CALIFORNIA HEALTH INTERVIEW SURVEY

CHIS 2003 METHODOLOGY SERIES

REPORT 4

RESPONSE RATES

November 2005

This report was prepared for the California Health Interview Survey by Ismael Flores-Cervantes, Brandon Barrett, J. Michael Brick, and W. Sherman Edwards of Westat

www.chis.ucla.edu

This report provides analysts with information about the response rates in CHIS 2003. The response rates are estimates of the percentage of sampled persons that participated in the survey, where the sample may be across the entire state, restricted to a county, or some other subgroup. To estimate response rates, the probability of sampling persons is taken into account. Thus, the response rates are weighted percentages of the number responding rather than simple unweighted percentages. Procedures used to increase the response rates are also discussed and, where possible, evaluated.

Suggested citation:

California Health Interview Survey. *CHIS 2003 Methodology Series: Report 4 – Response Rates*. Los Angeles, CA: UCLA Center for Health Policy Research, 2005.

The California Health Interview Survey is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. Funding for CHIS 2003 came from multiple sources: the California Department of Health Services, the California Endowment, the National Cancer Institute, the Robert Wood Johnson Foundation, the California Office of the Patient Advocate, the U.S. Centers for Disease Control and Prevention, Kaiser Permanente, the Alameda County Health Care Agency, and L.A. Care Health Plan.

PREFACE

Response Rates is the fourth in a series of methodological reports describing the 2003 California Health Interview Survey (CHIS 2003). The other reports are listed below.

CHIS is a collaborative project of the University of California, Los Angeles (UCLA) Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. Westat was responsible for the data collection and the preparation of five methodological reports from the 2003 survey. The survey examines public health and health care access issues in California. The 2003 survey is the second implementation of CHIS, the first was done in 2001. The plan is to monitor the health of Californians and examine changes over time by conducting periodic surveys in the future.

Methodological Reports

The first five methodological reports for the 2003 CHIS are as follows:

- Report 1: Sample Design;
- Report 2: Data Collection Methods;
- Report 3: Data Processing Procedures;
- Report 4: Response Rates; and
- Report 5: Weighting and Variance Estimation.

The reports are interrelated and contain many references to each other. For ease of presentation, the references are simply labeled by the report numbers given above.

This report describes the response rates from CHIS 2003. Response rates are the ratio of the number of units interviewed to the number of eligible sampled units. However, the computation of response rates for CHIS 2003 is involved because of the complexity of the survey. This report presents the rates and explains the rationale for the procedures used in computing the response rates from CHIS 2003.

The primary purpose of presenting these response rates is to provide information for analysts of the data. As a result, the response rates are reported separately for the main analysis subgroups—adults (ages 18 and older), children (age less than 12), and adolescents (ages 12 to 17). The response rates are estimates of the percentage of sampled persons that participated in the survey, where the sample may be across the entire state, or it may be restricted to a county or another subgroup. To estimate response rates, the probability of sampling persons is taken into account. Thus, the response rates are weighted percentages of the number responding rather than simple unweighted percentages.

A secondary goal of this report is to examine procedures used in the survey to increase the response rates. The specific operational methods are described more completely in *CHIS 2003 Methodology Series: Report 2 – Data Collection Methods*, but the methods are summarized here briefly to provide some context for the examination.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

List of Tables

TABLE OF CONTENTS (continued)

List of Figures

1. CHIS 2003 DESIGN AND METHODOLOGY SUMMARY

1.1 Overview

The California Health Interview Survey (CHIS) is a population-based random-digit dial telephone survey of California's population that is conducted every two years. First conducted in 2001, CHIS is the largest health survey ever conducted in any state and one of the largest health surveys in the nation. CHIS is a collaborative project of the UCLA Center for Health Policy Research, the California Department of Health Services, and the Public Health Institute. CHIS collects extensive information for all age groups on health status, health conditions, health-related behaviors, health insurance coverage, access to health care services, and other health and development issues.

The CHIS sample is designed to provide population-based estimates for most California counties, all major ethnic groups, and several ethnic subgroups. The sample is designed to meet and optimize two goals: provide estimates for large- and medium-sized population counties in the state, and for groups of the smallest population counties; and provide statewide estimates for California's overall population, its major race/ethnic groups, as well as for several Asian ethnic groups. The resulting CHIS sample is representative of California's non-institutionalized population living in households.

This series of reports describes the methods used in collecting data for the 2003 California Health Interview Survey (CHIS 2003). CHIS 2001 is described in a series of methodology reports.[1] These reports describe the second CHIS data collection cycle, which was conducted between August 2003 and February 2004.

CHIS data and results are used extensively by many State agencies, local public health agencies and organizations, federal agencies, advocacy and community organizations and agencies, foundations, and researchers. They use these data in their own analyses and publications to assess public health and health care needs, to develop health policies, and to develop and advocate policies to meet those needs.

[1] California Health Interview Survey, CHIS 2001 Methodology Series: Report 1 - Sample Design, Report 2 – Data Collection Methods, Report 3 – Data Processing Procedures, Report 4 – Response Rates, and Report 5 – Weighting and Variance Estimation, Los Angeles, CA: UCLA Center for Health Policy Research, 2002.

1.2 Sample Design Objectives

The CHIS sample is designed to meet two objectives: (1) provide estimates for counties and groupings of counties with populations of 100,000 or more; and (2) provide estimates for California's overall population and its larger race/ethnic groups, as well as for several smaller ethnic groups. To achieve these objectives, CHIS relied on a multi-stage sample design. First, the state was divided into 41 geographic sampling strata, including 33 single-county strata and 8 groups that included the 25 other counties. Second, within each geographic stratum, households were selected through random-digit dial (RDD), and within each household, an adult (age 18 and over) respondent was randomly selected. In addition, in those households with adolescents (ages 12-17) and/or children (under age 12), one adolescent was randomly selected for interview and one child was randomly selected and the most knowledgeable parent of the child interviewed.

Table 1-1 shows the 41 sampling strata (i.e., counties and groups of counties that were identified in the sample design as domains for which separate estimates would be produced). A sufficient amount of sample was allocated to each of these domains to support the first sample design objective. These strata were also used for the CHIS 2001 sample; because of funding limitations, the sample sizes allocated to most strata for CHIS 2003 were smaller than in 2001.

Table 1-1. California county and county group strata used in the CHIS 2003 sample design

1. Los Angeles	15. San Joaquin	29. El Dorado
2. San Diego	16. Sonoma	30. Imperial
3. Orange	17. Stanislaus	31. Napa
4. Santa Clara	18. Santa Barbara	32. Kings
5. San Bernardino	19. Solano	33. Madera
6. Riverside	20. Tulare	34. Monterey, San Benito
7. Alameda	21. Santa Cruz	35. Del Norte, Humboldt
8. Sacramento	22. Marin	36. Lassen, Modoc, Siskiyou, Trinity
9. Contra Costa	23. San Luis Obispo	37. Lake, Mendocino
10. Fresno	24. Placer	38. Colusa, Glen, Tehama
11. San Francisco	25. Merced	39. Sutter, Yuba
12. Ventura	26. Butte	40. Plumas, Nevada, Sierra
13. San Mateo	27. Shasta	41. Alpine, Amador, Calaveras, Inyo,
14. Kern	28. Yolo	Mariposa, Mono, Tuolumne

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey.

The samples in Los Angeles and Alameda Counties were enhanced with additional funding to allow sub-county geographic estimates, in Los Angeles at the Service Planning Area (SPA) level and in Alameda for the cities of Oakland and Hayward as well as the remainder of the county. These samples were implemented with and incorporated into the original statewide RDD sample.

To accomplish the second objective, larger sample sizes were allocated to the more urban counties where a significant portion of the state's Latino, African American and Asian ethnic populations reside. To increase the precision of the estimates for Koreans and Vietnamese, areas with relatively high concentrations of these groups were sampled at higher rates; these geographic samples were supplemented by phone numbers for group-specific surnames drawn from listed telephone directories to increase the sample size and precision of the estimates for these two groups.

1.3 Data Collection

To capture the rich diversity of the California population, interviews were conducted in five languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Vietnamese, and Korean. These languages were chosen based on research that identified the languages that would cover the largest number of Californians in the CHIS sample that either did not speak English or did not speak English well enough to otherwise participate.

Westat, a private firm that specializes in statistical research and large-scale sample surveys, conducted the CHIS 2003 data collection. Westat staff interviewed one randomly selected adult in each sampled household. In those households with children (under age 12) or adolescents (ages 12-17) associated with the sampled adult[2], one child and one adolescent were randomly sampled, so up to three interviews could have been completed in each sampled household. The sampled adult was interviewed, and the parent or guardian most knowledgeable about the health and care of the sampled child was interviewed. The sampled adolescent responded for him or herself, but only after a parent or guardian gave permission for the interview. Table 1-2 shows the number of completed adult, child, and adolescent interviews in CHIS 2003, by the type of sample (RDD or supplemental sample).

[2] Only children for whom the sampled adult was parent or legal guardian were sampled. The CHIS 2003 sample weights account for this sampling procedure.

Table 1-2. Number of completed CHIS 2003 interviews by type of sample, instrument

Type of sample	Adult	Child	Adolescent
Total RDD + supplemental cases	42,044	8,526	4,010
RDD	41,818	8,480	3,996
Supplemental samples:			
Korean	112	24	6
Vietnamese	114	22	8

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey.

Interviews done in all languages were administered using Westat's computer-assisted telephone interviewing (CATI) system. The average adult interview took 33 minutes to complete. The average child and adolescent interviews took 14 minutes and 21 minutes, respectively. Interviews in the non-English languages generally took longer to complete. Approximately 11 percent of the adult interviews were completed in a language other than English, as were 21 percent of all child (parent proxy) interviews and 7 percent of all adolescent interviews.

Table 1-3 shows the major topic areas for each of the three survey instruments (adult, child, and adolescent).

1.4 Response Rate

The overall response rate for CHIS 2003 is a composite of the screener completion rate (i.e., success in introducing the survey to a household and randomly selecting an adult to be interviewed), and the extended interview completion rate (i.e., success in getting the selected person to complete the full interview). To maximize the response rate, especially at the screener stage, an advance letter (in five languages) was mailed to all sampled telephone numbers for which an address could be obtained from reverse directory services. An advance letter was mailed for approximately 72 percent of the sampled telephone numbers. In 2003, the screener completion rate was 55.9 percent[3], and the rate was higher for those households that could be sent the advance letter. The extended interview completion rate was 60.0 percent for the adult survey. Multiplying the screener and extended rates gives an overall response rate of 33.5 percent. Response rates vary by sampling stratum.

[3] In CHIS 2003, households that refused at the screener level were subsampled and only the subsampled households were called again in an attempt to convert them to respondents. The response rates are weighted to account for this subsampling.

