Table II-2

**ELIGIBILITY RATES IN LOW-INCOME SAMPLE
BY TYPE OF AREA SEGMENT**

	<u>Type of Segment</u>		
	<u>Low Poverty</u>	<u>Medium Poverty</u>	<u>High Poverty</u>
	%	%	%
Screened, eligible households (interviewed, plus refusals to interview)	4.8	13.8	25.8
Screened, ineligible households	77.8	64.6	49.0
Other (e.g., vacant, refused screening)	17.5	21.8	25.15
(N) =	(1,236)	(2,473)	(1,873)

For the first wave of Year 1 of CSFII, information from prior like surveys was used to determine the number of sample housing units to be selected. In subsequent waves, the empirical results from earlier Year 1 efforts guided the determination of the number of sample HUs drawn. That is, the sampling rate for each succeeding wave was adjusted up or down, depending upon the field experience in preceding waves.

Once the total sample size was chosen for a wave for the basic survey, the sample was allocated to the segments in such a way that, within round-off error, all households in the segments assigned to that wave had the same total probability of selection. For the low-income survey, the sample for each wave was allocated to segments in such a way that the designed differential segments sampling rates were preserved.

For each segment, the sample housing units assigned to a given wave were selected systematically with a random start (an every *n*th unit sampling procedure). If additional HUs (not previously listed) were identified in the field at the time of the interviewer's visit, the HUs were brought into the sample using the *half-open interval method*. That is, any unlisted HUs occurring after a sample housing unit but before the next listed unit were added to the sample and identified for screening and, potentially, interview.

5. Intake information was sought from household members

In every qualifying household in the CSFII survey, individual household members were targeted to complete three days of intake recording. All permanent residents (i.e., those who regularly lived at the address) except roomers, boarders and live-in employees were eligible to complete records. Because of the difficulty in capturing information from traveling family members who were not available for the recording period, HNIS also chose to exempt these persons from the intake task.

6. All households participating in the CSFII portion of the survey were also eligible to take part in the DHKS survey as long as they continued to reside in their communities

In addition to completing a CSFII interview, a household was invited to complete a DHKS interview. In principle, all households were eligible for this survey if they qualified for the basic or low-income CSFII interview. In practice, only households completing a CSFII interview and still in residence in their communities at the point of the follow up DHKS contact were part of the DHKS sample frame.

B. All CSFII and DHKS Weights Were Based on Probabilities of Selection, Response Rates, and Census Bureau Total Household and Population Estimates

As is true for most well-designed weighting procedures, the weights for Year 1 were composed of the following three components:

- Weight factor(s) incorporating the probability of selection into the sample.

This component, which is the fundamental component of the weight, provides the statistical basis for making unbiased estimates. That is, it is a fundamental result in sampling theory that if there were perfect population coverage and no nonresponse, this weight component would allow for unbiased estimates of all population totals. This component is often referred to as the (basic) *sampling weight*.

- Weight factor(s) accounting for survey nonresponse.

Nonresponse creates biases of an unknown magnitude in the survey estimates. Although these biases cannot be eliminated, a sensible non-response weight adjustment procedure can reduce them.

- Weight adjustments to "known" population totals.

If known population totals, or good independent estimates of population totals are available, and if these are correlated to the survey variables, it is worthwhile to adjust weights to align with them. These types of adjustments can simultaneously reduce sampling errors, coverage bias, and nonresponse bias.

All weights computed for the CSFII were developed in the context of the three components listed above. Even though the sample was controlled in three-month quarterly draws, weights derived for respondents from all four waves simultaneously. This approach allows for the maximum flexibility in developing weight adjustment cells to account for nonresponse and to align with known totals.

The specific formulas and detailed discussions for the various weights that were assigned to CSFII and DHKS respondents are given in the following sequence: household (HH) weights for the basic and low-income surveys, person weights for the basic and low-income surveys, and weights for DHKS respondents.

1. Household weights for the basic survey were the most straightforward and were comprised of three components

The weight, WT(HH), for each cooperating household in the basic survey (i.e., each household that provided a satisfactory household questionnaire and a satisfactory 24-hour intake record for at least one of its members) was computed as follows:

$$WT(HH) = W(PR) \times W(NR) \times W(CEN) \quad (1)$$

where

W (PR) = the inverse of the household selection probability (i.e., the sampling weight)

W (NR) = the inverse of the household response rate for a segment or group of segments

W (CEN) = a weight adjustment factor based on household counts provided by the Census Bureau via HNIS for 27 cells defined by 9 geographic divisions and 3 levels of urbanization

The sampling weight, W(PR), in equation (1) is completely dictated by the sample selection procedures since they determine the probabilities of selection. This factor incorporates the HH selection probabilities at all three stages: (1) PSU selection, which was proportional to the projected 1985 HH counts, (2) segment selection which was proportional to the 1980 Census HH counts for the blocks, EDs, or ED chunks in the segment, and (3) the systematic selection of HHs from the listings within selected segments. The sample was designed so that, within round-off error, each household in the population would have an equal probability of selection (i.e., a self-weighting sample).

The nonresponse (NR) weight factor, W(NR), adjusts the weights of the respondents to account for the nonrespondents, by cell, which is generally a segment. The weight factor was calculated by dividing the number of *eligible* (i.e., occupied) HHs in the cell by the number of respondent HHs in the cell. Cells were defined in an attempt to put together respondent and nonrespondent HHs that have similar survey characteristics since the non-response weight factor essentially substitutes, for the missing survey characteristics of the nonrespondents in a cell, the average of the survey characteristics of the respondents in the cell.

The best available way to define nonresponse weight cells is in terms of segments since HHs in the same segment are usually close together geographically, and may have similar survey characteristics. However, in

some cases in which there were only one or two respondents in a segment, the nonresponse weight factor exceeded 2.5. In order to avoid the impact on survey variances of high weight adjustments, we generally limited this factor to 2.5 by combining two or more nearby segments to form a nonresponse cell. In a few instances this factor was allowed to go as high as 3.2 in order to keep the nonresponse cell within one PSU. When segments had to be combined from different PSUs to keep the nonresponse adjustment factor at or below 2.5, segments within the same Census division were grouped together.

Fortunately, good estimates of HH counts were available from the Census Bureau (via HNIS) to use to derive weight adjustments, $W(\text{CEN})$, which aligned the HH weights with these Census counts. These HH counts were from the Bureau's Current Population Survey (CPS) and were obtained by HNIS for 27 cells defined by the 9 Census divisions and 3 urbanization levels (central city, suburban and non-metro). The adjustment factor for each cell was computed by dividing the CPS HH estimate for the cell by the sum of the initial respondent weights based on the first two weight factors [i.e., $W(\text{PR}) \times W(\text{NR})$]. All 27 of these factors were 2.5 or less.

After this weight adjustment was applied, the sum of the final weights [i.e., $W(\text{PR}) \times W(\text{NR}) \times W(\text{CEN})$] was equal the CPS HH counts for each of the 27 cells. This adjustment attempts to correct for (1) sampling error and (2) coverage errors in terms of the field listing of HHs.

2. Household weights for the low-income survey were computed in a way that was very similar to that of the basic survey

The weighting expression given in equation (1) to compute HH weights for the basic survey were also used to compute HH weights for the low-income survey. However, the first two terms, $W(\text{PR})$ and $W(\text{NR})$, had to be modified slightly for the low-income sample.

First, the sampling weight, $W(\text{PR})$, had to take into account the differential selection probabilities applied to segments in terms of poverty-level designation. Sample segments in low-poverty level Census tracts were subsampled at the rate of 25%; therefore, respondent HHs in these segments had their sampling weights increased by a factor of 4. Sample segments in medium-poverty level tracts were subsampled at the rate of 40%; therefore, respondent HHs in these segments had their sampling weights increased by a factor of 2.5. Sampled segments in high-poverty level tracts were not subjected to subsampling; therefore, no additional weight factor was applied to respondent HHs in these segments. Because of the differential sampling by poverty level, the HH sample for the low-income survey was, by design, not "self-weighting."

Nonresponse adjustments are generally computed among *eligible* sample members. For the low-income survey, eligibility (i.e., household income at or below 130% of the poverty income criterion) could only be determined by screening. Therefore, the nonresponse adjustment weight factor, $W(NR)$, required the inclusion of an additional NR factor that reflected the NR to screening. (For the basic survey, we assumed that eligibility -- i.e., occupancy -- could be determined even for HHs that were not screened.) This additional nonresponse weight factor was simply the number of households eligible for screening (i.e., occupied) in a segment divided by the number of households screened in the segment. Implicit in this adjustment is the assumption that the eligibility rate among screened households in a segment equals that among those households not screened in the segment.

The nonresponse adjustment factor, $W(NR)$, for the low-income sample was the product of the screening nonresponse factor, described above, and the interview nonresponse factor. The second factor was similar to the NR adjustment factor used for the basic survey: the ratio of the number of screened *eligible* sample HHs to participating sample HHs within non-response cells defined by segments or groups of segments. As was done for the basic survey, segments were combined, as necessary, to keep the adjustment factor below 2.5, unless the cell could be restricted to a single PSU by allowing the factor to be as high as 3.0.

Since the response rate was higher for the low-income sample than for the basic sample, fewer segments had to be combined than for the basic sample to keep the nonresponse adjustment factors below 2.5. When segments did have to be combined, they were combined within the same poverty level, whenever possible. In those few cases where segments had to be combined across poverty levels, the interview nonresponse adjustment factor was computed using *weighted* counts of eligible and interviewed cases in order to account for the differing segment selection probabilities across poverty levels. That is, segment counts were weighted by 4, 2.5, or 1, depending on whether the segment was classified as a low, medium, or high poverty-level segment.

The third component, $W(CEN)$, of the HH weight for the low-income sample was computed the same way as that for the basic sample. In this case, however, the control total for each of the 27 cells was the CPS estimate of the number of low-income households, rather than of all households (see Table II-3). Several of these adjustments exceeded 2.5 requiring the combining of some of the 27 cells to develop Census adjustment factors that were under 2.5.

Table II-3

**HOUSEHOLD WEIGHT ADJUSTMENT FACTORS W(CEN) TO
CENSUS (1989 CPS) TOTALS FOR THE LOW-INCOME SAMPLE**

(Numbers in parentheses indicate cells that have been combined)

<u>Division</u>	<u>Zone</u>	<u>W(CEN)</u>
1	1	1.809912 (1)
1	2	1.647752 (2)
1	3	0.430872 (3)
2	1	1.809912 (1)
2	2	1.647752 (2)
2	3	0.430872 (3)
3	1	2.182806
3	2	0.702910 (4)
3	3	0.932692
4	1	1.157931
4	2	0.702910 (4)
4	3	0.952075
5	1	2.255640
5	2	0.789144
5	3	0.477415
6	1	1.930655
6	2	2.159828
6	3	1.599979
7	1	1.590024
7	2	1.045125
7	3	1.085558
8	1	1.407276
8	2	1.996997 (5)
8	3	1.996997 (5)
9	1	1.420397
9	2	1.742327 (6)
9	3	1.742327 (6)

3. Two sets of person weights were created to analyze intake data for both surveys

For households participating in the CSFII, interviewers attempted to obtain food intake data for three consecutive days -- the day before the interview, the day of the interview, and the day after the interview -- for all eligible household members. In order to allow for separate analysis of the Day 1 intakes and of the three-day intakes, a set of weights was computed for each of these subsamples.

The first set of weights, *the Day 1 weights*, apply to the subsample of all persons who provided recall intake data for the day before the CSFII interview. The second set of weights, *the 3-day weights*, apply only to that subsample of persons who provided food intake information for all three days.

Since all persons in a respondent household were automatically included in the survey, each person "inherits" the HH weight as his/her initial weight. In addition, a weight component is included to adjust each set of weights to "known" totals. This component attempts to simultaneously correct for person nonresponse, sampling error in the person sample, and person coverage errors. Specifically, the weight for each subsample of persons was computed as follows for both the basic and low-income surveys:

$$WT(PER) = WT(HH) \times W(CEN) \quad (2)$$

where

WT(HH) = the final weight of the person's HH

W(CEN) = a weight adjustment to 1989 CPS estimates of persons by cells defined by several age-race-sex categories.

The decision to use the adjustment to Census estimates as the nonresponse adjustment was not an obvious one to the National Analysts' project team. The use of an adjustment that weights up respondents in a household to all household members was considered, since persons in a household eat from basically the same food stock and may share eating habits. However, such an approach would, for example, sometimes substitute women's and children's intake data for men's missing intake records. To avoid this and other substitutions HNIS deemed undesirable, nonresponse weight adjustments were computed across a much larger base (i.e., several age-race-sex cells).

For either the Day 1 or three-day subsamples, the weight adjustment, W(CEN), was computed in each cell by dividing the CPS estimate of the number of persons in the cell by the sum of the initial (household) weights of

the respondent persons in the cell. For the basic sample the CPS person totals used to compute either the Day 1 or the three-day weights were the total person estimates for the cell. For the low-income sample, the appropriate CPS estimates for computing weights for either subsample were the number of persons residing in households at or below 130% poverty threshold.

For either the basic or low-income sample, the subsample of persons receiving the 3-day weights is smaller than the Day 1 subsample. Consequently, the weight adjustments to Census Bureau estimates, $W(\text{CEN})$, are larger for the three-day subsample. However, for each cell, the sum of the person weights are the same, subject to rounding error, for the two subsamples (i.e., the Census estimate of total persons or of low-income persons).

In defining the age-race-sex cells to use for computing this adjustment, two sex categories and two race categories (i.e., black and non-black) were used. As requested by HNIS, there were several categories used for age and these varied by black and non-black, by male and female, and by the basic and low-income surveys. The full set of age categories specified by HNIS consisted of the following 14 classes:

Under 1 year	15 to 19 years	50 to 59 years
1 to 2 years	20 to 24 years	60 to 69 years
3 to 5 years	25 to 29 years	70 to 79 years
6 to 11 years	30 to 39 years	80+ years
12 to 14 years	40 to 49 years	

Although the full set of 56 cells (i.e., $2 \times 2 \times 14$) was not used for creating the $W(\text{CEN})$ adjustment for any of the subsamples, only a minimal amount of collapsing was needed for the non-black cells for either the basic or low-income samples. The specific weight adjustment factors and collapsing patterns that were used for all four subsamples are given in Table II-4. Cells that were combined can be identified by the numbers in parentheses.

4. Weights for the DHKS were based on a household universe, rather than a person universe

For the DHKS, an attempt was made, after completion of the regular CSFII survey, to conduct an interview with the meal planner/preparer in each household. Since, by definition, there is one meal planner/preparer per household, the meal planners/preparers inherit the HH weight as their initial weight. Weight adjusts to universe estimates of the meal planner/preparers cannot be made since good estimates of this universe are not available.

Table II-4

**PERSON WEIGHT ADJUSTMENTS FACTORS TO CENSUS (1989 CPS)
TOTALS, W(CEN)* -- BASIC SAMPLE**

One Day Subsample

<u>Age Category</u>	<u>Black</u>		<u>Non-Black</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Under 1	1.03 (1)	1.03 (1)	1.48 (10)	1.48 (10)
1 to 2	1.03 (1)	1.03 (1)	0.89	0.78
3 to 5	1.34 (2)	1.34 (2)	1.04	1.05
6 to 11	1.62 (3)	1.75 (6)	0.92	0.92
12 to 14	1.62 (3)	1.75 (6)	1.08	1.23
15 to 19	1.62 (3)	1.75 (6)	1.18	1.07
20 to 24	2.46 (4)	1.42 (7)	1.58	1.17
25 to 29	2.46 (4)	1.42 (7)	1.16	0.99
30 to 39	2.46 (4)	1.48	1.34	1.09
40 to 49	2.46 (4)	1.36 (8)	1.14	1.09
50 to 59	2.46 (4)	1.36 (8)	1.19	0.95
60 to 69	1.22 (5)	1.71 (9)	1.01	0.89
70 to 79	1.22 (5)	1.71 (9)	1.03	1.01
80 and over	1.22 (5)	1.71 (9)	1.28	1.19

Three Day Subsample

<u>Age Category</u>	<u>Black</u>		<u>Non-Black</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Under 1	2.34 (11)	2.34 (11)	1.79 (19)	1.79 (19)
1 to 2	2.34 (11)	2.34 (11)	1.01	0.97
3 to 5	2.34 (11)	2.34 (11)	1.21	1.54
6 to 11	1.95 (12)	2.46 (15)	1.13	1.05
12 to 14	1.95 (12)	2.46 (15)	1.21	1.40
15 to 19	1.95 (12)	2.46 (15)	1.41	1.20
20 to 24	2.63 (13)	2.35 (16)	1.89	1.50
25 to 29	2.63 (13)	2.35 (16)	1.41	1.14
30 to 39	2.63 (13)	1.60	1.58	1.38
40 to 49	2.63 (13)	2.11 (17)	1.38	1.30
50 to 59	2.63 (13)	2.11 (17)	1.41	1.23
60 to 69	1.69 (14)	1.85 (18)	1.24	1.04
70 to 79	1.69 (14)	1.85 (18)	1.24	1.17
80 and over	1.69 (14)	1.85 (18)	1.60	1.56

*Numbers in parentheses indicate cells that have been combined.

