METHODOLOGY FOR LARGE-SCALE SURVEYS OF HOUSEHOLD AND INDIVIDUAL DIETS

Home Economics Research Report No. 40

Agricultural Research Service
UNITED STATES DEPARTMENT OF AGRICULTURE

METHODOLOGY FOR LARGE-SCALE SURVEYS OF HOUSEHOLD AND INDIVIDUAL DIETS

By Marguerite C. Burk and Eleanor M. Pao

Home Economics Research Report No. 40

Agricultural Research Service

UNITED STATES DEPARTMENT OF AGRICULTURE

Washington, D.C. Issued November 1976

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402 - Price $1.50
Stock No. 001-000-03620-3

CONTENTS

Company names are used in this publication solely for the purpose of providing specific information. Mention of a commercial organization does not constitute a guarantee or warranty by the U.S. Department of Agriculture or an endorsement by the Department over other companies not mentioned.

METHODOLOGY FOR LARGE-SCALE SURVEYS OF HOUSEHOLD AND INDIVIDUAL DIETS

By Marguerite C. Burk[1/] and Eleanor M. Pao
Consumer and Food Economics Institute
Northeastern Region
Agricultural Research Service

1. INTRODUCTION

The U.S. Department of Agriculture (USDA) is responsible for measuring and appraising trends and variations in U.S. food consumption. Large-scale surveys have been made to study variations in household food consumption and dietary levels in 1936-37, spring 1942, 1948 (urban only), spring 1955, and 1965-66. In spring 1965 the food intake of a nationwide sample of individuals was also surveyed. The Agricultural Research Service (ARS) has been the key agency for this series of one-time studies, which have provided both a cross-section view of food consumption in this country and a historical perspective on changes in food patterns. Year-to-year changes in food consumption by the entire U.S. civilian population are measured by the official estimates of per capita food consumption, developed by the Economic Research Service (ERS) from data on food supplies and distribution, often referred to as "disappearance data."[2/] From these per capita food data, ARS calculates the average nutritive values of the annual food supply.

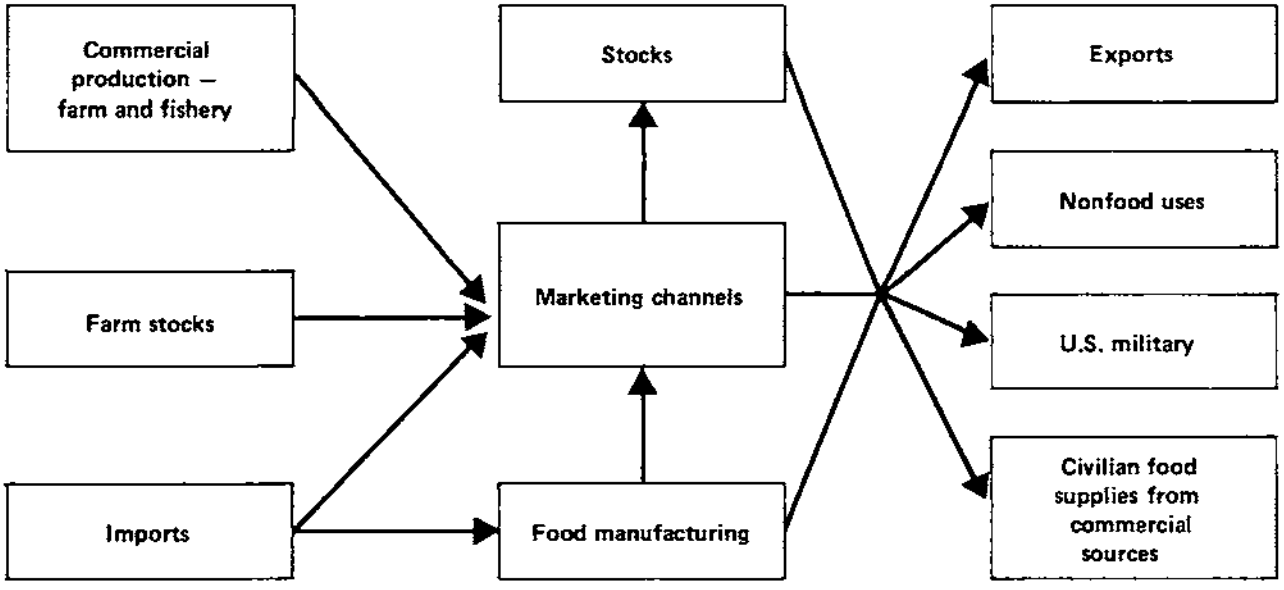
The general relationships between the household and individual survey measurements and between these survey measures and the ERS per capita food consumption estimates are diagramed in figure 1. Although most of the ERS estimates for individual foods approximate quantities at the wholesale level of distribution, since retail stock data are not available, they are regularly converted to retail weight equivalents and published on both bases. The USDA household food consumption surveys yield measurements of

---

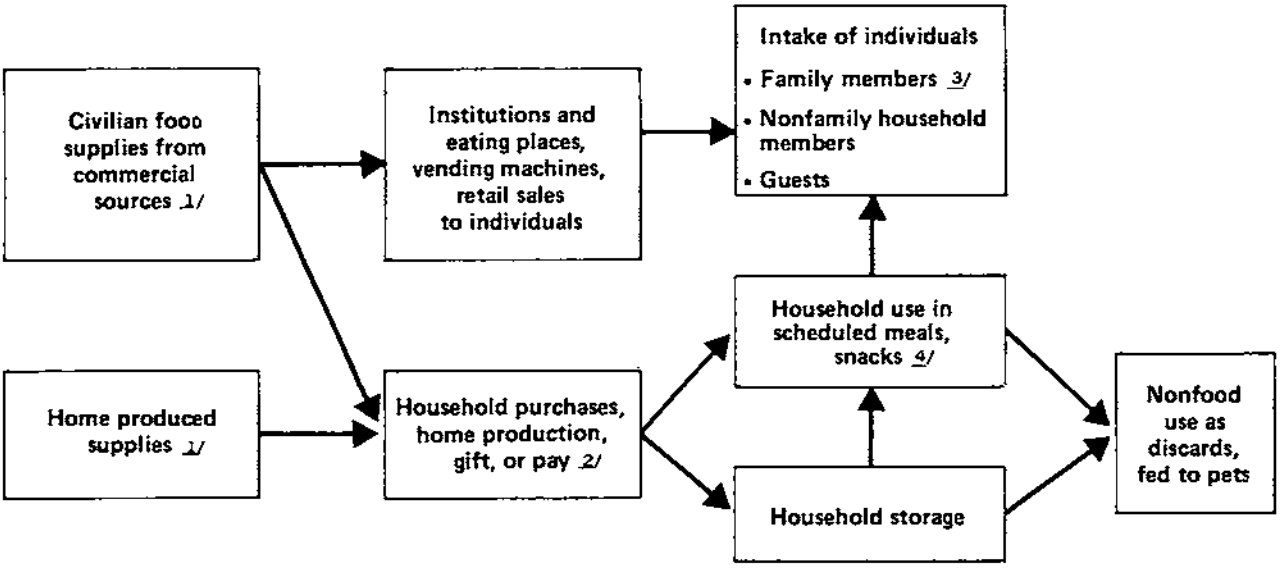
1/ Retired July 1975.

2/ Described in Hiemstra (1968). The year (underlined) after the author's name refers to Literature Cited, p. 76.

PART 1 — Food supplies and distribution



PART 2 — Movement to household and individuals



1/ Measured by Economic Research Service disappearance data.

2/ Measured by food accounts data like those of Great Britain.

3/ Measured by individual dietary surveys.

4/ Measured by Consumer and Food Economics Institute household food consumption surveys

Figure 1. — Diagram of the flow of food from producer to consumer.

household food consumption in terms of the retail weights and money values, mostly at retail prices, of the quantities of all foods actually used at home during a week. Measurement of individuals' food consumption in spring 1965 was based on the quantities of foods eaten by members during a 24-hour period, both at home and away.

In formulating plans for the next nationwide surveys, the authors have reviewed and appraised both the methods used in the 1965-66 surveys and several alternatives used in smaller scale surveys in the United States and other countries. The study of USDA household food surveys reported here is based partly on the historical review by Janet Murray (1970), on various papers by Faith Clark (e.g., 1969, 1974), and on working papers and technical suggestions of Evelyn Grossman and Doretta Popka (1976). This report is designed to make these findings available to other researchers. Because the major emphasis is on measurement and evaluation of food consumption and nutrient content, the surveys are often called "dietary surveys" to differentiate them from consumer expenditure surveys and market surveys for particular foods. Consumer expenditure surveys do not obtain sufficient detail on foods for calculation of the nutrient content of diets.

Since 1900, researchers have used both household and individual dietary surveys, conducted with varying procedures and yielding measures of different concepts (Marr, 1971). Until about 1930 the emphasis was on family food consumption studies, using the procedure of weighing beginning and ending inventories, then adding the difference to recorded acquisitions. Early estimates of individual diets within the family were based on meager evidence from very small so-called samples. In the 1930's, along with metabolic studies, studies of individual diets were undertaken in several countries to appraise food requirements, followed by even greater interest aroused by the need to study the association of diet with disease and health, particularly in epidemiological studies. Because of the wide variety of requirements for dietary data, no single approach can be used for all purposes. In fact the authors' research indicates the analytical potential of interrelating household and individual dietary data, which measure different concepts.

The objectives of the USDA surveys have changed over the years. Originally the emphasis was primarily on appraising how well fed the various segments of the population were. During World War II, food quantity and nutrient data from the Consumer Purchases Study (1936-37) and the spring 1942 survey were used extensively in national food planning and administration. Since 1945, the food survey data have been an important source of marketing information. The wide range of uses or problem areas for which the survey data are now sought includes public information, food market research and demand analysis, formulation of food plans at alternative cost levels, analysis of USDA food policies and programs, fishery and game problems relevant to health and marketing, and health-related problems for which the U.S. Food and Drug Administration and the Public Health Service have responsibilities. Also included are scientific problems in the areas of food and nutrition, food technology, and consumer education; survey methodology; and income-maintenance problems.

This report summarizes technical information relevant to planning large-scale surveys of the food consumption of households and the food intakes of individuals from which detailed data on nutrient content can be calculated. It includes (1) specification of concepts of food consumption measured by several types of data; (2) brief descriptions of alternative procedures used to measure concepts just specified, including those used in the nationwide 1965-66 household food consumption survey and the spring 1965 survey of food intake of individuals; (3) identification of several problems in evaluating alternative survey methods; (4) statement of criteria to be used here in evaluating alternative methods; (5) a summary of research findings relevant to evaluating alternative ways of measuring household food consumption; and (6) a comparable summary for individual food surveys.

2. CONCEPTS OF FOOD CONSUMPTION TO BE MEASURED

The concept of food consumption to be measured varies with the objectives of the researcher. Agricultural economists are interested in the amounts of food taken from different distribution channels by different categories of users at different times, because such data are essential for study of the demand and price of food and of farm incomes. Ordinarily home-produced supplies of food are included along with supplies obtained from marketing channels. Economists are much less concerned than nutritionists with what users do with food after it is bought. In contrast, home economists and nutritionists focus more on the use of food in the home and on the food intake of individuals in the household.

Several concepts of food quantities are used in economic analysis of food consumption. One concept refers to the total amount of all food from all sources moving into civilian consumption in a specified time period. Another refers only to current consumption of food from all sources within households in a limited period. A third economic approach to food consumption is to consider the flow of food into the households during a specified period without taking into account any changes in household stocks.

2.1 Total Consumption in the Country

The first of these quantitative concepts is measured by the annual per capita food consumption estimates of the Economic Research Service. They encompass the total amounts of all foods consumed by the entire U.S. civilian population during each calendar year at home, in eating places, as snacks from retail stores and vending machines, and in institutions measured in quantitative terms (e.g., pounds of carcass meat) at approximately the wholesale distribution level. These annual estimates, which include wastes and losses in distribution and in homes and eating places, are reported in the "National Food Situation."[3/] Approximations

3/ Published quarterly since 1947 by the Economic Research Service, U.S. Department of Agriculture.

for wastes and losses in distribution are used to convert these estimates to retail weights, which are then used by ARS along with food composition data to calculate the nutritive value data published each year in the "National Food Situation." These highly aggregated data are derived from total U.S. production estimates, foreign trade data, and changes in farm and wholesale stocks. Little information on movement of food across State lines is available; therefore economists must turn to survey data to obtain measures of food consumption in different places and in different parts of the country.

Several value concepts and measures related to this quantitative concept must be mentioned here. Farm value of food refers to the value in terms of farm prices of those proportions of farm commodities used as food. Retail value is the value of food priced at the retail-store level, assuming that all food was purchased at retail-store prices. Market value is used to represent the value of food at the prices paid by final consumers at several distribution levels, whether the food was purchased from farmers, retailers, or various types of eating places. Food expenditures ordinarily mean dollar outlays for food by consumers, excluding the imputed value of home-produced food. However, the U.S. Department of Commerce series on food expenditures includes the value of most food consumed on farms where produced and excludes food bought by business firms as meals for clients. For further information, see Burk (1961).

2.2 Household Food Consumption

Consumption of food by households during a specified period has been conceptualized and measured in terms of quantities, money value (mostly at retail prices), and nutrient content. Different approaches have been used. The first approach yields a measure of food use in private homes that is conceptually close to the measure of per capita food consumption just described. The second approach of measuring the flow of food into private homes is closer to the food expenditure concept. The third approach does not yield a very useful economic measure of consumption, because it excludes all wastes and losses of food in homes that must be taken into account for most economic analyses.

(1) Food Use From All Sources.--When consumption is conceived as actual use in homes by households and changes in household stocks are expected, actual use in the specified time period may be measured in equivalent retail weights from data on stocks, current purchases, current home production, and food received as gift or pay. Such household use includes food fed to animals and discards in the home, as well as food eaten by family and nonfamily members of the household at home or carried from home, by guests, and by employees. This is the concept measured by the 1965-66 household food survey.

(2) Quantities Purchased or Obtained From Other Sources.--On the assumption of no significant change in household stocks, consumption is viewed as being equal to current purchases of food from civilian distribution channels plus home-produced supplies plus food received as gift or pay. The food accounts procedure is based on this concept.

(3) Sum of Foods as Eaten.--Actual use in the home may also be conceptualized as the total amount of food actually prepared for each day's meals and eaten, with subtractions for discards in preparation, plate waste, or both. Such estimates may be derived by aggregating individual food intakes in order to measure shares of family members in the family food totals.

2.3 Food Consumption by Individuals

Food consumption by individuals is usually referred to as food intake or dietary intake and is generally appraised in terms of the forms in which foods are eaten--for example, fried potatoes, reconstituted orange juice, and beef stew. Differences in survey objectives determine the choice between concepts of individual food consumption to be measured: (1) The intake during a specified period being surveyed or (2) the intake representative of the usual diet of the individual. Concerning usual dietary intake, there is vagueness about the time span to be considered--week, month, year. Surveys concerned with relating variations in diets among individuals with socioeconomic characteristics need unbiased measurements of actual consumption.

Epidemiological surveys often seek measurements of usual consumption over several months or over years in order to relate diets to such matters as specific diseases, health status, or growth. Some use a series of observations to compute averages as measures of usual. Others rely on individuals' responses to questions about usual consumption and thus disregard variations that may be extremely low or extremely high. Usual intakes are subject to seasonal, cyclical, and longer range changes. Young and Trulson (1960) concluded from their methodological studies during the 1950's that there is no way to know the true usual dietary intake.

The key objective of cross-section surveys, such as the USDA survey of food intake of individuals in spring 1965, is to determine the kinds of foods actually eaten currently by individuals with varying demographic and socioeconomic characteristics. Accordingly they must sample actual diets both across individuals and across days of the week in the survey period.

3. ALTERNATIVE PROCEDURES USED TO MEASURE THE SEVERAL CONCEPTS OF FOOD CONSUMPTION

Differences among the various procedures to measure the several concepts of food consumption can be seen more clearly when the procedures are sorted by four levels of aggregation. The most highly aggregated data include all foods consumed by all people in the country and permit no breakdown by food commodity or population group. At the other end of the spectrum are data on separate food items consumed at specified times by individuals identified by sex, age, income, and so on.

3.1 Measure Derived From National Accounts

At the most highly aggregated level is the measure of change in overall food consumption in a country from one year to another. This measure is occasionally derived by economists from national accounts data on consumption expenditures for food, using a measure of changes in food prices in order to check other measures of change and to compare changes in food consumption with changes in consumption of nonfood goods and services. The national accounts data include information on gross national product, national income, personal income, and personal consumption expenditures subdivided among major categories, such as food, nonfoods, and services.

The U.S. source is the Department of Commerce's data on consumption expenditures for food, excluding business expenditures and including farm- and home-produced food. Such estimates in dollar terms are developed from sales data collected primarily in the censuses of wholesale and retail trade, supplemented by information from the Department of Agriculture, the Internal Revenue Service, and other agencies. These dollar data on a highly aggregated basis can be converted to measures of quantitative changes using special price deflators developed by the Department of Commerce from the Bureau of Labor Statistics and the Department of Agriculture's price indexes. Because of the many assumptions and fixed ratios built into such complex sets of data, they have proved to be less reliable than the Department of Agriculture's more direct measurement of food commodity flow. Descriptions of methods used to derive the national income and product accounts and detailed statistical tables are published as supplements to the "Survey of Current Business." 4/

3.2 Per Capita Consumption Data

The ERS per capita consumption data for all foods combined and for major food categories are derived from food balance sheets or supply and distribution tables by dividing estimates of the total disappearance into

---

4/ Published monthly by the U.S. Department of Commerce, Bureau of Economic Analysis.

civilian distribution (calculated as a residual) by the total civilian population.[5/] Because these data take into account the farm and wholesale stocks, not retail stocks, they may be viewed as being measured at the wholesale level. Approximate conversion factors are used to adjust the per capita quantities to retail equivalents in pounds. The poundage data are weighted by constant prices of a specified period to calculate index numbers for civilian per capita food consumption.

Civilian consumption or disappearance data include consumption of food in homes and away from home in eating places and institutions and as snacks, as well as nonfood use, as by pets, and wastes and losses in such places and in distribution. Because these estimates are derived as residuals from U.S. data only, there is no way to subdivide the aggregates among end users without special surveys. Moreover the lack of detailed data on interstate shipments precludes disaggregation by area.

3.3 Household Food Surveys

At the third level of aggregation are the household data obtained in surveys. Currently there are four types of data collection procedures used for such surveys. They include record or recall of purchases and other current acquisitions, record of acquisitions plus change in inventory or the inventory-record method, the list-recall procedure for obtaining current food use from all sources, and weighed or measured records of supplies to be used each day.

Food consumption in households has been measured by surveys in the United States and other countries for many years by asking homemakers to keep records of the amounts of foods brought into homes, whether purchased from retail stores or other sources, produced at home, or received as gift or pay. In the 1930's, as awareness of problems with nonresponse increased and the frequency of shopping trips decreased with more refrigerators and the beginning of supermarket shopping, the survey procedure used by the U.S. Department of Agriculture was changed after experimentation to interviewing homemakers with a list of major foods and asking them to recall use in their households during the preceding 7 days. During the 1930's in the United States, increasing attention was paid to sampling procedures and to problems related to representativeness of the final samples.

In highly urbanized countries like Great Britain, food consumption of households has been surveyed by having homemakers keep records of food purchases and of food from other sources entering the household during 7 days or longer. This measure assumes no significant changes in household inventories. It is apparently adequate in places where homes have limited

---

5/ For current data, see the annual supplements to Hiemstra (1968). For detailed discussion of procedures for estimating commodity data, see U.S. Dept. Agr., Bur. Agr. Econ. (1953).

storage facilities, especially refrigeration, and where food shopping occurs daily or several times a week. Because of the similarity to procedures followed in keeping diaries for family expenditure surveys, this survey procedure is sometimes called the food accounts method.

An example is the National Food Survey of Great Britain, a continuous sampling inquiry in which each private household selected participates voluntarily for 1 week of the year. A completely new sample, selected by means of a three-stage stratified random sampling scheme, is contacted each week of the year. An interviewer visits each home at least three times, leaving the logbook on the first visit, checking on the account keeping on the second visit, and picking it up on the last.[6/] The survey measures food quantities consumed at home in retail units and retail prices.

Market researchers have used extensively both record and recall procedures to obtain information on food purchases, particularly for a limited number of food items. Their emphasis has often been on quantities purchased or expenditures for particular commodities, including information on brands. The success of a survey using a recall procedure obviously depends on the skill of interviewers in helping homemakers to recall their purchases, on the importance of items to be recalled and the frequency of their purchase, and on a variety of other factors, such as bounding the time period. On the other hand, the success of recordkeeping procedures in market research depends also on the willingness and ability of the homemakers to keep the records and to write the purchases in their record books soon after they are made.

Both U.S. and British surveys of household food consumption formerly used the inventory-record method, which consisted of taking beginning and ending inventories and recording by the homemaker of foods brought into the home as purchased, home produced, gifts, or as payment in kind. In the mid-1930's the USDA shifted for most surveys to a list-recall procedure, whereas the National Food Survey of Great Britain retained the record of food purchases and dropped the use of inventories.

The 1965-66 USDA survey of household food consumption utilized the list-recall procedure to obtain information from a nationwide sample of housekeeping households. In this procedure the interviewer uses a list of major food items in a structured questionnaire to help the homemaker or another knowledgeable respondent to recall the amounts and prices or purchase values of all foods used in the household in a specified period, usually 7 days. Further information on the methods used in the USDA 1965-66 food consumption surveys is given in section 4.

The USDA's methodological survey in Cincinnati, Ohio, in 1969-70 measured household food use with both methods--the list-recall and the inventory-record of purchases and food acquired from other sources. The survey was conducted in two waves separated by 1 or 2 months. It consisted

---

6/ Appendix A of Min. of Agr., Fisheries and Food, Natl. Food Survey Committee (1974).

during the first wave of using the list-recall with one sample and inventory-record with a parallel sample. The participating households were reinterviewed during the second wave, reversing the schedule forms.

Records of food use during each day or each scheduled meal have been made in small-scale surveys in noncommercial food economies by weighing or measuring in household units the supplies for the day or meal in advance. As applied in Finland and in Central America, this procedure has usually included daily visits by field investigators, although some literate homemakers learn to do the weighing and recording themselves. Food survey specialists of the Food and Agriculture Organization recommend this procedure for use in primitive rural areas where foods are simple, the variety is limited, home production is very important, and units for food buying are not standardized. Foods not eaten by household members may or may not be deducted, depending on survey decisions. (Pekkarinen, 1970; Reh, 1962)

3.4 Surveys of Individual Diets

The least aggregated approach to measuring the consumption of food by a population, such as that of the United States, is to survey the amounts of different foods consumed by individuals. The six methods used to collect such data include (1) records in which data based on weighing food are entered currently, (2) records based on measuring or estimating in household units the amounts eaten, (3) recall over varying spans of time, such as 24 hours or several days, (4) dietary history based on recall of usual intake over a longer period, even a year, (5) frequency of intake estimated for either past or current consumption, and (6) combinations of these methods. Schedules may be self-administered or completed by an observer or interviewer. The respondent may give information for herself or for all members of her household.