Table 1-3. CHIS 2003 Survey topic areas by instrument

HEALTH STATUS	**ADULT**	**TEEN**	**CHILD**
General health status, height and weight	✓	✓	✓
Emotional health		✓	
Days missed from school due to health problems		✓	✓
HEALTH CONDITIONS	**ADULT**	**TEEN**	**CHILD**
Asthma	✓	✓	✓
Heart disease, high blood pressure, epilepsy	✓		
Diabetes	✓	✓	
Physical disability/need for special equipment	✓	✓	✓
Elder health (stroke, falls, incontinence)	✓		
Parental concerns with child development, attention deficit disorder (ADD)			✓
HEALTH BEHAVIORS	**ADULT**	**TEEN**	**CHILD**
Dietary intake		✓	✓
Physical activity and exercise		✓	✓
Walking for transportation and leisure	✓		
File and pneumonia immunization	✓		
Alcohol and tobacco use	✓	✓	
Drug use		✓	
Sexual behavior, STD testing, birth control practices	✓	✓	
WOMEN'S HEALTH	**ADULT**	**TEEN**	**CHILD**
Pap test screening, mammography screening, self-breast exam	✓		
Emergency contraception, pregnancy status	✓	✓	
Menopause, hormone replacement therapy (HRT)	✓		
CANCER HISTORY AND PREVENTION	**ADULT**	**TEEN**	**CHILD**
Cancer history of respondent	✓		
Colon cancer screening, prostrate cancer (PSA) test	✓		
DENTAL HEALTH	**ADULT**	**TEEN**	**CHILD**
Last dental visit, could not afford care, missed school/work days	✓	✓	✓
Dental insurance coverage	✓	✓	✓
INJURY/VIOLENCE	**ADULT**	**TEEN**	**CHILD**
Serious injuries (frequency, cause)		✓	✓
Injury prevention behaviors (bike helmets, seatbelts)		✓	✓
Infant-toddler home safety			✓
Interpersonal violence		✓	

Table 1-3. (Continued)

ACCESS TO AND USE OF HEALTH CARE	**ADULT**	**TEEN**	**CHILD**
Usual source of care, visits to medical doctor	✓	✓	✓
Emergency room visits	✓	✓	✓
Delays in getting care (prescriptions, tests, treatment)	✓	✓	✓
Health care discrimination due to race or ethnic group	✓		
Communication problems with doctor	✓	✓	✓
Ability and parental knowledge of teen contacting a doctor		✓	
Child immunization reminders			✓
HEALTH INSURANCE	**ADULT**	**TEEN**	**CHILD**
Current insurance coverage, spouse's coverage, who pays for it	✓	✓	✓
Health plan enrollment, characteristics and assessment of plan	✓	✓	✓
Whether employer offers coverage, respondent/spouse eligibility	✓		
Coverage over past 12 months	✓	✓	✓
Reasons for lack of insurance	✓	✓	✓
EMPLOYMENT	**ADULT**	**TEEN**	**CHILD**
Employment status, spouse's employment status	✓		
Work in last week, industry and occupation	✓		
Hours worked at all jobs	✓	✓	
INCOME	**ADULT**	**TEEN**	**CHILD**
Respondent and spouse's earnings last month before taxes	✓		
Household income (annual before taxes)	✓		
Number of persons supported by household income	✓		
Assets	✓		
PUBLIC PROGRAM ELIGIBILITY	**ADULT**	**TEEN**	**CHILD**
Household poverty level (100%, 130%, 200%, 300% FPL)	✓		
Program participation (TANF, CalWorks, Public Housing, Food Stamps, SSI, SSDI, WIC)	✓	✓	✓
Assets, alimony/child support/social security/pension	✓		
Reason for Medi-Cal non-participation among potential eligibles	✓	✓	✓
FOOD INSECURITY/HUNGER	**ADULT**	**TEEN**	**CHILD**
Availability of food in household over past 12 months	✓		
PARENTAL INVOLVEMENT	**ADULT**	**TEEN**	**CHILD**
Parental presence after school, parental knowledge of whereabouts and activities		✓	
Child's activities with family			✓
NEIGHBORHOOD AND HOUSING	**ADULT**	**TEEN**	**CHILD**
Neighborhood cohesion	✓		
Neighborhood safety	✓	✓	
Neighborhood characteristics for children			✓
Length of time at current address/neighborhood, type of housing	✓		
Home ownership, number of rooms, amount of mortgage/rent	✓		

Table 1-3. (Continued)

CHILD CARE	ADULT	TEEN	CHILD
Current child care arrangements			✓
Child care over past 12 months			✓
Reason for lack of childcare			✓
RESPONDENT CHARACTERISTICS	**ADULT**	**TEEN**	**CHILD**
Age, gender, height, weight, education	✓	✓	✓
Race and ethnicity	✓	✓	✓
Marital status	✓		
Sexual orientation	✓		
Citizenship, immigration status, country of birth, English language proficiency	✓	✓	✓

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey.

The CHIS response rate is comparable to response rates of other scientific telephone surveys in California, such as the California Behavioral Risk Factor Surveillance System (BRFSS) survey. California as a whole, and the state's urban areas in particular, are among the most difficult parts of the nation in which to conduct telephone interviews. Survey response rates tend to be lower in California than nationally, and over the past decade response rates have been declining both nationally and in California.

One way to judge the representativeness of a population survey is to "benchmark" its results against those of other reliable data sources. The CHIS 2001 sample yielded unweighted and weighted population distributions and rates that are comparable to those obtained from other sources. The demographic characteristics of the CHIS 2001 sample (such as race, ethnicity, and income) are very similar to those obtained from 2000 Census data. CHIS 2001 respondents also have health characteristics and behaviors that also are very similar to those found in other reliable surveys, such as the California BRFSS. An extensive benchmarking project is being undertaken for the 2003 California Health Interview Survey.

Adults who had completed at least 80 percent of the questionnaire (i.e., through Section I on health insurance) after all followup attempts were exhausted to complete the full questionnaire were counted as "complete." At least some items in the employment and income series or public program eligibility and food insecurity series are missing from these cases.

Proxy interviews were allowed for frail and ill persons over the age of 65 to avoid biases for health estimates for elderly persons that might otherwise result. Eligible selected persons were

recontacted and offered a proxy option. For 171 elderly adults, a proxy interview was completed by either a spouse/partner or adult child. Only a subset of questions identified as appropriate for a proxy respondent were administered. (Note: The questions not administered are identified in their response set as being skipped (denoted by a value of "-2") because a proxy is responding for the selected person.)

1.5 Weighting the Sample

To produce population estimates for the RDD CHIS results, weights are applied to the sample data to compensate for a variety of factors, some directly resulting from the design and administration of the survey. The sample is weighted to represent the non-institutionalized population for each sampling stratum and statewide. Sample weighting was carried out in CHIS 2003 to accomplish the following objectives:

- Compensate for differential probabilities of selection for households and persons (Note: telephone numbers for which addresses could be found and advance letters mailed were assigned a higher probability of selection than those without addresses);
- Reduce biases occurring because nonrespondents may have different characteristics than respondents;
- Adjust, to the extent possible, for undercoverage in the sampling frames and in the conduct of the survey; and
- Reduce the variance of the estimates by using auxiliary information.

As part of the weighting process, a household weight was created for all households that completed the screener interview. This household weight is the product of the "base weight" or the inverse of the probability of selection of the telephone number and adjustment factors computed for the following weight adjustments:

- Subsampling for numbers with addresses;
- Multiple chances of being selected in the RDD and supplemental samples;
- Unknown residential status;
- Subsampling screener refusals for conversion attempt;
- Screener interview nonresponse;

- Multiple telephone numbers; and
- Household poststratification.

The resulting poststratified household weight was used to compute a person-level weight. This person-level weight includes weight adjustments for the within-household sampling of persons and nonresponse. The final step is to adjust the person-level weight using a raking method so that the CHIS estimates are consistent with population control totals. Raking is an iterative procedure that forces the CHIS weights to sum to known totals from auxiliary data sources. The procedure requires iteration to make sure all the control totals or dimensions of raking are simultaneously satisfied (within a specified tolerance).

The control totals or raking dimensions used in CHIS 2003 were created primarily from the 2003 California Department of Finance estimates of the numbers of persons by age, race, and sex, and from the 2000 Census of Population counts from the U.S. Census Bureau. The 14 dimensions are combinations of demographic variables (age, sex, race, and ethnicity), geographic variables (county, city, and, in Los Angeles County, Service Planning Area), household composition (presence of children and adolescents in the household), and socio-economic variables (home ownership and education). The socio-economic variables are included to reduce biases associated with excluding households without a telephone number from the survey. One of the limitations of using the Department of Finance data is that it includes about 2.4 percent of the population of California who live in "group quarters" (i.e., persons living with 9 or more unrelated persons). These persons were excluded from the CHIS sample and, as a result, the number of persons living in group quarters had to be estimated and removed from the control totals prior to raking.

1.6 Imputation Methods

To enhance the utility of the CHIS 2003 data files, missing values were replaced through imputation for nearly every variable. This was a massive task designed to eliminate missing values in all source variables. Westat imputed values for variables used in the weighting process, and the UCLA staff imputed values where missing due to item nonresponse for nearly all other variables.

Two different imputation procedures were used by Westat prior to delivering the data to UCLA to fill in missing responses for items in CHIS 2003 that were essential for weighting the data. The

first imputation technique is a completely random selection from the observed distribution of the respondents. This method is used only for a few items when the percentage of the items that are missing is very small. For example, when imputing values for self-reported age which had a very low item non-response rate, the distributions of the responses for age by type of interview (adult, child, or adolescent) were used to randomly assign an age using probabilities associated with these distributions.

The second technique is hot deck imputation without replacement. The hot deck approach is probably the most commonly used method for assigning values for missing responses in large-scale household surveys. With a hot deck, a value reported by a respondent for a particular item is assigned or donated to a "similar" person who did not respond to that item. The characteristics defining "similar" vary for different variables. To carry out hot deck imputation, the respondents to an item form a pool of donors, while the nonrespondents are a group of recipients. A recipient is matched to the subset pool of donors based on household and individual characteristics. A value for the recipient is then randomly imputed from one of the donors in the pool. Once a donor is used, it is removed from the pool of donors for that variable. Hot deck imputation was used to impute race, ethnicity, home ownership, and education in CHIS 2003.

The UCLA staff imputed for missing values through a hierarchical sequential hot deck method with donor replacement. This method rank-orders the control variables from the most essential to the least essential, allowing the control variables to be dropped if the imputation conditions (such as minimal number of donors or no missingness in control variables) are not met in the imputation process. The control variables are dropped one at a time sequentially, starting from the least essential. CHIS incorporated an automated data quality control check both before and after the imputation process.

Imputation flags for CHIS source variables are included in separate data files to identify all imputed values.

1.7 Methodology Report Series

A series of five methodology reports are available with more detail about the methods used in CHIS 2003:

- Report 1 – Sample Design;
- Report 2 – Data Collection Methods;
- Report 3 – Data Processing Procedures;
- Report 4 – Response Rates; and
- Report 5 – Weighting and Variance Estimation.

For further information on CHIS data and the methods used in the survey, visit the California Health Interview Survey Web site at www.CHIS.ucla.edu or contact CHIS at CHIS@ucla.edu.

2. USE OF RESPONSE RATES

While the use of response rates as a single measure of the quality of a survey is unjustified, Madow et al. (1983) suggest response rates provide valuable information on the success of the survey at representing the population sampled. Keeter et al. (2000) and other researchers note that the response rate alone is not sufficient for this objective because the bias in an estimate is related to both the response rate and the characteristics of those responding. This relationship is discussed in more detail below.

The main objective of this report is to present response rates that can be used by analysts of the CHIS 2003 data to better understand how well the California population is represented. To accomplish this goal, the response rates are weighted so that the weighted response rate is an estimate of the proportion of the population responding to the survey. This procedure is consistent with the standards given in The American Association for Public Opinion Research (AAPOR) (2001). For example, since the sample was selected with differing sampling rates by county, the response rates are weighted so each county accounts for its appropriate fraction when the state response rate is reported.

The rationale for using weights in computing the response rate is that the bias of a simple estimate such as a mean, $\bar{y}_r$, is a function of the response rate and the difference in the means between the respondents and nonrespondents. A simple way of conceptualizing this is by assuming the population is partitioned into a stratum of respondents (*R*) and a stratum of nonrespondents (*NR*). The estimate from the survey is computed from the observations in the respondent stratum, where each observation is weighted by the inverse of its selection probability. In a probability sample, the bias attributable to nonresponse of a survey estimate of a mean is

$$bias(\bar{y}_r) = (1-r)(\bar{Y}_R - \bar{Y}_{NR}), \quad (1)$$

where *r* is the appropriately weighted response rate and the quantity on the right is the difference in the means between the respondent and nonrespondent strata (Lessler and Kalsbeek, 1992). This formula shows that the bias increases as the response rate decreases, provided the difference in characteristics between respondents and nonrespondents remains constant. If the response rates are not weighted, this relationship does not hold. Returning to the example, if the county samples are not weighted by their selection probabilities, then the response rate cannot be used in the bias equation (1).

The sample for CHIS 2003 includes both an RDD sample and a sample from lists of persons with Vietnamese and Korean surnames. The sample is described in *CHIS 2003 Methodology Series: Report 1 – Sample Design*. The weighted response rates are computed for the combined RDD and surname list samples in CHIS 2003. The Vietnamese and Korean surname list samples were jointly weighted with the RDD sample, and the sampling weights reflect the multiple probabilities of selection from the different sampling frames. In CHIS 2001, the supplemental race and ethnic samples were not jointly weighted with the RDD sample, so only unweighted response rates could be computed for the supplemental samples. The weighting procedures are described in detail in *CHIS 2003 Methodology Series: Report 5 – Weighting and Variance Estimation*.

3. DEFINING RESPONSE RATES

The first step is to define "response rate" because the term is used in so many different ways across surveys and organizations. Two organizations that describe response rates in a relatively consistent manner are the Council of American Survey Research Organizations (CASRO, 1982) and the American Association for Public Opinion Research (AAPOR, 2004). The AAPOR report is periodically updated and is available on the organization's Uniform Resource Locator (http://www.aapor.org).