Table II-4
(Continued)

**PERSON WEIGHT ADJUSTMENTS FACTORS TO CENSUS (1989 CPS)
TOTALS, W(CEN)* -- LOW-INCOME SAMPLE**

One Day Subsample

<u>Age Category</u>	<u>Black</u>		<u>Non-Black</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Under 1	1.02 (1)	1.94 (6)	1.24 (10)	1.24 (10)
1 to 2	1.02 (1)	1.94 (6)	1.15	0.99
3 to 5	1.15	1.23	1.11	1.48
6 to 11	1.75	0.82	1.33	1.45
12 to 14	1.02 (2)	0.96 (7)	0.96 (11)	1.09
15 to 19	1.02 (2)	0.96 (7)	0.96 (11)	1.07
20 to 24	1.01 (3)	1.17 (8)	0.77	0.87
25 to 29	1.01 (3)	1.17 (8)	1.07	1.18
30 to 39	1.07 (4)	0.84	1.12	0.91
40 to 49	1.07 (4)	1.30	1.22	1.14
50 to 59	1.05 (5)	1.02	1.05	1.07
60 to 69	1.05 (5)	1.19 (9)	1.17	0.98
70 to 79	1.05 (5)	1.19 (9)	0.98 (12)	1.20
80 and over	1.05 (5)	1.19 (9)	0.98 (12)	1.09

Three Day Subsample

<u>Age Category</u>	<u>Black</u>		<u>Non-Black</u>	
	<u>Male</u>	<u>Female</u>	<u>Male</u>	<u>Female</u>
Under 1	1.07 (13)	2.02 (18)	1.54 (22)	1.54 (22)
1 to 2	1.07 (13)	2.02 (18)	1.31	1.22
3 to 5	1.56	1.77	1.28	1.76
6 to 11	2.58	0.89	1.54	1.74
12 to 14	1.27 (14)	1.19 (19)	1.16 (23)	1.16
15 to 19	1.27 (14)	1.19 (19)	1.16 (23)	1.18
20 to 24	1.27 (15)	1.44 (20)	0.91	1.19
25 to 29	1.27 (15)	1.44 (20)	1.20	1.50
30 to 39	1.10	0.99	1.34	1.03
40 to 49	1.10	1.92	1.70	1.67
50 to 59	1.17 (17)	1.15	1.05	1.24
60 to 69	1.17 (17)	1.44 (21)	1.42	1.21
70 to 79	1.17 (17)	1.44 (21)	1.34 (24)	1.45
80 and over	1.17 (17)	1.44 (21)	1.34 (24)	1.27

*Numbers in parentheses indicate cells that have been combined.

As primarily a nonresponse weight adjustment, the project team did consider adjusting the initial weights to a *projected* meal planner/preparer universe based on screening responses. However, the team felt that the definition of a household's meal planner/preparer was not sharp enough, and someone other than the designated meal planner/preparer completed the DHKS interview for some households. Therefore, it was agreed to adjust the initial weights to a household universe, rather than to a projected person universe.

Therefore, the weight for each DHKS respondent was calculated as the respondent's household weight times a household-based nonresponse adjustment. Specifically, the weight for respondent meal planners/preparers, for both the basic and low-income surveys, was computed as follows:

$$WT(DHKSs) = WT(HH) \times W(CEN) \quad (3)$$

where

WT(HH) = the DHKSs respondent's final HH weight

W (CEN) = an adjustment to 1989 CPS household estimates by 27 cells defined by 9 Census divisions and 3 levels of urbanization

The 27 HH control totals used in the calculation of the W(CEN) factor were the same ones used for computing the household weights: total households for basic DHKS respondents and low-income households for low-income DHKS respondents. For each cell W(CEN) was calculated as the appropriate control divided by the sum of the initial (household) weights of all DHKS respondents in the cell.

For the basic DHKS sample, no collapsing of the 27 cells was needed to keep the W(CEN) factors below 2.5. For the low-income DHKS sample, no collapsing was needed beyond that which was done in calculating the low-income HH weights.

Two additional sets of DHKS weights were computed in response to analytical needs of the HNIS: one for the basic sample and one for the low-income sample. In both cases the second set of DHKS weights was calculated for the subset of DHKS respondents that also provided a full three days of intakes.

The methods used to compute weights for these two subsamples was the same as those used to compute weights for the full DHKS respondent samples from equation (3). For each survey (basic or low-income) the same CPS cell control counts were used for computing $W(\text{CEN})$ for the three-day/DHKS subsample as for the full DHKS subsample. Therefore, the sum of the respondent weights for the full DHKS subsample and the three-day/DHKS subsample will be the same.

Because the three-day/DHKS sample sizes were smaller than those for the full DHKS subsamples, the $W(\text{CEN})$ adjustment factors were larger for the three-day/DHKS subsamples. However, only two two-cell combinations had to be made to keep $W(\text{CEN})$ from exceeding 2.5: one for the basic sample and one for the low-income sample. (Actually, the combination that was made for the basic sample only brought the adjustment factor down to 2.65, but no other collapsing was judged to be advisable.)

Because the total number of weights calculated and the weight component information which was provided to HNIS exceeded data field sizes originally anticipated and approved in the final format files, a totally separate weight tape with documentation was provided to HNIS.

III. SURVEY PREPARATION

Preparation for the survey effort for the first year of the Continuing Survey of Food Intakes by Individuals and for all subsequent years began in 1988. An initial meeting was held between HNIS and National Analysts to discuss survey activities. The survey preparation activities involved document preparation and testing, a dress rehearsal, and the selection and training of survey field and coding personnel. Particular attention in the development phase was devoted to the preparation and assessment of the newly-incorporated data instrument, the Diet and Health Knowledge Survey (DHKS) questionnaire. These survey preparation efforts are the subject of this chapter.

A. Initial Preparation of Survey Instruments and Supporting Materials Began Prior to the Kick-off Meeting

On October 19 and 20, 1988, a kick-off meeting between key members of the National Analysts team and the HNIS project personnel was held. Activities central to the development and refinement of survey operations were discussed in this meeting. Attending that meeting from National Analysts were the Officer-In-Charge, the Project Director, Assistant Project Director, Data Collection Manager and Sampling Statistician.

Prior to the initial meeting between the key members of the research teams, HNIS supplied annotated draft copies of the two sets of major survey documents -- the CSFII and DHKS questionnaires with suggestions for revision. National Analysts reviewed the documents and prepared fresh instruments, incorporating the proposed changes. These questionnaires were presented and discussed in the two-day kick-off meeting. Minor modifications to the CSFII questionnaires and supporting documents continued up to the onset of the dress rehearsal.

The DHKS instrument underwent the most changes. This new addition to the Continuing Survey format was to be a 25-minute telephone-administered interview to be completed by every participating household several weeks after the capture of the CSFII interview. The DHKS questionnaire, designed to explore attitudes and knowledge of diet, nutrition, health, food safety and related issues, was extensively reworked following the kick-off meeting. Initial pretesting began with hot-house trials of the document among National Analysts' employees during the first half of November 1988. A workable but lengthy instrument was sent to HNIS by the middle of that month. Subsequently, three sets of revisions were proposed by the USDA team to hone the topics of interest and to shorten the instrument. A final version was created in late December for use in the dress rehearsal.

B. A Full-Scale Dress Rehearsal Was Conducted Prior to the Launch of the National Survey

On January 4, 1989 the full-scale dress rehearsal was begun which involved the training of interviewers, the completion of the CSFII and DHKS interviews, the reviewing and interactive coding of documents and the testing of the data processing system. The goal of this effort was threefold:

- To test all CSFII procedures and protocols as they were to be employed in the nationwide survey, ranging from interviewer recruitment to hand off of processed data in final format
- To pretest, refine and properly size the DHKS instrument
- To create and debug the new linkage/control system between the traditional CSFII activities and the DHKS survey efforts

The dress rehearsal was completed in the January-March, 1989 period and provided valuable feedback for the national endeavor. Key findings pointed to the need for greater emphasis, during interviewer training, on the recording of intake information and to the value of lengthening the interval between the CSFII and the DHKS interviews in order to achieve better DHKS participation rates.

1. The dress rehearsal effort identified areas to be addressed in the national survey

The dress rehearsal was conducted in three different venues -- a major urban, a suburban and a nonmetropolitan community. In each location, eight basic and eight low-income-qualifying households were targeted for interview. In-person interviewers were trained in the use of the national CSFII protocol and, then, conducted interviews with the main meal planners/preparers in the survey households. Sets of three-day intake records were sought from all household members.

Once the documents were received in our office, they were reviewed, edited and coded. The data were subjected to computer editing and standard CSFII data manipulation and expansion. A final dress rehearsal data file was prepared.

Follow-up DHKS interviews were attempted both in person and by telephone with cooperating households. In addition, random respondents were interviewed by telephone to further test and refine this survey instrument. Summary survey administration statistics from the DHKS effort were provided to HNIS.

2. Dress rehearsal interviewer training for the CSFII portion of the survey began in January, 1989

Five in-person interviewers were trained in Philadelphia on January 4 to 6, 1989. Two interviewers were from Philadelphia, representing the major urban center, two from Union Co., New Jersey, the suburban location and one from Anderson Co., South Carolina, the rural survey point. One interviewer was a bilingual, bicultural Spanish speaker.

Training was conducted by two experienced National Analysts survey personnel following the Agenda shown in Appendix B. In addition, an Interviewer's Manual was prepared and used to guide interviewers throughout their data collection activities (see Appendix C for cover and table of contents).

As a result of the dress rehearsal, both trainers and interviewers suggested more time be given in future training efforts to the recording of actual food and beverage intake information and the answering of related questions in the intake records. It was further suggested that more emphasis be placed upon the Food Instruction Booklet for determining the reporting of foods and quantities. Interviewer feedback at the debriefing session following the dress rehearsal indicated that all felt prepared for their tasks as a result of training, and that the problems they encountered in the field were readily resolved through telephone conferences with their field supervisor.

3. Forty-eight dress rehearsal CSFII interviews were completed early in 1989

The CSFII interviewing portion of the dress rehearsal was carried out between January 10 and February 14, 1989. A total of 48 households were interviewed, 16 in each location, half with basic sample households and half with low-income sample households.

The initial CSFII interviews averaged 100 minutes for completion of both the household and the DAY 1 intake records. The household portion ran about 25 minutes, while collection of the intake portion of the data averaged 75 minutes.

Follow-up visits to these households were made to retrieve the DAY 2 and 3 diaries. Interviewers collected 70% of the leave-behind records. Reasons most often given in the dress rehearsal for not completing the diaries were outright refusals to cooperate and reconsideration of the agreement to participate once the full burden of the effort became apparent.

4. All dress rehearsal CSFII documents were coded using existing and new systems

Completed sets of dress rehearsal CSFII questionnaires were coded using manual and interactive protocols. A coding manual was written and coders trained in its application for the manual section of the coding effort. The existing interactive food coding system was slightly modified to accommodate CSFII changes and experienced coding personnel were briefed on the changes. One three-day intake set was coded from each household.

Coders judged the coding manual to be thorough and complete. Based on coding of the 48 questionnaire sets, no new precodes were identified, although changes were made to the identification of data fields for the national survey.

5. Dress rehearsal CSFII data were manipulated successfully with little revision to in-place programs

Existing computerized data editing programs were updated to accommodate the revised Household Questionnaire and Individual Intake Record additions. All coded documents were processed, cleaned and expanded using the newly-revised nutrient data file provided by HNIS. The final format for the dress rehearsal data file was prepared and sent to HNIS.

6. DHKS interviewer training was completed for both telephone and in-person interviewers

In-person interviewers were trained in the administration of the DHKS questionnaire during their three-day briefing session. Telephone interviewers were given an initial four-hour face-to-face training by the National Analysts field manager. The survey document needed to be changed several times during the dress rehearsal (see below). With each change in the survey document, refresher briefings were conducted in person with the centralized telephone interviewing staff. A team of three English-speaking telephone interviewers were trained and worked on this phase of the study.

7. Four waves of data collection were required to refine and pretest the DHKS questionnaire

The first wave of DHKS telephone interviewing began approximately three weeks after the first CSFII interview was taken and continued through February 23, 1989. All households for which no telephone information was available (N = 10) were attempted in person. In the remaining households, contact was tried either by telephone or in person based on a randomized procedure. Approximately half of all the 48 households were to be

attempted by telephone and half in person. In addition, to bolster sample random digit dialings (RDD), interviews were attempted with meal planners/preparers across the country.

Interviewers debriefed after this phase of the dress rehearsal suggested that respondents appeared tired of the survey and worried that the follow up would require as much work by them as had the initial recordkeeping for CSFII. Moreover, interviewers felt that they had not adequately prepared respondents to expect follow-up contacts. Given that the DHKS contact was so close in time to the last interview, and that it was unexpected, there was a higher than anticipated refusal rate (only 22 of the 48 CSFII households provided DHKS interviews). It was recommended that the follow-up DHKS interview begin no earlier than six weeks after the CSFII.

In addition, the dress rehearsal revealed that this version of the DHKS interview was a lengthy effort; it ran longer than the planned 25 minutes with sample households. Reinterviews with CSFII households averaged 47 minutes, while those contacted at random by telephone ran 34 minutes.

Coupled with the longer administration time, several problems were noted in the questionnaire -- some items confused respondents, some produced many "other" responses, while other items required minor wording changes to ease administration.

Concurrent with the first pretest, outside reviewers from the Public Health Service, Agriculture Research Service, the Economic Research Service and the Extension Service of USDA studied the draft DHKS questionnaire and offered additional comments. Therefore, the questionnaire was revised to incorporate these inputs and the changes resulting from the initial stage of the dress rehearsal.

On February 24 and 25, 1989 a total of ten DHKS interviews were completed by telephone with a set of randomly selected meal planners/preparers using the revised questionnaire. The administration time for this questionnaire averaged 36 minutes and several questions required additional revision to elicit unambiguous responses. Hence, the document was revised again and a third wave of interviews completed.

On March 7, telephone interviewers completed eight more interviews the newly-revised survey instrument. This version ran 29 minutes, on average, and required only minor adjustments to ease the flow of questions and to create a more "respondent-friendly" questionnaire.

The fourth and final test of the DHKS was conducted on March 15 and 16, 1989. It involved a split-half technique where two versions of the questionnaire were used, each with a limited number of questions unique to that

version. Twelve interviews were taken, all by telephone, at an average length of 29 minutes, with a range of 20 to 37 minutes. Respondents appeared to understand the questions and fewer open-ended recordings were required by the interviewers.

Based on the results of this pretest, the DHKS questionnaire was finalized and prepared for administration in the national survey. Both a hard-copy version for administration by in-field interviewers and a CATI version were prepared.

8. Processing the dress rehearsal DHKS questionnaire proved the document was field ready

Because of the relatively small number of questionnaires and the multiple waves of document testing, only the first draft of the DHKS questionnaire was coded, cleaned and electronically processed. The remaining waves of pretest interviews were hand tabulated for key survey administration outcome variables (e.g., length of interview). The data processing effort proved to be straightforward.

9. Linkages between the two survey components -- CSFII and DHKS -- were created and tested in the dress rehearsal

The final goal of the dress rehearsal was to establish the means by which the two survey components -- the CSFII interviewing and the DHKS interviewing -- could be linked together. Two concerns were to be addressed. The first was that the system should prompt the follow-up effort in such a manner that the DHKS contact was timely. The second aspect of the system was the linkage between the two data sets. The system had to ensure that the same household was being recontacted, that the correct respondent in that household was being reinterviewed, and that the two data sets could be analyzed as a unit in the future.

To create these linkages, National Analysts developed subroutines in our existing field check-in programs. At the time a completed CSFII interview was logged into the system, it was assigned a unique identification number. This number linked the two parts of the survey and was used in a subroutine which printed out informational labels for the DHKS questionnaire. These labels not only identified the contact information for the household (e.g., name, address, telephone number, if available), but provided the earliest date for contact and a two-week window for initiating DHKS interviewing.