Combinations of at least two methods appear to be gaining in usage by dietary researchers and epidemiologists. For example, a recall of food eaten on the day up to the interview was combined with a record (diary) for the remainder of the first day and the following day for a total of 2 consecutive days in the national Preschool Nutrition Survey, 1968-70, sponsored by the former Children's Bureau (Owen et al., 1974, p. 600). Patterson (1971) combined a 24-hour recall and a 2-day record in a study of Phoenix, Ariz., area children. Futrell et al. (1971) used a combination of a 3-day recall and 4-day record to cover 7 consecutive days in a study of the nutritional status of black preschool children in Mississippi.

The record (diary) with recall to complete the dietary information as needed was the method used in the fall 1965 North Central Regional Survey of Preschool Children (Eppright et al., 1972). The interviewer visited each home at least twice. On the first visit she explained how to keep the record and demonstrated sizes of servings. The second visit was made after completion of the dietary record, which was checked and then completed by the homemaker recalling apparent omissions.

Information on the food intake of individual family members was collected for the USDA survey during April, May, and June 1965 from the basic sample of the USDA household food consumption survey. After the schedules for household food data were completed, the respondent was asked to provide information on the diets in the preceding 24-hour day for each family member for whom she was able to report. For the others, schedule forms were left to be completed and mailed. The information on individual food intake was in terms of food "as eaten," e.g., pancakes, not flour and milk.

4. METHODS OF THE USDA FOOD CONSUMPTION SURVEYS IN 1965-66

This brief summary of the methodology of the household and individual food surveys in 1965-66 covers sampling, questionnaires and field procedures, and data processing and reporting.

4.1 Sampling

The sample was designed to be representative of housekeeping households of one or more persons in the United States. A household was defined as housekeeping if at least 1 person had a minimum of 10 meals from home food supplies in the preceding 7 days. The 1965-66 survey covered four seasons, with the spring sample approximately three times as large as the other three samples.

The samples for the survey were selected by the contractor for the field work, National Analysts, Inc., according to USDA specifications. The 15,112 households interviewed were scientifically selected to represent those in metropolitan areas, cities of various sizes, and rural farm and nonfarm areas in all parts of the United States except Alaska and Hawaii. Excluded were the approximately 5 percent of the population that was not housekeeping; 1.5 percent had not been visited because they were living in group quarters, such as rooming houses, hospitals, and prisons, and the other 3 to 4 percent were excluded when an initial screening indicated that they were in households in which no member ate as many as 10 meals from the home food supplies. Interviewing began in all regions on April 3, 1965, and continued until April 2, 1966.

The sample design provided in each season for a national self-weighting basic sample plus a supplementary farm sample, which overweighted the number of farm households in the proportion of 5:1. The households to be interviewed were selected according to a multistage area sample design with added control by season. For the basic sample, 106 strata were categorized by geographic area and population density, 38 of them were subdivided, and 144 first-stage units were selected at random. Then 2,000 second-stage units of 30 expected housing units, distributed systematically within the first-stage units in proportion to population, were selected.

Each of the second-stage units was visited and a list of housing units prepared. By systematic selection, housing units were chosen for interview in the spring in sufficient number to yield an average of three schedules per second-stage unit after allowing for vacancies and ineligibles or otherwise nonparticipating households. The lists were updated in the summer, in the fall, and in the winter, and sufficient numbers of households were chosen to yield an average of one schedule per second-stage unit in each of these seasons. The four seasonal samples were selected independently by using different random starts in the second-stage units.

No substitutes were provided for households unable or unwilling to participate in the survey. Interviewers were instructed to call as many as 3 times if necessary to make the original contact in rural places, 4 times in urban places, and 6 times in 281 second-stage sample units in 15 large cities where collection difficulties were anticipated.

Further information on the household samples is available from the sample analysis in the report of the U.S. Department of Agriculture (1972b, pp. 209-213).

During the 1965-66 survey of household food consumption, 92 percent of the housing units visited had eligible households. This percentage varied only slightly by season and between urban and rural nonfarm categories; 97 percent of the farm households were eligible because of much less eating away from home. Of the eligible households, data from the contractor indicate 96 percent participation, but some internal inconsistencies indicate that the actual rate may have been lower. In the spring 1955 survey, conducted by the same firm with the same sample design and comparable collection procedures, 89 percent of the eligible households provided schedules. This rate varied from 85 to 92 percent by region and was 86, 93, and 91 percent for urban, rural nonfarm, and rural farm households, respectively.

Information on the food intake of individual family members was collected in April, May, and June 1965 from the basic probability sample of the household food consumption survey, not from the supplementary farm sample nor from the sample for the other three seasons.

The sample for the survey of individuals included those persons living in the household who were related to the head by blood, marriage, or adoption and excluded roomers, boarders, household and farm help, and guests. However, the survey did include individuals living alone and households in which all members were unrelated.

Individual food reports were requested from one-half of the systematically selected persons aged 20 to 65 years. To compensate for the subsampling and to provide proper representation in the population, information for the persons in these age groups was counted twice in the tabulations. There were 14,519 schedules collected, which when weighted yielded 19,245 cases, the sum of the number of persons shown in the tables

in the report of the U.S. Department of Agriculture (1972a). The nonresponse rate was 7 percent of the eligible family members in the participating basic sample of households.

4.2 Questionnaires and Field Procedures

The schedules were put in final form by the contractor, National Analysts, Inc., based on questionnaires used in earlier surveys and the 1963 pretest in the Washington, D.C., area by members of the former Consumer and Food Economics Research Division.

The basic field procedure was a personal interview with the homemaker or another person responsible for food management in the household to obtain by the list-recall method the amounts and prices or total value of all food used in the household in the preceding 7 days on an "as purchased" basis and including food wasted, spoiled, or fed to pets, but not pet food bought as such. A detailed list of foods helped respondents recall the foods and the amounts used up during the 7 days prior to the visit. Food from all sources was recorded--home produced, gift or pay, and federally donated, as well as purchased. Information about foods used was obtained in considerable detail in order to calculate the nutritive content of diets. Special attention was given to obtaining information on whether or not the foods were enriched or fortified. Also, information on the form in which the food was obtained, that is, fresh, frozen, canned, or dried, was recorded in order to meet the needs of marketing specialists.

Related information important to evaluating household food consumption was collected, for example, the age and sex of all persons who ate from home food supplies, the number of meals eaten at home and away from home by each family member, and the income and other socioeconomic characteristics of the family. No attitudinal data were included in the survey.

After the schedules for household food data were completed, the respondent was asked to provide the individual dietary data for 24-hour periods for each family member. Schedule forms were left for individuals for whom she could not report. They were to be completed and mailed by those family members who were expected home within 3 days. About 76 percent of the schedules left were received. The information on individual food intake was in terms of food "as eaten," for example, pancakes, not flour and milk. After completing the interview, the interviewer reviewed each schedule for completeness and mailed the schedules to the central office of the contractor for editing and data processing.

4.3 Editing and Other Data Processing

The field survey contract with National Analysts, Inc., covered processing of the data through the stage of preparing detailed data tapes with household and individual food quantities plus characteristics. After editing in the central office, wide area telephoning was used to check questionable entries with the interviewers. Because of the time lapse, nonavailability of the completed forms to the interviewers, and costs, this stage of the operation was not very satisfactory after the initial weeks of the survey. Coding and some conversions to common units were done at this stage, using information supplied by the former Consumer and Food Economics Research Division (CFERD). The data were key punched. Listings were prepared and sent to CFERD for review. Corrections were to have been made before data tapes were supplied to this agency, but some were overlooked.

CFERD contracted with another private organization for calculations of nutrient content, using detailed nutrient factors supplied by the agency, and for all tabulations. The contractor prepared preliminary drafts of all tables for detailed review, corrected the tables, and submitted computer output of the final tables for publication. Eighteen statistical reports were published. Because of extraordinary delays in completing the contract, several additional reports were not published but have been made available to interested people through the National Agricultural Library.

The computers used by the data processing contractor were unique and not compatible with any other. Nor were the computers used early in the contract compatible with replacements introduced during the life of the contract. In the analytical work performed later by USDA, the original tapes from the National Analysts had to be used. They had to be extensively cleaned up for the second time. This process was not completed until October 1974.

5. CONSIDERATIONS IN EVALUATING ALTERNATIVE DIETARY SURVEY METHODS

Evaluation of alternative methods of collecting data in nationwide surveys must take into account the reliability and the validity of the measurements of food consumption obtained, the burden on respondents, costs of field work and of data processing, as well as the usefulness of the data. The following review is limited to the first four considerations and does not encompass the reliability and validity of measuring socioeconomic concepts relevant to variations in food consumption.

5.1 Reliability

Reliability is often taken to mean reproducibility or repeatability. There are two aspects to be considered. One is related to sampling and the other to ability of the respondents to provide reliable data. The latter aspect partly overlaps the accuracy aspect of validity. Within-subject variation from time to time is very difficult to control or appraise. Therefore measurement of the differences between survey procedures and instruments has encountered serious statistical problems.

The respondents' ability to recall the considerable amount of detail on quantities, prices, and money value of individual foods consumed in a day or a week has been seriously questioned because of research findings on forgetting curves by experimental psychologists. Behavioral theory of forgetting holds that ". . . forgetting is a direct function of the degree to which substitute responses are associated with the original stimuli during the retention interval. This will be recognized as a definition of retroactive interference, hence the interference theory must specify the variables which determine the direction of retroaction. . . . We may say that identity between responses in original and interpolated activities yields facilitation, whereas difference between responses yields interference (forgetting). . . ." (Osgood, 1953, pp. 550-551). Osgood identified the following variables as affecting the character of forgetting curves: Method of measuring retention, whether materials to be learned are discriminated from other information by virtue of association with other distinctive reactions, change of "set" or "warm-up" effects, reminiscence, organization of materials, method of learning, intentional learning, and the nature of interpolated activity (ibid., pp. 556-597).

Most of the experimental research by psychologists has utilized nonsense materials, not meaningful materials associated with respondent activity and interest. Therefore it is difficult to draw conclusions relevant to the problem of whether homemaker respondents can recall the amounts and prices of individual foods the household has used during the preceding time period. Obviously the higher prices for food and the familiar tasks of shopping and preparing meals provide far more interest and "set" than was found in experiments with college sophomores learning nonsense syllables. But the question remains moot until findings from much more problem-specific research are available.

The requirements for probability sampling are rarely met in the final samples because of difficulties in obtaining cooperation of respondents and the high interviewing costs. In a few studies alternative survey procedures have been used sequentially or concurrently with the same subjects or supposedly matched samples. Most of the methodological findings relevant to individual diets reported here are derived from very small numbers of selected participants.

In appraising the reliability of a survey, one should consider the degree to which the sample design, the field survey procedures and instruments, and the data processing provided the basis for sound statistical inferences. Most discussions of the reliability of large-scale dietary surveys, as for economic surveys, have been concerned with sampling errors of means.

5.2 Validity

To appraise the validity of a set of dietary survey data, one considers the relevance of the food consumption responses to the concepts being measured, which are determined by survey objectives. There is no "true" measure of consumption with which the data derived with various survey procedures can be compared. Although actual intakes of selected individuals can be precisely measured by careful weighing, such measurements are frequently not relevant to some survey objectives because they fail to meet certain types of validity tests. Five aspects of validity enter into appraisals of dietary survey methods in this report--accuracy, concurrent validity, construct validity, content validity or representativeness, and predictive or criterion-related validity.

5.2.1 Accuracy

Under this aspect of validity we consider systematic error in measuring, whether by the respondent-consumer, another family member, or the investigator. For example, it has been hypothesized that there is a systematic upward bias in the recall of food consumption over several days because of telescoping, e.g., including two weekends in a week. Also, there have been investigations of the possibility of systematically different estimates of intakes when measured and recorded in household units from those based on precisely weighed records. Systematic errors in reporting may also arise because of inability of certain groups of subjects to estimate consumption accurately, changes in consumption because of survey participation, and unwillingness to report accurately the measurements or estimates actually made.

5.2.2 Concurrent Validity

Many individual dietary studies have compared the results of simplified, less expensive survey procedures with those from generally accepted but expensive survey techniques, e.g., food intakes measured with 24-hour recalls compared with data from 7-day records. Less common have been comparisons such as protein intake with lower leg muscle development (Reed and Burke, 1954, p. 1025) and ascorbic acid intake with blood ascorbic acid concentration (Eppright et al., 1952, p. 47; Huenemann and Turner, 1942, p. 563).

5.2.3 Construct Validity

Several approaches have been used in validation of constructs in food survey research. One is investigation of the degree to which the survey procedure measures the actual variability in the concepts to be measured. For example, the food energy content of the household diets computed from the USDA household food consumption survey of 1965-66 included only that supplied by food obtained from household supplies. However, it also included food energy from food fed to pets, plate waste, and extra trimmings of fat from meats, for example. In contrast, the food energy content calculated from the individual dietary surveys ordinarily covers food obtained from all sources and actually consumed by the individual. Variations in the two sets of measures obtained from the same families are being compared in research currently underway, recognizing that they are at different levels of distribution.

Nutrients calculated from data elicited from the same subjects with different survey procedures have frequently been compared, either the averages or the variations, using correlation analysis. Variations in food and nutrition measures have been correlated with measurements of other characteristics. An example is caloric intake with physical measurements and activity. However, variations among individuals' nutrient intake during a short period bear little relationship to variations in biochemical and physical measures (Abraham et al., 1974).

In behavioral research, construct validation may involve measurement of the proportions of total test variance that can be explained by several constructs. This type of construct validation is not directly relevant to establishing the validity of dietary survey instruments in measuring food consumption, because dietary survey methods must be appraised in terms of their coverage of the total consumption of food, not the validity of consumption of certain foods as indicators. However, analysis of dietary data does include comparisons of (1) the relative importance of food sources for particular nutrients among (a) diets of varying quality or (b) different sample groups and (2) the relative importance of independent variables in explaining variations in dietary quality.

5.2.4 Content Validity

Content validation includes evaluation of the representativeness of the concept's measure, and it also refers to analysis of the interrelationships between the parts and the whole. Three different aspects are of concern in evaluating food consumption survey procedures. One is the measurement of food intake in a short time period as an indicator of intake over a longer period--for example, 24-hour diets as indicators of usual diets. Second, variations in household food surveys that cover only food at home (quantities) are frequently used as indicators of variations in total food consumption. A third area for study is the degree to which the

variations in food consumption among subgroups in the final sample represent variations in food consumption in the total population. Nonresponse rates are critical in this type of evaluation.

5.2.5 Predictive or Criterion-Related Validity

Predictive or criterion-related validity enters into the evaluation of food survey procedures in several ways. Small-scale dietary studies frequently are made in the search for shortcuts to dietary measurements, for example, measuring the quantities of selected food groups as indicators of the overall quality of diets. Variations in diets by season with level of purchasing power, family size, and composition, by sex-age category, and the like are compared with expectations based on earlier research or other types of data to determine their predictability.

Predictive validity can also be applied in evaluating survey procedures in terms of how successful they are in eliciting data required for decision-making. If the survey technique yields internally consistent measurements of variations in consumption on which can be based sound decisions relevant to nutrition education programs, administration of food aid programs, and food marketing programs, it has predictive validity.

5.3 Respondent Burden

Consideration of the burden of the survey procedure on the participants must include time required for each interview, frequency of contact, time and effort between interviews required to develop or record information, and intrafamily complications that may arise as the homemaker attempts to get information from other family members about which they may be sensitive. Examples are drinks on the way home from work and purchase of cola drinks and hotdogs for lunch instead of the school lunch.

5.4 Field Survey Costs

Such costs include sampling, development and testing of survey instruments, and training of field supervisors and interviewers; interviewer and supervisor time, travel, and telephone calls; preparation of lists for sampling; requirements for special equipment; field editing and rechecks; transmission of questionnaires to central offices; quality control; and postage for mailings.

5.5 Data Processing Costs

The procedures and instruments used in data collection materially affect the costs of processing the data. Such costs include expenditures for programming and actual coding, conversions, and quality control; programming for

data revisions on tapes and reruns; and preparation of final detailed tapes, special summary tapes, and descriptions of tape formats. Later costs of programming and making analytical runs are also affected by data collection instruments and procedures.

6. EVALUATION OF ALTERNATIVE WAYS OF MEASURING HOUSEHOLD FOOD CONSUMPTION

This section contains summaries of research findings on principal survey procedures for measuring household food consumption.

6.1 Food Accounts

The reliability of food accounts has been appraised by comparing several weeks' data from panel studies and by calculating percentage standard errors.[7/]

From an experimental study with diary methods of obtaining food expenditures in retail stores and in restaurants (but not food quantity data), Sudman and Ferber (1971) found that first and third weeks' expenditures for grocery and food items averaged the same and that the second and fourth weeks' averages were significantly lower. This suggested initial stocking up in the first week of recordkeeping and subsequent reduction in purchasing while stocks were on hand.

The report on the 1972 United Kingdom survey of household food consumption and expenditure has this important reminder about the concept being measured: "The Survey thus records the quantity of food entering the household, not the amount actually consumed; it cannot therefore provide frequency distributions of households classified according to levels of food consumption or nutrition. . . . "(Ministry of Agriculture, Fisheries and Food, National Food Survey Committee, 1974, pp. 171-172).

Regarding the reliability of survey results, the report concludes: "The results obtained from the Survey are subject to chance variations as are all estimates from sampling investigations, but this sampling error will not normally be more than two or three times the standard error.[7/] (Estimates of the percentage standard errors of the yearly national averages of purchases for selected foods, chosen from table 15 of Appendix A, are given here in table 1.) Usually the standard errors (and the percentage standard errors) of the quarterly averages will be approximately double those for the annual

---

[7/] The percentage standard error is calculated as the ratio of the standard error to the mean times 100. This ratio is called the coefficient of relative variation by some statisticians and the relative standard error by others. This relative measure must be differentiated from the coefficient of variation, which is the ratio of the standard deviation to the mean.

averages, but for some foods which have a marked seasonality the standard errors can also vary throughout the year. The estimates of the standard errors were obtained by applying the formula for a single-stage random sample and take no account of the complex nature of the sample which incorporates a multi-stage, stratified design. The reduction in sampling variance gained from stratification is almost certainly more than offset by the increase in variance caused by the use of several stages in the sample design, especially by the limited number of first-stage units; the estimated standard errors may therefore be understated in some cases."

TABLE 1.--Percentage standard errors of yearly national averages of purchases, United Kingdom, 1972[1/]

Item[2/]	Percentage standard error
	Pct
Liquid milk	0.53
Cheese	1.18
Carcass meat	1.32
Fish	1.35
Fats	.80
Fresh potatoes	1.68
Fresh green vegetables	1.32
Vegetables	1.04
Fresh fruit	1.33
Fruit	1.14
Bread	.63
Bread and cereals	.53

1/ From Appendix A, table 15, of Ministry of Agriculture, Fisheries and Food, "Household Food Consumption and Expenditure, 1972." These percentage standard errors are computed by dividing the standard error of the mean by the mean and multiplying by 100. The quotient is also called the coefficient of relative variation.

2/ Each labeled as total in the report.

Various approaches have been used to establish the validity of food accounts or record of purchases. Bavly (1972, p. 21) compared data from the Israeli survey of 1968-69 (food accounts kept for 4 weeks) with averages from food balance sheets in table 10 of her report. "From Table 10 it may be concluded that in relation to food consumption, the figures given by the balance sheet are 20 percent to 100 percent higher than those provided by the dietary survey." She traced the major differences in eggs, peanuts and oilseeds, oils, margarine, sugar, sweets, and jam to use in bakeries and processed foods and the differences in fruits and vegetables to underestimates of spoilage assumed by the balance sheets. In terms of calories, the food balance sheet estimate was a third above the dietary survey average. Among the nutrients listed, only the average for vitamin A from the dietary survey exceeded that from the food balance estimate.

Comparison of the United Kingdom's 1972 National Food Survey averages for nutrient content with averages calculated from food balance sheets, i.e., disappearance data, indicates much closer agreement for the United Kingdom's survey data based on records of purchases for 1 week than for the Israeli 4-week data. The United Kingdom data follow:

Nutrient	Survey averages[1]	Disappearance averages[2]
Food energy (kcal)	[3]2,673	3,080
Total protein (g)	79.8	85.3
Fat (g)	123	143
Calcium (mg)	1,111	1,110
Iron (mg)	14.5	15.0
Thiamin (mg)	1.39	[4]1.90
Riboflavin (mg)	1.96	1.96
Vitamin C (mg)	57	[4]100
Vitamin A (µg) (retinol equivalent)	1,340	1,360
Vitamin D (mg)	3.20	3.15

[1] Calculated from table 27, Min. of Agr., Fisheries and Food, Natl. Food Survey Committee (1974), by adding back the 10-percent reduction made for waste in homes.

[2] Ibid., p. 210.

[3] Excluding sweets and alcoholic and soft drinks consumed away from home.

[4] No allowance made for cooking losses.

In an experiment with alternative types of consumer expenditure diaries in two Illinois areas, Sudman and Ferber (1971, p. 730) compared survey averages for grocery and restaurant expenditures with estimates derived from sales taxes. The estimated expenditures in a week varied from $18.50 (fourth week) to $22.78 (third week) compared with the concurrent average of $25.71 per household based on sales tax reports. The highest survey average for restaurant expenditures per household in a week was $7.92, markedly lower than the $12.50 tax-based figure. This indicates substantial underreporting.

Effects of nonresponse in diary keeping methods on the representativeness of the sample data are indicated by the conclusions of Sudman and Ferber from the study just cited. "The differences in the characteristics of diary keepers and non-cooperators are similar to those observed in other panel studies and reflect the willingness and ability of the household to keep a written record. . . . The lowest cooperation is found among one and two member households and households in which the head is over 55 years of age, has eight years of school or less and has an income under $2,000. These are not independent factors, but reflect a cluster of variables that characterize households suspicious of surveys or that find keeping written records difficult or too time consuming." (Ibid., p. 728)

Another measure of the nonresponse problem is provided by the United Kingdom's National Food Survey in 1972. The effective sample of respondent households was 53 percent. No contact was made with 15 percent, but the remaining 32 percent either refused to cooperate initially or failed to keep the 7-day logbooks. Response was relatively lower in the London area and larger towns and higher in semirural and rural areas. (Ministry of Agriculture, Fisheries and Food, National Food Survey Committee, 1974, p. 170)

The food accounts data have been collected in the United Kingdom for many years and have been very useful in public policy determination and food market research. In recent years each survey report has included some kind of special investigation, obviously in response to special needs. Although aggregates calculated from the survey data do not match aggregates calculated from the disappearance data, the regular users of the data are accustomed to making the necessary adjustments.[8/]

The effective sample of 2,431 families in Israel for the food survey based on 4-week records was about 70 percent of the original sample (Bavly, 1972, pp. 75-77).

No information on respondent burden has been found in the reports on the food consumption surveys in Great Britain or Israel. However, the procedure as described requires daily recording by the family of food purchases and supplies brought into the home from other sources.

Field survey costs are not reported, but a few indications are given. The British survey procedure incorporates a minimum of three visits to each household plus additional visits where needed. The Israeli survey required a visit about every 3 days. (Ibid., p. 75)

Data processing costs have not been evaluated in the British and Israeli reports. However, the effective sample sizes (7,578 and 2,431, respectively) provide some indication. Also, the times for processing data in several stages for the Cincinnati study, reported in section 6.2, are relevant.

8/ Conclusions drawn from the senior author's personal investigation of the National Food Survey on site in 1956 for the Administrator of the U.S. Agricultural Marketing Service and from several subsequent conversations with Ministry officials.