Both reports recommend that a survey response rate be defined as the ratio of completed interviews to eligible reporting units. However, the application of this recommendation is more difficult than it may appear, especially in RDD surveys. One problem in estimating response rates in RDD surveys is the determination of the eligibility of some of the sampled numbers. Some telephone numbers are never answered or are only picked up by answering machines, even after being called multiple times over a range of days. This outcome may occur for many reasons, as discussed by Shapiro et al. (1995). The eligibility of these numbers cannot be determined directly, adding ambiguity to the definition of a response rate.

We use procedures described in the AAPOR (2004) report, and the same method used in CHIS 2001 to resolve this ambiguity in determining the number of eligible telephone numbers. The AAPOR report has several different response rate definitions and the one used in this report is AAPOR's "RR4" equation (note that at the screener level RR4 is equal to AAPOR's RR3 equation since there are no partial screener interviews). Since the CHIS sample of telephone numbers was sampled with different probabilities of selection, we use the weighted number of telephone numbers rather than the unweighted number in the computation. This also compensates for the under-sampling and over-sampling that occurred in different geographic areas.

Telephone numbers with unknown eligibility are assigned to be either eligible (i.e., a residence) or ineligible based on the "survival method" (Brick, Montaquila, and Scheuren, 2002). This approach is a more empirically-based estimate of the percentage of the "unknown eligibility" telephone numbers that are likely to be eligible (residential). The method works by selecting a subsample of the telephone numbers with unknown eligibility and dialing these numbers additional times. In CHIS 2003, Westat selected a sample of 3,511 RDD telephone numbers with unknown residential status for the survival analysis (2,250 coded as "ring no answer," and 1,261 coded as "answering machine"). All "ring

no answer" phone numbers received 14 call attempts, and those subsampled for the survival analysis were dialed at least seven more times to determine their final status. All phone numbers with "answering machine" status received 18 call attempts, and those subsampled for the survival analysis were also dialed at least seven more times. The data are then analyzed using survival analysis to predict the percentage of telephone numbers that would be residential if dialing continued indefinitely. This estimation of the proportion of eligible telephone numbers is denoted as '***e***' in the AAPOR RR4 equation. After the eligibility proportion is established, the response rate can be computed as the weighted ratio of the responding telephone numbers to the total of known and estimated eligible numbers. We do compute some alternative response rates using the CASRO equation (CASRO estimates '*e*' as the proportion of the resolved telephone numbers that are residential) and other approaches for comparison purposes, but these are not used for most of the analysis.

A new procedure used in CHIS 2003 involved subsampling households that refused to participate in the initial screening interview. Only a subsample of these households were called again in a refusal conversion process. Since only the subsampled cases are retained for the analysis (along with those that did not refuse), the subsampled cases are weighted by the inverse of the subsampling rate. Refusal subsampling is described in Chapter 8.

The next step in computing response rates depends on the particular part of the interview process being analyzed, such as the adult interview. For example, to compute the response rate for the adult interview, the numerator of the rate is the weighted number of completed adult interviews, and the denominator is the weighted number of eligible adults sampled in households that completed the screening interview. An overall or joint response rate can be computed by multiplying the screening and adult interview rates.

Computing a response rate for a subgroup requires that all the units in both the numerator and denominator of the rate can be classified as members of the subgroup. To do this, data must be available to classify all sampled units, not just respondents. Because the screening interview identifies if any children were in the household, extended response rates can be computed separately for households with children and without children. However, the joint rate must be computed by multiplying the extended rate for the subgroup by the overall screener response rate because data on the presence of children are not available for every sampled telephone number.

At the screener level, data on the RDD and surname list samples of telephone numbers are limited; the telephone numbers can be classified by geography (county or group of counties used for sampling) and by whether there was an address for the telephone number that could be used to send an advance mailing. At the extended interview or person level, data from the screener can be used to classify households by characteristics that are known for all completed households. These data are used to compute the response rates in CHIS 2003 later in this report.

4. REVIEW OF CONTACT METHODS

CHIS 2003 Methodology Series: Report 2 – Data Collection Methods gives a detailed discussion of the methods used in CHIS 2003 to contact and interview persons. Here we briefly review the key procedures to provide some background on the response rates and evaluation measures presented later in this report.

As mentioned before, the survey contained both screening and extended interviews. In each household, one adult was sampled for an extended interview. In households with persons under age 17, one child and one adolescent were also sampled. The screening interview took, on average, about 2 to 3 minutes to conduct. A parent or guardian was interviewed about the sampled child and the sampled adolescent was interviewed if a percent or guardian parent gave permission. The adult extended interview averaged about 33 minutes in length, the child interview about 14 minutes, and the adolescent interview about 19 minutes. The interviews in languages other than English generally took longer than these averages. More detailed timings on the interviews are given in *CHIS 2003 Methodology Series: Report 2 – Data Collection Methods*.

Shortly before calling sampled telephone numbers, Westat mailed an advance or prenotification letter to those for which an address could be obtained from reverse directory services. The letter informed the household that they would be called to participate in CHIS 2003, that their participation was voluntary but important to the success of the survey, and that the survey was legitimate.

After the advance mailing, initial telephone calls were made to complete the screener interview with a household respondent who was at least 18 years old. Multiple attempts, at least 14 attempts if needed, were made to establish the initial contact with the household. If the household refused to participate, additional attempts were made to complete the screener after waiting at 1-3 weeks following the first refusal.[4] Prior to attempting to convert these refusals into participants, an express letter was sent to the household (if an address was available) informing them again about the validity of the study and the importance of their participation. If the household refused again, a second refusal conversion telephone attempt was made at least another 2 weeks later.

[4] A 60 percent subsample of all sample phone numbers was randomly assigned a refusal conversion flag. Additional calls were attempted only for initial refusals that were part of this 60 percent subsample. See Chapters 7 and 8 for additional details.

A similar process was used at the extended level for the sampled adult. The sampled adult was asked to participate in the study up to three times—an initial attempt and two attempts at refusal conversion.[5] If the adult refused, an express letter was sent (if an address was available) urging him or her to participate. A second refusal conversion attempt for both the screener and the adult extended interview was done only for the subset of those cases where the review of interviewer reports on the previous refusals indicated an additional attempt was warranted. For child and adolescent interviews, one refusal conversion attempt was made. No express letters were sent for either the child or adolescent interview. However, if the parent refused permission for the adolescent to be interviewed, then an express letter was mailed to the parent asking him or her to reconsider. Attempts at refusal conversion were stopped at any point if the respondent expressed hostility at being called or specifically requested that they not be called again.

A variety of other methods were used to increase response rates in CHIS 2003. A very important procedure involved translating and conducting the interview in Spanish, Chinese (Cantonese and Mandarin), Korean, and Vietnamese to accommodate households that did not speak English. Another response rate enhancement method was the use of some outreach programs to raise awareness of CHIS in the community and to encourage participation. Yet another method to increase response rates was the use of proxy interviews for adults who were over age 65 and unable to participate because of mental or physical limitations. Other adult household members knowledgeable about the sampled persons' health, almost always a spouse or child of the sampled adult, completed a proxy interview in these cases.

In addition to the efforts to encourage respondents to participate, other approaches were used to increase response rates. Interviewers were trained and given refresher training on methods to avoid refusals and to convert those who had refused. Only those interviewers who had above average response rates were trained and allowed to conduct the refusal conversions. Multiple call attempts were made to contact sampled household members to complete the extended interviews. On average, 14 call attempts were made to contact an adult before a case was classified as a nonrespondent.

[5] All cases were available for refusal conversion at the extended interview level. The 60 percent subsample applied only to the screening interview.

Later in this report, we discuss some of these methods and describe the increases in the number of interviews that resulted, where this is possible. Of course, some methods such as interviewer training cannot be assessed quantitatively without specially designed experiments.

5. RESPONSE RATE FORMULAS

This chapter describes the formulas used to compute the response rates for CHIS 2003. The types of response rates are screener response rates, extended interview or person response rates, and overall or joint response rates. It is important to note that in CHIS 2003, refusal conversion procedures were applied to a random subsample of screener interview refusals. As a result, unweighted response rates are not comparable to weighted rates and should not be used to assess response patterns because they do not reflect the subsampling of refusal conversion cases. See Chapter 8 for additional details of refusal conversion subsampling. We begin with the screener response rates.

A screener response rate is calculated for each sampled county or group of counties that were combined for sampling purposes. In the report we will often refer to these as counties, but we are referring to the groups of counties used in sampling as given in Table 1-1. In the tables of response rates, these groups are called sampling strata to avoid confusion. Response rates for a group of counties or the entire state can be computed in the same way. The formula for the screener response rate (rr_S) in a sample county is

$$rr_S = \frac{\sum_{i \in S_{resp}} w_i}{\sum_{i \in S_{resid}} w_i} , \qquad (2)$$

where w_i is the weight for telephone number i in the county after adjusting for differential sampling rates, refusal conversion subsampling, and the assignment of telephone numbers with unknown residential status; S_{resp} is the set of telephone numbers in the county that responded to the screening interview; and S_{resid} is the set of telephone numbers in the county that were residential. As noted earlier, the estimated residential rates were determined using the survival analysis method.

The screener response rate for the state is computed in exactly the same way, except the sum is over the whole state rather than in the specific county. The state screener response rate is thus a weighted average of the county screener response rates with weights equal to the population in the counties. As a result, the state response rate differs from what would be obtained from the unweighted average of the response rates of the counties.

The extended response rate for the adult interview in a county is the weighted percentage of the adults sampled in the screener who completed the adult extended interview. The weight in this case is the inverse of the probability of selecting the adult within the household. Because of this weighting, adults sampled from households with more than one adult have a larger effect on the response rate than those in households with only one adult. The extended adult response rate (rr_a) is

$$rr_a = \frac{\sum_{i \in A_{resp}} w_i'}{\sum_{i \in A_{eligsamp}} w_i'}, \tag{3}$$

where the numerator is summed over all adult respondents, and the denominator is summed over all eligible sampled adults. The weight being summed in this case, w', is the adult weight that accounts for selecting the adult within the household. The adult response rate is conditioned on the completion of the screener interview.

The extended response rate computation for children and adolescents is similar to the adult procedure; however, the method of sampling does add some complexity. Persons under 18 years of age are not enumerated in the screener, even though the screener did ascertain whether or not there were children in the household. The full enumeration of persons under 18 is done in the adult extended interview. As a result, the child and adolescent extended response rates are computed for only those households in which the adult extended interview is completed. In other words, the child response rate is a conditional rate like the adult rate, but it is conditional on both the screener and adult interviews being completed.

The extended child response rate (rr_C) is

$$rr_c = \frac{\sum_{i \in C_{resp}} w_i''}{\sum_{i \in C_{eligsamp}} w_i''}, \tag{4}$$

where the numerator is summed over all child respondents, and the denominator is summed over all eligible sampled children. The weight being summed in this case, w'', is the inverse of the probability of selecting the child within the household.

Exactly the same procedure is used for the adolescent extended interview response rate (rr_t) and it is

$$rr_t = \frac{\sum_{i \in T_{resp}} w_i'''}{\sum_{i \in T_{eligsamp}} w_i'''}, \tag{5}$$

where the numerator is summed over all adolescent respondents, and the denominator is summed over all eligible sampled adolescents. The weight being summed in this case, w''', is the inverse of the probability of selecting the adolescent within the household.

An important source of nonresponse for the adolescent interview was the parent denying permission to conduct the interview with the adolescent. The response rate given by (5) includes the parent permission as a source of nonresponse. Another response rate of interest is the adolescent response rate conditioned on the parent giving permission to interview the adolescent. This fully conditional adolescent response rate is

$$rr_{t-p} = \frac{\sum_{i \in T_{resp}} w_i'''}{\sum_{i \in T_{eligsamp-per}} w_i'''}, \tag{6}$$

where the only difference is that the denominator is summed over only those adolescents for whom the parents gave permission for the adolescent to be interviewed.

The response rates defined above, except for the screener response rate, are conditional rates in the sense that they depend on the household participating in a previous stage of CHIS. Overall response rates eliminate the conditioning. For example, since the adult response rate is conditioned on the completion of the screener, the product of the screener and adult response rate is an unconditional or overall adult response rate. Thus, the overall adult response is

$$orr_a = rr_s \cdot rr_a \tag{7}$$

Since the child response rate is conditioned on both the screener being completed and the adult interview in the household being completed, the overall response rate for the child is the product of the screener response rate, the adult extended response rate, and the child response rate. The overall response rate for the child is

$$orr_c = orr_a \cdot rr_c \,, \tag{8}$$

because orr_a is the product of the screener and adult response rates.

