

# **Surveys on Attitudes Toward Sexual Orientation: General Public Survey; and Survey of Gays, Lesbians and Bisexuals**

## **Methodology Report**

Charles E. Denk, PhD  
Princeton Survey Research  
September 15, 2000

The study, sponsored by the Kaiser Family Foundation, had two components. The General Public (GP) Survey comprised telephone interviews with a nationally representative sample of 2,283 adults 18 and older living in continental United States telephone households. The Gay, Lesbian and Bisexual (GLB) Survey comprised 405 telephone interviews in selected geographical areas of the U.S. Interviews were completed in both English and Spanish, according to the preference of the respondent. The interviews were conducted by Schulman, Ronca & Bucuvalas, Inc. (SRBI) from February 7 through September 4, 2000.

## **SAMPLE DESIGN**

### **General Public (GP) Survey**

The sample for the GP component was designed to generalize to the U.S. adult population in telephone households, and to allow separate analyses of responses by African-Americans and Hispanic respondents. We employed PSRA's customized "consolidated" oversample design to achieve these objectives in a cost effective manner. This design uses random-digit dialing (RDD) methods, but telephone numbers are drawn disproportionately from area code-exchange combinations with higher than average density of African-Americans and/or Hispanics households.<sup>1</sup> This method increases the proportion of respondents in these target categories, but special weighting adjustments are required to restore the overall representativeness of the sample. The specifics of the composition of survey returns and weighting factors is described below.

For this GP sample, standard *list-assisted* RDD methodology was used so that only working banks of telephone numbers were selected. A *bank* is defined as 100 contiguous telephone numbers, and a *working* bank contains one or more residential listings. This method allows the inclusion in the sample of households that have received a telephone number in a new exchange. Households assigned to new exchanges are likely to have recently moved or experienced a disruption in their telephone service due to some other cause.

### **Gay, Lesbian and Bisexual (GLB) Survey**

The GLB component of the survey was designed to provide a baseline sample of gay, lesbian and bisexual adults to which we might compare the attitudes of the general public. Scientifically sampling the GLB population is complicated by the lack of agreement on a precise definition of its membership and therefore knowledge of its absolute size. These factors make it difficult to evaluate the quality of any candidate sampling frame. The fact that the incidence of gay, lesbian and bisexual adults in the general adult population has been estimated as low as 2% makes RDD sampling a very expensive proposition.

After much discussion we decided that it was preferable to start with a random sample of households and ask for adults willing to identify themselves as gay, lesbian or bisexual, rather than to rely on any commercial GLB listing. For reasons of cost, we assembled a restricted RDD telephone sample geographically limited to 15 major U.S. metropolitan areas — the largest MSA from each of the 14 largest CMSAs, plus Oakland, CA.<sup>2</sup> We further limited our coverage to telephone exchanges associated with the 30 “central cities” in those metropolitan areas. A recent national survey<sup>3</sup> suggests that such a sampling strategy would cover somewhat less than 40% of the national GLB population, but over 50% of those living in urban areas.

During an initial phase we monitored GLB incidence experienced in each sampled city. In the second phase we made further cost-oriented adjustment. First, we discontinued sampling in cities that were not the core city of the MSA, except for Miami Beach, FL. Second, we oversampled the metropolitan areas with greater GLB incidence for the remainder of the study. Table 3 (in the section on weighting) describes the major geographic features of the GLB survey sample.

---

<sup>1</sup> The RDD telephone sample was provided by Survey Sampling, Inc. (SSI), who also maintain the minority group density information required by this design.

<sup>2</sup> An MSA is the traditional Metropolitan Statistical Area; a CMSA is a Consolidated Metropolitan Statistical Area, comprised of individual MSAs. Oakland was added on substantive grounds.

<sup>3</sup> Human Rights Campaign Survey, 1996, conducted by Lake, Snell, Perry and Associates.