Pekkarinen (1970, pp. 152-153) of the University of Helsinki summed up the advantages and disadvantages of the food accounts method in a methodological article:

"The food accounts method has both advantages and disadvantages. The major advantage is the large sample that can be obtained by it; it is also possible to collect food consumption data from a relatively long period of time as compared to many other methods. Since the data are, in addition, collected during different months of the year, as is the current practice in many countries, the method provides excellent information on the annual mean food consumption and the general food patterns and habits of the population. The method does not alter the diet of the families under survey, not to any large extent anyway, since the recording is not as detailed and time-consuming as in some other methods; these might involve e.g. a tendency to simplify the diet in order to reduce the amount of work caused by the recording. The collection of food consumption data is relatively cheap, since trained personnel is needed in the field work only for advising and controlling the recording.

"Of the disadvantages may be mentioned first of all that the families or households finally included in the survey are not always a representative sample of the whole population, although they would initially have been chosen by statistic methods. It has been said (TRULSON, 1959) that the method can be used among literate people, but in spite of being able to read and write, all people are still not able or willing to keep records. The extra work caused by the recording of consumption data may alone restrict the participation of some of the selected families. The families that finally participate are above average in many cases, which should be remembered when presenting the results of the survey.

"The data obtained by the food accounts method are not always very accurate. In some cases the respondents may forget to record the amounts of food consumed, and later it is difficult to recall exact amounts. The use of household measures for estimating the amounts may also cause inaccuracy, likewise when the amounts eaten must be assessed from the price of the item. The records do not always reveal how much food was e.g. given to animals, thrown away because of spoiled, or discarded as plate waste or for other reasons. The proportion of edible food discarded may be considerable; it is usually higher in well-off than in poor families. By this method it is thus impossible to estimate the amounts of food actually consumed by the family. When evaluating the results a reduction has to be made for discarded food. An often used estimation is 10% of the total energy of the diet. This assumption is supported by surveys made in Finland on the amounts of discarded food. In rural households it was found to be on an average of 10% of the total energy, in hospitals 7 to 13%."

6.2 Direct Comparison of Inventory-Record and List-Recall Procedures

Several studies comparing the inventory-record and list-recall procedures for surveying household consumption have been conducted by or in cooperation with the U.S. Department of Agriculture since 1936 (Murray, 1970). The latest, a 1969-70 study in Cincinnati, Ohio, provides the best basis for appraisal of the two methods (Grossman and Popka, 1974). The survey utilized 3 subsamples of 300 households each, selected so that they gave equal representation to 3 economic levels, based on low, medium, and high value of housing. Three interview patterns were used in two waves of interviewing: List-recall, list-recall with a memory aid, and inventory-record. From the first wave in October-November 1969, 544 completed schedules were obtained. These households supplying completed schedules were contacted on the second wave in January-February 1970 for interviews with alternative procedures, yielding 433 completed schedules. The nonresponse rate for interviews with the inventory-record procedure was much higher in both waves than for either of the list-recall procedures.

The food quantity, money value, and nutrient data obtained from the three procedures in two waves have been analyzed (1) without adjustment for differences in response rates, (2) for identical families, and (3) as adjusted or standardized by matching samples in terms of seven characteristics. All three analyses yielded substantially the same conclusion. The quantity of foods, their money value, and their nutritive values obtained with the inventory-record procedure averaged roughly 20 percent less than those derived from the two list-recall procedures. The last two sets of data were so similar that they were pooled for subsequent analyses. Much of the differential was contributed by the meat, dairy, and "other foods" categories. "Other foods" included alcoholic beverages, which were high in the list-recall data because New Year's Eve was covered, an error in interview scheduling.

Variance measures relevant to comparison of the reliability of the means have been derived from the quantities of major food groups obtained with the list-recall and inventory-record procedures. The standard errors for the standardized samples are similar, but the smaller means for several food groups reported by the inventory-record sample result in somewhat higher relationships for the standard errors to the means. All standard errors derived from this small sample survey are much larger than those based on the spring 1965 data for nonfarm households in the North Central Region in table 3. The percentage standard errors calculated from the standardized sets of quantity data for food groups obtained by means of the list-recall procedure varied from 14 to 20 percent, whereas those from the inventory-record procedure ranged between 18 and 28 percent.

For identical families, the percentage standard errors, based on the same quantity measures derived from the two survey procedures, are much smaller (5-12 percent) and generally similar. Accordingly we may conclude that replication of the surveys would yield food consumption data averaging close to the means reported.

The Cincinnati study has provided much information relevant to the validity of data obtained with the two survey methods. In using the inventory-record procedure, the interviewers and respondents encountered problems in supplying the desired information in the form of measuring the commodity content of mixed dishes, weighing foods excluding containers, and evaluating edible remainders of partly used foods like hams. There was evidence of changes in consumption to avoid the measurement task. Interviewers reported that some homemakers failed to keep complete records and tried to fill out the forms by recall during the final interview.

The ability of respondents to recall information on household use of all individual foods in the preceding week is frequently questioned. Several findings from the Cincinnati study are relevant. The prices per pound of all major food groups calculated from data obtained by inventory-record and list-recall were within 10 percent. The fact that the number of items consumed in a week averaged 43 for list-recall and 41 for inventory-record indicated little telescoping. Major shopping was done only once a week in 63 percent of the Cincinnati households surveyed, whereas 27 percent shopped once or twice a month. The homemakers did 90 percent of the major shopping and made 65 percent of minor trips. Nonmajor trips averaged 1.3 per household per week. The use of a memory aid in the form of preliminary recall of menus for each day of the preceding week had little effect on the amounts recalled, according to comparison of data for two subsamples. But the memory aid procedure annoyed many respondents, who objected to repeating the procedure of recalling their week's food activities in different forms.

The next aspect of validity to be considered is the possibility of conditioning food consumption behavior, which introduces systematic errors. Information from the inventory-record procedure of the Cincinnati study indicated some changes in homemakers' food practices. The number of different items in the household inventories at the end of the inventory-record periods averaged 51.5 compared with 61.8 at the beginning. (However, this comparison may underestimate changes. A decrease in the number of cans of peas on the shelf from the beginning of the week, for example, three cans to one can at the end, would not change the item count of one for canned peas.) The decrease in the item account amounted to 25 percent for milk, meat, and other protein food; 20 percent for enriched or whole grain; 13-15 percent for vegetables, fruit, other grain, sugar, and miscellaneous foods; and 10 percent for fats. The decrease was 20 percent in the low rental value sample and 15 percent in the middle and high rental samples. In the beginning inventories, the content of several mixed dishes appeared to be a nuisance for the respondents to estimate. No mixed dish was found in the closing inventories. Interviewers reported that respondents had said they used up food on hand or discarded leftovers before the closing inventory was to be taken.

Several authors of survey reports have noted that the week for record-keeping affords homemakers the opportunity to consider the desirability of reporting certain foods, whereas the momentum and time pressure of the interview using the list-recall procedure provide little chance for such consideration.

Perhaps the greatest concern of survey specialists with record-type procedures is the representativeness of the measures because of higher nonresponse rates. Nonresponse takes several forms--initial refusals to participate, failure to keep records as arranged, and incompleteness of reporting on key items. In the Cincinnati study, 42 percent of the households identified for the inventory-record subsample were reported as not having been contacted by the interviewers in three visits compared with 24 percent for the two recall subsamples. This raises the possibility of interviewer bias against the inventory-record schedule. In addition to this initial sample loss, 27 percent of the eligible households refused to participate in the inventory-record. This figure is to be compared with 3 percent for the list-recall, indicating that respondents also were less willing to participate in an inventory-record procedure. The first wave of visits resulted in schedules from 70 percent of the list-recall samples and from 41 percent of the inventory-record sample. No effort was made to contact households not reached in three visits nor to gain cooperation of eligibles not willing or able to participate.

Households participating in the first wave of interviews were visited again in January-February. This time the 214 households that had provided list-recall data with the memory aid were to use the inventory-record procedure; the 122 inventory-record households were asked for list-recalls with memory aid; and the 209 list-recall households were asked only questions on related food practices. Whereas 86 percent of the households asked to provide list-recall schedules participated, only 66 percent of households identified for the inventory-record in the second wave did so. The 34 percent not participating included 8 percent reported as having moved, 19 percent as not contacted, and 7 percent as refusals.

The sample loss for the inventory-record procedure resulted in the participating group having proportionately more of the largest households and fewer home owners and elderly homemakers than the list-recall group.

Evaluation of the validity of responses also entails consideration of the precision with which characteristics of the foods consumed are reported. The Cincinnati survey apparently encountered fewer problems when the homemakers recorded purchases of concentrated foods like frozen orange juice than was the experience with list-recall.

Averages of the responses from all households obtained with the two procedures have been compared. Other information from the study is relevant to the predictive validity of the data. The two sets of responses from one-person households were close, whereas those from the largest households were far apart. Regression analysis of the two sets of data indicates that the relationships between food consumption (measured in

several ways) and socioeconomic factors are very similar. The only significant difference is the relatively greater importance of household size in the variations in money value of food consumed estimated from the list-recall data than for the inventory-record data.

The Cincinnati study provides only part of the information needed to measure respondent burden. The interviewers noted the time required to take the inventories and for interviewing during their two visits. But no data were sought on time spent by respondents in recording information in the record books. An average of 3 hours was required in the first wave for the beginning inventory and 2 hours for the ending inventory. Time spent by the interviewers with respondents on other parts of the inventory-records averaged 45 minutes. In addition to the two personal interviews, the interviewers contacted the respondents at midweek by telephone to remind them of their recordkeeping. For comparison, the single interviews based on the list-recall procedure without the so-called memory aid averaged 2 hours, whereas those with the memory aid took 2-1/2 hours.

These time requirements and the two visits plus one telephone call for the inventory-records versus one visit for the list-recalls indicate the substantial differences in field survey costs. The list-recall procedure is a much cheaper way to collect data.

The costs of processing data from the inventory-record were also much higher. The key facts are that 12 hours were required, on the average, to prepare the inventory-record data for computer processing compared with 5-1/4 hours per list-recall schedule. Also, there were 3.7 times as many editing and processing errors per set of inventory-record schedules as per list-recall schedule.

6.3 Other Evaluations of the List-Recall Procedure

Several sets of measures of percentage standard errors have been developed from the USDA survey data on household food consumption. Percentage standard errors calculated from unpublished spring 1955 survey data for nonfarm households are given on both the per household and per person basis in table 2. Those calculated from per person data are fractionally lower than those per household. Comparable household data for all U.S. households in spring 1965 are also given in the table. These indications of reliability based on the U.S. spring surveys using list-recall procedures vary little more than one point from the percentage standard errors for the British 1972 food accounts data (table 1). As noted in the discussion of the yearly data for Great Britain, the British statisticians found the quarterly percentages to be about two times the yearly figures.

TABLE 2.--Percentage standard errors for quantities of food consumed in United States, spring 1955 and 1965[1/]

Food group	Spring 1955, nonfarm households, based on averages		Spring 1965, all households, based on averages per household
	Per household	Per person[2/]	
	Pct	Pct	Pct
Meats, poultry, fish	1.0	1.0	---
Meats	---	---	0.9
Poultry and fish	---	---	1.2
Fats	1.1	.9	1.1
Milk products excluding butter	1.1	.9	1.3
Eggs	1.1	1.3	1.1
Grain products	1.3	1.0	---
Flour and cereals	---	---	1.5
Bakery products	---	---	1.0
Sugars and sweets	1.3	1.1	1.2
Fruits and vegetables:			
Potatoes	1.4	1.2	1.3
Citrus and tomatoes	1.4	.9	---
Sweetpotatoes, dark-green and yellow vegetables	1.9	1.9	---
Dry beans and peas	2.5	2.2	---
Other vegetables and fruits	1.1	1.3	---
Fresh vegetables	---	---	1.1
Fresh fruits	---	---	1.4
Canned fruits and vegetables	---	---	1.3
Frozen fruits and vegetables	---	---	2.4
Fruit and vegetable juices	---	---	1.6
Dried fruits and vegetables	---	---	2.9

1/ Based on unpublished data in files of the Survey Stat. Group, Consumer and Food Econ. Inst., ARS. Data not currently available to calculate percentages for precisely matching food groups.

2/ Using number of persons calculated in terms of 21 meals at home per person.

Data on food consumption of low-income households are of considerable scientific and administrative importance. Therefore percentage standard errors for major foods consumed by nonfarm North Central Region and southern households classified as "poor" and "near poor" are compared in table 3 with those for all nonfarm households in each of these regions.[9/] The percentages for all nonfarm households in the North Central Region are between 2 and 2.5 (N=1,594) compared with 1-1.5 for all U.S. households (weighted N=6,174). The differences appear to arise from sample size because the means and standard deviations are rather close. The percentage standard errors for all nonfarm households in the two regions, reported in table 3, are similar. Except for tomatoes, percentages for 529 poor and near poor households in the North Central Region average about 3.5-4.0, a little less than twice the percentage errors for all 1,594 of the nonfarm households. In the South where the 935 poor and near poor households made up about half of the total 1,818 nonfarm households, their percentage standard errors for major categories were only about 1 percent higher than those for all nonfarm households.

Several types of research data are focused next on the validity of food consumption data obtained from surveys utilizing list-recall procedures. Researchers, such as Janet Murray, formerly of USDA, and Charlotte Young, of Cornell University, who are experienced in surveying homemakers to elicit food data, insist that most homemakers are able to recall household food use between two time boundaries if they are related to the cycle of food activity. The Cincinnati study revealed that only one major shopping trip per week was made by 63 percent of the homemakers. Mason and Madden (1971) found that 55 percent of low-income, black, inner-city homemakers made only one major shopping trip in a week. Accordingly a 7-day period appears to encompass a natural cycle of food activity.

In using the list-recall method, interviewers begin by asking the homemakers about major categories of foods used in the preceding 7 days. Then they use the list of major items within each category to probe for information on individual items, both on the list and in addition. The list helps the interviewers check on details about the characteristics of each food, such as canned or fresh, enrichment, or fortification with vitamins. Such details are very essential to calculating nutrient content because one form, variety, or brand may contain several times as much of a particular nutrient as another. Review of survey data has indicated that 60 percent of the North Central urban homemakers surveyed in spring 1965 were able to recall between 33 and 69 different items. The minimum number of items used in a week was 6, the maximum 139, and the median 50.

Research by Cannell and others at the Michigan Survey Research Center underlines the importance of two interviewer-related phenomena in collecting dietary data with the list-recall procedure (Lansing and Morgan, 1971, pp. 114 ff. and 150 ff.). In the list-recall interview the interviewer can evaluate the responses she receives and probe for additional information,

---

9/ The criteria for categorizing households as poor and near poor were income, family size, and urbanization. For example, four-person urban and rural nonfarm households with less than $3,000 after tax income were identified as poor and those with $3,000-$5,000 incomes as near poor.

TABLE 3.--Means and reliability measured for quantity of selected food groups consumed in a week in spring 1965 by all nonfarm households and by those categorized as poor and near poor, North Central Region and South[1/]

Food group	North Central Region						South					
	All households (N=1,594)			Poor and near poor (N=529)			All households (N=1,818)			Poor and near poor (N=935)		
	Mean	Standard deviation	Percentage standard error	Mean	Standard deviation	Percentage standard error	Mean	Standard deviation	Percentage standard error	Mean	Standard deviation	Percentage standard error
	Lb	Lb	Pct	Lb	Lb	Pct	Lb	Lb	Pct	Lb	Lb	Pct
Meat, poultry, fish	15.5	10.1	1.6	14.3	11.4	3.5	14.4	10.0	1.6	13.3	10.3	2.5
Milk products, excluding butter, on calcium-equivalent basis	29.8	21.6	1.8	27.9	23.6	3.7	25.8	18.5	1.7	24.6	19.5	2.6
Eggs	2.6	2.0	2.0	2.6	2.2	3.7	2.8	2.2	1.8	2.8	2.4	2.7
Fats and oils	2.6	1.9	1.9	2.5	2.2	3.8	3.0	2.2	1.7	3.1	2.6	2.7
Flour, cereals, bakery products	7.9	6.3	2.0	8.8	7.8	3.9	9.6	8.0	2.0	11.0	9.8	2.9
Potatoes	7.3	6.6	2.2	7.2	7.1	4.3	5.1	5.1	2.3	4.7	5.4	3.8
Tomatoes	2.5	2.7	2.7	2.2	2.8	5.6	2.5	2.8	2.7	2.0	2.8	4.6
Vegetables excluding potatoes, tomatoes, dark green and yellow	7.8	5.4	1.7	7.1	5.7	3.5	8.1	6.1	1.8	6.9	6.0	2.8
Fruits excluding citrus and others high in vitamin C	7.5	8.1	2.7	6.7	8.0	5.2	6.3	8.4	3.2	5.1	7.6	4.9
Sugars and sweets	3.4	3.2	2.3	3.8	3.6	4.2	4.1	3.5	2.0	4.4	3.9	2.9

[1/] Family size and household size in number of 21 meal persons were approximately the same for all households and for those categorized as poor and near poor in both regions. See text footnote 9 for criteria for these categories. Poor and near poor households were approximately 4 percent larger than all nonfarm households in both regions.

including joint examination of labels on foods on the cupboard shelves. The quality and completeness of the information are often improved by the responses of the interviewer to the respondent that reward her for supplying detailed information and obvious attempts to be accurate.

Responses obtained with the list-recall method have been compared with inventory-record data (e.g., the Cincinnati study previously discussed), with a consumer panel, and with food balance or disappearance data. One comparison related the spring 1942 all-U.S. household survey data to data for comparable food items collected by means of diaries of purchases kept by the Consumer Panel of the U.S. Office of Price Administration in 1943. This survey program was administered by the U.S. Bureau of the Census. The conclusion drawn from the comparison was that the differing collection methods rather than real declines accounted for the markedly lower figures from the panel in 1943 compared with the spring 1942 averages. The differences shown for rationed foods and items known to be in short supply in 1943 were no greater than those for other food items. (Murray, 1970, p. 23)

The following statements about the 1955 survey data are taken from an article by Burk and Lanahan (1958, pp. 89-90):

"How do the estimates of food consumption derived from the 1955 Survey of Household Food Consumption check with other measures? Some critics of one-time surveys argue that surveys of this kind yield gross overestimates. Because such survey data provide the principal basis for analysis of the cross-section of our national food market in terms of its buyers, they would be useful for many purposes even if their levels were out of line.

"We have carried through a variety of checks on the overall dollar figures, on overall measures of per capita food consumption, and on quantities of major foods consumed. Before going into the findings, these facts need emphasis: A range of error is to be expected in these survey data as well as in the aggregate figures for food expenditures and food disappearance. Neither set of data proves or disproves the validity or accuracy of the other.

"In brief, these are our findings to date:

"1. The survey data on market value of all farm food commodities consumed, adjusted to United States aggregates for the year, are 5 or 6 percent higher than our estimates of the market value of all farm foods and meals consumed by the civilian population. About half of the difference arises from the disparity between the amount of home food production as estimated for the disappearance data and that reported by housekeeping households, both for a week of spring 1955 and for the year 1954.

"2. A comparable degree of difference was found between the overall level of use per person of farm food commodities by the sample of housekeeping households in a week of spring 1955 and the level indicated by the

index of per capita use of farm foods in the year 1955. Again, about half of the difference arose from the estimation of home production. The small discrepancy remaining seems to indicate that seasonal variations for individual foods balance out in the total for all foods.

"3. Among commodities, there is wider variation between averages computed from survey data for the housekeeping population's use of food at home and those derived from disappearance data. Average use of sugar at home in all forms, adjusted to a yearly total from the survey data, was much lower than average annual per capita consumption. But use at home excludes all the candy, soft drinks, and desserts consumed away from home.

"At the other extreme, survey data on eggs appear to average substantially higher than AMS estimates of per capita consumption. The procedure by which equivalent persons are calculated apparently leads to upward or downward bias for foods consumed primarily at one meal of the day. When allowance is made for seasonal variations in food consumption, the survey data for meats and for fats and oils were found to be close to the levels indicated by annual per capita consumption data. Study of data for other commodities is still in progress."

LeBovit of the Economic Research Service (ERS) has matched detailed data for 1965-66 from the household food consumption survey with ERS disappearance data for major food groups. The following tabulation provides data for all food groups for which the gross survey averages differed by more than 5 percent from the disappearance data:

Food group	Gross difference (survey/disappearance)	Home production in survey as percent of total consumption	Overestimate arising from household size calculation
	Pct	Pct	Pct
Meat	+12	5	5-10
Poultry	+16	7	N.A.
Eggs	+8	8	5-10
Fruits (fresh equivalent)	+51	(Fresh, 8)	(Citrus juice, 5)
Melons	+12	18	N.A.
Potatoes, sweetpotatoes	-8	(1/)	N.A.
Sugars and sweets	-27	---	N.A.
Coffee, tea, cocoa	+13	---	N.A.

1/ Farm garden output included in disappearance data.

Differences between the survey and disappearance data have been traced to several factors. The major factor is known exclusion or underestimated production for own home use in the production data used in the food balance sheets, particularly for meats, poultry, eggs, fruits, and melons. Commercial estimates of apple production are limited by statute to specified commercial areas. Therefore some commercial production and all production for home use are omitted from the disappearance figures. The fruit comparison is complicated further by consumer confusion between canned fruit juices and the fruit ades, which contain little fruit juice, and by confusion between home-preserved fruits and preserves processed from purchased fruits and those from home-produced fruits.

The popularity of french-fried potatoes, candy, and soft drinks as away-from-home snacks contributes to the negative difference between per capita disappearance into all distribution channels and survey data on food consumption at home. Also measurement of home use of coffee, tea, and cocoa presents serious problems.

The percentages in the last column of the tabulation indicate the degree to which the conventional estimates of household size lead to overestimation of per person averages when divided into the household survey averages. Recall first that the household data pertain only to use of food at home. Because family members often consume some meals away from home during the survey week and guests may eat meals in the home, the number of persons eating out of the household food supply has to be adjusted. In brief, the adjustment process includes counting the number of main meals, not snacks, eaten at home in the 7 days and dividing by 21. Several critical assumptions underlie this procedure. They have been carefully investigated and the effects measured for several foods, using spring 1965 individual food intakes. (Burk and Ward, 1973) The estimates in the last column are based on this research.

One aspect of the predictive validity of food consumption survey data can be evaluated by examining the consistency in food and nutrition levels, patterns, and relationships to socioeconomic factors derived from the survey data. By consistency one may refer to consistency between surveys taken at different times or between population groups within the sample of one survey. Both types are considered here.

The text of the first report on the 1965-66 household food consumption survey compared the average money value of all food consumed at home per U.S. household in spring 1965 with that for spring 1955. The 15-percent increase was close to the 13-percent increase in the Bureau of Labor Statistics retail price index for food at home in urban areas. During this 10-year period a higher proportion of U.S. households encountered urban prices because of the off-farm movement of the population. The increase in expenditures per household for meals and snacks away from home approximated the increase in prices of such foods. The civilian per capita consumption index, based on ERS disappearance data and constant retail prices, was practically the same in 1965 as in 1955.