For adolescents, the overall response rate accounting for all levels of conditioning (completion of the screener, the adult interview, and the permission request) is

$$orr_t = orr_a \cdot rr_t \,. \tag{9}$$

We do not present an overall response rate for the adolescent excluding the permission request because it is not of much interest as an overall rate.

Note that this calculation of the child and adolescent response rates assumes that the response rate for the adult interview is the same in households where children and/or adolescents are sampled as in those with no child or adolescent sampling. This is a necessary assumption, since we do not know the household composition for much of the nonresponse that goes into calculating the adult response rate. The response rate formulas are applied in the next chapter.

6. RESPONSE RATE TABLES

This chapter gives tables of response rates from the combined RDD and surname list samples for CHIS 2003. The first tables are the response rates for the specific interviews: the screener, the adult, the child, and the adolescent interviews. The overall response rates for adults, children, and adolescents are then presented. All of the rates in the tables in this chapter are weighted and use the formulas presented in the previous section.

6.1 Screener Response Rates

The screener response rates for each county (sampling stratum) are given in Table 6-1. The first column in the table gives the number of households that completed the screening interview. Overall, 66,547 households across the state cooperated with this first step of the CHIS 2003 interview. In all of these households, one adult was sampled.

The overall screener response rate for the state is 55.9 percent. As discussed in Chapter 3, the response rate is computed using the survival method to allocate the undetermined numbers (those for which every call was not answered or only answered by an answering machine). Alternative definitions for allocating these undetermined numbers are used in some other surveys and may give different response rates. One approach used by some is to ignore the undetermined numbers in the computation of response rates. This approach gives a ***cooperation rate***. Dropping all the undetermined numbers for CHIS 2003 gives an overall state-level cooperation rate of 61.4 percent. Another approach is to use what is called the CASRO rate. The CASRO screener response rate for the entire state is 55.6 percent, which is just slightly lower than the survival method. For the remainder of the report, we use the survival method for all response rates.[6]

[6] The CASRO method was still used for the part of the sample drawn from the surname lists.

Table 6-1. Number of completed screeners and response rates by sampling stratum and whether an advance letter was sent

	Total		Letter		No letter	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
State Total	66,657	55.9	57,785	57.4	8,872	48.9
Los Angeles	17,770	51.0	15,429	52.5	2,341	44.5
San Diego	3,701	56.8	3,231	58.3	470	49.9
Orange	3,674	54.2	3,209	56.8	465	43.5
Santa Clara	2,042	57.1	1,854	58.5	188	47.7
San Bernardino	2,005	61.0	1,674	62.2	331	56.6
Riverside	1,901	59.1	1,605	60.3	296	54.2
Alameda	7,440	54.9	6,113	56.7	1,327	47.1
Sacramento	1,576	60.3	1,361	62.5	215	51.7
Contra Costa	1,199	58.8	1,098	59.9	101	50.8
Fresno	1,011	59.5	891	60.7	120	52.9
San Francisco	1,453	44.2	1,304	45.7	149	36.3
Ventura	962	57.1	846	58.9	116	48.0
San Mateo	945	54.6	868	55.3	77	49.3
Kern	790	62.9	717	65.0	73	50.4
San Joaquin	828	58.1	730	59.1	98	52.7
Sonoma	739	56.6	680	57.8	59	47.6
Stanislaus	858	61.0	770	61.7	88	57.0
Santa Barbara	754	61.0	635	62.2	119	56.4
Solano	790	61.9	705	63.0	85	55.5
Tulare	882	66.2	759	65.7	123	68.8
Santa Cruz	772	57.7	668	59.1	104	51.7
Marin	749	54.5	689	56.4	60	41.7
San Luis Obispo	751	64.4	666	65.8	85	57.1
Placer	761	60.9	623	63.0	138	54.6
Merced	847	61.4	745	62.3	102	57.0
Butte	789	63.8	714	64.8	75	57.2
Shasta	718	63.2	614	63.4	104	62.0
Yolo	733	64.4	648	66.0	85	56.7
El Dorado	751	59.4	628	60.0	123	57.3
Imperial	857	62.0	762	62.8	95	57.8
Napa	756	56.4	678	58.6	78	45.5
Kings	837	60.1	721	61.0	116	55.9
Madera	828	62.2	676	63.1	152	59.5
Monterey, San Benito	810	58.1	719	59.3	91	51.9
Del Norte, Humboldt	728	64.3	620	63.7	108	66.9
Lassen, Modoc, Siskiyou, Trinity	568	65.4	483	66.2	85	62.3
Lake, Mendocino	579	61.8	519	62.5	60	57.5
Colusa, Glen, Tehama	650	68.0	553	68.2	97	67.2
Sutter, Yuba	697	67.3	604	68.5	93	61.9

Table 6-1. Number of completed screeners and response rates by sampling stratum and whether advance letter was sent (continued)

	Total		Letter		No letter	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
Plumas, Nevada, Sierra	581	58.8	477	60.3	104	53.9
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	575	57.2	499	59.2	76	48.9

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

The table shows that the screener response rates vary by county, which is also portrayed by the distribution in Figure 6-1. The median response across all counties is 60.1 percent, and the highest response rate is 68.0 percent in the Colusa-Glen-Tehama stratum. San Francisco has the lowest response rate at 44.2 percent, which is clearly at the lower end of the scale in Figure 6-1. The next lowest response rate (Los Angeles) is about 7 percentage points higher than the San Francisco rate. The county rankings as shown in Figure 6-1 are relatively consistent from 2001 to 2003, as will be discussed later.

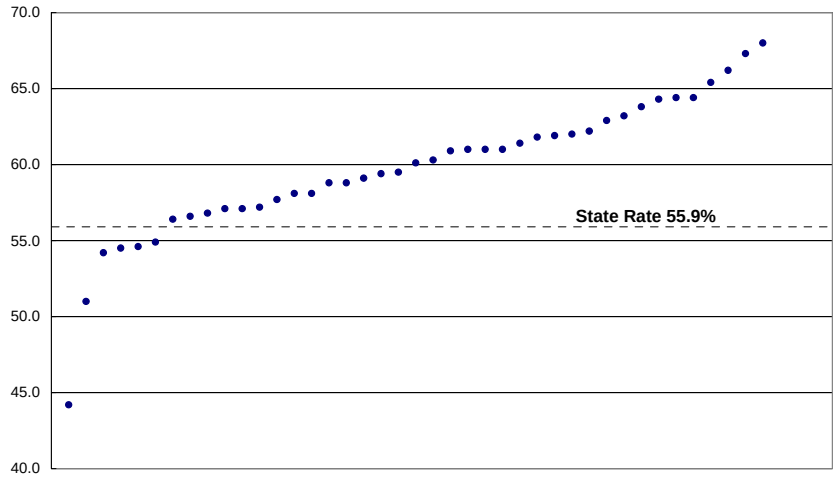

Figure 6-1. Screener response rate distribution by sampling stratum

Another interesting finding is that the median response rate for counties with a population of more than 500,000 persons (the counties from Los Angeles through San Joaquin in Table 6-1) is 56.6 percent. This is five percentage points lower than the 61.7 percent median response rate for the smaller counties. Looking at the individual counties suggests that this difference may be a function of proximity

to a metropolitan area or population density rather than the population size of the county. Small, highly urban counties have rates similar to those of the more populous counties. This differential was also observed in CHIS 2001 stratum-level response rates.

Table 6-1 also has tabulations on the response rates by whether an advance letter could be mailed to the household. We discuss these rates in Chapter 9. Next, we examine the response rates for the extended adult, child, and adolescent interviews.

6.2 Person Response Rates

The adult, child, and adolescent extended interview response rates for each stratum in CHIS 2003 are given in Table 6-2, along with the number of completed interviews for each of the instruments. A total of 42,044 adult interviews, 8,526 interviews about children, and 4,010 adolescent interviews were completed in this very comprehensive survey of the residents of California.

The statewide response rate shown in Table 3-2 for the adult interview was 60.0 percent, a decrease of almost four percentage points from CHIS 2001. As with the screener response rate, counties with larger populations tended to have lower adult extended interview response rates. The median adult response rate for the counties with a population of more than 500,000 is 60.7 percent, while for counties with less than 500,000 the median adult response rate is 64.7 percent. This difference may be attributed to a variety of reasons, including the different distribution of persons by age, education, etc., by county.

Data collected in the screener interview about the household and the sampled adult can be used to examine the adult extended response rates since the data are available for all sampled adults. Table 6-3 shows the adult response rates by these screener data items.[7] There was substantial variation in the response rates by the characteristics known in the screening interview. Women responded at a higher rate than men, older adults were more likely to respond than younger adults. In CHIS 2001, the rates were about the same for households with and without children; however, adults in households without children had a five percentage point higher response rate in 2003.

[7] In some cases the data from the screener interview and the adult interview may differ. For example, the age of the adult reported by the household member in the screener may be different from the age reported by the sampled adult. All of the data used in these tabulations are the screener data because no other data are available for the nonresponding adults.

Table 6-2. Number of completed extended interviews and response rates by sampling stratum and type of interview

	Adult		Child		Adolescent	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
State Total	42,044	60.0	8,526	81.4	4,010	57.3
Los Angeles	10,350	55.1	2,112	80.2	925	56.5
San Diego	2,310	60.7	457	84.2	208	59.8
Orange	2,231	58.0	466	77.5	201	49.1
Santa Clara	1,340	64.3	279	80.7	123	60.0
San Bernardino	1,238	59.5	285	80.3	149	55.4
Riverside	1,180	58.7	265	83.2	136	55.2
Alameda	4,734	62.1	950	81.1	403	56.2
Sacramento	1,062	63.0	201	77.8	81	53.3
Contra Costa	820	66.3	163	79.7	87	64.8
Fresno	626	61.6	178	86.2	66	57.5
San Francisco	917	59.9	115	79.4	36	58.0
Ventura	617	60.3	127	88.7	59	60.8
San Mateo	609	61.4	110	80.6	54	51.1
Kern	537	65.5	124	79.9	64	58.1
San Joaquin	521	59.2	114	86.7	62	52.3
Sonoma	507	67.0	96	91.1	39	56.7
Stanislaus	549	62.4	119	84.7	65	60.9
Santa Barbara	504	64.6	107	86.2	59	67.3
Solano	510	60.8	113	73.3	67	60.3
Tulare	575	64.7	142	77.2	82	62.4
Santa Cruz	512	64.0	87	80.2	48	68.6
Marin	521	65.2	94	88.3	31	58.4
San Luis Obispo	503	64.9	83	87.6	46	63.0
Placer	507	63.0	97	79.4	56	67.0
Merced	520	57.7	141	80.9	69	64.8
Butte	564	69.5	98	93.2	53	60.7
Shasta	506	66.7	81	86.9	43	54.5
Yolo	517	66.3	102	82.1	56	58.7
El Dorado	503	64.4	94	81.6	55	57.9
Imperial	529	61.9	124	72.1	85	66.4
Napa	505	65.4	87	89.1	43	68.5
Kings	531	61.7	161	88.2	72	64.4
Madera	512	59.9	104	85.1	68	68.6
Monterey, San Benito	520	63.1	122	81.8	44	56.0
Del Norte, Humboldt	529	71.0	91	84.9	44	60.9
Lassen, Modoc, Siskiyou, Trinity	419	72.3	65	92.0	37	69.7
Lake, Mendocino	409	67.8	71	87.5	32	62.4
Colusa, Glen, Tehama	425	63.0	90	80.0	41	57.0

Table 6-2. Number of completed extended interviews and response rates by sampling stratum and type of interview (continued)

	Adult		Child		Adolescent	
	Complete	Response rate	Complete	Response rate	Complete	Response rate
Sutter, Yuba	460	64.7	105	92.1	52	70.8
Plumas, Nevada, Sierra	403	66.1	53	82.0	38	72.0
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	412	69.1	53	89.8	31	62.5

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

Table 6-3. Adult response rates by characteristics of the sampled adult

Characteristic	Response rate
Total	59.9
Sex	
Male	54.7
Female	64.8
Age	
18 to 30 years	53.5
31 to 45 years	58.4
46 to 65 years	63.3
Over 65 years	68.7
Type of household	
With children	57.2
Without children	62.1
Adults in household	
1	76.2
2	61.3
3 or more	52.5

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

A somewhat surprising finding is that adults in households with fewer adults were more likely to respond than adults in households with more adults. A concern in many RDD surveys is undercoverage of women in households with no other adults, but this does not appear to be the case with CHIS 2003[8]. Of course, it must be noted that these rates are conditional on the screener being completed, so there are no data to examine this completely.