## **DATA COLLECTION PROCEDURES**

### **Testing**

The questionnaire was developed in collaboration between staff of the Kaiser Family Foundation and PSRA. In order to improve the quality of the data, the questionnaire was pretested with a small number of respondents; the GP component was tested using RDD telephone numbers; respondents for testing the GLB component were selected from the Gay America Megafile maintained by Metamorphics Media. The pretest interviews were monitored by PSRA staff and conducted using experienced interviewers who could best judge the quality of the answers given and the degree to which respondents understood the questions. Some changes were made to screening procedures in the GLB component and question wording and order based on the monitored pretest interviews.

### **Contact Procedures**

Interviews for both survey components were conducted during the period February 7 through September 4, 2000. As many as 20 attempts were made to contact every sampled telephone number. Sample was released for interviewing in replicates, which are random subsamples of the larger sample. Using replicates to control the release of sample ensures that complete call procedures are followed for the entire sample. The use of replicates also ensures that the regional distribution of numbers called is appropriate. Again, this works to increase the representativeness of the final sample.

Calls were staggered over times of day and days of the week to maximize the chance of making contact with potential respondents. Each household received at least one daytime call in an attempt to find someone at home. In each contacted household, interviewers asked to speak with the youngest male currently at home. If no male was available, interviewers asked to speak with the oldest female at home. This systematic respondent selection technique has been shown to produce samples that closely mirror the population in terms of age and gender.

The initial section of the GLB interview identified households with gay, lesbian and bisexual adults. In order to help avoid inaccurate self-identification that would artificially increase or decrease GLB incidence, the question was embedded in a larger inventory of household member characteristics:

- S1. I'm going to start with a few questions about the adults your household...In answering, please keep in mind that this interview is completely confidential. (First,) is there any adult in your household age 18 or older ...**(INSERT—READ ITEMS IN ORDER)**  
**(FOR EACH “YES”, ASK: Is that one adult, two adults, or more than two adults?)**
- a. Who is now attending a college or university?
  - b. Who is now unemployed and looking for work?
  - c. Who is physically disabled?
  - d. Who is gay, lesbian or bisexual?

If no person in the household met condition [d] the interview was terminated. If more than one adult was gay, lesbian or bisexual, one of them was selected randomly to respond to the rest of the interview.

## **WEIGHTING AND ANALYSIS**

### **Weighting the GP Survey Component**

Weighting is generally used in survey analysis to adjust for the effects of the sample design and to compensate for patterns of nonresponse that might bias results. The weighting for this survey component was accomplished in two stages: a first stage sampling weight to adjust for the designed oversampling discussed above, and a second stage adjustment to account for demographic distortions due to non-response.

#### ***First Stage Sample Design Weight***

All completed interviews from the consolidated sample were given a first stage sample weight based on the level of disproportionality imposed by the sample design. This weight brings the distribution of total phone numbers dialed into alignment with the actual distribution of telephone numbers across the strata defined by minority density. It is proportional to the true proportion in each stratum divided by the sample proportion. Table 1 documents design parameters and survey returns across strata.

**Table 1: GP Sample Design and Weights; Survey Returns**

<i>Stratum</i>	<i>Estimated households</i>	<i>RDD sample</i>	<i>Design weight</i>	<i>Un-weighted completes</i>	<i>By race/Hispanic origin</i>		
					<i>white/other</i>	<i>black</i>	<i>Hispanic</i>
1	9.9%	52.1%	1.00	1134	329	401	404
2	17.9%	30.3%	3.11	695	429	134	132
3	32.0%	9.1%	18.42	201	173	20	8
4	25.7%	7.1%	18.95	171	154	11	6
5	14.5%	3.7%	20.47	82	77	3	2
Total	100.0%	100.0%		2283	1162	569	552

### ***Second Stage -- Demographic Adjustment***

In the second weighting stage, the demographic composition of each racial/ethnic subsample was weighted to match national parameters for sex, age, education, and region (U.S. Census four-level) for that group. These parameters came from a special analysis of the March 1999 Current Population Survey (CPS) that included all households in the continental United States that had a telephone. For the larger subsample comprising those neither Hispanic nor black (Whites/Others), complete interviews were weighted to the distributions of age by sex, education by sex, age by education, and region. For the smaller samples of African-Americans and Hispanics, complete interviews were weighted to marginal distributions of sex, age, education, and Census region. The correct incidence of African-American and Hispanic persons was also enforced in this stage.