Furthermore, the proportion of food dollars allocated to major food groups changed very little from spring 1955 to spring 1965. Intensive analysis of changes in spring 1955 and spring 1965 data for dairy products by households grouped by region, urbanization, and income indicated that they were consistent with changes in per capita consumption rates derived from disappearance data. (Burk, 1969a, p. 38)

Numerous regression analyses of spring 1942, spring 1955, and spring 1965 household food consumption survey data for major food groups have revealed internal consistency in the relationships of money value and consumption data to income, age, composition, urbanization shifts, and consumption of other food groups (Burk, 1958; 1969b, chap. 4, 10, 12; George and King, 1971). Such analyses have been used extensively in food demand analyses, in other types of food market research, and in decision-making on agricultural and food policies and programs.

The only available information on respondent burden of the 1965-66 household food consumption survey is that similar list-recall household questionnaires used in a 1963 Washington, D.C., pilot study and in the Cincinnati study required about 2 hours per interview. Administration of the individual questionnaire in the Washington pilot study averaged about 30 minutes, whereas the family characteristics and household list-recall schedules required a little less than 1-1/2 hours.

Based on the contract with National Analysts, Inc., the average costs for field surveys and data processing to the detailed data tape stage for the 1965-66 surveys have been estimated at $40 per set of household schedules and $10 per individual schedule. The cost of tabulating about 20 statistical reports from the 1965-66 data tapes by the General Electric Company has been apportioned among household and individual cases. The resulting estimates are $15.30 per household and $6.30 per individual case.

6.4 Record of Food Use After Weighing

The procedure of recording food use by weighing, either by an investigator or by family members, has been used for many years with several variations. In some rural, less developed areas with relatively few and mostly staple foods in the diet, the supplies of individual food commodities to be prepared during the day may be weighed each morning by an investigator, or the investigator may visit the home at any time in the day to check on the homemaker's weighing of foods brought into the house since the preceding day with or without taking a beginning and ending inventory, as was done in the U.S. wage-earner study of 1934-37. Some epidemiological studies concerned with actual intake of all or several family members have required weighing of foods to be prepared for each meal, weighing of individual portions, and weighing of discards. This procedure probably provides the most accurate measurement of actual family intake, but serious disadvantages are reported in the literature.

The size of the sample is usually small because of the time and skill required of investigators, e.g., 24 families in each of 2 areas of Finland. Pekkarinen (1970, p. 155) wrote, "It is seldom representative. Random sampling cannot be employed, and volunteers are selected. The results thus obtained cannot be generalized." She concluded, "Despite its many limitations the weighing method has been used in a large number of surveys, especially in such surveys where the accuracy of the data (i.e., physical measurement of foods ingested) is more important than the size of the sample. For this reason it is restricted to relatively small surveys. . . ." (Ibid, p. 158)

The validity of data obtained by this method as measures of actual consumption patterns of households is questionable on other grounds. Most reports on studies involving weighing procedures remark on the problem of changes in food patterns because of the burden of measuring (Pekkarinen, 1970, p. 154). The usual reaction of homemakers being surveyed is to simplify the diet, although there is some evidence of attempts to impress the investigator with particular dietary practices. However, the sample designs and sizes have not provided definitive evidence of the latter bias in particular. Weighing of meals and snacks eaten away from home is rarely possible. Furthermore, the sample losses are so large that the representativeness of the sample is always problematical.

Use of the weighing method in surveys is very expensive because of the frequency of visits by interviewers for supervision and control (Pekkarinen, 1970, p. 155). In surveys conducted in Finland and Italy only two families could be handled by one investigator in a week, but as many as five could be handled if the family members did their own weighing.

The weighing method has definite advantages for the surveys in areas where shopping or other procurement of food is frequent; where foods are not distributed in labeled, consumer packages; where units for recipes are in weights, as in England, not cups or other volume measures, as in the United States, or where measuring devices are rarely used in food preparation, as in rural areas of less developed countries; and where labor costs are low.

6.5 Other Research Relevant to Household Food Consumption Surveys

Several other approaches to surveying household food consumption have been used or suggested, e.g., daily telephone interviews, consumer panels, diary keeping of individual household members, and split-schedules.

The possibility of using daily telephone interviews to obtain food expenditure data was explored by Sudman and Ferber (1974) in an experiment conducted in the Chicago area that compared the responses with diary data. There were substantial differences in the number of food purchases in total and for commodity groups. Although the difference between the averages for

total food expenditures reported by telephone and in the diaries was not statistically significant, the expenditures for major commodity groups reported by telephone were significantly lower. The investigators concluded that the data obtained by telephone were not as complete or as accurate as those from diaries.

Consumer panels are used extensively in food market research. The Market Research Corporation of America (MRCA) has used its National Consumer Panel for many years to obtain information on consumer purchases of selected manufactured food products, not fresh foods like meats, fruits, vegetables, and intermittently to obtain menus, i.e., lists of dishes and their content, of meals eaten by individuals in the households. The U.S. Department of Agriculture purchased data from MRCA for several years after World War II on the purchases of selected dairy products and selected citrus products. Its statisticians observed the great difficulty in maintaining a probability sample (despite compensation for respondents) for the panel survey because of the very high rate of refusals (close to 50 percent) to participate even at the beginning and the gradual increase in dropouts, particularly of low-income, less educated families. Furthermore, the USDA concluded that the costs were relatively high for the quality of the data obtained for research purposes.

Much of the literature on consumer panels deals with response errors and emphasizes the greater accuracy in supposedly daily recording of purchases of a limited list of branded items compared to recall (Sudman, 1964). Large food manufacturing firms have found the data on trends in volume, shares of the market, and other brand information to be very useful.

Two single-city consumer panels have been sponsored by universities to obtain food purchase data--the Lansing panel of the Michigan State University (1951-58) and the Atlanta panel of the University of Georgia (1958-62). Quackenbush and Shaffer (1960) evaluated the data from the Lansing panel in various ways, including special surveys run concurrently.

In one study, Quackenbush and Shaffer compared expenditures (1) recorded by the panel, (2) recalled by a special sample of families, and (3) recorded by a subsample of the special sample, using the regular panel's questionnaire. An example of the problem of sample losses is the fact that out of 300 families asked to keep a copy of the diary for 1 week and to mail it in to the university, only 195 actually returned the diaries. These researchers found evidence that the recall procedure resulted in the reporting of nearly 20 percent greater expenditures than the panel. When they compared the reports of the 195 families who were interviewed to obtain recalls and who kept diaries, they found that the products that varied the most were the cooking aids and the vitamins, beverages, meats, poultry, and fish. Those that varied the least were bakery and cereal products, fruits, and vegetables.

From further study of the differences between the recall interview and the panel method, Quackenbush and Shaffer suggested that at least a partial explanation of the difference in expenditures is the higher percentage of families recalling the purchase of the products. Based on analysis of the

particular products recalled as bought by a higher proportion of families, these researchers concluded that there may very well have been a problem of "telescoping." This term refers to unconscious extending of the time period backward or forward. From their research Quackenbush and Shaffer concluded:

"The analysis indicates that the two methods of collecting purchase data gave widely different results for some products, some groups of products, and all food at home. The analysis does not prove which gives the most accurate results. Panel members probably underreported and they may have been conditioned in their buying. . . ." (Ibid., p. 39)

In an experimental study sponsored by the U.S. Bureau of Labor Statistics on collecting consumer expenditures, Sudman and Ferber (1971, p. 726) attempted to get individual family members to keep diaries of expenditures, hoping to derive household totals by aggregating. They concluded that their effort was a "conspicuous failure." ". . . Repeatedly, interviewers reported that individual household members were willing to report purchases to the recordkeeper, but would not bother to keep individual diaries. . . ."

Another suggestion for improving the reporting on household food surveys has been to use split-schedules. A 1945 survey of rural levels of living in two Mississippi counties experimented with split-schedules (Reagan and Grossman, 1951).

Janet Murray appraised the results as follows:

"In the appraisal of the split-schedule technique consideration was given to costs, quality of the data obtained, and problems of analyses. It was found that the split-schedule was a relatively expensive procedure, since it required a larger sample and increased travel and supervisory costs. Interview time for the sum of the splits was about fifty percent greater than for the complete schedule. However, this did mean that the time with a family in one of the splits was about half of the time taken for the complete schedule (three hours). The question of whether response rate would therefore be increased by adopting this procedure was not really tested--the refusal rate was very low--only 2 percent--and those who did not cooperate did so before knowing the length of the interview. It might be argued, however, that better data would be obtained in the shorter interview. There was a tendency for the average expenditure for the food, clothing and housing categories to be somewhat higher on the splits than on the complete schedules, but the analysts thought that differences in family characteristics rather than the methods could have accounted for such differences.

"This experience provides no evidence to suggest that the split-schedule procedure would be fruitful in solving problems of collection methods in food consumption surveys--particularly as analytical problems would be incurred." (Murray, 1970, pp. 29-30) The analytical problems of matching

households' socioeconomic characteristics and their food preference patterns in order to measure various aspects of the total diet appear to be insurmountable.

Researchers at the Michigan Survey Research Center routinely use advance letters to obtain cooperation. They were used successfully in the University of Minnesota survey of upper income families and in the Cincinnati study previously described. However, the only experimental studies of their effects reported in the professional literature have pertained to mail surveys and are not definitive for personal interview surveys. (Ford, 1967)

7. SURVEY ALTERNATIVES IN MEASURING INDIVIDUAL FOOD CONSUMPTION

Differences between the objectives of dietary surveys and of nutritional status surveys affect the conceptualization and measurement of food intakes. Dietary surveys are designed to obtain information about current food consumption behavior of individuals with varying socioeconomic and demographic characteristics; they measure actual food intakes during a specified current period. Nutritional status surveys are concerned with interrelating information on diets and biochemical, clinical, or both kinds of data relevant to health. To assess factors related to health, measurements of usual dietary intake over longer periods are often needed. Because of cost and response problems in obtaining measurements of diets over extended periods in addition to the cost of measuring health indicators, large-scale surveys of nutritional status frequently use 24-hour recalls or 2- or 3-day recall-record combinations (Abraham et al., 1974; Owen et al., 1974).

The procedures and problems in design, collection, analysis, processing, and interpretation of dietary survey data for individuals have been discussed in numerous reviews (Becker et al., 1960; Christakis, 1973; Leitch and Aitken, 1950; Marr, 1971; Trulson, 1960; Young and Trulson, 1960) and manuals or guides (Bigwood, 1939; National Research Council, 1949; Norris, 1949).

This review includes (1) survey procedure, (2) sample, (3) food measurement technique, (4) how many days to include, (5) which days to include, (6) interviewer influence, (7) who should be the respondent, and (8) ability of respondent to give accurate information.

The principal survey procedures for measuring food intake of individuals to be evaluated here in terms of reliability and validity include (1) records of current food intake--(a) the weighed record method and (b) the estimated record method--often referred to as the household measure record method, (2) recalls of past food intake--(a) 24-hour recall method, (b) dietary history method, and (c) frequency method, and (3) a combination of the two principal methods--combined recall-record method.

7.1 Record of Current Food Intake: Weighed Record Method

A weighed record is a listing of all foods eaten by an individual in a specified period with the amounts of each, determined by weighing and usually entered in grams. This method is less commonly used in the United States than in Europe, where scales are more frequently utilized in food preparation.

7.1.1 Reliability

Reliability of data is often evaluated by reference to standard errors of the means. Unless a probability sample is available, the theoretical basis for sampling errors does not exist and statistical inferences cannot properly be made. However, some statisticians argue that complex surveys, yielding probability samples but not simple random samples, may result in serious underestimates of standard error when the usual formula ($S/\sqrt{n}$) is used (Petrasovits and Nargundkar, 1974). No satisfactory probability samples using weighed and analyzed records were found in the literature. Reasons given for failure to use probability samples with the weighed record were the great cost in time and money, need for literate and highly cooperative subjects, and the small number of persons who could be handled (Marr, 1971; Mongeau, 1974).

In one attempt to gather a probability sample of weighed records as well as aliquots or samples of the foods for chemical analysis, the effort to maintain a random sample failed, according to the investigators, because 10 percent of the 54 randomly selected men were unwilling to keep weighed records for 7 days (den Hartog et al., 1965). These 54 men were a subsample of a large cardiovascular study (N=1,088), which already had a nonparticipation rate of 17 percent.

However, there are well-known studies with carefully weighed records covering 4-10 days where inferences regarding reliability were drawn without a sound basis because of the failure to follow adequate sampling methods. Among the survey errors made were the use of volunteers, choices of subjects by the researchers as typical of the population, and, because of their cooperativeness, the combining of studies utilizing different methods of dietary data collection, and interrupting the usual food behavior of the subjects by requiring a basal metabolism and chemical tests during the period covered by the collection of dietary data (Burrill et al., 1959; Leverton and Pazur, 1957). The presentation of standard errors of the means for such data is deceptive.

Comparison of consumption in two time periods presumably is another way to appraise reliability or repeatability of the weighed record in the sense that repeated measurements of the same person's consumption might be expected to be similar. Attempts to measure time-to-time variability of actual nutrient consumption have been reported, as in the following six examples, where the weighed record was the method for collecting dietary data, and food composition tables were used to calculate the nutrient content of the food.

Evidence of actual day-to-day variation in food and nutrient intakes is provided by the Widdowson study (1947) of weighed records for 7 consecutive days supplied by a nonprobability sample of about 1,000 predominately middle class children, 1-18 years of age, in England. The coefficients of variation for caloric intake of the boys ranged from 10.6 for 9-year-olds to 20.5 for 16-year-olds. For girls the range was smaller, from 13.7 for 9-year-olds to 19.5 for 11-year-olds. No consistent pattern of day-to-day variation within or from group to group appeared, although the lowest coefficients of variation for both sexes were for the 9-year-olds. The maximum caloric intake for the group was often double the minimum. Day-to-day fluctuations in nutrients were frequently large, as might be expected from the day-to-day differences in the kinds of foods eaten. However, week-to-week fluctuations were smaller.

Widdowson (1947) and Wait and Roberts (1932) noted the following factors as affecting day-to-day variation in the intake of food energy: (1) Previous day's intake, with the lowest intake often following the largest, (2) serving of food liked or disliked, (3) physical activity, (4) variations in economic stress, (5) emotional state, (6) variability in health status, and (7) festive days. In view of such obvious reasons for differences between dietary measures for 2 successive days, these day-to-day measurements do not provide a sound basis for judging the reliability of the weighed record methods in terms of the absence of systematic error.

In contrast, the much smaller variability in empirical data from weighed records for 3- to 7-day periods suggests that this variability may be more random. This hypothesis has not been adequately tested because of difficulties in obtaining random samples of weighed records.

Group means of nutrient intakes calculated from weighed records of food eaten during 2 consecutive weeks by 39 business and professional men were remarkably close, except for vitamin A and ascorbic acid (Adelson, 1960). These two nutrients were higher the first week, when larger amounts of fruits and vegetables were consumed. Wide fluctuations in these two vitamins are to be expected from variations in the inclusion of particular fruits and vegetables in the family menus from week to week. When the 2-week records of individual men were paired, 94 percent of the records agreed within ±20 percent for food energy, 85 percent were within this range for protein, 64 percent for calcium, and 44 percent for vitamin C.

Reports on several studies with small numbers of subjects indicate the sampling problem encountered in weighed record surveys and its relevance to the reliability of this survey method. For Yudkin's study (1951), six young dietitians weighed and recorded their daily intakes during 4 consecutive weeks. One subject took 2 percent more food energy in the highest week than in the lowest, but another's differed by 68 percent. Vitamin A was two to five times greater in 1 week compared with another. Also, one subject was pregnant. Yudkin concluded from these six case studies that 1 week was too short a period to assess accurately an individual's intake.

Another group of researchers studied 5-day sets of weighed diet records, which had been kept for 25-120 consecutive days by six women, aged 68-80 years (Fry et al., 1963). They found a high degree of constancy for half of the women from period to period in 5-day averages of major nutrients and concluded that "variability perhaps depends more on the existence of a food pattern than on the length of time studied." (Ibid., p. 221) With this sample they could not explore the possibility that food patterns may be related to socioeconomic factors in such a way that the underlying variation in intakes may be random.

The Huenemann-Turner study (1942) utilized weighed records kept for several 10- to 14-day periods over a year by 25 dental clinic patients, 6-16 years of age. Their analysis of period-to-period variability based on means for individual cases reveals the likelihood of systematic differences caused by seasonal variations in intakes of such foods as fruits and vegetables. During the year only two children had vitamin A intakes that varied less than 20 percent from period to period. Only one child's vitamin C varied so narrowly. By way of contrast, 52 percent of the group showed less than ±20 percent variation for food energy; 56 percent were within this range for protein and 36 percent for calcium. With this small selected sample no analysis could be made of the degree of variability after the seasonal factor was taken into account.

The seasonality complication in comparing sets of weighed 7-day records is highlighted by another small-scale study. Chappell (1955), a 34-year-old lecturer, kept weighed records of her own daily intakes for 70 weeks and her father's for 13 weeks. For 7-day periods the percentage standard errors in her mean intakes of food energy, carbohydrate, protein, fat, calcium, iron, thiamin, riboflavin, and nicotinic acid were between 7 and 15 percent. The errors for vitamins A, C, and D were greater.

If the samples of weighed records were random, Pearsonian correlation analysis of separate week's intakes could be used to appraise the reliability of this method. Apparently the correlation analyses of nonrandom samples reported by Marr (1971) are based on this method. Despite these reservations and others noted by Marr, the correlation coefficients calculated by her (such as 0.78 for food energy in consecutive weeks in Adelson's (1960) study and 0.30 for food energy when data were 3 years apart as in Keys et al. (1966) investigation) do suggest the hypothesis that with the longer lapses of time between repetitions of dietary surveys, the actual intakes may change more for some nutrients than others. If so, appraisal of the reliability of weighed record surveys can be based on repetition of the random surveys only within a month or two.

7.1.2 Validity

Appraisal of the validity of a method requires consideration of several aspects of the measure (Kerlinger 1973). The weighed record is the commonly accepted measure against which other measures of individual diets are validated. However, several investigators have pointed out the difficulty of establishing the validity of a diet collection method when a reference method that yields absolutely true results does not exist (Keys, 1965; Marr, 1971; Samuelson, 1970).

The accuracy aspect of validity includes consideration of systematic error. Problems in separating such error from random variations have been previously identified. Thomson (1958) stated that means can be misleading when the errors of estimate are not random. One possible source of systematic error is indicated by the frequent finding in the small-scale, nonrandom studies that food intake in the first week exceeds that of later weeks. For example, Yudkin (1951) found food energy, protein, and fat averages 10-25 percent higher for the first week of diets reported by the six young dietitians than for the following weeks. This finding suggests the possibility that food behavior may be altered by participation in a survey using the weighed record method. Although the measurements obtained may be accurate, the weighed record method may not provide the unbiased data needed to measure relationships between diets and socioeconomic characteristics of the U.S. population.

Evidence of another kind of systematic bias introduced by the weighed record method comes from a study by Ohlson et al. (1950). They compared weighed diet records with recall data for the same 13 women and found that 7 had from 300 to 800 fewer calories of food energy per day during the period of weighing intake. They traced the differences to snacks and commented that snacking became inconvenient when it was necessary to weigh each mouthful and snacks were seldom found on the weighed diet records. den Hartog et al. (1965) noted that several housewives said their husbands had changed their food pattern during the period when their food was weighed in order to lessen the work and that simpler meals with less variations or fewer servings were eaten. Another research report based on estimated records indicated that some individuals were embarrassed to reveal all that they really had eaten (Paul et al., 1963).

Still greater effect on food behavior may be expected when the weighing and recording are done by or in the presence of an outsider in the home at meal preparation and serving times, as in the study by Leverton and Pazur (1957). Although the measurement of the portion eaten was undoubtedly very accurate, the disturbance of customary patterns was not studied.

A common survey practice is another source of possibly systematic error. When a meal is eaten away from home during the period of collecting weighed records, the portions of food eaten may be estimated by the same subject. In some studies, if the portion size is missing, an average size is inserted. The possible magnitude of this bias has been investigated by several researchers. In a large study of pregnant women's diets, Dieckmann et al.

(1951) tabulated weighed portion sizes and found the average serving did not conform to those postulated in printed tables. Variations from the calculated averages were wide and frequent. In another study, weighed diets collected from 242 North Central Region women, 30-92 years of age, were examined for portion weights of frequently selected foods (Beegle et al., 1954). Most women showed individual variation in weights of the same foods selected at different meals, and there was substantial variation among subjects. But mean and median amounts eaten of any given food were relatively close. (Portion size data from the USDA spring 1965 Survey of Food Intake of Individuals illustrate these points, although the portions were estimated, not weighed (Pao and Burk, 1975).)

Another source of error, possibly systematic, is the inconsistency in precision of steps taken in collecting and analyzing dietary data. In several studies, after obtaining precisely weighed records of food quantities, the nutrient contents were obtained partly from precise chemical analyses of some foods and partly with imprecise shortcut methods, such as applying one value for a group of foods. (Burrill et al., 1959; Fry et al., 1963; Leverton and Pazur, 1957) Unless the diets of the subjects contain food items in the same proportions within food groups as in the basic data from which the average factors were calculated, there is a possibility of introducing systematic bias. The current availability of data tapes with food composition values from Agriculture Handbook No. 8, "Composition of Foods--Raw, Processed, Prepared" (Watt and Merrill, 1963), and of large-scale computers in universities and government agencies greatly reduces the need for using group values.

From time to time since the 1930's nutritionists have investigated the possibilities of systematic error arising from using tables with average nutrient composition values for individual foods instead of concurrent chemical analysis of aliquots or samples of the foods. A few of these investigations are reviewed briefly in the appendix. The average food composition values from tables, such as in Handbook No. 8 (Watt and Merrill, 1963), are likely to result in less error in nutrient estimates from nationwide survey data than estimates based on samples, because the handbook data are adjusted for variations in nutrient content among seasons and areas, as well as varieties.

Concurrent validity consists of utilizing more than one method simultaneously to find out whether the survey results check. Obviously one's judgment enters into deciding which method comes closer to measuring the true value of the diets. Nutrient data calculated from weighed records have been compared in small-scale studies with data calculated from concurrent recorded estimates in terms of household units (in cups, tablespoons, units, cubic measures) and with biochemical measurements. For some nutrients the two sets of estimates were close; for others the differences were notable. In one such study, mothers of 25 school children recorded the foods eaten by the children during 1 day, first in estimated or measured servings, then after weighing them (Eppright et al., 1952). The mean food

energy and nutrient intakes calculated from the estimated diets exceeded the means as calculated from the weighed diets. Differences for carbohydrate, thiamin, and ascorbic acid were not significant, but differences for other nutrients were significant.

In other studies, Huenemann and Turner (1942) found dietary ascorbic acid to be significantly correlated with plasma ascorbic acid, and Dieckmann et al. (1951) identified a significant relationship between urinary nitrogen excretion and protein intake.

Construct validity of the measurements of food intake obtained from weighed records must be approached from several directions. If one is seeking the most precise measure of food eaten, the weighed record method is obviously the best survey procedure. But if one seeks the measurement of customary dietary intake unaffected by the method of collecting data, the evidence cited earlier about the effects of the survey procedure on behavior indicates this method to be unsatisfactory.

Evaluation of the validity of data obtained from weighed records kept for 1, 3, or 5 consecutive days as measures of the construct, dietary intake over 7 consecutive days or more, leads us back to the research on day-to-day variation cited in the discussion of random versus systematic variability and reliability. Although the studies referenced in that section indicate that for some nutrients weighed records for 3-5 successive days are adequate measurements of an individual's dietary quality, for others they are not. Supporting the case for adequacy of the measure are the conclusions from two European studies reported in the 1960's. Fidanza and Fidanza-Alberti (1967) concluded that about 80-90 percent of the information obtained in weighed records for 7 successive days was captured in records for 3 successive days. Heady (1961) also found little information was lost by using 3 consecutive days rather than more consecutive days. Other researchers consider even 7 days to be too short a period to accurately assess the intake of nutrients by an individual (Chappell, 1955; Huenemann and Turner, 1942; Yudkin, 1951).