Now, we examine the child extended interview response rates. Overall, Table 6-2 shows that, across the state, the child-level response rate is 81.4 percent, which is high but still about six percentage points lower than the same rates computed for CHIS 2001. The median rate in the more populous counties (80.6 percent) is about four percentage points lower than the rate in smaller counties (84.9 percent).

Table 6-4 gives the child response rates by the characteristics of the child and household using data collected in the adult interview where the children were enumerated for sampling. The child rates do not show much variation by sex, age, or number of children in the household.

Table 6-4. Child response rates by characteristics of the sampled child

Characteristic	Response rate
Total	81.4
Sex	
Male	80.5
Female	82.8
Age	
Less than 4 years	82.5
4 to 7 years	81.9
8 to 11 years	80.9
Children in household	
1	81.2
2	82.4
3	79.8
4 or more	81.0

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

[8] Another explanation of this finding is that in smaller households the person completing the screening interview is more likely to be the sampled adult and there is evidence that extended response rates for persons completing the screening interview are higher than for other adults in the household.

The last interview-level response rates presented are for the adolescent interview. An important fact to remember is that the adolescent could not be interviewed unless a parent or guardian gave verbal permission to conduct the interview. This requirement resulted in lower response rates for the adolescent interviews than for the child instrument. Table 6-2 shows the state adolescent response rate is 57.3 percent. If we exclude the nonresponse due to parents not giving permission to interview the adolescent, the cooperation rate rises 26 percentage points to 83.3 percent. We discuss the differences in greater detail below.

As with the adult and child interviews, we find that there is a difference in response rates for the adolescent interviews by the size of the county. The more heavily populated counties have a median response rate of 56.5 percent and the counties with less than 500,000 persons have a median response rate of 62.5 percent.

Table 6-5 gives the adolescent response rates by the characteristics of the adolescent and household using data collected in the adult interview. These rates, like the corresponding child rates, are not very different by sex, age, and the number of adolescents in the household.

Table 6-5. Adolescent response rates, by characteristics of the sampled adolescent

Characteristic	Response rate
Total	57.3
Sex	
Male	57.0
Female	58.2
Age	
12 to 14 years	56.8
15 to 17 years	58.7
Adolescents in household	
1	57.3
2	57.2
3 or more	57.5

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

To better understand the success rate for interviewing adolescents, we examine the response rates for the adolescent interview including only those adolescents the parents gave permission to interview. This rate is more indicative of the ability of the survey operations to contact and interview the adolescents. These rates are given in Table 6-6. Table 6-6 is similar to Table 6-5, but the sampled adolescents without parental permission are excluded from the denominator of the computation. Even

though the rates in Table 6-6 are about 25 percentage points greater than those in Table 6-5, the relative rates by the characteristics are relatively consistent across the tables with a couple of exceptions. One noticeable difference is the drop in the rate for households with more than two adolescents. Also, in Table 6-5, the younger group (12 to 14 years) has a rate about two percentage points lower than the older group (15 to 17 years), but this is reversed in Table 6-6 when we exclude parental permission nonresponse. In Table 6-6, the younger group has a cooperation rate larger by one and a half percentage points. Clearly, parents were less likely to grant permission for the interview for younger adolescents than they were for older adolescents.

Table 6-6. Adolescent cooperation rates excluding parental permission nonresponse by characteristics of the sampled adolescent

Characteristic	Response rate
Total	83.3
Sex	
Male	83.0
Female	83.8
Age	
12 to 14 years	84.1
15 to 17 years	82.6
Adolescents in household	
1	83.4
2	84.1
3 or more	80.5

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

Differences in response rates can lead to nonresponse bias as suggested by bias equation (1). To reduce the potential for this bias, these characteristics were taken into account in the development of the estimation weights as described in *CHIS 2003 Methodology Series: Report 5 – Weighting and Variance Estimation*. For example, nonresponse adjustments were done separately by county, thus accounting for the differences noted above by the size and urbanicity of the counties. In addition, the weights were also adjusted to be consistent with data from the control totals so that other residual biases could be reduced.

6.3 Overall Response Rates

This section presents the overall, or unconditional, response rates for the adult, child, and adolescent interviews. Table 6-7 gives these response rates for the entire state and by county. As discussed in Chapter 5, the overall rates are the product of screener and extended response rates. For the adult interview, the overall adult response rate is the screener response rate (from Table 6-1) multiplied by the adult response rate (from Table 6-2). This rate is computed using equation (7). The child and adolescent overall rates are the product of the overall rate for the adult (from Table 6-7), which accounts for both the screener and adult interview participation, multiplied by the appropriate child or adolescent extended response rate (from Table 6-2). These are the rates given by equations (8) and (9), respectively.

Since the response rates in these tables are the product of two or more rates at the interview level, the previously described issues surrounding the differences in rates by county, type of household, and characteristic of the sampled person also apply here. The overall adult response rate is about four percentage points lower than it was in CHIS 2001.

Table 6-7. Overall response rates by sampling stratum and type of interview

	Adult overall response rate	Child overall response rate	Adolescent overall response rate
State Total	33.5	27.3	19.2
Los Angeles	28.1	22.5	15.9
San Diego	34.5	29.0	20.6
Orange	31.4	24.4	15.4
Santa Clara	36.7	29.6	22.0
San Bernardino	36.3	29.1	20.1
Riverside	34.7	28.9	19.1
Alameda	34.1	27.6	19.2
Sacramento	38.0	29.6	20.2
Contra Costa	39.0	31.1	25.3
Fresno	36.7	31.6	21.1
San Francisco	26.5	21.0	15.4
Ventura	34.4	30.5	20.9
San Mateo	33.5	27.0	17.1
Kern	41.2	32.9	23.9
San Joaquin	34.4	29.8	18.0
Sonoma	37.9	34.5	21.5
Stanislaus	38.1	32.2	23.2
Santa Barbara	39.4	34.0	26.5
Solano	37.6	27.6	22.7
Tulare	42.8	33.1	26.7
Santa Cruz	36.9	29.6	25.3
Marin	35.5	31.4	20.8
San Luis Obispo	41.8	36.6	26.3
Placer	38.4	30.5	25.7
Merced	35.4	28.7	23.0
Butte	44.3	41.3	26.9
Shasta	42.2	36.6	23.0
Yolo	42.7	35.1	25.1
El Dorado	38.3	31.2	22.1
Imperial	38.4	27.7	25.5
Napa	36.9	32.9	25.3
Kings	37.1	32.7	23.9
Madera	37.3	31.7	25.6
Monterey, San Benito	36.7	30.0	20.5
Del Norte, Humboldt	45.7	38.8	27.8
Lassen, Modoc, Siskiyou, Trinity	47.3	43.5	33.0
Lake, Mendocino	41.9	36.7	26.1
Colusa, Glen, Tehama	42.8	34.3	24.4
Sutter, Yuba	43.5	40.1	30.8
Plumas, Nevada, Sierra	38.9	31.9	28.0
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	39.5	35.5	24.7

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

7. COMPARISONS TO OTHER SURVEYS

In this chapter the response rates obtained in CHIS 2003 are compared to rates from other surveys. The first section compares the response rates from the 2001 and 2003 CHIS surveys. In the second section, response rates observed in CHIS 2003 are contrasted with other RDD surveys conducted after 2001. For comparisons to earlier surveys, see the *CHIS 2001 Methodology Series: Report 4 – Response Rates* (UCLA Center for Health Policy Research, 2002).

7.1 Comparisons of Response Rates between CHIS 2003 and CHIS 2001

Although there are some variations in sampling and content, CHIS 2003 and 2001 are very similar. One adult is sampled from each household and asked to complete a relatively lengthy interview, and then other household members are interviewed if there are children and/or adolescents present in the household.

The response disposition codes and formulas used to compute the response rates in CHIS 2003 are similar to the ones used in CHIS 2001. As a result, differences in response rates in these two surveys may be viewed as an indicator of public cooperation in California between 2001 and 2003. Table 7-1 presents a summary of the differences of response rates for the screener interview, extended interviews and overall response rates by type of interview. Table 7-1 shows that state-level response rates declined from 2001 to 2003, with decreases between 4 and 11 percentage points for the overall rates. This decrease in response rate over two cycles is consistent with the decline in RDD response rates observed by Curtin, Presser, and Singer (2003). Some of this downward shift may be explained by the increase in refusal rates in California following September 11, 2001, as described by DiSogra et al. (2003). The appendix has tables showing the rates for each county from 2001 and 2003.

7.2 Comparisons of Response Rates between CHIS 2003 and Other Surveys

In this section we compare the CHIS 2003 to response rates from other surveys conducted after 2001. These comparisons are difficult to make because other surveys are different in terms of type

Table 7-1. Comparison of state-level response rates between CHIS 2001 and CHIS 2003

Type	CHIS 2001	CHIS 2003	Difference
Screener Interview	59.2%	55.9%	3.3%
Extended Interview			
Adult	63.7%	59.9%	3.8%
Child	87.6%	81.4%	6.2%
Adolescent	63.5%	57.3%	6.2%
Adolescent[1]	84.5%	83.3%	1.2%
Overall			
Adult	37.7%	33.5%	9.0%
Child	33.0%	27.3%	7.8%
Adolescent	23.9%	19.2%	10.8%

[1]Adolescent response rate with cases where permission was not granted removed from the denominator

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

of person selected for the interview, length of interview, and the fact that CHIS is conducted in California. A more generic reason for the difficulty of comparisons to other surveys has to do with the lack of detailed information on disposition codes available for most RDD surveys conducted in the United States as noted in several places such as by McCarthy (2003). A publication by AAPOR (2004) is a recent attempt to address this problem. We mention some specific comparability problems below. Another factor complicating comparability is the recency of the data collection as response rates in RDD surveys have been decreasing over the years. This section includes surveys conducted after 2001 only.

One RDD survey that can be compared to CHIS 2003 is the 2003 California Behavioral Risk Factor Surveillance System (BRFSS) (www.cdc.gov/brfss/technical_infodata/2003QualityReport.htm). In this survey, one adult in each household is sampled and asked to complete an interview of about 20 minutes on health-related topics. The BRFSS interview is shorter than CHIS 2003 and does not have multiple interviews within the household, but has many other similarities with CHIS.

The Centers for Disease Control and Prevention (CDC) (2004) publishes response rates for the BRFSS. The CDC report for the survey shows detailed disposition codes, very much in the spirit of the AAPOR (2004) recommendations. Despite the detail given, it is very difficult to map the 2003 California BRFSS disposition codes unambiguously to the corresponding disposition codes used in CHIS 2003 because different survey organizations use different classification schemes to create the disposition codes. The codes from both systems provide much needed information for the conduct of the operations of the survey, but they are not the same. This difference highlights the difficulty of making direct comparisons between surveys.

Several cooperation and response rates are reported for the 2003 California BRFSS in Table 6 of the CDC Quality Report. The response rates of interest are the 39.0 percent CASRO response rate and 31.5 percent overall response rate for California. We are not sure which rate is more comparable to the CHIS overall adult response rate of 33.5 percent. Nevertheless, the overall rates for the two surveys are not very different, given the potential discrepancies in the definitions and methods. For example, the CDC report suggests that the rates for some states do not include persons who speak a language other than English (or English and Spanish) as eligible. A lower rate for CHIS 2003 would be anticipated given the longer interview and the multiple interviews per household. We also note that the CASRO response rate for the California BRFSS declined some six percentage points from 2002 to 2003, while the overall response rate was unchanged.

Another survey that can be compared to CHIS 2003 is the 2002 National Survey of America's Families (NSAF) (Brick, et al., 2003). One advantage of using this study for comparison is that Westat also conducted the 2002 NSAF, so the methods of computing response rates are similar to those used for CHIS. The 2002 NSAF also had a large enough sample size in California to provide reliable estimates of the response rates. A major difference between the two studies that has a large effect on the response rates is the use of monetary incentives. The 2002 NSAF used monetary incentives while CHIS 2003 did not. As discussed in more detail in Chapter 8, monetary incentives are typically very effective in raising response rates. Another difference between CHIS 2003 and the 2002 NSAF is the way the sampling and interviewing was done. The 2002 NSAF used the approach of only enumerating some persons in the screener, but children rather than adults were enumerated in the NSAF screener. With the method used to sample adults in CHIS 2003, there was no enumeration for most cases. The method used to sample adults during the screening process is described in Chapter 8. The difference is related to the focus of the survey. Another difference is that households with children and with low income were sampled at a higher rate than other households in NSAF, so a substantial fraction of the households had to complete only the screener interview. In CHIS 2003, an adult was sampled in every household.