C. A Total of Eight Final Survey Documents and Related Materials Were Used in the National Survey

Following the dress rehearsal, a final set of survey instruments and supporting materials were crafted and given OMB approval. These documents include:

- **Household Questionnaire** -- This brief questionnaire establishes information about the household and its members. The questions are directed to the main meal planner/preparer who is typically the female head of household and addresses several areas of concern, specifically: typical shopping patterns and food expenditures; household size and composition; educational attainment and employment status of all members over the age of 14; pregnancy and lactation status of age-appropriate female household members; incidence of being breast-fed among children; participation in food benefits programs including WIC, school lunch/breakfast, food stamp and food distribution programs; amount and sources of income and assets; household food sufficiency; sources of household drinking water.
- **Individual Intake Record** -- This, the core of the CSFII interview, is divided into two parts, both of which cover the details of all foods and beverages consumed during specified time periods. In the first part, the interviewer administers a DAY 1 intake record for the 24-hour calendar day immediately prior to the CSFII interview, using the standard food measurement aids (i.e., measuring cups, spoons, ruler and instruction book). This information is collected with the meal planner/preparer for her/himself and for all children under the age of 12 years. In addition, other permanent members of the household (other than roomers, boarders and employees) present at the time may be asked to complete the DAY 1 record as well. If they are not present, the meal planner/preparer is asked to report for them or a record is to be left for completion at a later time. In the second part of the effort, all eligible household respondents complete intake information for the day of the interview and the following days (DAY 2 and 3), first under the direction of the interviewer and then on their own. The type of information in the intake records is:

DAY 1

- Detailed description of the type and quantities of food ingested at and away from home
- Sources of foods/beverages consumed
- Use of fats and salt in food preparation (meal planner/preparer only)

- Quantity of water consumed in 24-hour period
- Typicality of this day's intake
- Healthfulness of diet
- Typical salt usage
- Food supplements, including vitamins, minerals, fiber and fish oils
- Height and weight
- Incidence of physical disorders and diseases
- Activity levels (for those over 17 years of age)
- Television viewing habits
- Smoking patterns (for those over 17 years of age)
- Handling of purchased precooked food dishes

DAYS 2/3

- Same first four data elements as in DAY 1
 - Food frequency for 13 foods
- **Food Instruction Booklet** – This multi-page booklet is a newly-revised version of the food reporting aid used by both interviewers and respondents in conjunction with the intake records. This guidebook is organized by 12 food groupings and identifies for each the nature and level of detail of the information sought for every food line item recorded in the intake record. Specific probe questions are given for the description of foods and the units of measurement to be reported.
 - **Screening/Call Report Form/Nonresponse Questionnaire** – This document is a short questionnaire used to guide the interviewer to the correct respondent in the household (e.g., the meal planner/preparer) and to identify survey-qualifying households in the low-income sample. In addition, minimal household composition information is gathered for sample dwellings in which contact is made and additional descriptive information is recorded for households not cooperating with the survey effort.

In the low-income portion of the survey, all households were screened for eligibility based on household size and income. The monthly income cutoff levels shown below are 130% of the poverty income guidelines published in the Federal Register of February 16, 1989:

<u>Number of People</u>	<u>Monthly Income Cutoff Levels</u>
1	\$648
2	\$869
3	\$1,090
4	\$1,311
5	\$1,532
6	\$1,753
7	\$1,974
8	\$2,195
9	\$2,416
10	\$2,637
11	\$2,858
12	\$3,079
13	\$3,300
14	\$3,521
15	\$3,742
16	\$3,963
17	\$4,184
18	\$4,405
19	\$4,626
20	\$4,847

These cutoffs were applied throughout the entire twelve months of data collection for Year 1. Household size was reported in terms of number of persons who regularly reside in the household and income was for the calendar month prior to interview.

This document also serves as a permanent record of calls to the household. The Call Report Form portion of the document identifies all attempts to reach the household and to attempt an interview. Both successful and unsuccessful contacts are captured. In addition, the form reports the number of intake records sought, the number retrieved and the reason for not securing a record, if known.

- **Respondent Letter** -- This introductory letter was either mailed or hand-delivered by the interviewer to the sample households at the time of initial contact. It was designed to invite participation in the survey and to explain the purpose and conditions under which the survey was being taken. The letter, signed by the HNIS Director of this project, explains that the survey participation is voluntary and that participants will be modestly compensated for their cooperation.
- **Facts Sheet** -- This new two-page information sheet explains the importance of participation in the CSFII survey and describes the auspices of the research, some of the uses of the information from the survey and gives the household a point of contact at HNIS for receiving further information.

- **The Diet and Health Knowledge Survey (DHKS)** -- This questionnaire, designed to capture current issues in food knowledge and attitudes, is a 13-page document appearing in two versions, each with a unique question. In Year 1 the topics covered are:
 - Knowledge of dietary guidelines and personal diet performance
 - Attitudes toward dietary guidance
 - Attitudes toward dietary guidance
 - Knowledge of diet and health problems
 - General nutrition knowledge
 - Food preparation practices
 - Food shopping attitudes and practices
 - Food labeling usage
 - Food handling and storage practices
 - Food safety knowledge

Both versions of the questionnaire explore the consumers' knowledge of meat labeling. In half the questionnaires, the DHKS respondent is asked the difference between meats labeled "natural" and meats not so labeled. In the other version of the questionnaire, the difference between meats labeled "organic" and meats not so labeled, is explored.

- **Nonresponse Follow-up Survey Questionnaire** -- This brief (9-page) questionnaire gathers demographic and food benefits program participation information from noninterviewed households or their proxies. To be used twice a year with a subsample of nonparticipating households, the data gathered are designed to provide information about reasons for nonresponse and about characteristics of nonrespondents.

D. Newly Revised Interviewer Training Materials Were Prepared for the Year 1 Survey

All training materials were either newly developed or revised for the CSFII effort. For interviewers in the field, several sets of manuals were prepared. They include:

- **Preconference Manual** -- This is a 22-page training manual sent to all interviewers for review prior to attending the training conferences. It presents an overview of the survey effort and National Analysts; general interviewing

guidelines; a review of questionnaire completion and recording conventions; and an introduction to food reporting and measurement activities. A copy of the manual is found in Appendix D.

- **CSFII Interviewer Instruction Manual** – This lengthy document was given to interviewers at their training sessions and is designed to answer all questions regarding fieldwork. In addition to giving an overview of interviewers' activities, this manual addresses the topics of identifying sample households; conducting screenings, household and intake interviews; handling the leave-behind intake questionnaires; field editing procedures and other related field topics. Particular emphasis is placed in the manual on the completion of the 24-hour intake recording. A copy of the table of contents for this manual is given in Appendix E.
- **DHKS Interviewer Instruction Manual** -- This 11-page booklet covers the completion of the DHKS questionnaire by both in-person and telephone interviewers. It provides direction on when to take the interview; selection of the appropriate respondent; and completion of the questionnaire. This manual is shown in Appendix F.
- **Nonresponse Manual** -- This 12-page manual identifies the field activities involved in recontacting nonresponsive households and in completing the brief follow-up interview with either a member of the targeted household or a neighborhood proxy respondent (see Appendix G).

E. In-Person Interviewers Selected for this Project Were Prepared for Their Work in Three-Day Group Training Conferences. While the Telephone Interviewers Were Given Intense Half-Day Briefings

CSFII Interviewing Training

1. Eight formal interviewer training conferences were held to prepare both experienced and new in-person data collectors for their Year 1 assignments.

The cadre of National Analysts' interviewers, with food consumption survey experience, were reviewed and those whose prior work was deemed acceptable were invited to join the CSFII effort. In addition, general notices were sent to interviewers in many of our sampling points, alerting them to the assignment and inviting their expression of interest. From this group of interviewer candidates, approximately 120 were chosen for initial training.

Prior to attending three-day briefing conferences, all potential interviewers received the Preconference Manual for home study. In addition, a copy of the new DHKS questionnaire was sent with instructions that it should be administered to an acquaintance and that the completed document be brought to the training sessions.

As proposed in the kick-off meeting on October 19 and 20, 1988, interviewer training sessions were staggered to facilitate recruitment and to smooth the flow of schedules. Therefore, seven training sessions were conducted between the end of March and the first weeks of May. The dates and location of these sessions were:

<u>Location</u>	<u>Dates</u>	<u>Number of Interviewers Trained</u>
Philadelphia	March 29 -31	23
Chicago	April 10 -12	25
Los Angeles	April 3 - 5	17
	May 8 - 10	12
Nashville	April 3 -5	27
	May 3 - 5	13
Oklahoma City	April 10 - 12	16

Following these initial training sessions, we planned for one additional briefing conference to be held later in the year to bolster ranks and refresh interviewers as needed. We held this session on August 14 to 16, 1989 in Philadelphia in which six interviewers were trained.

In addition to these formal conferences, four training sessions with small groups of interviewers (two or three at a time) were conducted in Albany, Miami, Philadelphia, and Tampa.* These briefings were held to replace field personnel as they dropped from the study later in the year. In total, 149 interviewers were trained and conducted CSFII interviews. Several additional field personnel worked on the assignment in ancillary roles as escorts, listers, nonresponse follow-up interviewers, and the like. Eight Spanish bilingual, bicultural interviewers were trained for assignments in areas with concentrations of Hispanic households.

*In March 1990, a training conference was held in Atlanta to prepare new interviewers for Year 2 work. Some of these newly-trained interviewers helped finish Year 1 fieldwork, especially the in-person DHKS interviews.

2. The three-day training sessions provided comprehensive coverage of the CSFII Year 1 procedures and materials for in-person interviewing

The formal training sessions were attended by an average of 18 interviewers and were conducted by two-person teams of experienced National Analysts trainers. Most conferences were attended by one or more representatives from USDA as well as National Analysts coding personnel, all of whom took part in some portion of the training activities. Every session leader had experience with food consumption survey research and with interviewer training. A total of five different trainers conducted these sessions.

The conference format was built off of the knowledge gained from past food consumption trainings and the results of the dress rehearsal. The presentations combined role-playing, classroom lectures, written exercises and one-on-one coaching sessions. The first day's session began with a general introduction to food consumption research, the role of USDA in monitoring the nation's nutritional status and the history of the CSFII. Next, an overview of Year 1 data collection activities and the role of the interviewer were explored. Following this discussion, the Screening Form questionnaire was introduced and, in a round-robin format, interviewers practiced reading and completing four different iterations of this document. Presentation of the Household Questionnaire followed with discussion and practice in its completion.

The afternoon of the first day was turned over to study and drill on the intake portion of the survey. First, the necessary materials (e.g., DAY 1 Individual Intake Record, Food Instruction Booklet, measuring utensils) were introduced. Then an overview of the format and questions in the intake records was discussed. The trainers next spent time discussing food terminology, use of the Food Instruction Booklet, food measurement techniques and when to use which measurement convention (e.g., weight, volume, sizes and physical dimensions). In addition, practice was given in special recording activities, such as sandwiches, salads and mixtures. The session closed with the completion of a mock DAY 1 Individual Intake Record and the assignment of homework for the completion of a second intake with an acquaintance.

The second day of training began with a detailed review of intake recording and a presentation of DAY 2 and 3 records, including their pick-up and field review for completeness. An additional mock intake interview was completed using the DAY 2 and 3 instrument.

As a change of pace for the afternoon session, several paid respondents were brought into the conference setting and interviewers broke into small groups to conduct screening, household and intake interviews with each "live"

respondent. The second day's session ended with the introduction of the DHKS. All interviewers participated in the reading of questions and recording of answers to a mock DHKS interview. The evening's homework was the completion of another DHKS and DAY 2 intake record with a friend or family member.

One-on-one review and coaching of interviewers took up much of the morning of the third day. During this period, the trainers and assisting personnel gave individual attention to each interviewer, first in review of their intake homework and then in personally-tailored drills on areas in which interviewers were weakest. Concurrent with these coaching sessions, interviewers completed self-paced knowledge tests covering all CSFII interviewing materials. As their final supervised practice, interviewers listened to an audiotaped interview -- from the approach to the household and screening interview to the completion of the DAY 2 and 3 diaries -- and filled out the appropriate survey documents. Trainers observed all recordings, paying close attention to any person whose work may have been questionable.

The final session was devoted to discussion of sampling materials and the selection and identification of sample housing units. Issues of initial door approaches, handling nonresponse, callbacks and refusal conversion techniques were explored. The conference ended with a final review of procedures, materials, timetables and general field issues.

3. Most trained in-person interviewers took assignments and worked the entire year long survey

National Analysts' original plans called for training approximately 150 interviewers for the in-person CSFII survey. One hundred forty nine interviewers were trained to perform this work and all but four were given assignments to conduct CSFII screening and interviewing activities in Year 1. One hundred thirty nine of the field personnel were prepared in one of the eight formal training conferences, ten were trained in small group sessions.

The trained interviewers who worked on the assignment averaged six completed CSFII interviews and 12 attempted/nonresponse screenings. The average number of completed screenings and interviews is in keeping with the size of the field assignments for this first year of CSFII work. Most work throughout the entire 12-month period (see Table III-1). All but one interviewer worked both the basic and low-income samples, although some took very few low-income interviews due to the nature of the sample areas assigned to them (e.g., non-low-income areas). Of the 149 interviewers who performed work in Year 1, all but 29 continued with the study, taking assignments as available to them and were retained for work in Year 2.

Table III-1**LENGTH OF INTERVIEWER SERVICE IN YEAR 1 CSFII**

	<u>%</u>
One Wave	10.1
Two Waves	18.8*
Three Waves	14.8
Four Waves	56.4
(N) =	(149)

*Includes eight interviewers for whom work was available in their areas in two waves only.

DHKS Interviewer Training

4. Centralized telephone interviewers were trained in a four-hour briefing session

To prepare centralized telephone interviewers for their DHKS assignment, a four-hour briefing session was held on site at the WATS facility. Interviewers were presented with the DHKS Manual to study and to use as the guide for their survey activities. During the session, National Analysts' trainers reviewed the goals of this phase of the research and used hard-copy versions of the questionnaires to study the survey instrument on a question-by-question basis. Practice mock interviews were completed by recording answers on hard-copy questionnaires. In addition, responses to different types of interviewing situations were discussed (e.g., identifying the eligible respondent, avoiding interview break-offs) prior to interviewers going on the telephone lines to conduct "live" interviews.

Following the classroom study session, interviewers contacted sample households and attempted actual interviews using the CATI questionnaire. Each was silently monitored during the interview and feedback given on their performance by the trainers. An additional drill was instituted as needed.

In total, six WATS interviewers, including two Spanish speakers, were fully trained for the DHKS assignment, all of whom worked on a portion of the Year 1 DHKS interviewing effort.

F. Review and Interactive Coding Personnel Were Fully Briefed on the Year 1 Requirements by Experienced Coding Supervisors

All review personnel were prepared afresh for their CSFII assignment. Two seasoned coding supervisors (each with ten years experience) conducted two-day briefing sessions in early May and again later in the year to train a total of 13 field review staff members. The central focus of the session was communicating an understanding of what was required to code the documents, particularly the intake records. As it is the reviewers' function to "repair" any questionnaires that are incomplete, inconsistent or unscorable prior to the coding effort, it was necessary that they know the requirements for coding acceptability. Therefore, reviewers were given a detailed briefing on the interactive coding process, an overview of the reference materials used to support these coding activities and an introduction to post-coding cleaning efforts.

The detailed reviewer training activities covered the specific topics of:

- Document check-in and determination of completeness

- Listing and addressing questionnaire problems in the Review Summary Sheet for the Screening Form and Household Questionnaire
- Completing the Review Summary Sheet for the Intake Record
- Callback and recontact procedures for retrieving information from interviewers or respondents
- Validation activities
- Editing and coding of the screening forms and household questionnaires

A **CSFII 1989 Review Procedures Manual** was prepared to serve as a reference for reviewers (see Appendix H). Refresher trainings were conducted periodically throughout the year by coding personnel to keep the review efforts sharp.

At the time of the onset of Year 1 of CSFII, the twelve experienced intake coders (and three supervisors) who were working for National Analysts were judged to be a sufficient cadre to complete the interactive coding tasks. Four additional coders subsequently were added during the year to replace and reinforce the original work group.

Training of the experienced coders was completed in a one-day session in late May. The coding supervisor introduced all the CSFII documents and identified the differences between these and prior USDA survey questionnaires. The new interactive coding program changes were reviewed and new HNIS coding guidelines were presented.*

Training of the new coding staff combined classroom learning sessions with one-on-one coaching which extended over a two-week period. New personnel were given a general introduction to all the questionnaire documents, HNIS guidelines and coding reference materials. They were shown how to operate the computer and how to prepare food code and quantity request materials. Practice sessions were held on inputting intake questionnaire data.

Individual new trainees were assigned to an expert coder, who was an experienced trainer, for a two-week intensive coaching session. First the trainee observed the expert coder input intake records and listened to explanation of the decision-making process that went into the choice of food and quantity codings. Then the trainee inputted actual data under the direct supervision of the expert coder to whom he or she was attached. The trainee was initially coached in every decision until he or she established a basis for independent decision-making. Thereafter the trainee consulted the expert coder only on food line items about which he or she had questions.

*The HNIS guidelines required continued refinement throughout the first six months of the survey period and the revisions and updates were reviewed with the coding personnel on an ongoing basis.

For at least three months after the coaching period all the trainee's work was double coded by an expert coder and feedback on problems communicated on a regular basis. Double coding continued thereafter for all coders periodically to ensure consistency and accuracy.

IV. DATA COLLECTION

The subject of this chapter is the activities and outcomes associated with field data collection efforts. First, an overview of the tasks associated with all phases of data gathering is given. This is followed by a discussion of key outcomes of the field effort.

A. There Were Six Major Steps in the Fieldwork Activity for the First Year of the Continuing Survey

After the sample of housing units to be contacted had been established, interviewers in the field had six major tasks to accomplish in order to complete the CSFII fieldwork. These tasks were:

- Screen households to determine eligibility
- Conduct interviews, capturing DAY 1 intake records
- Instruct household members in completion of DAY 2 and 3 intake records
- Collect DAY 2 and 3 leave-behind intake records
- Complete DHKS interviews with appropriate members of previously interviewed CSFII households
- Conduct nonresponse follow-up interviews with a subsample of surveyed household respondents or proxy informants

Each of these activities is described below.