Validity is affected by the representativeness of weighed diets, which may in turn be affected by nonresponse rates. In one study of diets during pregnancy, Thomson (1958) tried to obtain a sample representative of the primigravida population in Aberdeen, although the lower class was oversampled in one period using the weighed record method. Results obtained from 713 subjects of 729 sampled showed an "apparent social-class gradient." Nonresponse was 4 percent in the upper class and 10 percent in the lowest class. Reliable data decreased from 93 percent of the cases in the upper class to 61 percent for the lower class based on the husband's occupation. Thomson (1958, p. 454) summarized his general impression thus: The less dependable subjects were often socially, intellectually, and clinically inferior to reliable subjects and their diets also tended to be inferior. Should this impression be correct, the agreement between the means for reliable and less dependable subjects in each class is spurious, causing the slope of the social gradients to be underestimated owing to a decreasing number of subjects with inferior diets at the lower end of the social scale.

This is a clear example of how poor sampling can lead to unreliable conclusions. This problem with nonresponse is also illustrated by Adelson's (1961) finding that of 63 businessmen who gave 7-day recalls only 59 would keep weighed food records for a week, a loss of 6 percent of the participants. She reported no analysis of the characteristics of nonparticipants.

The most important indication of the nonresponse problem, previously mentioned under reliability, is the great scarcity of attempts to use probability sampling in studies incorporating the weighed-record method of data collection.

The sampling problem also tremendously affects the predictive validity of data obtained with the weighed-record procedure. None of the weighed-record studies has provided even a part of the data needed for analysis of how U.S. diets vary with socioeconomic characteristics of individuals. The selectivity among cases resulting from nonresponse and the strong possibility of changed food behavior make this method a very poor choice for nationwide surveys with predictive objectives.

7.1.3 Respondent Burden

den Hartog et al. (1965) considered participation in a weighed-record study to be a burden on the participant and his family, and families with small or many children often did not wish to take part. This appraisal is supported by Eppright et al. (1952) in their statement that in mass studies it is impractical for subjects to weigh the food eaten, but they can be asked to record foods in terms of servings or estimated household measurement.

7.1.4 Field Survey Costs

Adelson (1960) reported that initial visits averaged 1 hour and that three shorter visits were required during the week of recordkeeping. den Hartog et al. (1965) reported that one dietitian can help about five persons a week in weighing records. Other researchers reported daily visits by dietitians (Marr, 1971). These interviewing requirements indicate heavy field costs.

7.1.5 Data Survey Costs

The only indications of data processing costs found in the literature are based on Adelson's data cited in Young and Trulson (1960, p. 812) where the number of hours required per schedule, from collection of data to hand calculation of nutritive values, totaled 34.

Costs of processing data from weighed records are probably smaller than for data in household units because of the need for the latter to be converted to equivalent weights. Obviously no substantial survey operation in the 1970's would include hand calculations of nutritive values, such as those for the Adelson study in the late 1950's.

7.2 Record of Current Food Intake: Estimated Record Method

The estimated record procedure is also often called the household measure record because whenever possible and appropriate, standard household measuring utensils, the 8-fluid ounce measuring cup and measuring spoons, are used in measuring and entering the amounts of each food. If a food item cannot be measured in a cup, it may be measured with a ruler and dimensions listed, as for a piece of meat or cake. For some items, counts are used, such as the number of eggs or slices of bread. For candy bars or prepackaged portions, weights are readily available and are used. When no relevant measuring device is available to the respondent, the portion size is estimated or described as accurately as possible. Accordingly several kinds of measurements, varying in their precision, usually appear on the estimated records.

7.2.1 Reliability

Just as for weighed records, the likelihood of low cooperation rates in this type of survey and pessimism regarding the possibility of maintaining the representativeness of samples have resulted in few attempts at probability samples. Marr (1971, p. 116) stated that subjects using records in surveys do not necessarily cooperate more than those in a weighed survey, yet the record survey is generally regarded as eliciting better response. On the same page a table by Marr shows that cooperation rates from several surveys, using estimated records, vary from 77 percent of reliable records (Guthrie, 1963) to 45 percent (Sprauve and Dodds, 1965), with the exception of one study for which the rate was 95 percent (Dierks and Morse, 1965), and here one might suspect that some kind of screening procedure was used. None of these samples was a probability sample. Despite the selectivity of prospective participants, the percentage of reliable records was not high except in the one case.

Quota samples of groups were obtained in the New York State Nutrition Survey of 1947, because the planners decided a random sample was impractical (Trulson et al., 1949). Chosen to represent a cross section of the public school population were 3.4 percent of the registered fourth graders (lowest grade able to write legibly, according to the authors) and 2.7 percent of the 10th graders (N=9,543). Tuesday, Wednesday, and Thursday were specified (probably introducing some bias) for children to keep estimated records so teachers could help collect the data. Some schools failed to cooperate, resulting in underrepresentation of large cities and overrepresentation of less populous areas. Application of a weighting system probably would have increased the representativeness of calculated means and standard deviations derived from the collected data. Despite this bias and that, which could have resulted from some shortcut methods in calculating the dietary intakes, relative or percentage standard errors (RSE) of the mean intake for protein were calculated for three groups of children by Trulson et al. (1949, p. 599) and are given here. They appear to indicate a satisfactory level of reliability. However, the undetermined amounts of bias just described make inferences from such data somewhat questionable.

4th graders (N=4,881, $\bar{X}$=74 g, S.D.=±17.3) RSE=3.3 percent
10th graders - boys (N=2,302, $\bar{X}$=93 g, S.D.=±25.6) RSE=5.7 percent
girls (N=2,360, $\bar{X}$=75 g, S.D.=±18.1) RSE=5.0 percent

Further evidence of the reliability of data obtained with the estimated record procedure comes from a smaller scale, but probability sample of the diets of 345 Iowa school children aged 9, 10, and 11 years in 1948-51. These data are from estimated records kept by the children for 3 specified days--Tuesday, Wednesday, and Thursday. (Eppright et al., 1955)

The following relative sampling errors (RSE) expressed in percentage standard errors for weighted mean nutrient intakes, taken from Eppright et al. (1955, p. 621), show good reliability in the data:

Nutrient	Boys			Girls		
	Weighted mean	Standard error of mean	RSE in percent	Weighted mean	Standard error of mean	RSE in percent
Food energy (kcal) ---	2,548	41	1.6	2,245	48	2.1
Protein (g) ----------	74.9	1.6	2.1	67.2	1.5	2.2
Calcium (mg) ---------	1,099	49	4.4	984	66	6.7
Vitamin A (IU) -------	8,007	668	8.3	6,813	515	7.5
Ascorbic acid (mg) ---	82.4	3.3	4.0	84.6	4.5	5.3

Some percentage standard errors, calculated by Pao and Burk from published means and standard deviations, indicate the reliability of the data for several nutrients derived from 3-day estimated records collected in the fall 1965 study of preschool children in the North Central Region by Eppright et al. (1972). The relevant measures for several sex-age groups are given in table 4. A self-weighting, stratified, multistage, cluster sample (drawn by the Iowa State University Statistical Laboratory) was used. The survey yielded 1,664 completed interviews, 81 percent of the original sample's eligible households. Substitutions were made for 336 of the 386 nonparticipants to reach a total final sample of 2,000 children.

The question of the reliability of the estimated record procedure can be approached from a different perspective, direct analysis of repeatability, with information from a study of the diets of 10 young dietitians (Celender, 1963). Their records kept during two 4-week periods about 6-8 months apart yielded maximum variation of 15 percent or less when a week's food energy was paired with the next in 13 out of the 20 observations, 25-30 percent for 6 instances. The greatest variation, 62 percent in one case, was traced to a special week of restaurant eating. Further examination of Celender's data supported the hypothesis of actual variability in food from day to day, week to week, month to month, and season to season.

TABLE 4.--Mean intakes and reliability measures for food energy and 4 nutrients for selected groups of young children, North Central Region, fall 1965[1/]

Food energy and nutrient	Boys, age group in months[2/]			Girls, age group in months[2/]		
	12-18 (166)	36-48 (308)	60-72 (360)	12-18 (113)	36-48 (253)	60-72 (268)
Food energy:						
Mean (kcal)	1,241	1,633	1,883	1,323	1,555	1,718
Standard deviation	326	415	480	389	360	414
Standard error	25.4	23.7	25.3	36.7	22.6	25.4
Percentage standard error	2.0	1.5	1.3	2.8	1.5	1.5
Protein:						
Mean (g)	53	64	76	58	61	68
Standard deviation	17	18	24	18	16	17
Standard error	1.3	1.0	1.3	1.7	1.0	1.1
Percentage standard error	2.5	1.6	1.7	2.9	1.6	1.6
Calcium:						
Mean (mg)	1,031	1,144	1,107	1,024	956	1,028
Standard deviation	336	322	341	345	282	334
Standard error	26.2	18.4	17.9	32.5	17.7	21.0
Percentage standard error	2.5	1.6	1.6	3.2	1.9	2.0
Vitamin A:						
Mean (IU)	6,814	6,006	6,473	7,751	5,657	5,850
Standard deviation	3,836	3,603	4,216	6,574	3,673	3,280
Standard error	298.8	205.9	221.9	620.2	231.0	206.3
Percentage standard error	4.2	3.4	3.4	8.0	4.1	3.5
Vitamin C:						
Mean (mg)	76	75	83	87	77	77
Standard deviation	42	49	51	53	52	50
Standard error	3.3	2.8	2.7	5.0	3.3	3.1
Percentage standard error	4.3	3.7	3.3	5.7	4.3	4.0

1/Vitamin and mineral supplements were included with foods in calculating mean intakes. Data from Eppright et al. (1972, pp. 282-285) (means and standard deviations). Standard errors and percentage standard errors were derived by the reviewers.

2/Number of cases in parentheses.

Random errors in measurement appear to have been minimized in the Celender study by obtaining the data from dietitians because they are accustomed to handling portion sizes. This is directly contrary to Celender's critical assumptions (1) that if weight of subject is maintained constant, caloric intake must be constant and (2) therefore variations in measurement are due to errors in measurements. Biochemists have found that intake of food energy may increase for a short time without a corresponding increase in body weight (Miller et al., 1967; Strong and Passmore, 1967; White et al., 1973). Hence the reviewers question Celender's assumption and instead assume the dietitians did estimate their intake rather accurately so that variations in measurements reflect actual variations in food intake.

Further evidence regarding the variability in actual intakes is provided by estimated records from a small heterogeneous sample of 16 women and 2 men, 23-49 years of age, for 28 consecutive days (Young et al., 1953b). About 50 percent of the subjects had weekly averages for all 4 weeks that were within ±10 percent of their 28-day average for food energy and six nutrients (protein, phosphorus, iron, thiamin, riboflavin, and niacin). Almost 100 percent were within ±20 percent for food energy and iron and 85 percent were within this range for the other nutrients studied. The researchers noted a disadvantage of measurements for short periods, since all the variations to which dietary intakes are normally subjected may not have been included for each individual. But the authors concluded that the daily means for the group were stable enough for less than a week's record to have provided an acceptable estimate of intake.

Correlation in intakes between consecutive days was found to be nearly zero, except for the college student group, by Chalmers et al. (1952). They treated the daily records as independent and separate observations in their analyses of variance to determine how many and which days should be included in the dietary record period. The authors concluded that "a dietary record need consist of only one day when characterizing the dietary intake of a group," and it is more effective to increase the number of subjects than the number of days. This conclusion has often been the justification for using 1-day data. However, other researchers have found significant correlations between days (Hankin et al., 1967; Trulson, 1951). The reviewers concluded that the independence of 1 day's intake from that preceding and following has not been sufficiently tested with random samples.

Two elements enter into comparisons of data from estimated records with those from weighed records, a frequent procedure for evaluating reliability. One is the reduction in precision arising from use of various volumetric measures, for example, cups and teaspoons, instead of weights measured with scales. The other is possible differences between servings weighed beforehand by experienced adults and those estimated in volumetric measures by the consumers.

Studies by Eppright et al. (1952) and Bransby et al. (1948) indicate some significant differences in precision of measurement. The Eppright study included records of 25 children's intakes that were estimated and then weighed by their mothers. The researchers found nutritive values

calculated from the two types of records to be correlated, but those from estimated records were consistently higher. Nevertheless they concluded that when records are being kept by untrained people, the information on food items is probably as satisfactorily obtained by servings as by weighed amounts.

The Bransby study used 3-day weighed and estimated records of intakes by 49 children in an orphanage. The average percentage differences between values for weighing and estimated (measured) records were for food energy 5, protein 8, fat 12, carbohydrate 0, calcium 6, and iron 9. For some foods average percentage differences for consumption found by weighing and by estimating (measuring) were rather large; for example, 26 percent for meat, bacon, and fish; 29 percent for fruit; and 31 percent for eggs.

The second element in the comparisons between weighed and estimated records of the food intake of the same individuals at the same time was introduced in the Chamberlain and Pyke research by having the foods weighed in the kitchen before serving them to 20 freshmen girls who estimated and recorded their own intakes over a 7-day period (Young et al., 1952b). The group averages for individual nutrients agreed closely, but the data for individual girls varied considerably. The values based on weighing were between +26 to -18 percent of caloric values derived from estimated portions. In protein the differences ranged from +43 to -14 percent. Ascorbic acid values calculated from the estimated records were consistently low.

These findings indicate the basic problem of estimating portion sizes which enter into both the estimated record and the recall procedures. Some researchers have used plastic measures and shapes to help respondents evaluate sizes and amounts of fruits, eggs, meat, salad vegetables, and other portions difficult to estimate. Others have used food models and concluded that they increased the accuracy of portion sizes (Abraham et al., 1974; Moore et al., 1967). The procedure of keeping estimated records has an advantage over the recall of sizes of portions in that the respondent before recording foods eaten meal by meal can and usually does some actual measuring of quantities comparable to those eaten.

Eppright et al. (1952) also reported comparison of records kept by the mother and by the child. There was no difference in the final evaluation in terms of food groups. Records kept by mothers and sons agreed more closely than those by mothers and daughters. Girls' records were more variable than the boys' records.

Trulson (1954) found the 7-day estimated records of 37 clinic patients, 7-12 years of age, resulted in higher mean values for selected food groups, as number of eggs and foods high in vitamin C, than those obtained a little earlier using recall or diet history methods. But this finding gives no clue as to which estimates were closer to the true values.

7.2.2 Validity

Accuracy of the estimated record may be affected by changes in consumption brought about by participation in the survey. Celender (1963) offered evidence that collection of dietary information altered the subjects' food intake systematically so that the dietitians recorded higher intakes for the first week of recordkeeping compared with the later 3 weeks in both periods of recordkeeping. Inspection of the data from Young et al. (1953b) shows the same relationship, that the mean caloric and nutrient intakes of 18 young adults were higher the first week than in the succeeding 3 weeks except for vitamin A.

The accuracy of estimated food portion sizes depends on the subject's ability to report accurately what he eats. Subjects' estimates of food portion sizes and omissions of food items are probably the greatest sources of error in keeping diet records (Young et al., 1952b). Young et al. (1953a) designed experiments to find (1) the difference between estimated and actual measured portion sizes, (2) the direction of bias in estimating portions, and (3) the variation in individual estimates and omissions. They found that for about two-thirds of the subjects the calculated nutritive intake from estimated diet records was within ±20 percent of that of the measured intake. Food types reported least accurately were puddings, sauces, gravies, and fruit (Young et al., 1952b). Whenever a food could be reported in a number of items or easily measured in cups or tablespoons, the mean reported estimate was closer to the actual mean. The coefficient of variation for meat reported in cubic inches was much greater than meat in ounces.

Subjects overestimated counting occurrences oftener than underestimated portion sizes. This bias was partially compensated by omissions of food eaten. In one study there was a tendency to underestimate as the portion size increased and tendency to overestimate as the portion size decreased. Regarding omission of foods, the smallest number of omissions occurred for meats, eggs, cookies, pies, and cakes, constituting a potential source of bias. Young et al. (1953a, p. 1220) also found that the best estimations "were made by subjects who were aware that they were being observed."

In the Total Diet Feasibllity Study (U.S. Bureau of the Census, 1964, p. 3), there was evidence that individuals from lower income families were able to estimate amounts better than some persons with higher incomes, but fewer of the lower income respondents could independently fill out the diary form.

Accuracy may also be affected by differences among interviewers who collect the estimated records. Estimated records need to be checked by interviewers for foods that may have been forgotten or inadequately described. When unchecked records were compared with checked records, discrepancies were found amounting to 10 percent or more for 16 percent of the subjects (Steele et al., 1951). However, if interviewers are well trained, differences are unlikely to exceed sampling variations (Church et al., 1954).

Several investigators found a systematic "day" effect. Cellier and Hankin (1963), Eppright et al. (1952), and Leverton and Marsh (1939) found diets on weekend days tended to differ in content from those on weekdays.

Concurrent validation of estimated records was achieved by using biochemical measurements of plasma ascorbic acid, which showed significant correlations with the ascorbic acid intakes as calculated from the diet records (Eppright et al., 1952; Owen et al., 1974).

Significant correlations were also found between plasma urea nitrogen and dietary protein and between urinary and dietary riboflavin in the Preschool Nutrition Survey (Owen et al., 1974). Results of concurrent validation by comparing results of two dietary collection methods are discussed under reliability.

Construct validity is concerned with the degree to which a measure actually captures the concept to be measured. The objective sought--whether the mean value for the group is the objective, whether values relevant to the food behavior of a specific individual are sought, or whether the objective falls between these two concepts--e.g., values predictive of the diets of individuals with specified characteristics--must affect the choice among dietary collection methods and detailed decisions on survey operations.

Young et al. (1952b, p. 48) studied this problem and concluded that for characterizing a group by its mean intake, increasing the number of 1-day records was more efficient than increasing the number of subjects. (These workers suggested that if it is feared the subjects may eat a better diet than usual when keeping the 1-day record, introducing a bias, the method of the 24-hour recall might be substituted.) However, the 7-day record and the 24-hour recall did not give the same information for an individual and hence they could not be used interchangeably for this objective (Young et al., 1952c).

Representativeness of estimated records for 1 day of the diet over a week is less than that of records collected for several days, according to results from several studies. Eppright et al. (1952) found that 3 weekdays did not give the same indications of dietary adequacy as 1 day or 7 days. In comparing distributions of 1-day dietary records with 3-day dietary records and 3-day with 7-day records, the 1-day diets appeared to be more adequate than those over 3 days, and the diets for 3 days in turn appeared more adequate than those from 7-day diet records. Certain combinations of days seem to vary as to measuring nutrient intakes if the representative diet is the construct sought.

Hankin et al. (1967) found that variances in daily intakes of 5 nutrients in 7-day measured records of 93 Japanese-American men differed significantly. Correlations of Sunday with each of the other days were lower than for combinations of other days. As mentioned previously, other researchers found dietary intake on weekend days often differed from intakes

on weekdays (Cellier and Hankin, 1963 ; Leverton and Marsh, 1939; McHenry et al., 1945). Cellier and Hankin (1963) concluded from analysis of both 4-day and 7-day records that the 4-day record retained about 90 percent of the information of the longer record when 1 weekend day was included and the group consisted of at least 100 women.

The gain in representativeness from collection of 3-day records (Tuesday, Wednesday, Thursday) in contrast to 1-day (Wednesday) and 7-day records is shown in another study of 132 children 10-12 years of age by the reduction in standard deviations of the group mean as well as the coefficients of variation (Trulson, 1951). Most of the potential increase in stability was captured in the 3-day average as compared with the 1-day and 7-day averages. For example, the coefficient of variation for protein decreased from 28.6 percent for 1 day to 22.5 percent for 3 days and to 20.4 percent for 7 days; the corresponding coefficients of variation for vitamin A were 107.1, 70.7, and 60.1 percent. However, a study of various 3-day and 2-day sequences scattered through the week, using random samples, would provide more information about how much increase in stability occurs under various conditions. Reliability of the data is obviously affected by whether the respondent knows what the subjects ate. Owen et al. (1974) found that 10 percent of the 4- to 6-year-olds prepared their own breakfast, but this was not related to any measurement differences.

Representativeness of data is also affected by the nonresponse rate in a sample. The nationwide Preschool Nutrition Survey, 1968-70, was unable to maintain a representative sample perhaps because it included clinical and biochemical measurements (Owen et al., 1974). Thirty-five percent of the eligible children (5,335) did not participate at all, whereas just over 40 percent took part in all phases. Dietary information was collected for 65 percent (3,441), consisting basically of 2 days of estimated record (cut back from 3 days in the original plan (Owen et al., 1969)) plus the diet intake recalled on the day of the clinic visit. Because of the low participation rate, standard errors were not given. The first record day included a recall of the time preceding the arrival of the interviewer at the home, and the method may more properly be regarded as a combination of recall and record.

The following tabulation, calculated from data by Owen et al. (1974, p. 603), shows the relationship between participants by race and the proportion of children in the two races estimated for the total child population in the region. No pattern of nonparticipation is revealed.

Region	Percent interviewed of those expected by race	
	Black	White
Northeast	69	58
North Central	49	71
South	77	74
West	67	74

Black families gave mother's employment outside the home as a frequent reason for not participating, whereas white families more often cited inconvenience or illness in the family. The authors commented that alerting communities to the project and explaining it were essential to the success of the survey.

In the Total Diet Feasibility Study (U.S. Bureau of the Census, 1964), a pilot test of a procedure including the mailing of a diary form with a letter asking the respondent to complete the form for a prescribed record day was highly unsuccessful, with only 5 respondents out of 27 even attempting to keep the diary. However, in personal interviews all 27 were finally completed.

Predictive validity of food and nutrient intake measured with estimated records is supported by most of the large-scale surveys mentioned previously. In the Preschool Nutrition Survey, 1968-70, Owen et al. (1974) found logical relationships among estimates of the dietary variables and the socioeconomic variables, using the Warner Index to rank the families. Despite the lack of a random sample, Owen et al. believed their findings would be helpful in making policy decisions and planning nutrition education programs. The same conclusion is reached by researchers on the fall 1965 North Central Regional Survey of Preschool Children (Eppright et al., 1972) and from several of the older surveys.

7.2.3 Respondent Burden

Regarding respondent burden, Mongeau (1974) stated that whether for 1, 3, or 7 days, food records impose a burden on the subjects, reducing the cooperation rate and introducing a bias in the sample, because the poorly educated and the low-intelligence people are underrepresented. Young et al. (1952b) also mentioned that of 55 industrial workers willing to be interviewed, only 9 cooperated by keeping 7-day records. However, greater cooperation could be expected for 2 days of record than for 7 days.

7.2.4 Field Survey Costs

In the Preschool Nutrition Survey, interviewers returned to the home on 2 successive days to review the respondents' records (Owen et al., 1974). In the fall 1965 North Central Regional Survey of Preschool Children, interviewers made at least two visits to each home (Eppright et al., 1972). The more visits the interviewer makes the greater are the field survey costs.

7.2.5 Data Processing Costs

Because household measures, counts of items, and so forth on the estimated record must be converted to a common measure (grams), data processing costs are higher for the estimated record procedure than for weighed records.