The 2002 NSAF overall response rate for adults in the California RDD sample was 44.2 percent. The NSAF rates are given in Table 5-14 in the report by Brick et al. (2002). The NSAF response rate is higher than the CHIS rate and much of the difference is probably due to the factors mentioned above, especially the use of monetary incentives in the NSAF.

The NSAF surveys also provide some additional information on bias as well as on response rates. A special followup study was conducted for the 1997 NSAF (See Groves and Wissoker, 1999). The study used intensive methods to recontact a subsample of households that would not participate in the NSAF and compared their characteristics to the NSAF respondents. The study also included other analytic investigations of indicators of nonresponse bias. The results of the study suggested that the nonresponse did not substantially bias the estimates from the survey. Since many of the same procedures were used in CHIS 2003 and 2001, the results are encouraging for CHIS. These results are consistent with the findings of Keeter et al. (2000), Curtin, Presser, and Singer (2000), and Keeter et al. (2004).

8. METHODS OF INCREASING RESPONSE RATES

In this chapter, we examine procedures used to increase response rates in CHIS 2003. The ability to evaluate the effects of the procedures is limited because no experiments were conducted especially for this purpose. Most of these methods were the result of observations and experiences gained in the first administration of CHIS in 2001 and in other RDD surveys.

One of the decisions made for CHIS 2003 that had important implications for the screener response rate was the use of a new procedure for selecting adults in the household. In CHIS 2001, only adult household members were enumerated in the screener interview. This decision was based on an experiment in the National Household Education Survey (NHES). See Brick and Collins (1997) for a description of the experiment. In CHIS 2003, a hybrid version of this approach and a new method for sampling adults called the Rizzo method (Rizzo et. al., 2004) was used to sample adults in the household. The advantage of this method is that the enumeration of adult household members is bypassed in most households, thus it is less intrusive but it results in a valid probability sample. A sampled adult is selected using the following algorithm after the number of adults in the household is ascertained. In households with one adult, the adult is selected with certainty. In households with two adults, a random number is used to determine if the screener respondent is selected or if the other adult is selected. In households with three or more adults, a random number is used to determine if the screener respondent is selected. If the screener respondent is not selected and there are more than two adults, the "most recent birthday" method is used to sample from the remaining adults. If the number of adults and/or birthday is not known then a full enumeration is required. See Chapter 4 of *CHIS 2003 Methodology Series: Report 1 – Sample Design* for additional details of the adult sample selection. Even though the CHIS 2003 screener response rate was lower than the 2001 screener rate, the effect of the Rizzo method for selecting an adult could not be assessed because it is confounded with other factors.

Another sampling decision consistent with the procedure used in CHIS 2001 is that only one adult is randomly selected from each household even when more are present. The restriction of one sampled adult per household is intended to limit the burden on the household and to increase response rates. Nevertheless, the burden on households in CHIS is substantial, especially if there are children and adolescents present. In this case, up to three interviews could be requested from the same household. If the sampled adult is the person most knowledgeable about the sampled child, then the same adult is asked to complete two interviews.

The within-household interviewing requirement is related to response rates because it requires asking the same person to participate more than once, and it increases the total interview length. Bogen (1996) found little experimental evidence on a relationship between the length of the interview and response rates, but her research did not cover multiple interviews within the same household. From an operational perspective and from monitoring the interviews during the conduction of CHIS 2003, there is wide agreement that the survey length and having multiple interviews within the household had a negative effect on response rates. However, limiting the sample size to one adult per household is clearly useful in preventing lower response rates.

Perhaps the most effective and the most controversial method of increasing response rates in RDD surveys is giving monetary incentives to respondents. The option of offering monetary incentives was carefully considered for CHIS but decided against. The literature on monetary incentives in RDD surveys is relatively recent. Singer, Van Hoewyk, and Maher (2000) discuss the issues associated with incentives in telephone surveys. Westat has conducted several experiments with monetary incentives, including in the NSAF, a large RDD survey that has many features similar to CHIS. See Cantor et al. (1998) for a discussion of the effect of these incentives. An experiment was also conducted in the 2003 NHES and these results are currently being evaluated. Since incentives were not offered in CHIS 2003, we do not discuss this topic further.

Another important procedure used to increase response rates that cannot be evaluated for CHIS 2003 is the interviewer training protocol. The interviewer training is discussed in detail in *CHIS 2003 Methodology Series: Report 2 – Data Collection Methods.* In particular, that report describes training each interviewer to help them to avoid refusals. Interviewers who were allowed to do refusal conversions were also given special training before they were permitted to attempt to contact households or persons who previously refused to cooperate. The interviewers were also given special training to alert them to the cultural issues that might affect the response rates for key demographic groups such as American Indians/Alaska Natives.

8.1 Advance Letter Mailing

We turn to procedures used to improve response rates that have associated quantitative outcomes, beginning with the analysis of the advance mailing. We stress that the data presented are indicators of the effect, but they are not experimental data and many of the effects are confounded with

other factors. For example, the advance mailing was sent to all households with addresses. We can examine the difference in response rate for those with an address and those without an address, but we cannot conclude the differences are due to the mailing. In fact, previous experimental research shows the actual effect is much smaller than the difference suggests (see Brick et al., 2000) because households with addresses are more likely to respond with or without an advance mailing. Despite this caveat, the results are informative.

Table 8-1 presents a summary of the state-level interview response rates for the screener, adult interview, child interview, and adolescent interview by whether or not an advance letter was mailed to the household. For each interview, the households that could be sent an advance letter had higher response rates. This pattern of response rates was also seen in CHIS 2001. The differences in response rates are especially large for the screener and adult interviews. See Table 6-1 for the county-level differences by mailable status for the screener. As noted above, much of the difference must be due to the different propensities of the households to respond, irrespective of whether a letter is mailed. We hypothesize that the effect of the letter on respondents is to legitimize the survey. If this were the case, we would expect the differences in response rates for the extended interviews (adult, child, or adolescent) to be small because the screening interview should have much the same effect (if it is completed). Nevertheless, there is a large difference between the group that was mailed the letter and the group that was not mailed the letter for the adult interview. It is reasonable to conclude that the differences are largely due to the attributes of the households rather than the effect of the advance mailing.

Table 8-1. Interview response rates by type of interview and advance letter

Type	Advance letter mailed		Difference
	Yes	No	
Screener	57.4	48.9	+8.5
Adult interview	61.4	53.1	+8.3
Child interview	81.9	79.6	+2.3
Adolescent interview	58.4	52.1	+6.3

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

Even if large increases in the response rates cannot be attributed to the advance letter mailing, the procedure is a relatively low-cost approach to increase response rates that has other desirable effects. For example, some respondents told interviewers that they read the letter and went to the survey web site to find out more about the survey before they were called. These respondents tended to be more

willing to cooperate, which reduces the cost of data collection. Furthermore, the research noted above found some increases in response rates as a result of the advance mailing.

8.2 Repeated Call Attempts

Another method used to increase response rates in CHIS 2003 was repeated attempts to contact households and sampled persons. In some surveys, the number of call attempts is limited in order to reduce the cost of data collection and to complete the survey in a short time period. These surveys will more likely exclude some households and "hard-to-reach" persons who are not usually at home. The procedures implemented in CHIS 2003 allowed many attempts to severely limit the bias from this source of nonresponse.

Figure 8-1 shows the percentage of all completed interviews by the number of call attempts for both the screener and the adult interview. A call attempt is a telephone call placed to the sampled household. The counts of attempts in CHIS 2003 were associated with the type of interview. The first calls were to complete the screener interview. Once the screener was completed, additional calls to the household may have been required to complete the adult interview, and still further attempts to complete the child and/or the adolescent interviews.

The patterns are similar for both the screener and adult interviews in Figure 8-1. Most interviews are completed within a few call attempts. The median number of call attempts for the completed screeners is three and for the adult interview is two. However, there is a long tail for each distribution. The 75th percentile of the number of completed interviews is not reached until the sixth attempt for the screener and until the fourth attempt for the adult interview. If the number of call attempts was not extended to at least 10 attempts, then the response rate would decrease by about 10 percent for both types of interviews. Similarly, the response rates were increased by about 5 percentage points by allowing for more than 14 call attempts for both the screener and the adult interviews. However, the cost for extra call attempts must be balanced against the gains in the number of completed interviews. The patterns for the child and adolescent interview are more compressed with fewer attempts needed to complete these interviews.

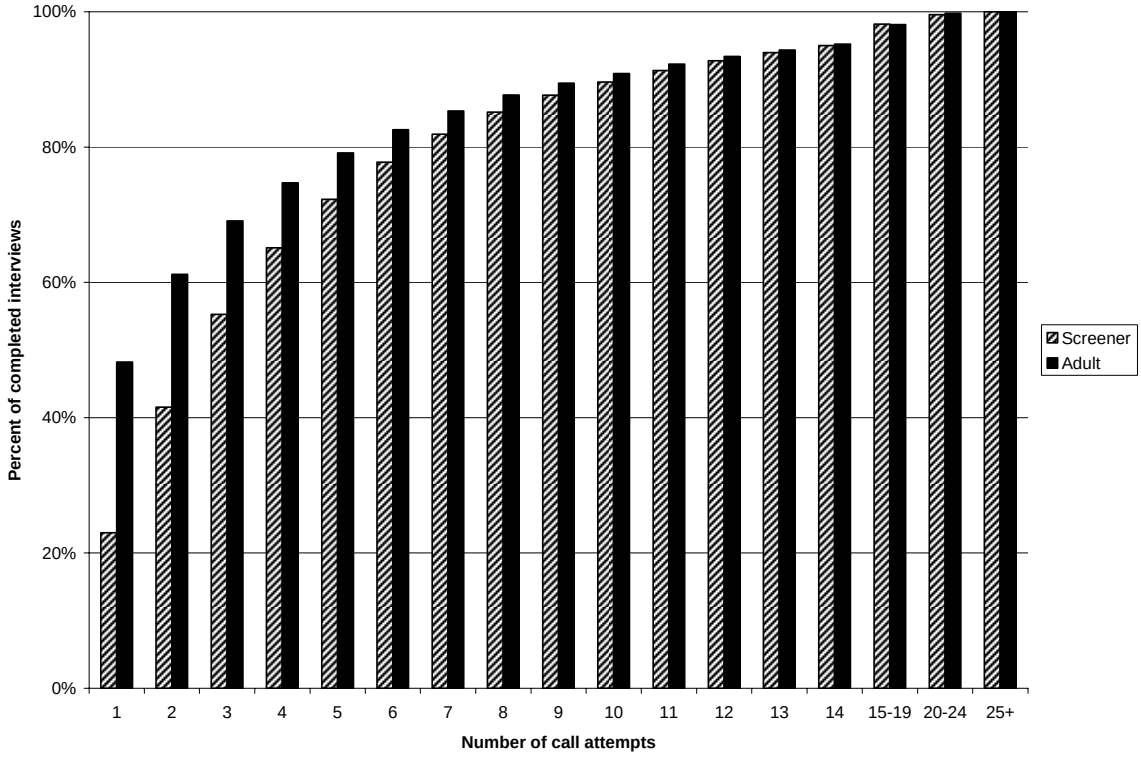

Figure 8-1. Percentage of completed screener and adult interviews by the number of call attempts

8.3 Refusal Conversion Attempts

An effective method of increasing response rates in an RDD survey is to recontact households and persons who refuse to participate in the initial interview and to ask them to reconsider and complete the interview. In CHIS 2001, attempts for refusal conversion were implemented for all screener nonrespondents. In CHIS 2003, these procedures were implemented only for respondents who refused to conduct the screener interview in a random subsample of 60 percent of the sample that was assigned during sample selection. If a household refused but was not selected for the subsample, no further calls were made to convert it. Refusal conversion subsampling has been used in other surveys, including the National Survey of America's Families in 2002 (Brick et al., 2003). In most cases, subsampling is done for cost reasons after the survey is already in the field rather than the "planned" refusal subsampling implemented in CHIS 2003. Hansen and Hurwitz (1946) originally proposed the idea and Srinath (1971) gave some extensions. More recently, Elliott, Little, and Lewitzky (2000) examined its use in two surveys, one conducted by the University of Michigan's Survey Research Center and the other conducted by the U.S. Census Bureau. Additional benefits of subsampling of refusals in CHIS 2003 included a faster

completion of the data collection because the fairly lengthy process of refusal conversion (hold cases for a time, release and make contact attempts, hold again for second refusal contact attempts) was largely "front loaded." This strategy avoided the need to process as many cases later in the data collection period. There were also some savings in data collection costs due to reducing the direct costs associated with refusal conversion mailings (an express mail was used in refusal conversion). A possible benefit that was noted in previous efforts was a slight increase in initial cooperation rates when predesignated cases for refusal conversion were fielded. As shown in Figure 8-2, there is some suggestion that this initially held in CHIS 2003, but there is no evidence of any real effect. With refusal conversion subsampling, only weighted response rates can be computed in order to reflect the subsampling of cases that are converted.