1. In-person screenings were conducted to determine household eligibility and to identify the appropriate household informant

Interviewers visited every sample address in person to visually inspect and to determine whether that location represented a residential housing unit. All locations that met this criterion were approached and invited either to participate, if the unit were part of the basic sample, or to complete a screening eligibility determination, if it were part of the low-income sample.

Interviewers were instructed to attempt screening interviews, if possible, with the household member who was responsible for planning and preparing the family meals. If the meal planner/preparer were not available, any knowledgeable adult member of the household who was 18 years old or older could serve as the screening respondent.

In the basic sample where all households qualified for interview, the screening contact was used to solicit survey cooperation. Once the potential respondent was identified, he or she was given information about the requirements of the survey, both orally and in writing, and asked to begin the interview immediately. The invitation letter from the HNIS Project Director and the Facts Sheet information were used as supporting documentation of the survey's authenticity and to motivate cooperation. Interviewers were also permitted to discuss the measuring utensils and the \$2.00 payment for completed DAY 2/3 records at this time if they thought it would help provoke cooperation.

Screening interviews with households in the low-income sample were taken to establish survey eligibility. Several questions regarding household size and monthly income from various sources were asked in the screening to determine if the household were at or below 130% of the poverty income guidelines. Those households meeting the income and size criterion were invited to participate in the survey.*

In situations where there were multiple meal planners/preparers in a household because the members kept separate food supplies – such as students sharing apartments – separate interviews were taken with each individual eating unit. Subsequently, data from the separate eating units were combined into a single interview to represent a Census-defined household.**

If a CSFII interview could not be scheduled conveniently to follow immediately after the screening interview, then an appointment was made for a later time. Up to six in-person attempts (or three in-person and eight telephone attempts, if a telephone number was known) were required to secure the screening information.

2. All CSFII interviews were completed in person across the country throughout the entire twelve-month period of the Year 1 survey

Consistent with prior USDA-sponsored food consumption surveys, an interview was taken with the primary meal planner/preparer for the sample

*If households did not provide interviewers with the necessary information to make eligibility determinations, follow-up efforts were made from the Philadelphia office in the form of telephone and mail contacts to gather the answers. The ineligibility determination was established for a portion of the households in the low-income sample through these types of contacts.

**By reference, the contract called for separate interviews for individual eating units.

household. The meal planner/preparer served as the informant for information about the household, as well as reporting on his or her own personal intake. The meal planner/preparer was the one person in the household judged to be most knowledgeable about the widest range of topics in the survey, although input from other household members was encouraged if the informant could not address any of the questions fully.

The CSFII interview was administered in person and began with the Household Questionnaire. This part of the interview established essential information about the composition and characteristics of the household and its food-related behaviors. This section was followed immediately by administration of the DAY 1 Individual Intake Record to the meal planner/preparer. The Individual Intake Record was completed with the aid of the standard measuring utensils, the Food Instruction Booklet and available home guides, such as coffee mugs, juice glasses and the like.

The meal planner/preparer reported first on his/her own consumption of foods and beverages for the 24-hour day prior to the interview and then that of children in the home under the age of 12. The meal planner/preparer was also permitted to report for any eligible household member about whose DAY 1 intake he or she was knowledgeable. Other eligible household members who were present at the interview were invited to report their own intake directly to the interviewer. Household members who were not available and for whom the meal planner/preparer could not provide accurate information were left intake forms to be completed later.

To be eligible for the intake portion of the survey, the household member was to be considered a regular member of the family but *not* a roomer, boarder or employee of the household. In addition, family members who were not expected to be in the home for the three days of potential reporting (e.g., traveling on business, in the hospital) were exempted from the intake reporting task.

The administration of the Household Questionnaire portion of the interview averaged 23 minutes, with the shortest interview running less than 10 minutes and the longest taking over 2 hours (see Table IV-1). There remarkable consistency between the two samples in the length of interview; this is true even though the low-income sample households are larger and have more household members about which to report information.

Completion of the Individual Intake Record averaged just over 27 minutes for both the basic and low-income sample households (see Table IV-2). In general, the first intake recording in a household took the longest time and following recordings were shorter. This represented a combination of respondent learning of the task requirements and interviewers becoming more efficient, the latter being exemplified by such notation on the subsequent records as "same recipe as on record #1".

Table IV-1

**LENGTH OF CSFII HOUSEHOLD INTERVIEW
BY SAMPLE TYPE**

<u>Number of Minutes</u>	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
10 minutes or less	10.5	10.3	10.4
11 to 15 minutes	23.1	20.1	22.1
16 to 20 minutes	24.1	25.2	24.5
21 to 30 minutes	27.3	30.3	28.3
31 to 40 minutes	8.1	8.8	8.4
41 minutes or more	6.9	5.3	6.4
Mean	23.4 minutes	23.2 minutes	23.3 minutes
	(N) = (1,490)	(725)	(2,215)

Table IV-2

**LENGTH OF DAY ONE INTAKE RECORDING*
BY SAMPLE TYPE**

	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
<u>Number of Minutes</u>			
15 minutes or less	6.5	10.5	7.8
15 to 19 minutes	15.7	20.6	17.3
20 to 24 minutes	23.8	20.5	22.8
25 to 34 minutes	31.9	26.3	30.1
35 to 59 minutes	18.3	18.4	18.3
60 minutes or more	3.7	3.7	3.7
Mean	27.6 minutes	26.6 minutes	27.6 minutes
	(N) = (3,260)	(1,567)	(4,827)

*Excludes known cases in which the interviewer reported only the total time for the intake task; that is, recorded the starting time as the time the first intake was begun and the ending time as the time the final intake was completed

The number of in-person attempts to complete interviews with eligible households was set at five.

3. Interviewers instructed household members who were willing to cooperate in the completion of DAY 2 and 3 Individual Intake Records

While recording individual food consumption for the day before the interview, interviewers were training respondents in the proper entry of food line item information, that is, the names, descriptions and quantity measurement of foods consumed. Interviewers then began reporting intake for the day of the interview (DAY 2) and left booklets for all eligible household members to finish reporting their own intake information for DAY 2 and for DAY 3 -- the calendar day following the interview. Meal planners/preparers were to continue reporting for themselves and for children under the age of 12. They were also encouraged to assist other members of the household as necessary. The measuring cups, spoons, ruler and Food Instruction Booklet were left in the household for use in recording quantities of foods and beverages consumed.

An appointment time was established for the interviewer's return to pick up the leave-behind records.

4. Interviewers returned to collect the DAYS 2 and 3 records and to pay households for their efforts

Interviewers returned to cooperating households after the two-day recording period to retrieve as many of the completed intake forms as were available. If the household members were present and willing, interviewers reviewed the documents with them and edited them on the spot. On some occasions, interviewers assisted in the completion of records at the time of pick-up. Most frequently, however, interviewers were simply able to retrieve the records, sometimes without household members present, and completed the editing at a later time.

Each household was compensated \$2.00 per intake record retrieved, up to a total of \$20.00 per household.

5. A second interview was attempted with every household that completed a CSFII survey

In all sample households where food consumption surveys were completed, DHKS interviews were attempted. Follow-up contact was first attempted by

telephone, if a number were available, and then in-person contacts were tried. If there were no telephone number in the record for the CSFII household, then only personal interviewing was attempted. Recontacts for the DHKS interview were scheduled for not earlier than six weeks after the date of the CSFII interview. Up to four calls were attempted before determining that the household was not available.

The targeted interview respondent was the same person who served as the household informant for the CSFII interview, that is, the primary meal planner/preparer. If that respondent was no longer a member of the household or was to be unavailable for an extended period of time, then the interviewer identified which household member was serving as the main meal planner/preparer and sought to interview that person, regardless of whether or not that person had been a CSFII participant.

Table IV-3 shows that the average DHKS interview ran just over the 25 minutes as planned, with one quarter being completed in less than 20 minutes and one in five running more than 30 minutes.

6. Twice a year small samples of nonrespondents to the CSFII interview portion of the survey were recontacted and requested to complete short household interviews

The final field activity -- the nonresponse follow-up survey -- took place twice in Year 1, following the completion of Wave 2 and the completion of Wave 4 of CSFII interviewing. A sample of occupied eligible* households, where CSFII interviews had not been completed, were identified for these brief follow-up interviews.

In half of the households, any knowledgeable adult member was asked to provide information. Up to four in-person visits to the household were made to complete the interview (the *household-only condition*). In the other half of the households, if information could not be gathered directly from the family after three visits, the interviewer attempted to complete a very short, scaled-down version of the interview with anyone in the neighborhood who could supply basic information about the sample residents (the *proxy respondent condition*).

*In the low-income sample, households where the eligibility determination had not been established were included in the pool of households which could be selected for follow-up. Information about their eligibility status was important to gather in its own right. Moreover, reasons for their nonparticipation, even at the screening stage, could prove to be useful data in future efforts.

Table IV-3**LENGTH OF DHKS INTERVIEW BY SAMPLE TYPE**

<u>Number of Minutes</u>	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
15 minutes or less	2.5	2.9	2.6
16 to 20 minutes	25.1	17.5	22.6
21 to 25 minutes	34.1	28.1	32.1
26 to 30 minutes	20.7	25.5	22.3
31 minutes or more	17.6	26.0	20.4
Mean	25.8 minutes	28.0 minutes	26.6 minutes
	(N) = (1,263)	(623)	(1,887)

B. A Total of 2215 Households Were Contacted and Interviewed from April 1989 to March 1990 in the CSFII Portion of the Year 1 Survey

1. Fourteen hundred ninety basic and seven hundred twenty five low-income sample households completed the CSFII interview

Tables IV-4 and 5 show the distribution of interviews and result of attempted contacts by wave, first for the basic and then for the low-income sample. A total of 1490 usable CSFII basic household interviews were completed (this number is reduced from the 1504 separate interviews which were completed with independent eating units and which were later consolidated into 1490 unique interviews). The two major reasons for not completing an interview at an assigned address were outright refusal by the household to cooperate with the survey (n = 436) and housing unit vacancy (n = 360). Refusal to complete the screening interview and no contact after repeated calls also constituted significant reasons for failure to interview. Language barriers and limited access to the entire area each represented just over 1% of the total field outcomes in the basic sample.

In the low-income sample a total of 725 interviews were completed (741 separate interviews were taken with individuals functioning as separate eating units and were later combined into the 725 CSFII interviews). As expected, the single most frequent reason for not taking an interview was ineligibility of the household. More than 62.3% (n = 3476) of the households identified for contact did not meet the size-and-income criterion for interview. The second major reason for a noninterview outcome was that the housing unit was unoccupied; 653 of the 5582 units were vacant. Two hundred ninety four households refused to complete the screening interview, while 154 of those who were eligible to participate in the survey chose not to do so. Very few interviews in the low-income sample were not completed because of language problems (n = 29) or access limitations (n = 37). Across the two samples less than 1% of the housing units were not interviewed for each of these reasons.

In planning for this effort, we anticipated successfully completing the screening phase with between 87% and 90% of the occupied, eligible housing units across both the basic and low-income sample. In fact, we completed screening at 86% of the occupied sample housing units. The overall completion rate, however, was higher in the low-income (72.9%) than in the basic sample (62.7%), with completion being defined as at least one individual in the household supplying at least one usable day of intake.

Table IV-4

**TOTAL COUNTS FOR CSFII BASIC SAMPLE
BY RESULT OF CALL AND WAVE**

<u>Result of Call*</u>	<u>Wave 1</u> #	<u>Wave 2</u> #	<u>Wave 3</u> #	<u>Wave 4</u> #	<u>Total Basic</u> #
Participated (Code 1)	359	377	385	369	1,490
Refused interview (Code 5)	100	105	113	118	436
Refused screening (Code 8)	47	45	42	50	184
No answer (Code 10)	34	38	41	50	163
Language barrier (Code 11)	17	3	11	7	38
Vacant (Code 12)	89	83	98	90	360
No access (Code 14)	2	16	5	13	36
Other (Codes 13 and 15)	4	11	9	3	27
(N) =	(652)	(678)	(704)	(700)	(2,734)

*Codes 2, 3, 6, 7 and 9 are not final outcome codes

Table IV-5

**TOTAL COUNTS FOR CSFII LOW INCOME SAMPLE
BY RESULT OF CALL AND WAVE**

<u>Result of Call*</u>	<u>Wave 1</u> #	<u>Wave 2</u> #	<u>Wave 3</u> #	<u>Wave 4</u> #	<u>Total</u> <u>Low Income</u> #
Participated (Code 1)	237	198	124	166	725
Ineligible (Code 4)	1,077	981	592	826	3,476
Refused interview (Code 5)	53	38	27	36	154
Refused screening (Code 8)	89	95	49	61	294
No answer (Code 10)	33	56	38	64	191
Language barrier (Code 11)	10	4	4	11	29
Vacant (Code 12)	228	151	127	147	653
No access (Code 14)	8	10	7	12	37
Other (Codes 13 and 16)	7	6	4	6	23
(N) =	(1,742)	(1,539)	(972)	(1,329)	(5,582)

*Codes 2, 3, 6, 7 and 9 are not final outcome codes

2. Interviews were distributed fairly evenly across the months of the year and by day of the week in the Year 1 survey

Interviews were slated to be taken throughout the twelve-month field period with one quarter of the targeted number of interviews to be completed each three-month period. Table IV-6 shows that this distribution of interviews was largely achieved. Not more than 25.8% nor less than 23.3% of the interviews were completed in any one of the targeted three-month intervals. The largest number of interviews completed in a single month were taken in May and October, while December accounted for the fewest number of total interviews. Table IV-7 shows the distribution of interviews by month based on the wave to which the household was assigned.

Because patterns of food consumption have been noted to vary by day of the week, especially weekday vs. weekend, efforts were made to collect interviews across the days of the week. These results are shown in Table IV-8. As has been found in the past with surveys of this type, most days of the week were equally well represented with the exception of Sundays. Just over 8% of all interviews were completed on Sundays. The largest number of interviews were completed on Saturdays (21.6%).

3. Most interviews were successfully taken on the first or second visit

Interviewers proved to be most effective on their first or second visit as shown in Table IV-9. Over 60% of the completed interviews and over 70% of the ineligibility determinations were made within two visits. To achieve final resolution at noninterviewed households took more calls; only 43% of these outcomes were determined in one or two visits; 10% of these households were contacted on nine or more occasions. The average number of contacts at interviewed households was 2.6 and went up to an average of 3.8 at noninterviewed households. These data demonstrate that at many of the households, repeated calls were attempted to get resolution.

4. Over ninety percent of the household members provided a DAY 1 Individual Intake Record

In Tables IV-10 and 13 are found the results of the intake completion efforts. Included in Tables IV-10 and IV-11 are counts of individuals who attempted to provide some of the information called for in the intake record, regardless of whether that information ultimately was judged complete and sufficient for comprehensive intake coding. For example, the respondent may have given information about his or her breakfast consumption but nothing more or may have provided useful food frequency information only. In Tables IV-12 and IV-13 are counts of individuals supplying one or more complete days of intake data.