7.3 Recall of Past Food Intake: 24-Hour Recall Method

The recall procedure relies on memory. If the time period extends back only to the preceding meal, the recall and record in household measures may actually be the same. However, most references to recall methods refer to periods of 1 day or several preceding the time of interview. If longer periods in the past are to be covered, the technique used is really a dietary history. The 24-hour period may be the preceding day or it may refer to the preceding 24 hours and thus covers parts of 2 days. The instructions and report should specify the precise coverage.

7.3.1 Reliability

Reliability of the 24-hour recall for individuals can be evaluated with the sampling errors of the means when the data are from probability samples. The percentage (or relative) standard errors for several sex-age groups from four surveys listed in table 5 indicate satisfactory levels of reliability for the mean intakes of food energy, protein, and calcium, with less reliability for vitamin A and in some groups for vitamin C.

The Iowa and South Dakota data on women's 24-hour diets were collected in 1948 and 1949 from samples representative of each State (Swanson et al., 1959). The large Iowa sample yielded consistently more reliable data than the South Dakota sample, but the mean nutrient intakes for the several age groups are generally rather close with few exceptions.

The data from the spring 1965 survey by the U.S. Department of Agriculture (1972a) pertain to the same sex-age categories, but they represent the entire North Central Region. The samples for the region are generally smaller than those in the Iowa study. When one considers the standard errors and the differences in season, year, and area, the two sets of State and the regional means are similar.

The percentage standard errors for food energy and the four nutrients in the diets of people 60 years and over derived from the USDA spring 1965 survey data for the North Central Region are generally comparable with the all-U.S. data for 1971-72 from the first Health and Nutrition Examination Survey (HANES) (Abraham et al., 1974). The latter sample was more than three times as large. The means for food energy, protein, and calcium are reasonably close. The vitamin A means are within sampling range, but the vitamin C mean from the HANES is considerably out of line with the Iowa, South Dakota, and USDA data in the table as well as the 55- to 66-mg range of the all-U.S. means from the USDA survey for sex-age groupings of people over 64 years (U.S. Department of Agriculture, 1972a).

The well-known "Nutritional Status U.S.A." (Morgan, 1959) summarizes numerous dietary studies conducted between 1947 and 1958 in the United States, using mainly 7-day estimated records but also some 24-hour recall.

TABLE 5.--Mean food energy and nutrient intakes with standard errors for specified population groups from 4 surveys using 24-hour recall method

Sex-age (years) group	Cases	Food energy			Protein			Calcium			Vitamin A			Vitamin C		
		Mean	Standard error	Percent standard error	Mean	Standard error	Percent standard error	Mean	Standard error	Percent Standard error	Mean	Standard error	Percent standard error	Mean	Standard error	Percent standard error
	Number	Kcal	Kcal	Pct	G	G	Pct	Mg	Mg	Pct	IU	IU	Pct	Mg	Mg	Pct
Iowa women:[1]																
30-39	282	1,850	37.8	2.0	64	1.4	2.2	0.56	0.024	4.3	4,935	317	6.7	67	2.9	4.3
40-49	241	1,815	36.1	2.0	61	1.4	2.3	.51	.019	3.7	4,810	319	6.6	67	2.8	4.2
50-59	232	1,740	38.8	2.2	57	1.4	2.5	.46	.019	4.1	5,115	398	7.8	66	2.6	3.9
60-69	170	1,695	46.2	2.7	55	1.8	3.3	.49	.023	5.2	4,655	348	7.5	65	3.1	4.8
70 and over	147	1,425	42.1	3.0	44	1.4	3.2	.43	.024	5.6	3,630	251	6.9	57	3.8	6.7
South Dakota women:[1]																
30-39	91	1,840	58.4	3.2	64	2.3	3.6	.62	.042	6.8	4,780	556	11.6	52	2.8	5.4
40-49	105	1,770	52.5	3.0	59	2.1	3.6	.52	.029	5.6	5,230	415	7.9	59	3.9	6.6
50-59	73	1,610	57.6	3.6	55	1.9	3.5	.41	.024	5.9	4,060	408	10.0	56	4.5	7.6
60-69	45	1,630	83.3	5.1	52	2.7	5.2	.46	.045	9.8	4,690	514	11.1	62	5.0	8.1
70 and over	25	1,355	93.9	6.9	42	3.1	7.4	.40	.041	9.3	2,235	244	10.9	34	4.5	13.2
USDA survey in North Central Region:[2]																
Women:																
30-39	163	1,632	52.9	3.2	65	2.4	3.7	.51	.024	4.7	4,983	1,108	22.2	50	3.4	6.9
40-49	198	1,691	56.3	3.3	68	2.3	3.4	.50	.023	4.6	5,655	814	14.4	61	4.0	6.6
50-59	130	1,736	65.0	3.7	70	2.9	4.1	.49	.027	5.5	6,817	1,433	21.0	64	4.8	7.5
60-69	151	1,510	46.7	3.1	64	2.7	4.2	.45	.022	4.9	3,978	424	10.7	53	4.0	7.5
Both sexes:																
60 and over	563	1,784	34.2	2.0	70	1.4	2.0	.57	.015	2.6	4,989	409	8.2	58	2.7	4.7
HANES:[3]																
Both sexes:																
60 and over	1,938	1,641	32.6	2.0	68	1.5	2.2	.65	.019	3.0	6,507	770	11.8	96	4.4	4.6

[1] 1948-49 data from Swanson et al. (1959, p. 492).
[2] Unpublished spring 1965 data from Survey of Dietary Intakes of Individuals, U.S. Dept. Agr., Agr. Res. Serv.
[3] 1971-72 data from Abraham et al. (1974, pp. 40, 174).

Since results could not be generalized to the total population, the need for the 1965 survey of individuals was realized. This survey was designed to provide data from which could be drawn generalizations for the entire U.S. population and component sex-age groups.

Nonresponse rates for the Iowa and South Dakota surveys of women's diets have not been found in published reports. For the spring 1965 Survey of Food Intakes of Individuals (U.S. Department of Agriculture, 1972a), 7 percent of the eligible family members in the participating basic sample of households did not respond. The nonresponse rate on the dietary section of the HANES survey was 27 percent. Since the dietary information was obtained during the same visit to the mobile unit as clinical and biochemical examinations, participation may have been materially reduced by the unwillingness and inconvenience of participating in these examinations. Part way through the survey period of 18 months, remuneration of participants was begun and participation substantially rose. Because the final sample of 10,126 cases examined (of 14,147 selected) failed to meet requirements of the original probability design, the observations were weighted to obtain averages more representative of the total population. Some bias could have been introduced by this procedure. Some percentage standard errors are listed in table 5, but the 24-hour recall data appear reliable.

Standard errors of mean nutrients are not reported for the 24-hour recall dietary data collected in the Ten-State Nutrition Survey, 1968-1970, (U.S. Department of Health, Education and Welfare, 1972). The original sampling design provided a probability sample for assessment of the nutritional status and dietary practices of low-income groups within each selected State. However, medium- and high-income households were not properly sampled, and the 10 State samples were not designed to represent the whole country, either for low- or high-income households. Of the 29,935 families contacted, 23,846 (79 percent) participated. The interviewed families consisted of 86,352 individuals, and 47.3 percent visited specified clinics where the dietary information was obtained from "high-risk" groups.

Nutrition Canada (1973) also utilized the 24-hour recall in its recent nationwide survey (1970-72). Forty-six percent of the persons initially selected attended survey clinics. Such a low response rate may be the reason why no standard errors are reported. Men, 20-40 years of age, had the lowest response rates. Response rates were highest among those living in rural areas and lowest in urban areas. Among the Indian groups, only 30 percent of selected persons participated because a large number could not be contacted. Response rate for the Eskimos was 60 percent.

Comparison of data obtained with the same method in two (or more) time periods gives a clue to the reliability of the method. The magnitude of the variation, both random and systematic in origin, involved in measuring average intake was studied with data from a series of 24-hour recalls randomly repeated over 1 year with 1 in each month, initially with 100 volunteer participants in an Israeli heart study (Balogh et al., 1971). There was a tendency for larger coefficients of variation to be associated

with more recalls. The investigators explained that for most nutrients two or three monthly reports fairly close together did not reflect as much variation within the diet as when additional months of data were included. This apparent seasonal variation would be missed if only a few consecutive months are covered. The problem with dropouts is indicated by the fact that 71 percent supplied 8 recalls at least a month apart, 28 supplied 10, and only 12 carried on to supply 11 recalls.

Reliability of the 24-hour recall method has been investigated in three small studies using alternative measurements for the same subjects (school children) at the same time (school lunch). For one study, standard errors of the means were reported and the results led the investigator to conclude that what children report in the recall and their actual consumption agree well enough to use the recall technique for evaluating the nutritional intakes of groups, including children as young as 8 years of age. (Samuelson, 1970) Samuelson (1971) used the 24-hour recall in an epidemiological nutrition investigation to study the food consumption of 1,401 children in Sweden.

In the methodological study the school lunches consumed by 56 eight-year-olds and 43 thirteen-year-olds were observed and a duplicate portion weighed, for which nutrients were determined by both chemical analysis and calculation. The 8-year-olds had higher intakes of energy, fat, and iron according to their recall of school lunch than according to data from weighing. The medians of amounts of food consumed differed by only 2 percent, with the recall higher. For the 13-year-olds, the recall of amounts gave a 14-percent smaller amount for the group than the weighing. The main reason for the difference appeared to be that in the recall 10 girls omitted 300 grams of milk that each had consumed.

For the two tables from Samuelson (1970, pp. 330, 332), the reviewers calculated the relative standard errors for comparisons with other studies and surveys. Following are the mean values, standard errors, and relative standard errors for amounts of food, food energy, protein, fat, and iron for recall and for chemical analysis for 56 eight-year-olds and for 43 thirteen-year-olds:

	Recall			Chemical analysis		
	Mean	Standard error of mean	Relative standard error	Mean	Standard error of mean	Relative standard error
8-year-olds:						
Amount of food (g)-	395	21	5.3	378	20	5.3
Food energy (kcal)-	472	30	6.4	399	22	5.5
Protein (g) -------	17.3	1.0	5.8	17.5	.9	5.1
Fat (g) ----------	24.6	1.9	7.7	18.8	1.3	6.9
Iron (mg) ---------	4.1	1.0	24.4	3.4	.9	26.5
13-year-olds:						
Amount of food (g)-	409	20	4.9	491	12	2.4
Food energy (kcal)-	494	27	5.5	491	18	3.7
Protein (g) -------	19.6	1.1	5.6	23.2	.8	3.4
Fat (g) ----------	27.4	2.0	7.3	23.7	1.2	5.1
Iron (mg) ---------	3.2	.4	12.5	3.4	.6	17.6

The recalls appear to be only slightly more variable for most nutrients than the weighing method, which is taken to be most accurate, but for iron the weighed portion showed greater variation. As to the iron data, chemical analyses can give variable results owing to differences in food samples or even recording errors.

Another group studied the reliability of the 24-hour recall method with even younger children (Emmons and Hayes, 1973). By comparing observed consumption of lunch with the child's recall and also by comparing the child's recall of food eaten at home with the mother's recall of food eaten by the child at home (two respondents, observer and subject, for the same subject and the same time), the authors suggested that young children can provide information on their diet as accurately as their mothers. The children were in grades 1-4 (N=431, 6-12 years of age). Their ability to recall the lunch they had eaten improved with age; children in grade 1 remembered an average of 60 percent of the foods and children in grade 4 an average of 81 percent. Comparison of recalls of mother and child showed less agreement.

A third study comparing recall by the child of his school lunch with his observed consumption included 94 children, 9-18 years of age, and 3 groups on 3 succeeding days (Meredith et al., 1951). The authors reported that the children on the third day did better than children on the first 2 days, possibly because they knew they were going to be interviewed. They were obviously more aware of what they were eating. Forty-seven percent (45 children) omitted from 1 to 4 foods and 71 percent reported a quantity less than was eaten. There appeared to be a tendency for greater underreporting as the number of foods increased. Nevertheless the authors found that small differences occurred in calculated nutrient analyses, although there was lack of agreement when recorded items were compared.

Thus children appear to be able to report what they have eaten. Even first and second graders could help their mothers give a more reliable and accurate recall than the mothers alone could supply.

7.3.2 Validity

The critical question underlying evaluation of the validity of the recall procedure is the degree to which the respondent can remember kinds and amounts of all foods eaten in the period to be reported.

The ability to remember accurately what was eaten was investigated by comparing 24-hour recall data from 200 older (at least 65 years of age) and 100 younger (aged 20-40 years) adults (Campbell and Dodds, 1967). The menus eaten within the 24-hour period were known as some of the participants were patients with lung disorders in a hospital and others were from a retirement home. Any foods added to the recalled intake during probing were used as an indicator of forgetfulness and their caloric value was used in evaluating memory loss.

The results indicated that much information can be picked up by careful probing by interviewers, possible only in the recall procedure. A mean of 537 kilocalories of food energy was probed from the younger subjects compared with 675 kilocalories from the older persons, a significant difference at the 0.05 level. For the older subjects the mean caloric value of food energy obtained by probing of the men was 785, for women, 427. The food energy caloric values obtained by probing the memory of institutionalized subjects (menu used as reference) were then compared with corresponding values of subjects living at home (no menu available as reference). For institutionalized older men and women, 35 and 28 percent respectively, of the total food energy was a result of probing; for older men and women living at home, corresponding values were 18 and 13 percent. The authors believed these figures indicated the underestimation that may exist when using the 24-hour recall with groups of older individuals. This generalization is risky because institutionalized people are not likely to be representative of all elderly people.

In upstate New York, interviewers used the list-recall method to obtain information on the kinds and amounts of food used in the household for the previous 7 days from 283 low-income elderly households of 1 or 2 persons (LeBovit and Baker, 1965). Following this the respondent was asked to recall meals and snacks eaten at home and away from home for the previous 2 days. The investigators found differences of up to 60 percent in the amounts of food energy calculated from food used in the households in a week as compared with the food eaten by the individuals in 2 days. For other nutrients the differences between the two recalls ranged from 40 to 80 percent. The authors concluded that "food reported as used by households often tends to be overestimated, whereas diets reported consumed by individuals may be underestimated." (Ibid., 1965, p. 64)

Among the several reasons for these differences were (1) the list-recall method doubled as a memory aid, whereas no such function was served with the 2-day recall schedule. Examination of the two sets of supposedly completed schedules revealed that on the 2-day recalls respondents apparently omitted sauces, salad dressings, beverages, and spreads on bread and supplied incomplete descriptions of foods eaten. (2) Men may be less able to recall, describe, and estimate amounts of food. Households made up of one male member (N=23) showed the largest gap between average nutrient values of foods used in households and food intakes of individuals. (3) Quantities of food as used in the household (as pounds, quarts, dozens) appear to be easier to estimate than quantities of food as served or left on the plate in mounds or pieces. (4) Discrepancies were also due to the computation method.

Further indications of the accuracy of recall data are shown in three studies where data were collected by the recall method and compared with data derived from 7-day records. One research team concluded that the 7-day record and the 24-hour recall tended to give the same mean estimate for most nutrient intakes and under some conditions the 24-hour recall could be substituted for the 7-day record in estimating group intakes. However, "for an individual, the 24-hour recall did not give the same estimate of intake as the 7-day record." (Young et al., 1952c, p. 220)

Similarity in group averages obtained with the 7-day recall and 7-day record procedures when the recall preceded the record kept by 59 businessmen suggested either method could have been used to calculate averages of food and nutrient intakes (Adelson, 1960). For individuals there were differences between the record and recall weeks, but the differences were no greater than between 2 weeks of records that had been studied 1 year earlier. There was no evidence that one method gave better results than the other, according to that researcher.

In contrast to this finding from a study of men's diets, Trulson (1954) found that recalls by children yielded lower mean values and larger standard deviations. She compared averages from 7-day estimated records with averages of at least three 24-hour recalls, as previously described.

The recall procedure is much less likely to change consumption behavior than the record procedure, particularly that using weighed records. Thus it appears to provide a more accurate basis for study of relationships among foods consumed. As noted earlier, Ohlson et al. (1950) found recalls gave higher mean values for food energy and protein, along with higher standard deviations, than did weighed records because of cutback on snacking when keeping the weighed record.

Concurrent validation--comparing results provided by two methods of measuring intake at the same time--was examined in the three studies of children's recall of school lunch, previously mentioned. The researchers concluded that the recall method appeared to give valid data (Emmons and Hayes, 1973; Meredith et al., 1951; Samuelson, 1970).

Simultaneous comparisons of weighed records, estimated (measured) records, recalls, and chemical analysis of food as served to 49 children, 10-15 years of age, in an English orphanage were conducted over a 3-day period (Bransby et al., 1948). Averages derived from weighed records and recalls were closer than averages from weighed and estimated (measured) records. Recalls gave small but significant underestimates of mean nutrients calculated from weighed records for all except fat. However, for a group of 88 boys, 12 years of age and living at home, Bransby et al. also found recalls gave higher estimates for food energy, protein, carbohydrate, and iron.

In another study, 20 pregnant women who were cooperating in a weighed diet survey were asked to recall in the middle of the survey week what they had eaten during the previous 24 hours (Thomson, 1958). Mean energy intake for recalls was 2,140 calories and for the weighed record, 2,574 calories or 17 percent higher. The subjects with more to remember were apt to omit more items on recall. Even memory aids did not contribute a reliable estimate of quantities. Thus there are conflicting reports in the literature about the validity of the 24-hour recall in comparison with records when the group of individuals includes fewer than 50 individuals. For larger groups the comparison did not show significant differences.

Next to be considered is construct validation for the recall method. The first approach is comparison of percentage standard errors for 1 day's diets, as recalled in the USDA spring 1965 survey, with those for 3 days, obtained with the estimated record procedure. The latter set, reported in table 4, was derived by the reviewers from published data for the fall 1965 North Central Regional Survey of Preschool Children (Eppright et al., 1972). Similar data were developed for several corresponding sex-age groups in the North Central Region of the USDA spring 1965 survey of individuals (U.S. Department of Agriculture, 1972a), where the 24-hour recall was used. The means and percentage standard errors for food energy and four nutrients for several sex-age groups in the USDA survey are presented in table 6.

The percentage standard errors for the North Central Regional Survey of Preschool Children (table 4) are lower than for the USDA survey (table 6), but the groups are three to five times larger in the preschool survey and the diet records were collected for 3 days, contributing to less variability of the means. The means of the two surveys for food energy and protein are comparable. For calcium, vitamin A, and vitamin C the means in the preschool survey were higher, but vitamin and mineral supplements were included in those nutrient calculations, whereas such supplements were not included in the USDA survey calculations. Mean intakes of vitamins A and C were much more variable for the 1 day in the USDA survey as compared with the 3 days of data in the preschool survey. However, the USDA data for older children yielded smaller relative standard errors because of their larger groups. Probably because of the more scientific sample, when compared with the New York State Nutrition Survey, the statistics for protein in the USDA survey showed greater reliability, despite the large number of children in the New York survey.

TABLE 6.--Means and percentage standard errors for food energy and 4 nutrients from 24-hour recall data for selected groups of children in North Central Region, spring 1965[1/]

Food energy and nutrient	Boys, age group in years[2/]				Girls, age group in years[2/]			
	3 (62)	5 (69)	9-11 (188)	14-16 (157)	3 (59)	5 (70)	9-11 (182)	14-16 (170)
Food energy:								
Mean (kcal)	1,738.0	1,858.0	2,469.0	3,084.0	1,544.0	1,787.0	2,029.0	2,173.0
Percentage standard error	4.4	4.1	2.6	2.8	4.4	4.6	2.2	3.0
Protein:								
Mean (g)	67.5	69.7	92.3	115.8	59.1	70.5	75.8	81.9
Percentage standard error	5.1	4.7	2.6	3.0	6.0	6.0	2.4	3.1
Calcium:								
Mean (mg)	948.0	920.0	1,145.0	1,359.0	871.0	871.0	994.0	962.0
Percentage standard error	5.0	5.5	3.0	3.7	6.5	6.5	3.0	4.2
Vitamin A:								
Mean (IU)	3,230.0	7,319.0	5,626.0	7,749.0	4,349.0	5,097.0	5,027.0	5,331.0
Percentage standard error	19.3	23.7	12.9	17.7	34.5	28.9	9.7	18.1
Vitamin C:								
Mean (mg)	45.0	55.0	68.0	74.0	44.0	54.0	62.0	68.0
Percentage standard error	12.0	16.0	5.9	6.6	13.2	10.2	6.5	6.2

[1/] Data from U.S. Dept. Agr., Agr. Res. Serv. (1972a).
[2/] Number of cases in parentheses.

Another way to evaluate the data obtained by the recall method is to consider the degree to which the percentage contributions of specified food groups to the total intake of a nutrient varies among population groups. A comparison was made of the proportions of food energy contributed by quantities of each of 12 food groups eaten by 9- to 11-year-old boys (N=88) and by girls in the same age group (N=74) in U.S. households with incomes under $3,000. The findings indicated differences of less than 0.5 percentage points for 7 out of 12 food groups (U.S. Department of Agriculture, 1972a). Boys obtained relatively more of their 2,028 calories from potatoes and sweets but relatively less from other vegetables, fruits, fats, oils, and beverages than the girls. The food energy content of the girls' food intake averaged 1,790 calories. These relationships of the parts of the food intake to the total are well in line with nutritionists' expectations for younger children's diets.

The representativeness of the data derived from 24-hour recalls appears to be greater than for data from other dietary collection methods, mainly because probability samples are easier to maintain. As Mongeau (1974, p. 17) stated, the 24-hour recall is applicable to wide population groups regardless of age, education, and intelligence. Cooperation required is minimal and "compared to more sophisticated methods the 24-hour recall is likely to get the highest number of participants."

The representativeness of the information derived from 24-hour recall data is likely to be improved by collecting data for 1 or 2 additional days, either by recall or record. Evidence in the literature cited indicates that this addition will reduce variability among individuals by averaging out the day-to-day fluctuations in actual intakes.

The representativeness of the data derived from 24-hour recalls contributes materially to their predictive validity for making many policy decisions. Several examples follow. Nutrition labeling requires information regarding portion sizes and their frequency as commonly used by various sex-age groups in the population. Such information from the spring 1965 USDA survey is being utilized by both business firms and government agencies. Decisions on enrichment, fortification (as for iron in flour), and genetic engineering (as for potatoes with high vitamin C content) of foods are currently taking into account the level of usage of the particular foods in the diets of the target groups with the objective of improving their diets. More precise targeting for food subsidy programs and for nutrition education depends on intensive statistical analysis of the actual food behavior of subgroups in the U.S. population, such as the elderly, ethnic, and poverty-stricken groups. The 1965 survey of food intake of individuals (U.S. Department of Agriculture, 1972a) has supplied data for such analyses. However, changes in food supplies and food prices in recent years have undoubtedly affected the food behavior of these vulnerable groups. Until another survey is made, predictions of relationships are risky.

7.3.3 Respondent Burden

Young et al. (1952c) found that the 24-hour recall involves considerably less participant time and cooperation than the 7-day record. Therefore a more representative population sample would be possible.