This stage of the weighting process, which incorporated each respondent's first stage weight, was accomplished using Sample Balancing, a special iterative sample weighting program that simultaneously balances the distributions of all variables using a statistical technique called the *Deming Algorithm*. The second stage weight adjusts for non-response to the extent that non-response is related to particular demographic characteristics of the sample. This weight ensures that the demographic characteristics of each racial/ethnic group in the sample closely approximate the demographic characteristics of that group's national population.

**Table 2: Demographics – Total Sample**

	CPS Parameter	Unweighted Sample	Weighted Sample
Sex			
Male	47.8	44.3	47.2
Female	52.2	55.7	52.8
Age			
18-24	12.7	15.0	13.3
25-34	19.0	22.9	19.4
35-44	22.4	24.1	22.7
45-54	17.8	17.2	18.3
55-64	11.7	10.1	11.4
65+	16.5	10.7	14.8
Education			
Less than HS	16.4	14.7	15.1
HS graduate	36.7	29.7	35.3
Some college	23.3	26.9	24.2
College graduate	23.6	28.8	25.4
Race			
White	84.5	59.8	82.7
Black	11.2	27.9	12.6
Other race	4.3	12.3	4.7
Hispanic Origin			
Hispanic	9.7	24.2	10.1
Not Hispanic	90.3	75.8	89.9
Region			
Northeast	19.6	16.1	18.8
Midwest	23.8	13.3	23.0
South	35.3	42.03	35.6
West	21.3	28.6	22.6

### Weighting the GLB Survey Component

The GLB survey sample cannot be adjusted with the precision of the GP survey, because of the many unknowns about the GLB population nationally and the sampling plan's restricted geographic coverage. It is therefore only sensible to speak of generalizing this sample to urban gays, lesbians and bisexuals who are willing to volunteer their sexual orientation in a comparable context.

One design feature, however, does require an adjustment – the oversampling of metropolitan areas with higher GLB incidence. This correction is straightforward. Since telephone numbers from each city were sampled at one of three fixed proportions relative to other cities (based on GLB incidence in early returns), the correct weighting factor is the inverse

of those proportional factors. The geographical composition of the GLB sample and weighting parameters are reported in Table 3.

**Table 3: GLB Sample Design, Returns, and Weights**

<i>Stratum</i>		<i>Relative Sampling Rate</i>	<i>Completed Interviews</i>	<i>Unweighted Percentage</i>	<i>Weight Adjustment</i>	<i>Weighted Percentage</i>
High Incidence	New York, San Francisco, Oakland	4.47	237	58.5	1.00	31.0
Moderate Incidence	Atlanta, Boston, Chicago, Phoenix, Seattle	2.11	95	23.5	2.12	26.3
Low Incidence	Cleveland, Dallas, Detroit, Houston, Los Angeles, Miami, Philadelphia, Washington, DC	1.00	73	18.0	4.47	42.7
Total			405	100.0	1.89	100.0

### Effects of Sample Design on Statistical Inference

Specialized sampling designs and post-data collection statistical adjustments require analysis procedures that reflect those departures from simple random sampling. PSRA calculates the effects of these design features so that an appropriate adjustment can be incorporated into tests of statistical significance when using these data. The so-called "design effect" or "*deff*" represents the loss in statistical efficiency that results from disproportional sample design and non-response. PSRA calculates the composite design effect for a sample of size  $n$ , with each case having a weight,  $w_i$  (with mean of one) as:

$$deff = \frac{\sum_{i=1}^n w_i^2}{\sum_{i=1}^n w_i}$$

In a wide range of situations, the adjusted standard error of a statistic should be calculated by multiplying the usual formula by the square root of the design effect. Thus, the formula for computing the 95 percent confidence interval around a percentage  $p$  is:

$$\hat{p} \pm \left( \sqrt{deff} \times 1.96 \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} \right)$$

where  $\hat{p}$  is the estimate of the true population proportion,  $p$ , and  $n$  is the unweighted number of sample cases in the group being considered.