Table IV -6

**DISTRIBUTION OF CSFII INTERVIEWS
BY MONTH AND BY SAMPLE TYPE**

<u>Month Interview Completed</u>	<u>Basic Sample %</u>	<u>Low- Income Sample %</u>	<u>Total Sample %</u>
April 1989	5.8	7.2	6.3
May	10.3	14.1	11.6
June	6.7	10.5	7.9
Spring	22.8	31.8	25.8
July	8.8	11.9	9.8
August	8.8	8.0	8.6
September	7.1	5.9	6.7
Summer	24.7	25.8	25.1
October	12.8	9.7	11.7
November	9.5	7.2	8.7
December	3.4	1.9	2.9
Fall	25.7	18.8	23.3
January 1990	9.5	10.5	9.8
February	10.3	6.8	9.2
March*	6.9	6.4	6.8
Winter	26.7	23.7	25.8

(N) = (1,490) (725) (2,215)

*Includes six interviews taken in April 1990

Table IV-7

**DISTRIBUTION OF CSFII INTERVIEWS
BY MONTH BY WAVE AND SAMPLE TYPE**

<u>Month Interview Completed</u>	<u>Wave 1</u>		<u>Wave 2</u>		<u>Wave 3</u>		<u>Wave 4</u>	
	<u>Basic</u> %	<u>Low Income</u> %	<u>Basic</u> %	<u>Low Income</u> %	<u>Basic</u> %	<u>Low Income</u> %	<u>Basic</u> %	<u>Low Income</u> %
April 1989	24	22						
May	43	43						
June	27	32	*					
July	4	3	31					
August	1		34					
September			27		1	2		
October			7		42	42		
November					37	41		
December					13	11		
January 1990					6	4	31	43
February	*1		*1		*1		41	29
March							27	26
April							*	2
(N) =	(359)	(237)	(377)	(198)	(385)	(124)	(369)	(166)

*.05% or less

¹Based on a misunderstanding of her responsibilities, one interviewer completed four CSFII interviews while attempting the nonresponse follow-up survey

Table IV-8

**DAY OF THE WEEK OF CSFII INTERVIEW
BY SAMPLE TYPE**

<u>Day</u>	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
Sunday	7.1	10.3	8.2
Monday	16.1	16.9	16.3
Tuesday	15.4	13.6	14.9
Wednesday	15.0	15.0	15.0
Thursday	12.1	11.8	12.0
Friday	11.7	12.5	12.0
Saturday	22.6	19.9	21.6
(N) =	(1,490)	(725)	(2,215)

Table IV-9

**NUMBER OF ATTEMPTED CONTACTS
TO COMPLETE A CSFII INTERVIEW
BY OUTCOME AND SAMPLE TYPE**

<u>Interview Completed</u>	<u>Basic Sample</u> %	<u>Low- Income Sample</u> %	<u>Total Sample</u> %
One attempt	36.7	42.5	38.6
Two attempts	21.9	25.4	23.0
Three attempts	15.9	14.3	15.3
Four attempts	10.1	7.0	9.1
Five to eight attempts	13.0	9.1	11.7
Nine or more	2.5	1.6	2.2
Mean	2.7 attempts	2.2 attempts	2.6 attempts
(N) =	(1,490)	(725)	(2,215)
<u>Interview Not Taken</u>			
One attempt	22.8	31.9	26.9
Two attempts	15.5	16.8	16.1
Three attempts	13.1	11.1	12.2
Four attempts	10.5	9.8	10.2
Five to eight attempts	26.2	22.2	24.4
Nine or more	11.8	8.2	10.2
Mean	3.9 attempts	3.6 attempts	3.8 attempts
(N) =	(884)	(728)	(1,612)
<u>Ineligible (Low-Income Sample Only)</u>			
One attempt		52.4	
Two attempts		21.3	
Three attempts		11.8	
Four attempts		6.4	
Five to eight attempts		6.6	
Nine or more		1.5	
Mean		2.1 attempts	
(N) =		(3,476)	

Table IV-10

**DISTRIBUTION OF INTAKE RECORDS BY DAY
BY WAVE FOR BASIC SAMPLE -- PARTIALS AND COMPLETES**

Number of Individuals Providing Data* for:	BASIC				
	<u>Wave 1</u> #	<u>Wave 2</u> #	<u>Wave 3</u> #	<u>Wave 4</u> #	<u>Total</u> #
Day 1 intake	897	893	901	853	3,544
Day 2 intake	770	755	783	715	3,023
Day 3 intake	751	741	760	696	2,948
All three days of intake	748	734	754	694	2,930
Day 1 and 2 intakes only	19	15	23	20	77
Day 1 and 3 intakes only	0	1	1	0	2
Day 2 and 3 intakes only	3	6	5	1	15
Day 1 intake only	130	143	123	139	535
Day 2 intake only	0	0	1	0	1
Day 3 intake only	0	0	0	1	1

*Includes partials

Table IV-11

**DISTRIBUTION OF INTAKE RECORDS BY DAY
BY WAVE FOR LOW-INCOME SAMPLE -- PARTIALS AND COMPLETES**

Number of Individuals Providing Data* for:	LOW-INCOME				
	<u>Wave 1</u> #	<u>Wave 2</u> #	<u>Wave 3</u> #	<u>Wave 4</u> #	<u>Total</u> #
Day 1 intake	574	466	287	352	1,679
Day 2 intake	476	411	254	296	1,437
Day 3 intake	465	398	249	283	1,395
All three days of intake	461	397	248	281	1,387
Day 1 and 2 intakes only	12	13	5	13	43
Day 1 and 3 intakes only	1	0	0	0	1
Day 2 and 3 intakes only	3	1	1	2	7
Day 1 intake only	100	56	34	58	248
Day 2 intake only	0	0	0	0	0
Day 3 intake only	0	0	0	0	0

*Includes partial

Table IV-12

**DISTRIBUTION OF INTAKE RECORDS BY DAY
BY WAVE FOR BASIC SAMPLE -- COMPLETES ONLY**

<u>Number of Individuals Providing Data for:</u>	BASIC				
	<u>Wave 1</u> #	<u>Wave 2</u> #	<u>Wave 3</u> #	<u>Wave 4</u> #	<u>Total</u> #
Day 1 intake	887	883	891	841	3,502
Day 2 intake	757	742	773	702	2,974
Day 3 intake	749	738	748	690	2,925
Day 1 and 2 intakes only	8	8	24	14	54
Day 1 and 3 intakes only	0	2	1	1	4
Day 2 and 3 intakes only	9	8	7	6	30
Day 1 intake only	139	147	126	145	557
Day 2 intake only	0	0	2	1	3
Day 3 intake only	0	2	0	2	4
Completed three days of intake	740	726	740	681	2,887
Completed two days of intake only	17	18	32	21	88
Completed one day of intake only	139	149	128	148	564

Table IV-13

**DISTRIBUTION OF INTAKE RECORDS BY DAY
BY WAVE FOR LOW-INCOME SAMPLE -- COMPLETES ONLY**

<u>Number of Individuals Providing Data for:</u>	<u>LOW-INCOME</u>				
	<u>Wave 1</u> #	<u>Wave 2</u> #	<u>Wave 3</u> #	<u>Wave 4</u> #	<u>Total</u> #
Day 1 intake	559	464	276	349	1,648
Day 2 intake	460	404	246	288	1,398
Day 3 intake	453	395	244	280	1,372
Day 1 and 2 intakes only	11	9	4	8	32
Day 1 and 3 intakes only	3	0	1	0	4
Day 2 and 3 intakes only	7	1	8	3	19
Day 1 intake only	103	61	37	64	265
Day 2 intake only	0	0	0	0	0
Day 3 intake only	1	0	1	0	2
Completed three days of intake	442	394	234	277	1,347
Completed five days of intake only	21	10	13	11	55
Completed one day of intake only	104	61	38	64	267

Very few intakes were lost because they were only partially completed. Less than 2% of the DAY 1 recordings in both the basic and low-income samples were partially completed, while less than 3% of the three-day sets of intake were judged to be partials.

As with previous intake surveys, the largest number of Individual Intake Records were completed for the day before the interview. As the recording burden increased, the number of cooperating record providers declined. As shown in Table IV-10, 3544 basic Individual Intake Records (including partials) were retrieved for DAY 1, which represents a record for 90.4% of all individuals residing in the sample households.* Over three-quarters of these individuals provided something in a DAY 2/3 record, while total of 2930, or 74.8%, of the household members provided some information for all three days of intake reporting.

In the low-income sample, 1679 individuals gave some DAY 1 input which is equal to 91.8% of the total household members in the low-income sample.** Nearly 80% of the individuals gave some information in a DAY 2/3 record, while three days of recordings were received from 1387, or 75.8%, of the members of low-income sample households.

5. Outright refusals to cooperate were the primary reason for not providing records

Table IV-14 shows the reasons for not retrieving DAY 1 and DAY 2/3 booklets in the field (regardless of whether or not the data provided by the respondents were fully complete and scorable). Refusal to engage in the recording task was the primary reason for not retrieving a record; one in twenty gave this response for the DAY 1 record, while one in six refused the DAY 2/3 task. Just under 3% of the eligible household members were not available for the recording period (and, hence, not expected to produce intake records).

The overall participation rates at the household and individual levels are shown in Tables IV-15 and IV-16 for the basic and low-income samples.

*Not all household members were required to provide records. Roomers, boarders, employees who regularly lived in the household (n = 6) and all other members away for the reporting period (n = 102) were exempt from the intake recording activity. If these household members are excluded from the base, this rate is increased to 93%.

**If roomers, boarders, employees and travelers are dropped from the base, this rate is increased to 94.4%.

Table IV-14

**REASONS FOR NOT OBTAINING INTAKE RECORDS
IN THE FIELD -- PARTIALS AND COMPLETES**

<u>Result of Call</u>	<u>Basic Sample</u>		<u>Low Income Sample</u>	
	<u>Day 1</u> %	<u>Day 2/3</u> %	<u>Day 1</u> %	<u>Day 2/3</u> %
Record obtained	90.4	77.2	91.8	78.6
Participation refused	6.2	18.0	5.0	16.8
Person away from home	2.3	2.6	2.1	2.7
Record not obtained no reason given	0.7	1.1	0.6	1.0
Person too sick	0.2	0.7	0.2	0.6
Intake voided/Other	0.2	0.4	0.3	0.3
	(N) = (3,918)	(3,918)	(1,829)	(1,829)

Table IV-15

**PARTICIPATION AT THE HOUSEHOLD AND INDIVIDUAL LEVEL
BY WAVE FOR BASIC SAMPLE**

	<u>Wave 1</u>	<u>Wave 2</u>	<u>Wave 3</u>	<u>Wave 4</u>	<u>Total</u>
	#	#	#	#	#
Sample housing units	652	678	704	700	2,734
Occupied housing units	563	595	606	610	2,374
Contacted housing units	523	530	551	544	2,148
Screened housing units	459	482	498	487	1,926
Eligible housing units	359	377	385	369	1,490
Participating housing units	359	377	385	369	1,490
Individuals in participating housing units	969	1,005	983	961	3,918
Individuals completing DAY 1	887	883	891	841	3,502
Individuals completing DAYS 1, 2 and 3	740	726	740	681	2,887
Reasons for occupied housing units not contacted:					
No one home; no answer	34	38	41	50	163
No access	2	16	5	13	36
Other	4	11	9	3	27
Reasons for contacted housing units not screened:					
Refused screening	47	45	42	50	184
Language barrier	17	3	11	7	38
Reasons for screened housing units not participating:					
Refused interview	100	105	113	118	436

Table IV-16

**PARTICIPATION AT THE HOUSEHOLD AND INDIVIDUAL LEVEL
BY WAVE FOR LOW-INCOME SAMPLE**

	<u>Wave 1</u>	<u>Wave 2</u>	<u>Wave 3</u>	<u>Wave 4</u>	<u>Total</u>
	<u>#</u>	<u>#</u>	<u>#</u>	<u>#</u>	<u>#</u>
Sample housing units	1,742	1,539	972	1,329	5,582
Occupied housing units	1,514	1,388	845	1,182	4,929
Contacted housing units	1,466	1,316	796	1,100	4,678
Screened housing units	1,367	1,217	743	1,028	4,355
Eligible housing units	290	236	151	202	879
Participating housing units	237	198	124	166	725
Individuals in participating housing units	616	503	324	386	1,829
Individuals completing DAY 1	559	464	276	349	1,648
Individuals completing DAYS 1, 2 and 3	442	394	234	277	1,347
Reasons for occupied housing units not contacted:					
No one home; no answer	33	56	38	64	191
No access	8	10	7	12	37
Other	7	6	4	6	23
Reasons for contacted housing units not screened:					
Refused screening	89	95	49	61	294
Language barrier	10	4	4	11	29
Reasons for screened housing units not participating:					
Refused interview	53	38	27	36	154

C. Cooperation with the DHKS Portion of the Survey among CSFII Interviewed Households Was High and Was the Result of Significant Follow up Efforts

1. More than nineteen hundred DHKS interviews were completed

We attempted to complete DHKS interviews with every household that participated in the CSFII intake portion of the survey and which still resided in the neighborhood where they were first contacted.* The maximum number of DHKS interviews that could have been captured, therefore, was 2215 – 1490 basic and 725 low income.

As Table IV-17 shows, we garnered 1906 DHKS interviews, or 86% of the original sample. If the 73 out-of-scope households (i.e., vacants, moved -- no forwarding information and deceased) are removed from the base, then the completion rate is 88.8%. Both of these rates are better than the 85% expected prior to the survey's onset.

The major reason for not completing a DHKS interview was failure to reach the respondent at home (n = 94) after repeated call attempts on different days of the week and at different times of the day. Only 87, or 3.9%, refused to cooperate, while only 52, or 2.3%, could not be surveyed for some other reason.

2. Most DHKS interviews were taken within ten weeks of the completion of CSFII interviews

Over half of the DHKS interviews were completed within seven weeks after the CSFII interviews and three quarters were accomplished in twelve weeks (see Table IV-18). The average time span between the two was ten weeks. Because so many of the telephone attempts were followed up with in-person contacts, the length of time between the two interviews was extended considerably in some cases.

3. Unexpectedly, less than sixty percent of the DHKS interviews could be completed by telephone during the yearlong survey

The distribution of interviews by wave is shown in Table IV-19. Approximately the same number of interviews were taken each wave, with the fewest number being completed in Wave 3 (n = 434) and the largest in Wave 2 (n = 499). As a percent of the total possible DHKS interviews for a wave, proportionally, the fewest DHKS interviews were completed in Wave 1 (82.6%) and the largest in Wave 4 (90.0%).

*CSFII interviewed households that moved out of their communities or to any new address without forwarding information were deemed out of scope for the follow-on DHKS interview.

Table IV-17**FINAL COUNTS OF DHKS BY RESULT OF CALL**

<u>Result of Call</u>	<u>Basic Sample #</u>	<u>Low-Income Sample #</u>	<u>Total Sample #</u>
Interviewed	1,280	626	1,906
Interview refused	61	26	87
Telephone out of order	1	1	2
No one home/Respondent not available	73	21	94
Language barrier	0	1	1
Vacant/moved/deceased	43	30	73
Other	32	20	52
(N) =	(1,490)	(725)	(2,215)

Table IV-18

**LENGTH OF TIME BETWEEN CSFII AND DHKS INTERVIEWS
BY SAMPLE TYPE**

<u>Number of Days</u>	<u>Basic Sample %</u>	<u>Low Income Sample %</u>	<u>Total Sample %</u>
42 days or less	12.2	11.5	12.0
43 to 49 days	24.2	19.3	22.6
50 to 63 days	22.0	23.2	22.3
64 to 84 days	15.7	16.9	16.1
85 to 105 days	10.6	11.0	10.8
106 days or more	15.3	18.1	16.2
Mean	69.5 days	73.6 days	70.8 days
(N) =	(1,276)	(626)	(1,902)

Table IV-19

**DISTRIBUTION OF DHKS INTERVIEWS
BY WAVE AND BY SAMPLE TYPE**

	<u>Basic Sample</u>	<u>Low-Income Sample</u>	<u>Total Sample</u>
Wave 1	81.3%*	84.4%	82.6%
	N = 292	N = 200	N = 492
Wave 2	86.7%	86.9%	86.8%
	N = 327	N = 172	N = 499
Wave 3	86.0%	83.1%	85.3%
	N = 331	N = 103	N = 434
Wave 4	89.4%	91.0%	90.0%
	N = 330	N = 151	N = 481
	(N) = (1,280)	(626)	(1,906)

*Percent of completed CSFII interviews

The breakdown of DHKS interviews by mode -- telephone vs. in person -- is given in Table IV-20. The original intent of the survey design was that most interviews would be taken by telephone with in-person interviewing used as a back up only for households without phones and with unlisted numbers. It was expected, based on prior research, that 10% of the basic and 35% of the low-income sample households would be approached in person. However, in each of the four waves, between 40% and 50% of the completed DHKS interviews were captured in person.

This unusually large number of nontelephone interviews is due to the extensive follow-up efforts exerted to retrieve information from households initially assigned to the telephone mode. One in five of all DHKS interviews were the result of personal follow up with households which could not be reached or would not cooperate with the telephone effort.

Table IV-21 gives the reasons why the attempted contacts did not yield completed interviews. In the basic sample, lack of telephone service and respondents' refusals to be interviewed by phone were major, and both equally important, reasons for not obtaining interviews. In the low-income sample, the telephone numbers given for over 40% of the households were not working at the time of DHKS contact. In addition, one in five households refused to be interviewed on the telephone. In spite of having bilingual English- and Spanish-speaking interviewers, language problems posed another impediment to telephone interviewing (13%) in both samples. Finally, in spite of attempting telephone calls in the evenings, on weekends and during the weekday (up to six attempts at each number) 1 in 14 respondents could not be reached for interview within the time frame allowed for the initial DHKS contact. Therefore, these telephone nonresponse households were sent to the field for in-person follow-ups.*

4. Significant follow up efforts were made to gain the number of DHKS interviews achieved

Many attempts were made by telephone and, often, in person to capture the DHKS interview (see Table IV-22). The average number of calls to successfully interviewed households was 3.4, while noninterviewed households were pursued for an average of 5.1 calls.