In the Washington, D.C., pilot survey (U.S. Department of Agriculture, 1963), interviewers spent about 1-3/4 hours per family when they collected a household schedule and the individuals' reports for 1 day. Without the household schedule but with demographic information, the interviewer averaged 54 minutes. Following the household schedule, reporting the individual members' 24-hour intake required 29 minutes. When the household schedule was omitted, the respondent was not yet oriented to reporting the food data and 43 minutes were required for the individual 24-hour intake reports. For the individual schedules, the first schedule in the household took longest and the next three about 30 percent as long.

7.3.4 Field Survey Costs

The 24-hour recall averaged about 30 minutes in the Washington, D.C., pilot study (U.S. Department of Agriculture, 1963). One visit to 70 percent of the households was sufficient to complete the individual schedules, and 21 percent of the households required two calls.

Adelson (1960) stated that the cost of her study would have been less per diet if done by the recall method. One home visit of about 1 hour was required by interviewers to obtain information for a recall, whereas just the initial visit to start the weighed record lasted an hour. Interviewers collecting recall data require longer instruction and apparently need a better background in food habits and greater judgment than those using the record method, according to Adelson.

7.3.5 Data Processing Costs

Adelson (1960) compared processing steps for the recall and weighed record. Summarizing food intakes from recalls took longer than weighed data because quantities of food had to be converted from household measures to grams. Time required for such other operations as editing, coding, tabulating, and calculating was not affected by the choice of the data collection method.

7.4 Recall of Past Food Intake: Dietary History Method

The dietary history method was originally developed to measure diets for research studies of human growth and development. The usual or average food and nutrient intakes over a period of time were sought in order that they might be correlated with clinical and biochemical measurements to evaluate nutritional status. The rationale was that clinical and laboratory signs and findings may appear as a result of earlier, longtime food habits. Current intake may not reflect usual intake and so may have less value in evaluating the nutritional status of a particular individual. (Burke, 1947; Young, 1965)

The taking of dietary history records requires a highly trained, mature interviewer with a nutrition background. The interviews average about 1 hour, yet the recorded data are not exact. If an individual does not have an eating pattern, a dietary history cannot be obtained. A large probability sample appears impossible to obtain.

7.4.1 Reliability

Dietary history methodology is not standardized. The methods described in the literature are each different and usually simpler modifications of the Burke method. Few have been utilized in surveys designed to use probability samples. Accordingly the reliability of data from dietary histories cannot be appraised by using standard errors.

One nutritional epidemiologic study came close to meeting the requirements for a probability sample. It used the fifth and sixth graders in six schools in New York City as subjects. About 80 percent (N=642) participated. A simplified modification of the dietary history was developed, requiring about 20 minutes for a nutritionist to administer. Biochemical, physical, and clinical relationships to dietary intakes were sought. Reliability was judged to exist when relationships had empirical support. (Baker et al., 1967; Christakis et al., 1968; Ziffer et al., 1967)

A dietary history method was chosen for the Framingham study of coronary heart disease, because long-term dietary characteristics of the individuals (volunteers) were desired for correlation with the possible development of heart disease. All interviewers were nutritionists. Reliability was investigated by repeat interviews, 2 years apart, by the same nutritionist and by two different nutritionists. For the same nutritionist, correlations of data obtained 2 years apart ranged from 0.52 to 0.92; for two different nutritionists, the range of correlation coefficients was from 0.27 to 0.89. Most sets of correlations between nutrients were not significantly different whether done by the same or different nutritionists. (Dawber et al., 1962; Mann et al., 1962)

A group of Israeli researchers purported to study the reliability (repeatability) of yet another modification of the dietary history procedure, using the same nutritionist to do repeat interviews about 6-8 months apart (Reshef and Epstein, 1972). They considered correlations of 0.63-0.89 satisfactory. Because the interviewing is so critical to the repeatability of the results, use of the same nutritionist provides partly a test of the interviewer and partly a test of the method. Thus there is some research evidence that nutritionists who are highly trained and experienced in taking dietary histories can achieve a satisfactory degree of repeatability.

> The repeatability of diet histories of 10 dietitians was studied. The fact that histories were reproducible within about 10 percent for groups and individual averages does not imply that they were necessarily also good estimates of actual intake. (Celender, 1963, p. 58)

Repeat dietary histories of 39 Italian-American men, taken 2 years apart, showed for the group means no significant differences for food energy, fat, and protein according to paired t-tests (Trulson and McCann, 1959). However, the correlations were 0.59, 0.49, and 0.43, respectively, and the correlation for percentage of calories coming from protein was poor, 0.25, revealing considerable differences for individuals between the 2 years, although there was no significant difference for the group as a whole. McCann et al. (1961) commented that this may indicate the interview can be grossly inaccurate or that natural variation in eating patterns from one period to another may be rather large. Reed and Burke (1954), in evaluating reliability of the dietary history, found it to be only fair in measuring the average daily intake of a child over a year.

Tests of the reliability of dietary histories by comparing results with those from alternative measures for the same time with the same subjects indicate the likelihood of overestimation. Comparison of dietary history data to data obtained with the 7-day record and 24-hour recall procedures showed that the dietary history gave higher values for the mean nutrients of a group than either the record or the recall (Young et al., 1952c). The dietary history was inconsistent in estimation of quantitative intakes from population to population and the bias might vary with the type of subject or be due to the skill of the interviewer. Some of the dietary histories of children overestimated results by 25-60 percent as compared with the 7-day record (Young et al., 1952a).

Other researchers have found that dietary histories overestimate intakes as compared to more exact methods (Celender, 1963; Stevens et al., 1963; Trulson, 1951). One group attributed some of the overestimation to missed meals, which were forgotten by subjects (Stevens et al., 1963). Huenemann and Turner (1942) compared dietary histories and weighed records of actual intake and concluded the histories had little quantitative value.

7.4.2 Validity

The apparent systematic upward bias in data from dietary histories raises doubts regarding the accuracy of the method (Celender, 1963; den Hartog et al., 1965; Trulson, 1951; Young et al., 1952a). The chief reason for the difference between dietary histories and records seemed to be that the patients actually did not know what or how much they ate (Huenemann and Turner, 1942, p. 563).

Several researchers have investigated the validity of this method by comparing concurrent studies with other methods (Celender, 1963; Huenemann and Turner, 1942; Trulson, 1951; Young et al., 1952a). Young et al. (1952a, p. 127) concluded that it was impossible to predict intake measured by the 7-day record from a dietary history. They (1952c, p. 219) also found that the nutrient measures derived from the 24-hour recall differed from those obtained from the dietary history for individuals.

Construct validity, ability to capture actual variability in the concept measured, may be satisfactory for the dietary history method when used to assess food habits for insights into health status, but in U.S. Department of Agriculture surveys a different concept is being measured.

The dietary history procedure rates poorly in terms of the representativeness of the results because most researchers have not attempted to use probability samples with any but simple modifications of the method.

7.4.3 Respondent Burden

Dietary histories place relatively little burden on respondents because they require only 20-60 minutes when administered by well-trained nutritionists. In effect the burden is shifted to the nutritionist interviewer (den Hartog et al., 1965).

7.4.4 Field Survey Costs

Field survey costs would be affected by the need for interviewers educated in nutrition and having special skills and extra training (den Hartog et al., 1965; Young et al., 1952a).

7.4.5 Data Processing Costs

No information on data processing costs was found. Without standardized procedures, each study would vary depending on the amount of calculation from detailed data required.

7.5 Recall of Past Food Intake: Frequency Method

For certain purposes a quantitative appraisal of dietary habits, describing how often specified foods are eaten in a given time period, may be sufficient. This is especially true when quantities of nutrients are not required. The frequency recall is a rather simple procedure. By interview or self-administered questionnaire is recorded the number of times certain foods or combination of foods are eaten per day, week, month, or some other period of time. This method has been used to study the variations in dietary characteristics of a group of Irish and Italian men (Stefanik and Trulson, 1962). Abramson et al. (1963) considered the main use of the food frequency interview in epidemiological studies to be a simple and economical tool, not sharp but adequate for some purposes.

Shortcut methods for scoring diets may use the frequency of a food occurrence and derive scores by assigning standard quantities for servings, possibly derived from a tabulation of amounts actually eaten (Marr et al., 1961).

The frequency method provides descriptive data, but it is not useful for collecting detailed food intake data.

7.6 Combination Recall-Record Method

The combination of record and recall procedures appears to be the pragmatic approach that squares best with the real life situation. Respondents keeping records may neglect to record their intake immediately after eating. As a result, when the interviewer comes to collect and review the record, she frequently has to assist the respondent to complete it by recall; hence the purported record is in actuality a record-recall combination.

This has been the experience of Eppright et al. (1972) and Owen et al. (1974), who considered their method to be the record. Other researchers have planned specifically and apparently with success to use a combination recall-record. Using a random sample of black children, Futrell et al. (1971) obtained during a home interview a 3-day recall and then gave instructions to the homemaker about keeping a 4-day record. Patterson (1971) collected dietary information from children in the Phoenix, Ariz., area using a 24-hour recall and a 2-day record.

The U.S. Bureau of the Census (1964) in the Total Diet Feasibility Study found the most satisfactory method to be a combination of recall and record, a compromise to obtain the satisfactory participation rate of the recall with the supposedly more accurate information of the record.

No tests of reliability or validity of the combination methods were found in the literature.

7.7 Conclusions

Objectives and the population of interest are uppermost among factors governing the selection of a data collection method for large-scale surveys of food intake of individuals. In addition to these considerations, the review of literature for this section revealed several areas of methodology needing further study. When the four methods listed in table 7 were evaluated using the specified criteria, no one method was consistently advantageous over all others. Consequently, researchers will have to decide which trade-offs are most relevant to their objectives.

For random samples in large-scale surveys the 24-hour recall method had high response rates and seemed to have the highest reliability. This was not invariably true as was manifest in the less than desirable participation rates in Nutrition Canada (1973). Light respondent burden, along with the ease in taking a 24-hour recall from most population groups, appears to be among the reasons for the higher response rates for this method.

Regarding data reliability or repeatability, which overlaps the accuracy aspect of validity, no one method was clearly the best choice. The dilemma is due to the inability to measure actual variation apart from error variation, which exists in dietary intake over time. Estimated records for 1 day (Chalmers et al., 1952) or for 3 days (Eppright et al., 1952) were considered sufficient to give reliable average dietary intakes for groups but not for individuals. The 24-hour recall was recommended as a substitute for the estimated record if investigators wished to avoid the risk of respondents changing their customary food patterns (Young et al., 1952c). Several researchers (Emmons and Hayes, 1973; Samuelson, 1970) concluded that the 24-hour recall gave acceptable agreement with actual consumption. However, these two methods could not be depended on to give the same results.

Some evidence indicated the existence of upward bias in both the weighed and estimated record methods for the first day or week of a survey. Upward bias in the dietary history and downward bias in the recall method were also frequently reported. A combination method of recall with the estimated record may offset the opposing weaknesses of these two methods and would seem to warrant investigation. The number, spacing, and selection of days for a survey affect the reliability of data, but further research is needed to measure their effects, whether jointly or separately.

Five aspects of validity were reviewed. To obtain absolute accuracy of the quantity of food eaten, weighing would be the choice. However, weighing food and keeping a record disrupt customary routines and systematic bias may result. The recall method has the disadvantage of relying on memory. All these methods except weighed records may be biased because respondents are unable to estimate portion sizes accurately. Whether respondents availed themselves of the chance to measure portions when keeping an estimated record was not clear. Probing by skilled interviewers helped respondents remember,

TABLE 7.--Summary of evaluations of 4 quantitative dietary collection methods used in surveys of food intake of individuals with references

Criterion for evaluation[1/]	Dietary collection method			
	Weighed record	Estimated record	24-hour recall[2/]	Dietary history
Reliability:				
Related to sampling ----	Random sample very difficult to obtain because of low response rate (Burrill et al., 1959; den Hartog et al., 1965; Leverton and Pazur, 1957).	Random sample difficult to obtain because of low response rate (Marr, 1971; Trulson et al., 1949). Considered less difficult than weighed record (Marr, 1971).	Satisfactory samples obtained (Abraham et al., 1974; Swanson et al., 1959; U.S. Department of Agriculture, 1972a). Satisfactory samples not obtained (Morgan, 1959; Nutrition Canada, 1973; U.S. Department of Health, Education and Welfare, 1972).	Satisfactory sample of school children obtained (Christakis et al., 1968). Used with volunteers (Burke, 1947; Dawber et al., 1962; Mann et al., 1962).
Repeatability of data -----	Greatest variation in day-to-day records, but 7 days considered sufficiently reliable for group (Adelson, 1960; Widdowson, 1947). 1 week not considered reliable estimate of individual's intake (Chappell, 1955; Pry et al., 1963; Huenemann and Turner, 1942; Keys et al., 1966; Yudkin, 1951).	Satisfactory for group if 3 consecutive days are included (Eppright et al., 1955; 1972; Trulson et al., 1949). 1 day is satisfactory if for group (Chalmers et al., 1952). Estimated record gave higher nutrient values than weighed record (Eppright et al., 1952), but results were inconsistent in other studies (Bransby et al., 1948; Young et al., 1952b). Correlation between days (Hankin et al., 1967; Trulson, 1951) and no correlation (Chalmers et al., 1952). Evidence of actual variability (Celender, 1963; Young et al., 1953b).	Less repeatability if time lapse exceeded 3-4 months (Balogh et al., 1971). Good agreement with actual consumption for groups of children (Emmons and Hayes, 1973; Samuelson, 1970). Advance notice may help recall (Meredith et al., 1951). More underreporting with increased number of items (Meredith et al., 1951).	Correlation for repeat interviews using same nutritionist was 0.63-0.89 (Reshef and Epstein, 1972). Good repeatability of group averages does not necessarily imply good estimates for individuals (Celender, 1963; Reed and Burke, 1954; Trulson and McCann, 1959). Little quantitative value (Huenemann and Turner, 1942). Likely to give higher estimate than other methods (Celender, 1963; Stevens et al., 1963; Trulson, 1951; Young et al., 1952c).

TABLE 7.--Summary of evaluations of 4 quantitative dietary collection methods used in surveys of food intake of individuals with references--Con.

Criterion for evaluation[1]	Dietary collection method			
	Weighed record	Estimated record	24-hour recall[2]	Dietary history
Validity:				
Accurate (absence of systematic error)	Most accurate absolute measure (Marr, 1971). Upward bias at start (Yudkin, 1951). Downward bias if too much trouble (den Hartog et al., 1965; Ohlson et al., 1950). Entry for missing data (Beegle et al., 1954; Dieckmann et al., 1951). Use of composition tables for nutrient content of food gave acceptable accuracy (Widdowson and McCance, 1943) and some less than acceptable accuracy (Leverton, 1937; Manalo and Jones, 1966; Walberg and Adams, 1965).	Upward bias at start (Celender, 1963; Young et al., 1953b). Some respondents unwilling to report accurately (Paul et al., 1963). Bias because of inability to estimate portion sizes accurately (U.S. Bureau of Census, 1964; Young et al., 1952b; 1953a). Influenced by interviewers (Church et al., 1954; Steele et al., 1951). "Day" effect (Cellier and Hankin, 1963; Eppright et al., 1952; Leverton and Marsh, 1939).	Ability to remember increased, by probing, but evidence of underestimation (Campbell and Dodds, 1967). For group average, 24-hour recall may substitute for estimated 7-day record (Young et al., 1952c). 7-day recall by adult men similar to 7-day weighed records (Adelson, 1960). Recalls by children gave lower averages than 7-day estimated records (Trulson, 1954). Avoids change in usual consumption pattern (Ohlson et al., 1950).	Upward bias (Celender, 1963; den Hartog et al., 1965; Trulson, 1951; Young et al., 1952a).
Concurrent (2 measures of same concept)	Averages for estimated records exceeded means from weighed records (Eppright et al., 1952). Positive correlation with biochemical measures (Dieckmann et al., 1951; Huenemann and Turner, 1942).	Positive correlation with biochemical measures (Eppright et al., 1952; Owen et al., 1974).	Compared with weighed records recall gave lower estimates (Bransby et al., 1948; Thomson, 1958), higher (Bransby et al., 1948), and acceptable (Adelson, 1960; Samuelson, 1970).	Dietary history resulted in higher estimates than other methods (Celender, 1963; den Hartog et al., 1965; Huenemann and Turner, 1942; Young et al., 1952a, 1952c).

Construct (degree to which variability in concept is measured) ---	Gives precise measure of intake, but not necessarily of customary consumption (Marr, 1971). Records for 3-5 days capture most of information (Fidanza and Fidanza-Alberti, 1967; Heady, 1961). Less than 1 week is too short a period (Keys et al., 1966; Yudkin, 1951).	For group average, 1 day is sufficient, but not for estimate of individual's intake (Young et al., 1952b).	Relative standard errors depend somewhat on group size, but 24-hour recall compared favorably with estimated record (tables 4 and 6, sections 7.2.1 and 7.3.2).	Uncertain.
Representative (measure on level taken as an indicator on more general level) ------	Low response rates adversely affect generalizability of results when weighed record is chosen method (Adelson, 1961; Thomson, 1958).	Consecutive 3-day records did not give same results as 1 day, 7 days, or combinations of nonconsecutive days (Eppright et al., 1952; Hankin et al., 1967; Trulson, 1951). Weekend days different from weekdays (Cellier and Hankin, 1963; Leverton and Marsh, 1939; McHenry et al., 1945). Less than desirable response rate (Owen et al., 1974; U.S. Bureau of Census, 1964).	More representativeness than for other methods because response rates are higher (Mongeau, 1974; U.S. Bureau of Census, 1964).	Uncertain.
Predictive ----	Low for large-scale surveys.	Good support from large-scale surveys (Eppright et al., 1972; Owen et al., 1974).	Good response rate contributes to usefulness for predictive validity (U.S. Department of Agriculture, 1972a).	Uncertain.

TABLE 7.--Summary of evaluations of 4 quantitative dietary collection methods used in surveys of food intake of individuals with references--Con.

Criterion for evaluation[1]	Dietary collection method			
	Weighed record	Estimated record	24-hour recall[2]	Dietary history
Respondent burden	Heavy burden (den Hartog, et al., 1965; Eppright et al., 1952).	Heavy burden (Mongeau, 1974; Young et al., 1952b).	Light burden (U.S. Department of Agriculture, 1963; Young et al., 1952c).	Light burden (den Hartog et al., 1965).
Field survey costs	Higher costs if more than 1 interviewer visit is required (Adelson, 1960; den Hartog et al., 1965; Marr, 1971).	Higher costs if more than 1 interviewer visit is required (Eppright et al., 1972; Owen et al., 1974).	Lower cost if only 1 visit is necessary (Adelson, 1960).	Higher cost because interviewer must be highly skilled nutritionist (den Hartog et al., 1965; Young et al., 1952a).
Data processing costs	Decreased costs because foods do not require conversion to common weight (Young and Trulson, 1960).	Increased costs if household measures must be converted to common measure.	Increased costs when household measures must be converted to weight (Adelson, 1960).	Increased costs if household measures must be converted to common measure.

[1] Kerlinger (1973).
[2] Several references refer to longer periods of recall.

but underestimation and overestimation still remained a problem for the recall method. There was considerable agreement that the dietary history method showed upward bias compared with the other three methods.

Concurrent validity, or correlation between two measures of the same concept, was shown for both the weighed and the estimated method with a biochemical measurement. Although Young et al. (1952c) stated that the 24-hour recall may substitute for the 7-day estimated record when group averages are the objective, the results were not the same. Further investigation might clarify this contradiction.

Construct validity, or degree to which variability in a concept is measured, judging from relative or percentage standard errors for sizable groups, appeared to be about equally satisfactory for the 3-day estimated record and the 24-hour recall. Representative and predictive validity of the 24-hour data appeared to be greater than for data gathered by the other three methods because of more favorable response rate.

Field survey costs are heavily affected by the skill and education of the interviewer and the number of visits required. The 24-hour recall requires only one interview, whereas the weighed and estimated methods may require more unless mailing is relied on. Data processing costs would be less for the weighed record because quantity data are obtained in grams and conversion from household measures is not necessary.

LITERATURE CITED

Abraham, S., Lowenstein, F., and Johnson, C.
1974. Preliminary Findings of the First Health and Nutrition Examination Survey, United States, 1971-72: Dietary Intake and Biochemical Findings. U.S. Dept. Health, Educ. and Welfare, Natl. Cent. for Health Statis., Pub. Health Serv. DHEW Pub. No. (HRA) 74-1219-1, 183 pp.

Abramson, J. H., Slome, C., and Kosovsky, C.
1963. Food Frequency Interview as an Epidemiological Tool. Amer. Jour. Pub. Health 53: 1093-1101.

Adelson, S. F.
1960. Some Problems in Collecting Dietary Data From Individuals. Amer. Dietet. Assoc. Jour. 36 (5): 453-461.

_____
1961. Practical Procedures for Dietary Surveys. 3d Internatl. Cong. Dietet. Proc. [London] 1961, pp. 158-163.

Baker, H., Frank, O., Feingold, S., and others.
1967. Vitamins, Total Cholesterol, and Triglycerides in 642 New York City School Children. Amer. Jour. Clin. Nutr. 20 (8): 850-857.

Balogh, M., Kahn, H. A., and Medalie, J. H.
1971. Random Repeat 24-Hour Dietary Recalls. Amer. Jour. Clin. Nutr. 24: 304-310.

Bavly, S.
1972. Levels of Nutrition in Israel, 1968-69. Family Expenditure Survey Part II. 102 pp. Min. Educ. and Cult., Col. Nutr. and Home Econ., Cent. Bur. Statis., Jerusalem.

Becker, B. G., Indik, B. P., and Beeuwkes, A. M.
1960. Dietary Intake Methodologies - A Review. 123 pp. Mich. Univ. Sch. Pub. Health, East Lansing.

Beegle, R. M., Roberts, P. H., Howard, N., and others.
1954. Weights of Foods Eaten per Meal by 242 Women 30 to 92 Years of Age. Mich. Agr. Expt. Sta. Tech. Bul. 244, No.-Cent. Region. Pub. 37, 32 pp.

Bigwood, E. J.
1939. Guiding Principles for Studies on the Nutrition of Populations. 281 pp. League of Nations, Health Organ., Tech. Comn. Nutr., Geneva.

Bransby, E. R., Daubney, C. G., and King, J.
1948. Comparison of Results Obtained by Different Methods of Individual Dietary Survey. Brit. Jour. Nutr. 2: 89-110.

Burk, M. C.
1958. Some Analyses of Income-Food Relationships. Amer. Statis. Assoc. Jour. 53: 905-927.

_____
1961. Measures and Procedures for Analysis of U.S. Food Consumption. U.S. Dept. Agr. Agr. Handb. 206, 118 pp., illus.

_____
1969a. Consumption of Dairy Products: An Analysis of Trends, Variability, and Prospects. Minn. Univ. Agr. Expt. Sta. Tech. Bul. 268, 130 pp.

_____
1969b. Food Expenditures by Upper Income Families. Minn. Univ. Agr. Expt. Sta. Tech. Bul. 269, 323 pp.

_____ and Lanahan, T. J., Jr.
1958. Use of 1955 Food Survey Data for Research in Agricultural Economics. Agr. Econ. Res. X (3): 79-98.

Burke, B. S.
1947. The Dietary History as a Tool in Research. Amer. Dietet. Assoc. Jour. 23: 1041-1046.

Burrill, L. M., Alsup, B., Schuck, C., and others.
1959. Evaluation of the Self-Chosen Weighted Diets of 402 Women 30-97 Years of Age in Seven North Central States. S. Dak. State Col., Home Econ. Dept., Agr. Expt. Sta. Bul. 478, No.-Cent. Region. Pub. 105, 19 pp.