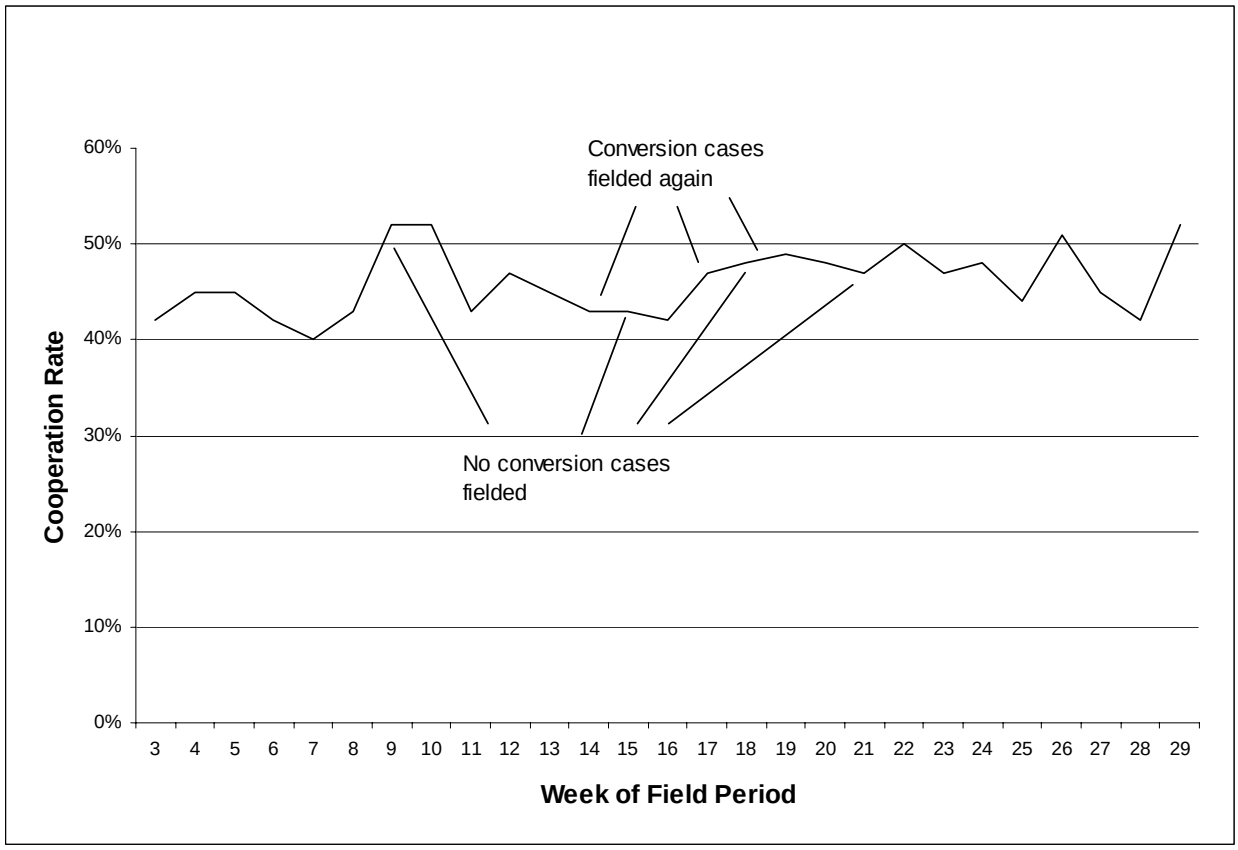

Figure 8-2. Screener cooperation rate by week of field period

In CHIS 2003, the success in converting households using the refusal conversion attempts are given in Table 8-2. The table shows the cooperation rate (the ratio of the number completed to the number that were either completed or refused expressed as a percentage) by the level of effort and the type of interview for the RDD and the surname list samples. The cooperation rates are greater than the response

rates presented earlier because other types of nonresponse such as not contacting the household or person after multiple attempts are not included in the denominator of these rates. More details on the cooperation rates and other sources of nonresponse are discussed in *CHIS 2003 Methodology Series: Report 2 – Data Collection Methods*.

Table 8-2. Completion rates by level of effort and type of interview

Level	Percentage Completed at Each Level of Effort			
	Screener	Adult[1]	Child[2]	Adolescent[2]
Initial Contact	44.5	69.4	90.4	86.0
First Conversion Attempt	35.9	30.1	---	---
Second Conversion Attempt	20.7	33.6	---	---

[1] Conditional on the number of completed screen interviews

[2] Conditional on the number of completed adult interviews

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

If the initial attempt to complete the screener or adult interview was not successful, an express letter was sent prior to attempting refusal conversion[9]. The rationale for sending the express letter is previous experimental evidence of the substantial positive effect such high visibility letters have on the conversion rates. See Brick et al. (1999). The cooperation rate in CHIS 2003 for the first refusal conversion effort for the screener interview is about 36 percent, while for the adult, child, and adolescent interview the conversion rates are around 30 percent. The screener cooperation rate for the first refusal conversion effort is about four percentage points lower than it was for CHIS 2001.

No refusal conversion attempt was originally planned for CHIS 2001 if a parent refused permission for interviewing the sampled adolescent. Because the permission rate was relatively low, a decision was made to send an express mail letter explaining the purpose and content of the adolescent interview to parents who did not give permission. It also indicated that they would be asked again if they would give permission to interview the adolescent. This conversion effort was repeated in CHIS 2003 and proved somewhat successful, as 26 percent of the recontacted parents gave permission.

As noted in Chapter 4, we also attempted a second refusal conversion for screener and adult interviews that were not completed in the first conversion attempt and deemed as not definite refusals. The results of these attempts were relatively successful. Table 8-2 shows, for example, that 44.5 percent

[9] Note that refusal letters were sent only for those telephone numbers for which an address was obtained from reverse directory services. For screener refusals, letters were sent only for those numbers designated for conversion.

of the screener interviews were completed during the initial contact, and in 69.4 percent of these completed screener interviews the adult interview was finalized. Of the balance that refused to cooperate but were eligible for a conversion attempt, 35.9 percent of the screener interviews were completed after the first conversion attempt, and in 30.1 percent of those completed screener interviews the adult interview was finalized. Finally, 20.7 percent of the screener interviews were completed during the second conversion attempt, and in only 33.6 percent of these screener interviews were adult interviews completed. Table 8-2 shows that about one-fifth (20.7%) of the screener refusals and about one-third (33.6%) of the adult interview refusals completed on a second refusal conversion attempt. The conversion rate on the second attempt for screeners was about 15 percentage points lower than the first attempt. However, for the adult extended interview, the second attempt resulted in a slightly larger conversion rate than the first attempt.

The refusal conversion effort greatly increased the number of completed interviews and, as a result, improved the survey response rates. Overall, considering all the interviews completed in both the RDD sample and the supplemental samples, converted refusals accounted for 23 percent of the completed screeners, 9 percent of the adult extended interviews, 2 percent of the child interviews, and 5 percent of the adolescent interviews.

As noted earlier, express letters were sent to every household and adult in the RDD sample who refused to be interviewed if an address was available for them. The effect of the express letter can be examined by comparing the first refusal conversion rates by whether an express letter was sent, but differences cannot be attributed to the letters because no experimental data exists. This situation is similar to the advance letter analysis because the express letter mailing is confounded with other attributes of telephone numbers associated with an address

Table 8-3 shows the first refusal conversion rates by express letter status for the RDD sample only. Overall, the conversion rates for the RDD sample cases only are lower than the rates for the full sample given in Table 8-2. The difference by express letter status for the screener interviews is about 11 percentage points. The earlier research suggests that, at most, five to eight percentage points might be attributable to the express letter. Nevertheless, the express letter appears to be effective for the screener interview. The rate for the adult extended interviews is nine percentage points higher for those sent an express mailing. However, the adolescent permission request rates are only two percentage points higher for those mailed an express letter. No experimental data on the effectiveness of express mailings at the extended level have been published, so it is difficult to draw any conclusions from these data.

Table 8-3. First refusal conversion rates by type of interview and express mailing status, RDD sample only

	Screener	Adult	Adolescent Permission
Express letter	38.0	31.0	26.0
No express letter	27.0	22.0	24.0

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

8.4 Proxy Reports

A special provision was made to permit other adults to respond for sampled adults who were over 65 and unable to participate because of mental or physical disabilities. No other adult proxy interviews were permitted in CHIS 2003. A total of 171 adult proxy interviews were done in the RDD sample. Proxy respondents had to be adult household members who were knowledgeable about the sampled person's health. The proxy respondent was almost always a spouse or child of the sampled adult. While the number of interviews completed using the proxy interviews is relatively small, it does provide coverage for a group of adults with very different health characteristics that would not otherwise be included in the survey.

8.5 In-Language Interviews

A very important procedure incorporated to enhance the response rate in CHIS 2003 was translating and conducting the interviews in the language requested by the sampled person. The languages included were: Spanish, Chinese (Cantonese and Mandarin), Korean, and Vietnamese. In many cases, households that did not speak English would not have been included in CHIS had it not been for the additional languages. In some cases, the respondents would have tried to respond in English but the quality of the interviews would have been much lower if the other languages were not provided. The translation of the instruments provides a common basis for the interviewers that would not be available otherwise.

Table 8-4 gives the number of interviews that were completed by language. The use of languages other than English had a dramatic effect on the response rates. Just under 9,000 households completed the screener using a language other than English. This accounts for about 13 percent of all the

completed interviews in CHIS 2003. Spanish is the most frequently used language, with about 81 percent of the non-English screeners being completed in Spanish. Vietnamese was the second most frequently used language in the interviews, although it was primarily required for the Vietnamese surname list sample.

Table 8-4. Number of completed interviews by language and sample type

Sample Type	English	Non-English	Spanish	Vietnamese	Korean	Cantonese	Mandarin	Total
Screener								
Total	57,731	8,926	7,229	482	513	347	355	66,657
RDD	57,659	8,584	7,229	305	348	347	355	66,243
Korean	43	170	-	7	163	-	-	213
Vietnamese	29	172	-	170	2	-	-	201
Adult								
Total	37,136	4,908	3,737	322	326	277	246	42,044
RDD	37,103	4,715	3,737	223	232	277	246	41,818
Korean	18	94	-	-	94	-	-	112
Vietnamese	15	99	-	99	-	-	-	114
Child								
Total	6,695	1,831	1,595	82	73	42	39	8,526
RDD	6,691	1,789	1,595	61	52	42	39	8,480
Korean	3	21	-	-	21	-	-	24
Vietnamese	1	21	-	21	-	-	-	22
Permission								
Total	4,013	1,082	936	54	32	28	32	5,095
RDD	4,007	1,062	936	42	24	28	32	5,069
Korean	1	8	-	-	8	-	-	9
Vietnamese	5	12	-	12	-	-	-	17
Adolescent								
Total	3,723	287	261	8	5	6	7	4,010
RDD	3,713	283	261	7	2	6	7	3,996
Korean	3	3	-	-	3	-	-	6
Vietnamese	7	1	-	1	-	-	-	8

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

The use of a language other than English varied by the type of interview. In both the screener and the adult interviews, between 11 and 13 percent of the total completed interviews were done in languages other than English. The child interview had the largest percentage of all interviews

completed in languages other than English, about 21 percent. The lowest percentage was for the adolescent interview; about 7 percent of these interviews were done in a language other than English.

The other main source of variation in the use of other languages is the type of sample. For example, 87 percent of the adult Vietnamese surname list sample interviews were conducted in Vietnamese. (Table 8-4 shows 99 adult interviews completed in Vietnamese; there was a total of 114 adult interviews in the Vietnamese surname list sample.) Similarly, about 84 percent of the Korean adult interviews were done in a language other than English.

Overall the effect of including these languages on the response rates was substantial. As with the proxy interviews, the reduction of nonresponse bias due to the use of the multiple language interviews is probably even greater than the simple response rate computations suggest. The non-English-speaking population is likely to have different health and other characteristics measured in CHIS 2003. As the bias equation (1) indicates, the combination of reducing nonresponse and removing systematic differences between the respondents and nonrespondents is very effective for reducing bias.