5. The household meal planner/preparer served as the DHKS respondent in nearly all the cases

More than nine out of ten DHKS interviews were successfully completed with the original CSFII respondent (see Table IV-23). Ninety four percent of

*If a telephone number were given by respondents, the first attempt to complete a DHKS interview was by telephone. Households were not given the choice of whether to be contacted by phone or in person

Table IV-20

DHKS INTERVIEW MODE BY WAVE AND BY SAMPLE TYPE

	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
<u>Wave 1</u>			
Telephone	67.5	49.5	60.1
In person	32.5	50.5	39.8
• Nontelephone household	16.4	24.5	19.7
• Nonresponsive to telephone contact	16.1	26.0	20.1
(N) =	(292)	(200)	(492)
<u>Wave 2</u>			
Telephone	57.8	50.0	55.1
In person	44.2	50.0	44.9
• Nontelephone household	21.1	25.6	22.6
• Nonresponsive to telephone contact	21.1	24.4	22.2
(N) =	(327)	(172)	(499)
<u>Wave 3</u>			
Telephone	53.5	44.7	51.4
In person	46.5	55.3	48.6
• Nontelephone household	30.5	37.9	32.2
• Nonresponsive to telephone contact	16.0	17.5	16.4
(N) =	(331)	(103)	(434)
<u>Wave 4</u>			
Telephone	56.4	49.7	54.3
In person	43.6	50.3	45.7
• Nontelephone household	18.2	31.1	22.2
• Nonresponsive to telephone contact	25.4	19.2	23.5
(N) =	(330)	(151)	(481)

Table IV-21

**REASONS FOR NOT COMPLETING A DHKS INTERVIEW
BY TELEPHONE WHEN A TELEPHONE NUMBER WAS AVAILABLE**

	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
Interview refused	30	22	27
Interview aborted	4	5	4
Language barrier	12	14	13
Respondent hearing problems	2	3	2
Not reached within time frame	8	6	7
Telephone out of order/disconnected	34	41	37
Other	11	8	10
	(N) = (284)	(196)	(480)

Table IV-22

**NUMBER OF ATTEMPTED CONTACTS TO COMPLETE
A DHKS INTERVIEW BY OUTCOME AND SAMPLE TYPE**

	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
<u>Interview Completed</u>			
One attempt	28.7	30.8	29.4
Two attempts	20.0	22.0	20.7
Three attempts	14.7	12.8	14.1
Four attempts	10.9	10.9	10.9
Five to eight attempts	19.1	18.0	18.8
Nine or more attempts	6.6	5.4	6.2
Mean	3.4	3.3	3.4
	attempts	attempts	attempts
	(N) = (1,280)	(626)	(1,906)
<u>Interview Not Taken</u>			
One attempt	15.7	18.2	16.5
Two attempts	15.7	10.1	13.9
Three attempts	8.6	16.2	11.0
Four attempts	10.5	16.2	12.3
Five to eight attempts	28.5	25.2	27.6
Nine or more attempts	20.9	14.1	18.8
Mean	5.3	4.8	5.1
	attempts	attempts	attempts
	(N) = (210)	(99)	(309)

Table IV-23**DHKS RESPONDENT BY SAMPLE TYPE**

	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
<u>DHKS respondent</u>			
Meal planner/preparer (CSFII respondent)	93.5	95.5	94.2
Other household member	6.5	4.5	5.8
(N) =	(1,280)	(626)	(1,906)

the DHKS respondents were the main meal planner/preparer in the household. Reasons for not reinterviewing the CSFII respondent include his or her extended absence from the household, deceased, and misidentification of proper respondent.

D. Information Was Captured for More than Half of the Households Pursued in the Nonresponse Follow-up Survey

1. Over five hundred questionnaires were completed across the two phases of the Nonresponse Follow-up Survey

Table IV-24 reports the results of attempted calls to households in the nonresponse follow-up surveys. Nearly 60%, or 517, of all households participated in the two follow-up surveys, with more taking part in the first follow-up (63.2%) than in the second (55.6%). Not surprisingly, the major reason for nonparticipation was refusal to interview (15.9%), with vacancies and no one home being the other primary reasons for not completing an interview.

2. Most interviews were taken by the second or third attempt

If a nonresponse interview was completed it was captured, on average, in the first or second visit (see Table IV-25). Over eighty percent of the interviews were completed by the third call to a household. On average, somewhat more calls were made to households that, ultimately, proved to be nonresponsive (2.4 vs. 2.0, respectively).

3. Neighborhood proxies provided information in fewer than one in five cases where nonresponse interviews were captured

About ten percent more interviews were completed in the proxy-respondent method than in the household-only approach (56.9% vs 43.1%) as seen in Table IV-26. This is due, primarily, to the increased number of interviews taken with nonhousehold informants. A shorter, less complete, questionnaire was completed with a neighborhood proxy for 17% of all the households for which nonresponse information was retrieved.

Table IV-24

**NONRESPONSE FOLLOW UP SURVEY RESULT OF CALL
BY SAMPLE TYPE BY PERIOD**

<u>Result of Call</u>	<u>Basic Sample</u>			<u>Low-Income Sample</u>			<u>Total Sample</u>		
	<u>Waves 1/2</u> %	<u>Waves 3/4</u> %	<u>Total</u> %	<u>Waves 1/2</u> %	<u>Waves 3/4</u> %	<u>Total</u> %	<u>Waves 1/2</u> %	<u>Waves 3/4</u> %	<u>Total</u> %
Interviewed	65.4	59.5	62.5	61.2	51.8	56.6	63.2	55.6	59.4
Interview refused	9.5	20.0	14.7	12.3	21.6	16.9	11.0	20.8	15.9
No one home	9.0	8.1	8.6	8.8	11.3	10.0	8.9	9.7	9.3
Language barrier	0.0	1.4	1.8	1.8	1.8	1.8	0.9	2.3	1.6
Vacant	6.2	8.1	7.1	10.6	11.3	10.9	8.5	9.7	9.1
Other	4.7	1.4	3.1	4.0	0.0	2.0	4.3	7.2	2.5
Lost in mail	5.2	0.0	2.6	1.3	2.2	1.8	3.2	1.2	2.2
(N) =	(211)	(210)	(421)	(227)	(222)	(449)	(438)	(432)	(870)

Table IV-25

**NUMBER OF ATTEMPTED CONTACTS TO COMPLETE
A NONRESPONSE FOLLOW UP INTERVIEW
BY OUTCOME AND SAMPLE TYPE**

<u>Interview Completed</u>	<u>Basic Sample</u> %	<u>Low-Income Sample</u> %	<u>Total Sample</u> %
One attempt	47.5	54.3	50.9
Two attempts	21.7	19.3	20.5
Three attempts	15.6	10.2	13.0
Four attempts	9.1	11.8	10.4
Five or more attempts	6.1	4.3	5.2
Mean	2.1 attempts	1.9 attempts	2.0 attempts
	(N) = (263)	(254)	(517)
 <u>Interview Not Taken</u>			
One attempt	52.5	53.3	53.0
Two attempts	10.1	13.3	11.9
Three attempts	8.2	8.7	8.5
Four attempts	12.7	9.7	11.0
Five or more attempts	16.4	14.9	15.6
Mean	2.4 attempts	2.3 attempts	2.4 attempts
	(N) = (158)	(195)	(353)

Table IV-26

**NONRESPONSE FOLLOW UP SURVEY INTERVIEWS
BY SAMPLE TYPE AND SOURCE OF INTERVIEW**

<u>Method</u>	<u>Basic Sample</u>			<u>Low Income Sample</u>			<u>Total Sample</u>		
	<u>Waves 1/2</u>	<u>Waves 3/4</u>	<u>Total</u>	<u>Waves 1/2</u>	<u>Waves 3/4</u>	<u>Total</u>	<u>Waves 1/2</u>	<u>Waves 3/4</u>	<u>Total</u>
	%	%	%	%	%	%	%	%	%
Household-only method	42.0	44.0	43.0	43.2	43.5	43.3	42.6	43.8	43.1
Proxy respondent method	58.0	56.0	57.0	56.8	56.5	56.7	57.4	56.2	56.9
Household member	40.6	38.4	39.5	41.7	38.3	40.2	41.2	38.3	39.9
Neighborhood proxy	17.4	17.6	17.5	15.1	18.3	16.5	16.2	17.9	17.0
(N) =	(138)	(125)	(263)	(139)	(115)	(254)	(277)	(240)	(517)

V. DATA REDUCTION AND FILE PREPARATION

This chapter covers the creation and handling of the several databases that resulted from the CSFII Year 1 survey. It addresses the processing effort beginning with the return of hard-copy questionnaires from the field or electronic files from the centralized WATS facility through to the documentation of the final format for each file. Four different data sets have been created for the Year 1 survey data -- the household and intake interview file, the CSFII screener/call report file for all sample housing units, the DHKS interview and nonresponse file, and the nonresponse follow-up survey files. The processing and handling of each of these different data sets is discussed, in turn, in this chapter.

A. Preparation of the CSFII Household and Intake Data Sets Builds From Existing Data Handling Protocols Used in Prior USDA Food Consumption Surveys

Many of the existing protocols and procedures used by National Analysts for other USDA surveys provided the foundation for processing the first year of the Continuing Survey data. This continuity allowed many of the previously created data systems, processing procedures and materials to be used with little or no modification and assured compatibility with USDA databases generated in earlier food consumption survey efforts. An overview of the process of preparing the CSFII survey interview data and screener data files is shown in Exhibit V-1 on the next page.

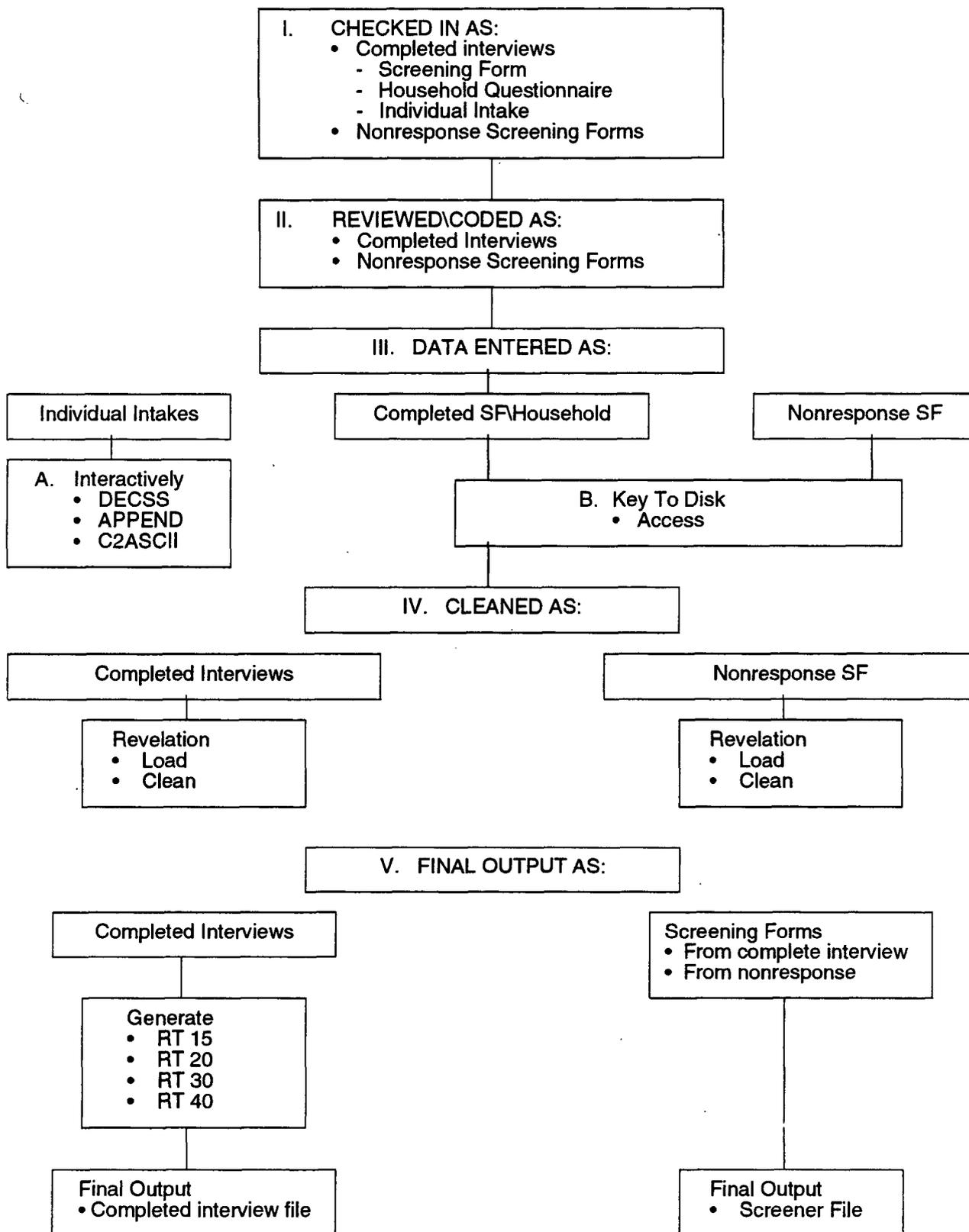
There are five separate steps in the preparation of these data -- logging in of materials, review and editing of data, coding and entry of data input, data cleaning and expansion, and preparation of final data files. Each of these steps is described below, first for the interview database and then for the screener file.

1. The process began with logging in of interviewer-generated materials into the central interview control file

The first step in the post-field processing of survey documents was the checking in of completed materials from the field. A central control file known as the *Field Monitoring File* was created to keep account of each wave of basic and low-income sample information from the field and to flag interviews for validations. Interviewers were instructed to mail completed interviews and any DAY 2/3 intake records captured as soon as they had been retrieved and edited. These completed materials arrived in the Philadelphia office almost daily throughout the year and were logged into the computer on the day of their arrival.

Exhibit V-1

SAMPLED HOUSING UNITS



The first processing task was an integrity check of the interviewer-mailed packages to assure that all the expected documents listed on the transmittal form were present. Specifically, each packet was checked for the presence of a completed Screening Form, Household Questionnaire, and Individual Intake Records. At least one DAY 1 intake record was to be present for the interview to be considered complete. If these materials were not all present, the package was set aside and the interviewer contacted to rectify the situation before the materials were logged into the system.

Completed interviews were checked into the control system by their pre-assigned unique segment and housing unit numbers. As a check on the entry effort, the system was programmed to accept only valid and active segment and housing unit numbers. When these numbers were entered, a five-digit sequence number was automatically assigned by the computer to each completed interview. These unique household identification numbers became the control numbers for the documents for the remainder of the processing effort. The sequence numbers used in CSFII Year 1 are:

**INTERVIEW IDENTIFICATION
NUMBER SEQUENCE**

<u>WAVE</u>	<u>BASIC</u>	<u>LOW INCOME</u>
1	10000-10370	20000-20260
2	11000-11398	21000-21221
3	12000-12400	22000-22132
4	16000-16391	26000-26176

This range is somewhat greater than the total count of interviews recorded in the final data file because some interviews assigned sequence numbers were subsequently combined, voided or invalidated.

The information contained in the check-in system was used to support several post-field activities. Weekly and *ad hoc* monitoring reports were generated from the check-in file (see Exhibits V-2 through V-5). The number and results of call of documents received in house were tracked in the report titled *Household Screeners Received*. Completion rate information was available from the *Household Screener Summary* report. The *Individual Intakes Received* report provided information on the number of intake records received from the field for potential processing. Production in the

HOUSEHOLD SCREENERS RECEIVED

RESULT CODE	BASIC		POVERTY		TOTAL	
	NUMBER	PER CENT	NUMBER	PER CENT	NUMBER	PER CENT
1 - PARTICIPATED	123	59.7	32	7.7	155	25.0
2 - ELIG. & APPT	0	0.0	0	0.0	0	0.0
3 - ELIG. & NO APPT	0	0.0	0	0.0	0	0.0
4 - INCOME INELIG	0	0.0	285	69.0	285	46.0
5 - ELIG. & REFUSED	17	8.3	8	1.9	25	4.0
6 - TELEPHONE BUSY	0	0.0	0	0.0	0	0.0
7 - TELEPHONE BROKE	0	0.0	0	0.0	0	0.0
8 - REFUSED SCREEN	20	9.7	12	2.9	32	5.2
9 - SCR APPT. MADE	0	0.0	0	0.0	0	0.0
10 - NO ANSWER	7	3.4	7	1.7	14	2.3
11 - LANG. BARRIER	2	1.0	2	.5	4	.6
12 - VACANT/ NOT HU	34	16.5	51	12.3	85	13.7
13 - OTHER	3	1.5	15	3.6	18	2.9
14 - NO ACCESS TO HU	0	0.0	1	.2	1	.2
15 -	0	0.0	0	0.0	0	0.0
16 -	0	0.0	0	0.0	0	0.0
17 -	0	0.0	0	0.0	0	0.0
TOTAL RECEIVED	206		413		619	
TOTAL ASSIGNED	687		968		1,655	
PCT. RECEIVED	30.0		42.7		37.4	

Exhibit 2

HOUSEHOLD SCREENER SUMMARY

CATEGORY	BASIC		POVERTY		TOTAL	
	NUMBER	PER CENT	NUMBER	PER CENT	NUMBER	PER CENT
SCREENERS RECEIVED	206		413		619	
TOTAL VACANTS	34	16.5	51	12.3	85	13.7
TOTAL OCCUPIED	172	83.5	362	87.7	534	86.3
TOTAL NOT SCREENED	32	18.6	37	10.2	69	12.9
TOTAL SCREENED	140	81.4	325	89.8	465	87.1
TOTAL NOT ELIGIBLE	0	0.0	285	87.7	285	61.3
TOTAL ELIGIBLE	140	100.0	40	12.3	180	38.7
TOTAL NOT INTERVIEWED	17	12.1	8	20.0	25	13.9
TOTAL INTERVIEWED	123	87.9	32	80.0	155	86.1