Campbell, V. A., and Dodds, M. L.
1967. Collecting Dietary Information From Groups of Older People. Amer. Dietet. Assoc. Jour. 51 (1): 29-33.

Celender, I. H.
1963. Reliability of Methods of Recording Dietary Intake of Women at Constant Weight. 161 pp. Univ. Microfilms, Ann Arbor, Mich.

Cellier, K. M., and Hankin, M. E.
1963. Studies of Nutrition in Pregnancy I. Some Considerations in Collecting Dietary Information. Amer. Jour. Clin. Nutr. 13: 55-62.

Chalmers, F. W., Clayton, M. M., Gates, L. O., and others.
1952. The Dietary Record - How Many and Which Days? Amer. Dietet. Assoc. Jour. 28: 711-717.

Chappell, G. M.
1955. Long-Term Individual Dietary Surveys. Brit. Jour. Nutr. 9: 323-339.

Christakis, G., ed.
1973. Nutritional Assessment in Health Programs. Amer. Jour. Pub. Health 63 (pt. 2), 82 pp. (Based on Nutr. Assessment Conf. Proc., Oct. 18-20, 1972.)

_____ Miridjanian, A., Nath, L., and others.
1968. A Nutritional Epidemiologic Investigation of 642 New York City Children. Amer. Jour. Clin. Nutr. 21 (1): 107-126.

Church, H. N., Clayton, M. M., Young, C. M., and Foster, W. D.
1954. Can Different Interviewers Obtain Comparable Dietary Survey Data? Amer. Dietet. Assoc. Jour. 30: 777-779.

Clark, F.
1969. The 1965-66 Food Consumption Survey: Scope, Methodology and Highlights. In N.C. State Univ. Agr. Policy Inst., Using Food Surveys in Consumer Education, pp. 43-56. Raleigh.

_____
1974. Recent Food Consumption Surveys and Their Uses. Fed. Proc. 33 (11): 2270-2274.

Dawber, T., Pearson, G., Anderson, P., and others.
1962. Dietary Assessment in the Epidemiologic Study of Coronary Heart Disease: The Framingham Study II. Reliability of Measurement. Amer. Jour. Clin. Nutr. 11: 226-234.

den Hartog, C., Van Schaik, Th. F. S. M., Dalderup, L. M., and others.
1965. The Diet of Volunteers Participating in a Long Term Epidemiological Field Survey on Coronary Heart Disease at Zutphen, the Netherlands. Voeding 26 (4): 184-208.

Dieckmann, W. J., Turner, D. F., Meiller, E. J., and others.
1951. Observations on Protein Intake and the Health of the Mother and Baby. II. Food Intake. Amer. Dietet. Assoc. Jour. 27: 1053-1058.

Dierks, E. C., and Morse, L. M.
1965. Food Habits and Nutrient Intakes of Preschool Children. Amer. Dietet. Assoc. Jour. 47: 292-296.

Emmons, L., and Hayes, M.
1973. Accuracy of 24-Hour Recalls of Young Children. Amer. Dietet. Assoc. Jour. 62: 409-415.

Eppright, E. S., Fox, H. M., Fryer, B. A., and others.
1972. Nutrition of Infants and Preschool Children in the North Central Region of the United States of America. World Rev. Nutr. and Dietet. 14: 269-332.

_____ Jebe, E., Sidwell, V., and others.
1955. Nutrition of 9-, 10- and 11-Year-Old Public School Children in Iowa, Kansas and Ohio. Iowa State Col., Iowa Agr. Expt. Sta. Bul. 434, No.-Cent. Region. Pub. 59, pp. 616-631.

_____ Patton, M. B., Marlatt, A., and Hathaway, M.
1952. Dietary Study Methods. V. Some Problems in Collecting Dietary Information About Groups of Children. Amer. Dietet. Assoc. Jour. 28: 43-48.

Fidanza, F., and Fidanza-Alberti, A.
1967. Dietary Surveys in Connection With the Epidemiology of Heart Disease: Reliability, Sources of Variation and Other Data From Nine Surveys in Italy. Voeding 28 (6): 244-252.

Ford, N. M.
1967. The Advance Letter in Mail Surveys. Jour. Mktg. Res. IV: 202-204.

Fry, P. C., Fox, H. M., and Linkswiler, H.
1963. Nutrient Intakes of Healthy Older Women. Amer. Dietet. Assoc. Jour. 42 (3): 218-222.

Futrell, M. F., Kilgore, L., and Windham, F.
1971. Nutritional Status of Negro Preschool Children in Mississippi. Evaluation of HOP Index. Amer. Dietet. Assoc. Jour. 59 (3): 218-223.

George, P. S., and King, G. A.
1971. Consumer Demand for Food Commodities in the United States With Projections for 1980. Calif. Univ. Div. Agr. Sci., Giannini Found. Monog. 26 (Mar.), 161 pp.

Guthrie, H. A.
1963. Nutritional Intake of Infants. Amer. Dietet. Assoc. Jour. 43: 120-124.

Hankin, J. H., Reynolds, W. E., and Margen, S.
1967. A Short Dietary Method for Epidemiologic Studies II. Variability of Measured Nutrient Intakes. Amer. Jour. Clin. Nutr. 20 (9): 935-945.

Heady, J. A.
1961. Diets of Bank Clerks. Development of a Method of Classifying the Diets of Individuals for Use in Epidemiological Studies. Roy. Statis. Soc. Jour., Ser. A (Gen.), 124: 336-361.

Hiemstra, S.
1968. Food Consumption, Prices, Expenditures. U.S. Dept. Agr., Econ. Res. Serv., Agr. Econ. Rpt. 138, 193 pp.

Huenemann, R., and Turner, D.
1942. Methods of Dietary Investigation. Amer. Dietet. Assoc. Jour. 18: 562-568.

Kerlinger, F. N.
1973. Foundations of Behavioral Research. Ed. 2, 741 pp. Holt, Rinehart and Winston, Inc., New York.

Keys, A.
1965. Dietary Survey Methods in Studies on Cardiovascular Epidemiology. Voeding 26 (7): 464-483.

_____ Aravanis, D., and Sdrin, H.
1966. The Diets of Middle-Aged Men in Two Rural Areas of Greece. Voeding 27: 575-586.

Lansing, J. B., and Morgan, J. N.
1971 Economic Survey Methods. 430 pp. Mich. Univ. Inst. Soc. Res., Ann Arbor.

LeBovit, C., and Baker, D.
1965. Food Consumption and Dietary Levels of Older Households in Rochester, New York. U.S. Dept. Agr., Home Econ. Res. Rpt. 25, 91 pp.

Leitch, I., and Aitken, F. C.
1950. Technique and Interpretation of Dietary Surveys. Nutr. Abs. Rev. 19 (3): 507-525.

Leverton, R.
1937. A Comparison of the Values Obtained by Calculation and by Analysis for the Iron Content of 85 Mixed Diets. Amer. Dietet. Assoc. Jour. 13: 139.

_____ and Marsh, A. G.
1939. Comparison of Food Intakes for Weekdays and for Saturday and Sunday. Jour. Home Econ. 31: 111-114.

_____ and Pazur, J.
1957. Food Practices and Nutritional Status of Typical Nebraska Families. Nebr. Univ. Agr. Expt. Sta. Misc. Pub. 5, 52 pp.

McCann, M. B., Trulson, M. F., and Stare, F. J.
1961. Follow-Up Study of Serum Cholesterol, Diet and Physical Findings of Italian-American Factory Workers. Amer. Jour. Clin. Nutr. 9: 351-355.

McHenry, E. W., Furguson, H. P., and Gurland, J.
1945. Sources of Error in Dietary Surveys. Canad. Jour. Pub. Health 36: 355-361.

Manalo, R., and Jones, J.
1966. The Content of Constant Diets: A Comparison Between Analyzed and Calculated Values. Amer. Jour. Clin. Nutr. 18: 339-342.

Mann, G., Pearson, G., Gordon, T., and Dawber, T.
1962. Diet and Cardiovascular Disease in the Framingham Study I. Measurement of Dietary Intake. Amer. Jour. Clin. Nutr. 11: 200-225.

Marr, J. W.
1971. Individual Dietary Surveys: Purposes and Methods. World Rev. Nutr. and Dietet. 13: 105-164.

_____ Heady, J. A., and Morris, J. N.
1961. Towards a Method for Large-Scale Individual Diet Surveys. 3d Internatl. Cong. Dietet. Proc. [London] 1961, pp. 85-91.

Mason, J. B., and Madden, C. S.
1971. Food Purchases in a Low-Income Negro Neighborhood: The Development of a Socio-Economic Behavorial Profile as Related to Movement and Patronage Farms. Amer. Mktg. Assoc. Spring and Fall Conf. Ser. 33, Combined Proc. 1971, 682 pp.

Meredith, A., Matthews, A., Zichefaase, M., and others.
1951. How Well Do School Children Recall What They Have Eaten? Amer. Dietet. Assoc. Jour. 27: 749-751.

Miller, D. S., Mumford, P., and Stock, M. J.
1967. Gluttony. 2. Thermogenesis in Overeating Man. Amer. Jour. Clin. Nutr. 20: 1223.

Ministry of Agriculture, Fisheries and Food, National Food Survey Committee.
1974. Household Food Consumption and Expenditure: 1972. 218 pp. Her Majesty's Stationery Office, London.

Mongeau, E.
1974. Dietary Assessment of Nutritional Status. In The Assessment of Nutritional Status, Miles Symp. Proc., June 25, 1973, pp. 13-27. Saskatchewan Univ., Canada.

Monsen, E., Kuhn, I., and Finch, C. A.
1967. Iron Status of Menstruating Women. Amer. Jour. Clin. Nutr. 20 (8): 842-849.

Moore, M. C., Judlin, B. C., and Kennemur, P. McA.
1967. Using Graduated Food Models in Taking Dietary Histories. Amer. Dietet. Assoc. Jour. 51: 447-450.

Morgan, A. F., ed.
1959. Nutritional Status U.S.A. Calif. Agr. Expt. Sta. Bul. 769, 131 pp.

National Research Council, Food and Nutrition Board, Committee on Nutrition Surveys.
1949. Nutrition Surveys: Their Techniques and Value. NRC Bul. 117, 144 pp.

Norris, T.
1949. Dietary Surveys, Their Technique and Interpretation. FAO Nutr. Studies No. 4, 108 pp.

Nutrition Canada.
1973. Nutrition Canada National Survey. 174 pp. Information Canada, Ottawa.

Ohlson, M. A., Jackson, L., Boek, J., and others.
1950. Nutrition and Dietary Habits of Aging Women. Amer. Jour. Pub. Health 40: 1101-1108.

Osgood, C. E.
1953. Method and Theory in Experimental Psychology. 800 pp. Oxford Univ. Press, New York.

Owen, G., Garry, P., Kram, K., and others.
1969. Nutritional Status of Mississippi Preschool Children. A Pilot Study. Amer. Jour. Clin. Nutr. 22 (11): 1444-1458.

_____ Kram, K. M., Garry, P. J., and others.
1974. A Study of Nutritional Status of Preschool Children in the United States, 1968-70. Pediatrics 53 (4), pt. II, sup., pp. 597-646.

Pao, E. M., and Burk, M. C.
1975. Portion Sizes and Day's Intakes of Selected Foods. U.S. Dept. Agr. ARS-NE-67, 70 pp.

Patterson, L.
1971. Dietary Intake and Physical Development of Phoenix Area Children. Amer. Dietet. Assoc. Jour. 59 (2): 106-110.

Paul, O., Lepper, M. H., Phelan, W. H., and others.
1963. A Longitudinal Study of Coronary Heart Disease. Circulation 28: 20-31.

Pekkarinen, M.
1970. Methodology in the Collection of Food Consumption Data. World Rev. Nutr. and Dietet. 12: 145-171.

Petrasovits, A., and Nargundkar, M.
1974. Statistical Considerations in the Assessment of Nutritional Status. In The Assessment of Nutritional Status, Miles Symp. Proc., June 25, 1973, pp. 65-74. Saskatchewan Univ., Canada.

Quackenbush, G. G., and Shaffer, J. D.
1960. Collecting Food Purchase Data by Consumer Panel--A Methodological Report on the MSU Consumer Panel, 1951-58. Mich. State Univ. Agr. Expt. Sta. Tech. Bul. 279, 73 pp.

Reagan, B., and Grossman, E.
1951. Rural Levels of Living in Lee and Jones Counties, Mississippi, 1945 and a Comparison of Two Methods of Data Collection. U.S. Dept. Agr. Inform. Bul. 41, 164 pp.

Reed, R. B., and Burke, B. S.
1954. Collection and Analysis of Dietary Intake Data. Amer. Jour. Pub. Health 44 (8): 1015-1026.

Reh, E.
1962. Manual on Household Food Consumption Surveys. FAO Nutr. Studies No. 18, 96 pp.

Reshef, A., and Epstein, L.
1972. Reliability of a Dietary Questionnaire. Amer. Jour. Clin. Nutr. 25: 91-95.

Samuelson, G.
1970. An Epidemiological Study of Child Health and Nutrition in a Northern Swedish County II. Methodological Study of the Recall Technique. Nutr. and Metabolism 12 (6): 321-340.

_____
1971. An Epidemiological Study of Child Health and Nutrition in a Northern Swedish County I. Food Consumption Survey. Acta Paediatrica Scand. Sup. 214, 44 pp.

Sprauve, M. E., and Dodds, M. L.
1965. Dietary Survey of Adolescents in the Virgin Islands. Amer. Dietet. Assoc. Jour. 47: 287-291.

Steele, B. F., Franklin, R. E., Smudski, V. L., and Young, C. M.
1951. Use of Checked Seven-Day Records in a Dietary Survey. Amer. Dietet. Assoc. Jour. 27: 957-959.

Stefanik, P. A., and Trulson, M. F.
1962. Determining the Frequency Intakes of Foods in Large Group Studies. Amer. Jour. Clin. Nutr. 11: 335-343.

Stevens, H., Bleiler, R., and Ohlson, M.
1963. Dietary Intake of Five Groups of Subjects. Amer. Dietet. Assoc. Jour. 42: 387-393.

Strong, J. A., and Passmore, R.
1967. Adaption to Overfeeding in Man. Nutr. Soc. Proc. 26 (2): 163.

Sudman, S.
1964. On the Accuracy of Recording of Consumer Panels. Jour. Mktg. Res. I (May): 14-19.

_____ and Ferber, R.
1971. Experiments in Obtaining Consumer Expenditures by Diary Methods. Amer. Statis. Assoc. Jour. 66 (336): 725-735.

_____ and Ferber, R.
1974. A Comparison of Alternative Procedures for Collecting Consumer Expenditures Data for Frequently Purchased Products. Jour. Mktg. Res. XI (May): 128-135.

Swanson, P., Willis, E. W., Jebe, E., and others.
1959. Food Intakes of 2,189 Women in Five North Central States. Iowa State Col., Iowa Agr. and Home Econ. Expt. Sta. Res. Bul. 468, No.-Cent. Region. Pub. 83, pp. 476-499.

Thomas, R., Rutledge, M., Beach, E., and others.
1950. Nutritional Status of Children XIII. Accuracy of Calculated Intakes of Food Components With Respect to Analytical Values. Amer. Dietet. Assoc. Jour. 26: 889-896.

Thomson, A. M.
1958. Diet in Pregnancy 1. Dietary Survey Technique and the Nutritive Value of Diets Taken by Primigravidae. Brit. Jour. Nutr. 12: 446-461.

Trulson, M. F.
1954. Assessment of Dietary Study Methods I. Comparison of Methods for Obtaining Data for Clinical Work. Amer. Dietet. Assoc. Jour. 30 (10): 991-995.

_____
1960. Cardiovascular Diseases. Subcommittee on Methodology for Diet Appraisal. Amer. Jour. Pub. Health 50 (10): 39-52.

_____ Hegsted, D. M., and Stare, F. J.
1949. New York State Nutrition Survey I. A Nutrition Survey of Public School Children. Amer. Dietet. Assoc. Jour. 25: 595-605.

_____ and McCann, M. B.
1959. Comparison of Dietary Survey Methods. Amer. Dietet. Assoc. Jour. 35: 672-676.

U.S. Department of Agriculture, Agricultural Research Service.
1972a. Food and Nutrient Intake of Individuals in the United States, Spring 1965. U.S. Dept. Agr. Household Food Consumption Survey 1965-66, Rpt. No. 11, 289 pp.

_____
1972b. Food Consumption of Households in the United States, Seasons and Year, 1965-66. U.S. Dept. Agr. Household Food Consumption Survey 1965-66, Rpt. No. 12, 217 pp.

U.S. Department of Agriculture, Bureau of Agricultural Economics.
1953. Consumption of Food in the U.S., 1909-52. U.S. Dept. Agr. Agr. Handb. 62, 249 pp.

U.S. Department of Health, Education and Welfare, Health Services and Mental Health Administration, Center for Disease Control.
1972. Ten-State Nutrition Survey, 1968-1970. DHEW Pub. No. (HSM) 72-8133, 340 pp.

Wait, B., and Roberts, L. J.
1932. Studies in the Food Requirement of Adolescent Girls II. Daily Food Requirements in the Energy Intake of the Individual. Amer. Dietetics Jour. 8: 323-331.

Walberg, S., and Adams, W.
1965. Calculated and Determined Values in Constant Diets. Amer. Dietet. Assoc. Jour. 47: 37-43.

Watt, B., and Merrill, A.
1963. Composition of Foods - Raw, Processed, Prepared. U.S. Dept. Agr. Agr. Handb. 8, 190 pp.

White, A., Handler, P., and Smith, E.
1973. Principles of Biochemistry. Ed. 5, 1296 pp. McGraw Hill Book Co., New York.

White, H. S.
1969. Inorganic Elements in Weighed Diets of Girls and Young Women. Amer. Dietet. Assoc. Jour. 55 (1): 38-43.

Widdowson, E. M.
1947. A Study of Individual Children's Diets. Med. Res. Council Spec. Rpt. Ser. No. 257, 196 pp. His Majesty's Stationery Office, London.

_____ and McCance, R. A.
1943. Food Tables: Their Scope and Limitations. Lancet 6: 230-232.

Young, C. M.
1965. Comparison of Results of Dietary Surveys Made by Different Methods. 4th Internatl. Cong. Dietet. Proc. (Sweden) 1965, pp. 119-126.

_____ Chalmers, F., Church, H., and others.
1952a. A Comparison of Dietary Study Methods I. Dietary History vs. Seven-Day-Record. Amer. Dietet. Assoc. Jour. 28 (2): 124-128.

_____ Chalmers, F. W., Church, H. N., and others.
1952b. Cooperative Nutritional Status Studies in the Northeast Region III. Contributions to Dietary Methodology Studies. Mass. Univ. Agr. Expt. Sta. Bul. 469, NE Region. Pub. 10, 95 pp.

_____ Chalmers, F., Church, H., and others.
1953a. Subjects' Estimation of Food Intake and Calculated Nutritive Value of the Diet. Amer. Dietet. Assoc. Jour. 29: 1216-1220.

_____ Franklin, R., Foster, W., and Steele, B.
1953b. Weekly Variation in Nutrient Intake of Young Adults. Amer. Dietet. Assoc. Jour. 29: 459-464.

_____ Hagan, G., Tucker, R., and Foster, W.
1952c. A Comparison of Dietary Study Methods II. Dietary History vs. Seven-Day Record vs. 24-Hour Recall. Amer Dietet. Assoc. Jour. 28 (3): 218-221.

_____ and Trulson, M.
1960. Methodology for Dietary Studies in Epidemiological Surveys II. Strengths and Weaknesses of Existing Methods. Amer. Jour. Pub. Health 50 (6): 803-814.

Yudkin, J.
1951. Dietary Surveys: Variation in the Weekly Intake of Nutrients. Brit. Jour. Nutr. 5: 177-194.

Ziffer, H., Frank, O., Christakis, G., and others.
1967. Data Analysis Strategy for Nutritional Survey of 642 New York City School Children. Amer. Jour. Clin. Nutr. 20 (8): 858-865.

Unpublished Material[1/]

Burk, M. C., and Ward, C.
1973. A Technical Analysis of Several Critical Problems in Appraising Individual and Household Diets. Paper presented at meeting of Res. Sect., Amer. Home Econ. Assoc., June 23, 1973, Atlantic City, N.J.

Grossman, E., and Popka, D.
1976. Methodology of a Household Food Consumption Survey in Cincinnati, Ohio, Nov. 1969-Jan. 1970.

Murray, J.
1970. USDA Food Consumption Surveys--A Historical Review With Special Reference to the Development and Testing of Collection Methods.

Trulson, M. F.
1951. Assessment of Methods for Dietary Intakes. [Ph. D. thesis. Copy on file Harvard Sch. Pub. Health, Boston, Mass.]

U.S. Bureau of the Census.
1964. Report on Total Diet Feasibility Study.

U.S. Department of Agriculture, Agricultural Research Service.
1963. Preliminary Report: Survey of Food Consumption of Households and Individuals, District of Columbia, June-July 1963. (Washington, D.C., pilot survey, Nov. 22.)

[1/] Unless otherwise indicated, copy on file Consumer and Food Econ. Inst., Agr. Res. Serv., U.S. Dept. Agr., Hyattsville, Md. 20782.

APPENDIX

Some Comparisons of Nutritive Values Obtained by Chemical Analysis and by Calculations With Food Consumption Tables

Several methodological studies relevant to individual diets have included investigations of possible errors in nutritive values related to the use of food composition tables instead of chemical analyses of the aliquots or samples of the foods actually eaten. The dates of the studies must be noted carefully because of the gradual accumulation of better data for use in preparing food composition tables.

Bransby et al. (1948) found weighed diets in their study tended to show lower amounts for most nutrients when food aliquots were chemically analyzed than when they were calculated from composition tables. The overestimate for food energy was 11-12 percent. Iron was an exception, probably because iron was picked up from cooking pots. Bransby et al. concluded that systematic errors might have been introduced because the sample was from one institution. Widdowson and McCance (1943) reported that up to 200 mg of calcium could be obtained from hard drinking water, and it is not usually recorded on diet records. However, they found that food composition tables gave results acceptably close to chemical analysis for mixed diets.

Various chemical methods employed in food analysis can yield different nutrient content as, for example, the lower iron values given by the spectroscopic method compared with usual chemical methods (White, 1969). Monsen et al. (1967) found good agreement between average daily iron intakes by colorimetric methods (9.2 mg) and calculation (9.9 mg). Leverton (1937) found the opposite for mixed diets, that colorimetric analysis gave higher values than calculation, possibly because iron was picked up from cooking utensils. Variation in analytical results of chemical analyses may, in addition to contamination, result from sampling, for example, as differences in vitamin C of apples from the north and south sides of the same tree, measuring, deterioration of food samples or reagents, and errors stemming from the method (Thomas et al., 1950). Walberg and Adams (1965) found in comparing calculated versus analyzed food values, a positive bias for nitrogen (thus protein), a slight negative bias for calcium (based on a sign test), but most percentage deviations were under 10 percent. Manalo and Jones (1966) reported that close approximations of actual diet content can be obtained from food tables for sodium, potassium, calcium, and nitrogen but not for phosphorus (underestimated in tables) and magnesium (moderately overestimated in tables).

* U. S. GOVERNMENT PRINTING OFFICE : 1976 O - 222-212