APPENDIX A

Table A-1. County Screener Response Rates from CHIS 2001 and CHIS 2003

	2001 Response rate	2003 Response rate	Difference
State Total	59.2	55.9	-3.3
Los Angeles	56.9	51.0	-5.9
San Diego	59.9	56.8	-3.1
Orange	59.0	54.2	-4.8
Santa Clara	57.1	57.1	0.0
San Bernardino	63.7	61.0	-2.7
Riverside	62.2	59.1	-3.1
Alameda	57.6	54.9	-2.7
Sacramento	61.3	60.3	-1.0
Contra Costa	57.6	58.8	1.2
Fresno	64.0	59.5	-4.5
San Francisco	50.7	44.2	-6.5
Ventura	59.4	57.1	-2.3
San Mateo	53.8	54.6	0.8
Kern	68.9	62.9	-6.0
San Joaquin	64.7	58.1	-6.6
Sonoma	61.3	56.6	-4.7
Stanislaus	65.7	61.0	-4.7
Santa Barbara	62.1	61.0	-1.1
Solano	61.5	61.9	0.4
Tulare	67.7	66.2	-1.5
Santa Cruz	57.7	57.7	0.0
Marin	54.7	54.5	-0.2
San Luis Obispo	61.6	64.4	2.8
Placer	60.3	60.9	0.6
Merced	66.2	61.4	-4.8
Butte	67.3	63.8	-3.5
Shasta	65.7	63.2	-2.5
Yolo	66.2	64.4	-1.8
El Dorado	57.8	59.4	1.6
Imperial	67.0	62.0	-5.0
Napa	59.0	56.4	-2.6
Kings	65.5	60.1	-5.4
Madera	67.8	62.2	-5.6
Monterey, San Benito	60.7	58.1	-2.6
Del Norte, Humboldt	65.4	64.3	-1.1
Lassen, Modoc, Siskiyou, Trinity	66.5	65.4	-1.1
Lake, Mendocino	60.9	61.8	0.9
Colusa, Glen, Tehama	68.9	68.0	-0.9
Sutter, Yuba	66.2	67.3	1.1
Plumas, Nevada, Sierra	59.5	58.8	-0.7
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	58.0	57.2	-0.8

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey.

Table A-2. County Adult Response Rates from CHIS 2001 and CHIS 2003

	2001 Response rate	2003 Response rate	Difference
State Total	63.7	60.0	-3.7
Los Angeles	60.0	55.1	-4.9
San Diego	63.3	60.7	-2.6
Orange	60.3	58.0	-2.3
Santa Clara	61.2	64.3	3.1
San Bernardino	64.0	59.5	-4.5
Riverside	64.7	58.7	-6.0
Alameda	65.2	62.1	-3.1
Sacramento	65.7	63.0	-2.7
Contra Costa	64.9	66.3	1.4
Fresno	59.8	61.6	1.8
San Francisco	59.1	59.9	0.8
Ventura	63.7	60.3	-3.4
San Mateo	60.4	61.4	1.0
Kern	66.6	65.5	-1.1
San Joaquin	63.7	59.2	-4.5
Sonoma	67.8	67.0	-0.8
Stanislaus	64.2	62.4	-1.8
Santa Barbara	66.1	64.6	-1.5
Solano	63.9	60.8	-3.1
Tulare	64.6	64.7	0.1
Santa Cruz	68.3	64.0	-4.3
Marin	70.4	65.2	-5.2
San Luis Obispo	69.7	64.9	-4.8
Placer	68.2	63.0	-5.2
Merced	64.0	57.7	-6.3
Butte	67.6	69.5	1.9
Shasta	69.4	66.7	-2.7
Yolo	69.3	66.3	-3.0
El Dorado	67.6	64.4	-3.2
Imperial	63.5	61.9	-1.6
Napa	66.6	65.4	-1.2
Kings	66.6	61.7	-4.9
Madera	67.3	59.9	-7.4
Monterey, San Benito	62.9	63.1	0.2
Del Norte, Humboldt	69.6	71.0	1.4
Lassen, Modoc, Siskiyou, Trinity	69.6	72.3	2.7
Lake, Mendocino	68.6	67.8	-0.8
Colusa, Glen, Tehama	65.9	63.0	-2.9
Sutter, Yuba	64.6	64.7	0.1
Plumas, Nevada, Sierra	70.5	66.1	-4.4
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	72.4	69.1	-3.3

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

Table A-3. County Child Response Rates from CHIS 2001 and CHIS 2003

	2001 Response rate	2003 Response rate	Difference
State Total	87.6	81.4	-6.2
Los Angeles	83.7	80.2	-3.5
San Diego	88.5	84.2	-4.3
Orange	84.5	77.5	-7.0
Santa Clara	92.2	80.7	-11.5
San Bernardino	91.2	80.3	-10.9
Riverside	90.8	83.2	-7.6
Alameda	90.3	81.1	-9.2
Sacramento	86.3	77.8	-8.5
Contra Costa	88.9	79.7	-9.2
Fresno	88.9	86.2	-2.7
San Francisco	88.5	79.4	-9.1
Ventura	85.4	88.7	3.3
San Mateo	84.5	80.6	-3.9
Kern	89.2	79.9	-9.3
San Joaquin	89.9	86.7	-3.2
Sonoma	95.0	91.1	-3.9
Stanislaus	85.8	84.7	-1.1
Santa Barbara	89.7	86.2	-3.5
Solano	87.0	73.3	-13.7
Tulare	91.0	77.2	-13.8
Santa Cruz	88.6	80.2	-8.4
Marin	89.1	88.3	-0.8
San Luis Obispo	93.1	87.6	-5.5
Placer	90.5	79.4	-11.1
Merced	86.7	80.9	-5.8
Butte	89.6	93.2	3.6
Shasta	87.0	86.9	-0.1
Yolo	95.2	82.1	-13.1
El Dorado	92.5	81.6	-10.9
Imperial	82.4	72.1	-10.3
Napa	84.0	89.1	5.1
Kings	89.5	88.2	-1.3
Madera	85.6	85.1	-0.5
Monterey, San Benito	87.2	81.8	-5.4
Del Norte, Humboldt	92.9	84.9	-8.0
Lassen, Modoc, Siskiyou, Trinity	96.1	92.0	-4.1
Lake, Mendocino	87.8	87.5	-0.3
Colusa, Glen, Tehama	90.7	80.0	-10.7
Sutter, Yuba	90.4	92.1	1.7
Plumas, Nevada, Sierra	90.0	82.0	-8.0
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	93.7	89.8	-3.9

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

Table A-4. County Adolescent Response Rates from CHIS 2001 and CHIS 2003

	2001 Response rate	2003 Response rate	Difference
State Total	63.5	57.3	-6.2
Los Angeles	58.5	56.5	-2.0
San Diego	62.1	59.8	-2.3
Orange	52.3	49.1	-3.2
Santa Clara	60.1	60.0	-0.1
San Bernardino	68.0	55.4	-12.6
Riverside	64.8	55.2	-9.6
Alameda	57.9	56.2	-1.7
Sacramento	65.3	53.3	-12.0
Contra Costa	64.1	64.8	0.7
Fresno	64.3	57.5	-6.8
San Francisco	51.4	58.0	6.6
Ventura	60.6	60.8	0.2
San Mateo	65.0	51.1	-13.9
Kern	66.2	58.1	-8.1
San Joaquin	65.7	52.3	-13.4
Sonoma	65.3	56.7	-8.6
Stanislaus	60.7	60.9	0.2
Santa Barbara	63.2	67.3	4.1
Solano	65.6	60.3	-5.3
Tulare	63.7	62.4	-1.3
Santa Cruz	70.5	68.6	-1.9
Marin	61.2	58.4	-2.8
San Luis Obispo	65.0	63.0	-2.0
Placer	70.1	67.0	-3.1
Merced	65.2	64.8	-0.4
Butte	64.5	60.7	-3.8
Shasta	63.2	54.5	-8.7
Yolo	68.8	58.7	-10.1
El Dorado	74.2	57.9	-16.3
Imperial	70.6	66.4	-4.2
Napa	61.1	68.5	7.4
Kings	70.1	64.4	-5.7
Madera	70.4	68.6	-1.8
Monterey, San Benito	66.4	56.0	-10.4
Del Norte, Humboldt	69.1	60.9	-8.2
Lassen, Modoc, Siskiyou, Trinity	68.1	69.7	1.6
Lake, Mendocino	67.9	62.4	-5.5
Colusa, Glen, Tehama	70.4	57.0	-13.4
Sutter, Yuba	65.9	70.8	4.9
Plumas, Nevada, Sierra	78.8	72.0	-6.8
Alpine, Amador, Calaveras, Inyo, Mariposa, Mono, Tuolumne	75.2	62.5	-12.7

Source: UCLA Center for Health Policy Research, 2003 California Health Interview Survey

REFERENCES

American Association for Public Opinion Research (AAPOR). (2004). Standard definitions: Final disposition of case codes and outcome rates for RDD telephone surveys and in-person household surveys: 3rd edition. *American Association for Public Opinion Research*, Lenexa, Kansas. Available from http://www.aapor.org.

Council of American Survey Research Organizations (CASRO). (1983). The report of the CASRO task force on response rates. In F. Wiseman (Ed.), *Improving data quality in a sample survey.* Cambridge, MA: Marketing Science Institute.

Bogen, K. (1996). The effect of questionnaire length on response rates—A review of the literature, *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 1020-1025). Washington DC.

Brick, J.M., Ferraro, D., Strickler, T. Rauch, C., (2003). *Report No. 8: 2002 NSAF Response Rates*. Washington, DC: Urban Institute.

Brick, J.M., and Collins, M. (1997). *An experiment in random-digit-dial screening*. U.S. Department of Education, Office of Educational Research and Improvement, NCES 98-255.

Brick, J.M., Flores-Cervantes, I., and Cantor, D. (1999). *Report No. 8: 1997 NSAF Response Rates and Methods Evaluation*. Washington DC: Urban Institute.

Brick, J.M., Montaquila, J., and Scheuren F. (2002). Estimating residency rates for undetermined telephone numbers. *Public Opinion Quarterly*, 66, 18-39.

Centers for Disease Control and Prevention (CDC). (2004). *BRFSS Summary Data Quality Report*. Available from www.cdc.gov/brfss/technical_infodata/2003QualityReport.htm.

Cantor, D., Allen, B., Cunningham, P., Brick, J.M., Slobasky, R., Giambo, P. and Kenney, G. (1998). Promised incentives on a RDD Survey, *Proceedings of the Eight International Workshop on Household Survey Nonresponse*, pp. 219-228.

Curtin, R., Presser, S. and Singer, E. (2000). The effects of response rate changes on the Index of Consumer Sentiment. *Public Opinion Quarterly*, 64, 413-428.

Curtin, R., Presser, S., and Singer, E. (2003). Recent response rate changes on the Michigan Survey of Consumer Attitudes. *Paper presented at 2003 AAPOR meetings.*

DiSogra, C., Grant, D., Ramirez, T., Robison, G., Hoegh, H., and Davis, B., (2003). Did the 9/11 terrorist attacks accelerate RDD refusal rates? A California Observation. *Paper presented at 2003 AAPOR meetings*.

Groves, R., and Wissoker, D. (1999). *Report No. 7: 1997 Early NSAF Nonresponse Studies*. Washington DC: Urban Institute.

Keeter, S., Miller, C., Kohut, A. Groves, R., and Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly*, 64, 125-148.

Keeter, S., Best, J., Dimock, M., and Craighill, P. (2004). The Pew Research Center of Survey Nonresponse: Implications for practice. *Paper presented at AAPOR meetings in Phoenix, AZ 2004*.

Lessler, J., and Kalsbeek, W. (1992). *Nonsampling errors in surveys.* NY: John Wiley and Sons.

Madow, W., Nisselson, H., and Olkin, I. (1983). Incomplete data on sample surveys, Volume 1 – Report and case studies, city, state?: Academic Press.

Massey, J., O'Connor, D., and Krotki, K. (1997). Response rates in random digit dialing telephone (RDD) surveys, *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 707-712.

McCarty, C. (2003). Differences in response rates using most recent versus final dispositions in telephone surveys. *Public Opinion Quarterly*, 67, 396-406.

Shapiro, G., Battaglia, M., Camburn, D., Massey, J., and Tompkins, L. (1995). Calling local telephone company business offices to determine the residential status of a wide class of unresolved telephone numbers in a random-digit-dialing sample. *Proceedings of the Survey Research Methods Section of the American Statistical Association* (pp. 975-980).

Singer, E., Van Hoewyk, J., and Maher, M. (2000). Experiments with incentives in telephone surveys. *Public Opinion Quarterly*, 64, 171-188.