Exhibit 3

Exhibit 4

USDA/CSFII FIELD SUMMARY REPORT WAVE 3

11-09-89 11:12AM

PAGE: 3

INDIVIDUAL SCREENERS RECEIVED

RESULT CODE	-----BASIC-----						-----POVERTY-----						-----TOTAL-----					
	---DAY 1---		---DAY 2&3---		---TOTAL---		---DAY 1---		---DAY 2&3---		---TOTAL---		---DAY 1---		---DAY 2&3---		---TOTAL---	
	NUM	PCT	NUM	PCT	NUM	PCT	NUM	PCT	NUM	PCT	NUM	PCT	NUM	PCT	NUM	PCT	NUM	PCT
1 - HAS INTAKE	279	89.4	232	74.4	511	81.9	79	89.8	72	81.8	151	85.8	358	89.5	304	76.0	662	82.8
2 - NO INTAKE	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
3 - REFUSED	23	7.4	67	21.5	90	14.4	8	9.1	15	17.0	23	13.1	31	7.8	82	20.5	113	14.1
4 - UNAVAILABLE	6	1.9	5	1.6	11	1.8	1	1.1	1	1.1	2	1.1	7	1.8	6	1.5	13	1.6
5 - ILL	3	1.0	5	1.6	8	1.3	0	0.0	0	0.0	0	0.0	3	.8	5	1.3	8	1.0
6 - RELIGION	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
7 - OTHER	1	.3	3	1.0	4	.6	0	0.0	0	0.0	0	0.0	1	.3	3	.8	4	.5
8 -	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
9 -	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
10 -	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0
TOTAL RECEIVED	312		312		624		88		88		176		400		400		800	

Exhibit 5

HOUSEHOLD SCREENERS RECEIVED BY SEGMENT

SEGMENT	RESULT CODES														TOTAL RECEIVED	TOTAL ASSIGNED	PER CENT RECEIVED
	1	2	3	4	5	6	7	8	9	10	11	12	13	14			
11112103	1	0	0	0	0	0	0	0	0	0	0	1	0	0	2	7	28.6
11112204	4	0	0	0	0	0	0	0	0	0	0	0	0	0	4	4	100.0
11113101	5	0	0	0	1	0	0	0	0	0	0	0	0	0	6	6	100.0
11113301	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	4	25.0
11113309	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	5	20.0
11114108	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	10	20.0
11117202	0	0	0	0	1	0	0	0	0	0	0	2	1	0	4	4	100.0
11119107	2	0	0	0	0	0	0	1	0	0	0	0	0	0	3	3	100.0
11119206	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	7	14.3
11119306	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	20.0
11122101	1	0	0	0	0	0	0	0	0	0	0	1	0	0	2	2	100.0
11122307	1	0	0	0	0	0	0	0	0	0	0	2	0	0	3	10	30.0
11122405	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	5	20.0
11122506	2	0	0	0	0	0	0	1	0	0	0	1	0	0	4	5	80.0
11123405	0	0	0	0	1	0	0	1	0	0	0	1	0	0	3	5	60.0
11124101	1	0	0	0	0	0	0	1	0	2	0	0	0	0	4	4	100.0
11124109	3	0	0	0	1	0	0	1	0	0	0	0	0	0	5	5	100.0
11125105	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	5	40.0
11125304	1	0	0	0	0	0	0	0	0	0	0	0	1	0	2	5	40.0
11126101	1	0	0	0	2	0	0	0	0	0	0	0	0	0	3	3	100.0
11126109	1	0	0	0	0	0	0	0	0	0	0	1	0	0	2	2	100.0
11127107	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	5	60.0
11129108	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	7	42.9
11129207	4	0	0	0	0	0	0	0	0	0	0	0	0	0	4	7	57.1
11129307	4	0	0	0	0	0	0	1	0	1	0	3	0	0	9	10	90.0
11129408	4	0	0	0	1	0	0	0	0	0	0	0	0	0	5	5	100.0
11134101	2	0	0	0	0	0	0	0	0	0	0	1	0	0	3	3	100.0
11134109	0	0	0	0	0	0	0	0	0	0	0	3	0	0	3	4	75.0
11134208	1	0	0	0	0	0	0	0	0	0	0	1	0	0	2	2	100.0
11135107	6	0	0	0	0	0	0	0	0	0	0	1	0	0	7	7	100.0
11135207	3	0	0	0	1	0	0	0	0	0	0	1	0	0	5	8	62.5
11135305	1	0	0	0	0	0	0	0	0	0	0	1	0	0	2	6	33.3
11136105	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	50.0
11136204	3	0	0	0	0	0	0	1	0	0	0	0	0	0	4	5	80.0
11212115	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	6	33.3
11212303	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	5	60.0
11213305	4	0	0	0	3	0	0	0	0	0	0	1	0	0	8	8	100.0
11217206	2	0	0	0	0	0	0	1	0	0	0	5	0	0	8	8	100.0
11218103	0	0	0	0	0	0	0	3	0	0	0	1	0	0	4	5	80.0
11219103	2	0	0	0	0	0	0	0	0	0	0	0	0	0	2	10	20.0
11219202	5	0	0	0	0	0	0	0	0	0	0	0	0	0	5	5	100.0
11222105	1	0	0	0	0	0	0	0	0	0	0	1	0	0	2	6	33.3
11223102	5	0	0	0	0	0	0	1	0	0	0	0	0	0	6	7	85.7
11223401	1	0	0	0	0	0	0	1	0	0	0	0	0	0	2	4	50.0
11225508	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	8	12.5
11227103	2	0	0	0	0	0	0	0	0	0	0	1	0	0	3	3	100.0
11229104	2	0	0	0	0	0	0	1	0	0	0	0	0	0	3	5	60.0
11229203	2	0	0	0	2	0	0	0	0	0	0	2	0	0	6	6	100.0
11229404	1	0	0	0	1	0	0	0	0	0	0	0	0	0	2	5	40.0
11229503	1	0	0	0	1	0	0	1	0	2	2	0	0	0	7	7	100.0

In addition to assigning sequence numbers and providing input for the monitoring effort, the check-in system identified a portion of the interviews for validation.

Each interviewer's work was subjected to validation efforts. The third interview completed by an interviewer was the first to be validated. Thereafter, every tenth interview from that interviewer was flagged and a validation attempted regardless of whether or not some other step in the processing effort triggered recontact with the household. A portion of nonresponse screenings (e.g., ineligible households) were also validated.

Validations first were attempted by telephone, if a phone number was given. Up to six attempts were made to reach respondents after which point mail validation forms were sent. If no phone number was available, the mail validation form was sent directly. Validation questions covered factual matters, such as number of people in the household, monthly income and food benefits program participation, as well as information about the interview process, such as the length of interview, the method of data collection, and the use of food reporting aids. Any interviewers whose work did not validate was subjected to further checking. All of his or her interviews were reviewed and validated.

2. Reviewers next examined documents in detail to determine completeness and their capacity to be coded

A team of trained editors accepted the completed questionnaires from the check-in effort and reviewed them for completeness, depth of information and integrity. Reviewers followed the steps outlined in the **CSFII 1989 Review Procedures Manual** using forms created for this phase of the research. The results of their reviews were recorded on **Review Summary Sheets** and, depending upon the nature of the item, the time period in the study and level of effort needed to edit the data, recontact with either the interviewer or respondent, may have been mandated. Reviewers were trained to exercise their judgment in the review process. If they could resolve the problem through careful edit and use of information elsewhere in the documents, they were advised to do so and to minimize unnecessary follow up with the interviewer or additional intrusion into the sample households, the latter being especially critical in this survey where the follow-on DHKS survey was still to take place in most households.

In the first wave of CSFII Year 1, *all* missing, incomplete or inconsistent information spotted by reviewers provoked callbacks to the interviewers. In addition, photocopies of their questionnaire recording problems were mailed to interviewers as a means of reinforcing proper data collection activities and of training them in the standard recording conventions of the study. Thereafter, callbacks to interviewers or respondents were attempted if:

- **Significant missing or questionable data were noted** – According to the guidelines established for this study, key data elements needed to be intact. This meant that: 1) the number of people in the household listed in the questionnaire was to match other sources of information about household size, such as the screening form data, 2) personal information from the cover of the intake record was to match the same information in the household listing, 3) work or schooling information for any appropriate household member needed to be complete, 4) school meal participation needed to be present (if appropriate), and 5) key income information was to be present. If any one of these conditions did not hold, a callback was to be attempted.
- **The Household Questionnaire was incomplete** – If fewer than 80% of the applicable questionnaire items were answered, then a callback was to be tried.
- **Intake records were deficient in codable information** -- If six or more of the food/quantity descriptions were incomplete or were difficult to fit into the existing HNIS code structure, then follow up was initiated.

In general, the interviewer was the first source of information retrieval, followed by the respondent. If the interviewer could not be contacted on the first or second attempt, then the household respondent was attempted. In some cases, where the reviewer judged that the information was more likely to be gathered from the respondent (e.g., identifying whether the quantity of meat reported in the intake record was raw or cooked), then he/she would reach first for the household. At the time a household contact was made for missing information, a validation interview was often attempted also.

Table V-1 shows the results of the validation efforts in the Year 1 survey. More than one in eight of the completed interviews were validated with respondents. These validations were divided two-thirds/one-third between the basic and the low-income sample households. In an additional five percent of the households, validations were attempted (most by mail to nontelephone households) but no response was obtained.

Unlimited calls could be attempted to retrieve information, although generally three were tried to complete the review stage. Answers to questions from either interviewers or respondents were recorded directly onto the documents in a distinctive color pencil (i.e., green) to denote the post-field inputting of data. In addition, the information was recorded on the Review Summary Sheets. If the information could not be gleaned from the callbacks or no contacts were successful in reaching the knowledgeable parties, then the documents were forwarded, as is, to the next stage of processing.

Table V-1

CSFII VALIDATION ATTEMPTED AND COMPLETED

	<u>Basic Sample #</u>	<u>Low-Income Sample #</u>
<u>CSFII Validations</u>		
Attempted and completed	198	89
Attempted and not completed	55	37
(N) =	(253)	(126)

3. Intake data were interactively coded and entered into the DECSS system, while the remainder of the questionnaire information was keyed using standard survey research data entry systems

The entry of the information from the CSFII interview was accomplished using two separate entry systems. The information contained in the intake record booklets was coded and entered interactively through a custom-tailored Data Entry and Coding Support System (DECSS) developed to handle USDA food intake surveys. The remainder of the questionnaire data -- the screening form and household questionnaire -- were manually coded and inputted using ACCESS, a commercially available data entry software key-to-disk system. Each of these activities is described in more detail below.

- a. Coding and entry of individual intake data was performed using DECSS*

The process of turning the verbal food descriptions and related information collected in the DAY 1 and DAY 2/3 Individual Intake Records into computer cleanable data was a three-step process:

- Interactive coding of the individual intakes
- Cumulating all coded individual intake files
- Converting the coded individual intake files into ASCII data files for transporting into cleaning format files

Data Entry and Coding Support System

The Data Entry and Coding Support System (DECSS) was developed in a computer language known as *C English*. It is a menu-driven, user-friendly support system designed to aid in the food code and gram weight assignment process. The computerized system helps eliminate many of the problems associated with the complicated, error-prone manual efforts of assigning food codes and gram weights to respondent-provided verbal descriptions of foods and beverages consumed. It also allows for the direct entry of all supplementary data contained in the DAY 1 and DAY 2/3 Intake Records.

The interactive coding system permits easy entry; it accepts only allowable codes for each data field and it forces the enterer to follow the appropriate skip patterns in the intake documents. The foundation of this process is a computerized version of the USDA Codebook for individual intakes. The information in the Codebook has been encoded into the system in two files.

*For further information about this stage of processing, please see the Computer Programs and Quality Control – Year 1 report

- *The food code description file*

The food code description file contains all the seven-digit food codes identified by USDA by the name of the food accompanied by the complete descriptions of each. This file is unique for two reasons:

- It is flexible. There may be many different descriptions assigned to a food code to accommodate all the variations embedded in a USDA code. For example, food code 631-2301 has the following descriptions in the file:

631-2301 GRAPE EMPEROR ADHERENT SKIN RAW

631-2301 GRAPE EUROPEAN ADHERENT SKIN RAW

631-2301 GRAPE TOKAY ADHERENT SKIN RAW

631-2301 GRAPE THOMPSON ADHERENT SKIN RAW

631-2301 GRAPE RED FLAME ADHERENT SKIN RAW

631-2301 EMPEROR GRAPE ADHERENT SKIN RAW

631-2301 EUROPEAN GRAPE ADHERENT SKIN RAW

631-2301 TOKAY GRAPE ADHERENT SKIN RAW

631-2301 THOMPSON GRAPE ADHERENT SKIN RAW

631-2301 RED FLAME GRAPE ADHERENT SKIN RAW

Any one of these choices may appear as a match to the data inputted by the coder.

- It is dynamic. This file was continually updated as "replacement pages" to the Codebook were received from USDA. More than 6,250 unique USDA food codes were employed in this study, for which there were more than 21,250 different descriptive statements that existed in the DECSS file. Many of these were additions made to the file during the course of this assignment.

- *Gram conversion file*

The gram conversion file contained all the food codes and all the existing allowable units of measurement for that code and the gram weight associated with each unit. Again, this file was updated with each set of replacement pages received from USDA.

To initiate the interactive food coding process, the coder entered the name of a food or beverage as given by the respondent and typed in a short string of descriptive statements about the item. The computer searched the USDA Codebook file for all available food names and descriptions that matched the information inputted. All potential seven-digit food code numbers and verbal descriptions that matched the input were ordered (from the largest to the fewest number of common key word matches) and arrayed on the screen. The coder selected the most appropriate match, if there was one, highlighted it on the screen and entered it into the file.

Next, the coder entered into computer the quantity consumed and the units of measurement recorded in the intake record for that food line item. If these were standard measurement units for that item, the computer accepted the input, linked the information to the *gram conversion file* and assigned a gram value to the line item. The computer then prompted the coder to enter information about the next food line item.

The interactive system also handled default situations as well as matches. If there was no match with the food name and/or descriptive information or if none of the choices were judged by the coder as appropriate, the coder could modify the description of the food entry and begin the search again. If a suitable match was still not found, the coder terminated the search for that item. No entry of that food line item was made at that time but a request for food code information was prepared for HNIS review.

If the food quantity units reported for the respondent did not match exactly the units acceptable for that food code, the computer screen displayed all of the units of measurements that were associated with that seven-digit food or beverage code. If the coder could convert the quantity into one of these allowable units of measurement, then he or she selected the appropriate values and entered the number of those units consumed. The computer automatically converted these values into gram amounts.

However, if the food units reported by the respondent could not be translated into the allowable units in the file, the gram amount was temporarily given a "missing" value in the grams field in the associated record. This situation could arise because the HNIS-provided gram conversion file was missing information needed to link a food code to gram amounts or because the unit entered by the coder was not found in among the "standard units of measurements" in the file (i.e., "inches" were entered and only "cup" units are available in the file to be converted to grams).

When the gram conversion error messages occurred, coders reviewed their entries (to determine if the entry was correct) and consulted any supplemental data from USDA. On occasion, there was information available, off line, which would allow the conversion of unusual

quantities into grams. If the gram linkage still could not be established, then a request for information was submitted to HNIS.

b. USDA regularly updated the dynamic food nutrient database and supported National Analysts' requests for food coding information

Requests to USDA were generated for two reasons, no food code matched the description given by the respondent and/or no gram amount could be assigned for whatever reason. The need for a request was flagged by the coder and the document was forwarded to the request writer. First, the request writer verified the necessity of a request and then filled out the USDA request form (see Exhibit V-6). Each request was kept in numerical order by request number for ease of retrieval upon return of information from USDA. Returned request information was inserted into the missing individual intake file by a cleaner/updater during the cleaning process.

Over the period of the Year 1 coding effort, the request process resulted in more than 2500 requests for food code and weight information being generated. A breakdown of these requests by wave and sample type is shown in Table V-2. As has been typical with these types of studies, the greatest number of requests are generated in the initial waves of the survey and proportionally more come from the basic than from the low-income sample. Not only are there more households in the basic sample but the variety of foods eaten and experimented with is usually greater.

With each request, the HNIS staff made the determination whether or not to generalize their response to other occurrences of that food or beverage item. If this were a single-case application, only the specified individual intake record was amended. If the information was to be put into general usage (i.e., pages to the Codebook were provided by USDA), the DECSS and general cleaning systems were also modified to accept the new values.

These requests led to approximately 95 new food codes being introduced into the file by HNIS. New food product introductions and home recipes, especially unusual mixtures, were the most common cause of food requests.