Table 4 shows the design effects and margins of error overall and for the major target subgroups of the GP sample. The design effect of 2.9 for the total sample indicates that the margin of error for national estimates based on all interviews is  $\sqrt{deff} = 1.72$  times as wide as a simple random sample of this size. The benefits are seen in the margins of error for African-American and Hispanic respondents, which would have been quite larger in a proportional sample, i.e., one that ignore density information across strata. It is interesting to note that the margin of error for respondents that are neither African-American nor Hispanic is larger than for the other subsamples, in spite of the larger number of interviews. This is due to the fact that the weighting has a greater impact on whites across strata, leading to greater *relative* inefficiency. The overall effect of our design has been to trade some precision in overall estimates for greater balance in precision among the subgroups, as desired.

**Table 4: Design Effects and Margins of Error, GP Survey**

	<i>Design Effect</i>	<i>Margin of Error</i> <sup>4</sup>
Total Sample	2.96	± 3.5%
White, not Hispanic	2.03	± 4.4%
Black, not Hispanic	3.09	± 7.2%
Hispanic	2.63	± 6.8%

The GLB sample does not allow such precise quantification of sampling error, again because of its ambiguous generalizability. Nevertheless it is important to think about the variability of the survey estimates. The analogous estimates for the GLB survey would be a

<sup>4</sup> The margin of error represents the size of the confidence interval around 50% and is the largest confidence interval for any estimated proportion.



design effect, due to incidence-based oversampling, of 1.47 and a margin of error of  $\pm 5.9\%$ .

## RESPONSE RATE

Table 5 reports the disposition of all sampled telephone numbers ever dialed from the GP and GLB survey samples. The response rate is intended to convey the fraction that were ultimately interviewed out of all eligible respondents who were included in the sample. It is calculated by taking the product of the contact<sup>5</sup>, cooperation and completion rates. PSRA's disposition codes and rate formulas are consistent with standards of the American Association for Public Opinion Research.

**Table 5: Sample Disposition Report**

	GP Survey		GLB Survey	
	Number	Rate	Number	Rate
<b>Numbers Dialed</b>	<b>9989</b>		48850	
<i>Non-working Numbers:</i>				
<i>Not In service</i>	2039		7366	
<i>Business/Non-residential</i>	1484		7876	
<i>Compter/Fax tone</i>	542		2677	
<i>Estimated additional non-residential*</i>	783		2819	
<b>Working Numbers</b>	<b>5141</b>	<b>51.5%</b>	<b>28113</b>	<b>57.5%</b>
<i>Numbers Never Contacted:</i>				
<i>Language problem</i>	123		990	
<i>Health problem, deaf, deceased</i>	98		267	
<i>All other no contact</i>	1214		12398	
<b>Total Contacted Numbers</b>	<b>3706</b>	<b>72.1%</b>	<b>14458</b>	<b>51.4%</b>
<i>Refusal</i>	1090		3041	
<b>Total Cooperating</b>	<b>2616</b>	<b>70.6%</b>	<b>11417</b>	<b>79.0%</b>
<i>Ineligible</i>	99		10985	
<b>Total Eligible</b>	<b>2517</b>	<b>96.2%</b>	<b>432</b>	<b>3.8%</b>
<i>Interrupted</i>	234		27	
<b>Total Complete</b>	<b>2283</b>	<b>90.7%</b>	<b>405</b>	<b>93.8%</b>
<b>Overall Response Rate</b>		<b>46.2%</b>		<b>38.1%</b>

<sup>5</sup> It is assumed that 90 percent of cases that result in a constant disposition of "No answer all attempts" over all 10 attempts are actually not working numbers.