YEAR 1

Page #: _____

INDIVIDUAL INTAKE REQUEST -- CSFII-1989
(Intake Record To Be Attached)

Request #: _____ Date Sent to USDA: _____

Quarter 1 2 3 4 Date Returned from USDA: _____

Household Sequence #: _____ Interviewer #: _____

Respondent Line #: _____ Coder #: _____

Food Item Line #: _____ (Include all line #'s if same item is repeated)

Name of Food/Drink (Q.4): _____

Food Source (Q.7): 1 2 3 >>> Where Obtained (Q.11): _____

Problem:
(Circle all that apply) Need Code Need Weight Other _____

Specific Reason for Submitting Request: _____

Suggested Code: _____
(Indicate "none" if code can not be suggested)

Callback Made: Yes _____ >>> Explain Result: _____
No _____

FOR USDA USE ONLY

Code(s)/Weight(s): _____

Notes: _____

Replacement page will follow _____

Table V-2

**NUMBER OF REQUESTS FOR FOOD CODE
AND QUANTITY INFORMATION
BY WAVE AND SAMPLE TYPE**

	<u>Base Sample #</u>	<u>Low-Income Sample #</u>	<u>Total Sample #</u>
Wave 1	708	349	1,057
Wave 2	547	145	692
Wave 3	303	66	369
Wave 4	301	85	386
(N) =	(1,859)	(645)	(2,504)

- c. All information contained in the intake record was interactively entered and cumulated

In addition to the intake data, coders interactively entered the supplemental data contained in each intake booklet (e.g., food frequency data). During the entry effort, range and skip pattern edits were applied. When a coder completed entering information from an intake record, he/she began the process again with the next record from that individual, from the next person in the household, or from another individual in the next household. Periodically, the work of coders, particularly new ones, was coded a second time by experienced personnel. One hundred fifteen intake records were double-coded.

DECSS is a diskette-based system; therefore, multiple intake records were entered and outputted to each diskette. Each output diskette contained approximately 125 food line items (or the equivalent of one to two household sets) of data. All coder-generated diskettes were cumulated into data files by wave and by sample type for subsequent processing.

- d. Data entry of the Screening Form and Household Questionnaire was accomplished using standard entry programs

The key entry of the Screening Form and Household Questionnaire data associated with completed CSFII interviews was accomplished using ACCESS, a proprietary data entry system. This system minimized data entry errors by accepting only valid entries as data inputting occurred. All data key entered were 100% verified, that is, the data were keyed a second time with any differences in entry being flagged for further inspection by the verifier. These data sets were then cumulated into larger files -- one for each wave by sample type -- and prepared for further cleaning along with the data from the intake records.

4. The database for each wave was subjected to extensive data edits and cleaning procedures to create files which were complete and logical representations of the survey results

These coded raw data which were associated with the CSFII intake and household/screening data sets were subjected to a series of cleaning and verifying procedures to detect problems in the file. The cleaning application programs were written using REVELATION, a database management system, and were designed to function as an iterative editing process. Data files, organized by wave and sample type, were run through the cleaning programs in batch mode. Each data set (e.g., all the data associated with a particular household) was tested against the array of edit specifications and all the errors or problems associated with it were identified on a printout.

Trained data cleaning personnel referenced the original source documents to resolve problems, noted their corrections on the printouts and amended the data file directly. Any time a line item of food was reviewed for unusual quantity, a notation was entered into the file to register the event. The outcome could be coded:

- 2 = Extreme value, verified as accurate
- 3 = Extreme value, corrected and still extreme
- 4 = Initially an extreme value, but corrected and no longer extreme

The corrected data unit would be cycled through the cleaning programs repeatedly until no fatal errors existed in the data set. By the end of data cleaning for a wave, all the related household and intake data were resident in the cleaned data file, organized by wave and by sample type.

Several types of specialized computer edits were adapted for the CSFII survey data. Entries outside established ranges or inconsistent with other entries were flagged as *fatal errors* or *warnings*. Examples of the type of situation which would be diagnosed as *fatal errors* include: 1) finding that the length of time a child was breast-fed is greater than the age of the child, or 2) determining that age given in the Household Questionnaire is not the same (or within a year) of the age calculated from the date of birth. Examples of conditions which would trigger *warnings* include highly unusual but feasible values. Cases such as weekly food expenditures per household member of less than \$5 or more than \$75, or persons older than 97 years of age constitute examples of warning edit checks. Appendix I lists the types of range and logic checks in this stage of the cleaning process.

Particular attention was given to the intake information in the cleaning effort. To ensure that unrealistic summary data were not being derived from the calculated nutritional variables, several tiers of parameter checking were employed. Many of these "reasonableness" checks were based upon empirical values arising from other USDA surveys. These special checks included:

- Gram warnings by individual food and day totals

Unusually large consumption of individual food items or of total intake for an individual, relative to his or her age and sex peers, triggered a lookup of the data recorded in the primary documents (see Appendix J) for the age-calibrated upper cut-off values for individual food items). The empirical cut-offs for total grams of food consumed in one day were:

<u>Age Category</u>	<u>Total Maximum Grams</u>
Children 0 to 11 years	2274.76
Males 12+ years	3806.51
Females 12+ years	2835.19

- Extreme value warnings for key nutrition elements

Empirically-derived extreme values were set for seven major nutrients and food energy for three age groupings – children (0 to 11 years), males (12 years and over) and females (12 years and over). The cut-off values against which each day totals of intake are checked are as follows:

<u>Nutrients</u>	<u>Units</u>	<u>Children 0 to 11 Years</u>	<u>Males 12+ Years</u>	<u>Females 12+ Years</u>
Food energy	Kilocalories	2,761.000	4,108.950	2,806.000
Protein	Grams	110.815	172.400	120.200
Fat	Grams	129.100	209.585	142.700
Vitamin A	IU	11,727.700	17,411.250	14,772.000
Vitamin E	Alpha-TE	30.000	50.000	50.000
Vitamin C	Milligrams	200.000	242.000	221.000
Vitamin B-12	Milligrams	7.911	11.649	7.970
Calcium	Milligrams	1,582.150	1,953.900	1,372.000

- Height and weight warnings

Another set of reasonableness reviews involved checking unusual height/weight configurations within ten age groupings. Any height measurements or weight measurements outside the ranges specified below were looked up in the source documents and corrected in the file, if in error.

	<u>Height in Inches</u>		<u>Weight in Pounds</u>	
	<u>Low</u>	<u>High</u>	<u>Low</u>	<u>High</u>
Under 1 year	14	33	6	30
1 to 3 years	18	43	20	49
4 to 6 years	28	54	30	72
7 to 10 years	36	61	40	111
11 to 14 years	49	71	61	167
15 to 18 years	59	76	92	208
19 to 22 years	59	76	95	229
23 to 50 years	60	76	97	249
51 to 75 years	59	76	100	239
76+ years	59	74	85	204

- Extreme value warnings for calcium consumptions

Any large amounts of typical intake in Q. 17 of the DAY 3 record were visually reviewed. The amounts, reported in cups, used to prompt a data review are shown below:

<u>Description</u>	<u>Cups per</u>		
	<u>Day</u>	<u>Week</u>	<u>Month</u>
Milk as a beverage	8.000	56.000	240.000
Milk on cereal	3.000	21.000	90.000
Milk in coffee, tea, other	3.000	21.000	90.000

5. The cleaned and expanded data were then ordered in the final format and given a final review

The last stage of CSFII data processing was preparation of the final file for completed interviews. This involved expansion and ordering of the survey database. First, the coded food information for each participating individual in a household was expanded into basic nutritional elements and summarized as day totals and as three-day averages. There were 29 nutritional components associated with every food line item of data. These are shown in Table V-3. The data for each of the 29 elements were totaled (across line items) for every completed daily intake record for every individual in the household. In addition, the daily totals for selected nutrients were reported as ratios to the RDAs (calibrated for age and sex) for each individual (see Table V-4).

Next, the final output was ordered hierarchically in a convention compatible with USDA requirements. Data were arrayed in five-digit sequence number order by wave and sample type. Within a household the lead record represented the information from the household questionnaire. Nested under that was individual-specific data followed by up to three days of intake information. The intake records themselves were nested by day and within day by line item nutrient information. This information was followed by day total summary nutrient data. These data were organized by "record type" as follows:

- **Record Type 15**

This record type consists of general household characteristic data. There is one Record Type 15 per household.

- **Record Type 20**

Table V-3**NUTRITIONAL ELEMENTS AND THEIR UNITS
OF MEASUREMENT REPORTED FOR
EACH COMPLETED INDIVIDUAL INTAKE RECORD**

Food energy -- kilocalories
Protein -- grams
Total Fat -- grams
Saturated Fatty Acids -- grams
Monounsaturated Fatty Acids -- grams
Polyunsaturated Fatty Acids -- grams
Cholesterol -- milligrams
Carbohydrate -- grams
Total Dietary Fiber -- grams
Alcohol -- grams
Vitamin A -- IU (international units)
Vitamin A -- RE (micrograms retinol equivalents)
Carotenes -- RE (micrograms retinol equivalents)
Vitamin E -- Alpha-TE (milligrams alpha-tocopherol equivalents)
Vitamin C -- milligrams
Thiamin -- milligrams
Riboflavin -- milligrams
Niacin -- milligrams
Vitamin B-6 -- milligrams
Folate -- micrograms
Vitamin B-12 -- micrograms
Calcium -- milligrams
Phosphorus -- milligrams
Magnesium -- milligrams
Iron -- milligrams
Zinc -- milligrams
Copper -- milligrams
Sodium -- milligrams
Potassium -- milligrams

Table V-4**NUTRIENTS FOR WHICH PERFORMANCE
AGAINST RDAs CALCULATED**

Food energy – kilocalories

Protein -- grams

Vitamin A -- IU (International units)

Vitamin A -- RE (Micrograms Retinol Equivalents)

Vitamin E -- Alpha-TE (alpha-tocopherol equivalent)

Vitamin C – milligrams

Thiamin -- milligrams

Riboflavin -- milligrams

Niacin -- milligrams

Vitamin B-6 -- milligrams

Folate -- micrograms

Vitamin B-12 -- micrograms

Calcium -- milligrams

Phosphorus -- milligrams

Magnesium – milligrams

Iron -- milligrams

Zinc -- milligrams

This record type contains general household characteristic and personal data, but no intake data. There is one Record Type 20 for every member of the household.

- **Record Type 30**

This record type contains individual intake data that describe each food line. There is one Record Type 30 for *each individual food item* listed in the intake file.

- **Record Type 40**

This record type contains the individual's daily nutrient intake expressed as a percentage of the Recommended Daily Allowances for his/her age group and as an absolute level of intake. There is one Record Type 40 for each day of intake per each individual, plus one record which is an average for all intake days, when multiple days of intake were reported.

Three types of data checks were performed before the individual record type files were merged together. These were final range, logic and reasonableness checks and are described below:

- **Range Checks**

A range check of all data fields with defined values was performed by reviewing printouts of frequency distributions. All fields were verified to ensure that only valid answers appear in the output data fields.

- **Logic Checks**

A logic check of all data fields which were dependent on information from other data fields was performed. This was done using a series of SAS programs that independently verified the output data. All fields were verified to ensure that the proper skip patterns were followed, that the appropriate number of records were present for each household and that the data were consistent across the file.

- **Reasonableness Checks**

A reasonable check of data fields with undefinable values was performed by identifying extremes of the frequency distributions and, then, inspecting all documents with the suspect data.

The final format for the output of the CSFII interview data set is found in Appendix K.

B. Creation of the Screening Data File Parallels the Work for the CSFII Interview File.

One of the critical data files emerging from this survey effort is the screening file. These data represent an accounting of all the sample housing units involved in the CSFII sample. Such information is necessary for determining completion rates and weighting factors. Information collected by the interviewers about every household, as well as the results of calls at those addresses, is contained in this file. As shown in Exhibit V-1, the processing of the CSFII screening/call report form data set moved in parallel with that of the handling of the interview data file.

The screening form information is comprised of several data elements. The data associated with completed interviews contain segment and housing unit information, household composition and eligibility data, outcome data identified for every attempted contact with the household and information about the presence (or absence) of intake records. The data present in the screening data set associated with housing units where interviews were not completed varied according to the reasons for noninterviews. For all noninterview screenings, there were data for segment and housing unit and for attempted results-of-call. Other data elements which may be present include household composition information, eligibility data and household descriptive data (e.g., race).

During different stages of post-field processing, the screening data associated with completed CSFII interviews and nonresponse screenings were handled separately. In the final stage, all screening data were brought together and a unitary file for each wave was prepared. These post-field handling and processing steps are described below.

1. Screening documents were first entered into the central interview control file

As previously described, completed CSFII interviews were logged into the control file with their attendant completed Screening Forms. Nonresponse Screening Forms were entered independently. Interviewers were advised to return these documents on a regular basis in small packs of 4 to 6.

As with the interviews, nonresponse screening documents generally were given initial attention within a day of their receipt in the Philadelphia office. Before entry they were examined and those that were incomplete were held for review with interviewers. Those that passed inspection were assigned unique sequence identification numbers as shown on the next page:

**NONINTERVIEW SCREENING FORM
NUMBER SEQUENCE**

<u>WAVE</u>	<u>BASIC</u>	<u>LOW INCOME</u>
1	30000-30309	40000-41800
2	32000-32324	42000-43999
3	34000-34329	44000-44862
4	36000-36354	46000-47800

In addition to assigning sequence numbers, the check-in system identified a portion of the nonresponse screenings to be validated.

2. Reviewers examined and coded screening documents according to procedures established for this survey

Reviewers coded the screeners associated with completed interviews when reviewing those data sets. Nonresponse Screening Forms were coded separately following established procedures, outlined in the review manual titled *CSFII and Nonresponse Screening Form* (See Appendix L). In addition to preparing the document for key entry, reviewers examined the reasons for nonresponse among eligible households. If there were conditions that indicated follow up might yield an interview, (e.g., the interviewer thought a central office call verifying the legitimate purpose of the survey would help), then the reviewer put the form aside for further field action. Any conversions to interviews were re-entered in the control system as part of completed interview data sets.

3. Nonresponse screenings were keyed using standard survey research entry systems

All Screening Form information was key entered and verified 100% using the ACCESS entry system. Screenings associated with completed interviews were entered with their relevant household data while nonresponse screening data were processed and verified in batches of 25 documents. The data were entered onto diskettes and then cumulated into a single file for each wave and sample type.

Based upon a misunderstanding of specifications, initially only the final result of call information was entered for every screening. Therefore, in Year 1 the results of call for all attempted contacts up to the last one was entered as a separate activity and then the data merged into the existing relevant files.

4. Nonresponse screening documents were edited for integrity and completeness

The data set for each wave was subjected to computerized edit checking as a whole. These checks addressed the logic of skip patterns, range checks and comprehensiveness of coverage. The latter was aimed at assuring all sample housing units were accounted for and that no duplicate entries were noted among the nonresponse screenings.

5. The last processing step wedded together the screening information from both the nonresponse and participating households into a single file

To prepare the screening data file for each wave, the screening data associated with completed interviews and the nonresponse screening data file were combined. This merged file was ordered by housing unit numbers within segment number order for each sample type.

A final set of edit checks was performed on the combined file: 1) to verify that all the defined data fields were within acceptable ranges, 2) to determine that skip patterns were followed, and 3) to ensure that all sample housing units were accounted for without duplication.

C. New Programs and Procedures Were Developed to Handle the Automated and In-Person DHKS Interview Data Files

As the DHKS interview activity was a totally new addition to the survey effort, new protocols and programs were developed to handle this portion of the research. DHKS questionnaires were collected both electronically by telephone and in hard copy in the field. A system was created to merge and process the two data sets together. Unlike the CSFII files where separate interview and screening data sets were created, for the DHKS a single merged file by wave was produced which contained both completed interviews and call report information from noninterviewed households. For each wave, a combined basic and low-income survey file was produced. These activities are described below.

1. The central field control file was adapted to track completed DHKS household contacts

The check-in system used to control the CSFII portion of the survey was adapted to maintain control over the DHKS survey effort. All completed hard-copy questionnaires and nonresponse call report forms from the field were logged into the system.

Precoded information from the DHKS interviews completed by telephone was recorded directly onto diskettes and cumulated in the "raw" data file. Verbatim responses to open-ended questions were printed out by sequence number and forwarded for coding. Hard-copy questionnaires were forwarded directly to review personnel for a complete edit and code.

2. Documents were coded using the framework developed from the first wave of interviews

National Analysts prepared a coding manual for handling the DHKS interview which was approved by HNIS (see Appendix M). Of particular importance was the development of coding structures for the open-ended and the partially-open responses. During the first wave, verbatim recordings of these types of responses were maintained. Based on these recording, National Analysts recommended, and HNIS approved, a code structure for the DHKS questionnaire. Thereafter, all verbatim responses to open-ended and partially-open questions were coded by a pair of expert coders.

3. DHKS data were key entered and cleaned using standard survey research data handling programs

The hard-copy DHKS questionnaires and nonresponse call report forms were key entered and 100% verified using the ACCESS entry system. Data from the hard-copy documents were merged with telephone questionnaire data to create one consolidated file. As with many of the other data files, the DHKS files were organized in sequence by the National Analysts assigned five-digit number for each wave and for each sample type.

The consolidated file was edited using the UNCLE survey data cleaning program, a proprietary survey/marketing research cleaning and tabulation package. The data were cleaned to assure that all the skip patterns were followed, that only legitimate responses were recorded, and that the information could be linked correctly to a previously-interviewed CSFII households.

D. The Nonresponse Follow-up Survey Data Were Processed Using Standard Survey Research Procedures and Systems

Handling of the nonresponse follow-up questionnaires paralleled that of the hard-copy DHKS documents. Reviewers visually inspected the completed forms as they were received and forwarded them to data entry personnel where they were keyed and verified 100%. The UNCLE data editing package was used to clean the file and then a single unified data file for each semi-annual survey was submitted to HNIS on diskettes ordered by a National Analysts assigned five-digit sequence number order.