Appendix A: Methodological Report

Boomers at Midlife:
The AARP Life Stage Study
Wave 2, 2003

Prepared by Princeton Survey Research Associates International for AARP
Summer 2003

Summary

The boomers@midlife survey, sponsored by AARP, conducted telephone interviews with a nationally representative sample of 3,500 adults living in the continental United States. Interviews were completed in both English and Spanish, according to the preference of the respondent. The interviews were conducted by Princeton Data Source, LLC (a subsidiary of Princeton Survey Research Associates International) located in Fredericksburg, Virginia during the period of April 3 through June 8, 2003. Quotas were set by age to ensure sufficient interviews among boomers[27] for statistical analysis. Results are weighted to correct for oversampling and to make the final total sample results representative of all adults living in the continental U.S. The margin of sampling error for the complete set of weighted data is ±3 percent.

Details on the design, execution and analysis of the survey are discussed below.

Design and Data Collection Procedures

Sample Design

The sample was designed to generalize to the U.S. adult population, and to allow separate analyses of boomers overall, and African-American and Hispanic boomers.

To oversample boomers, PSRAI screened standard list-assisted random digit dial (RDD) sample. To oversample minority groups, PSRAI supplemented the RDD sample with prescreened sample of African-American and Hispanic households.

RDD Sample

The RDD telephone sample was provided by Survey Sampling, Inc. (SSI) according to PSRAI specifications. To draw this sample, every active block of telephone numbers (area code + exchange + two-digit block number) that contained three or more residential directory listings is equally likely to be selected; after block selection, phone numbers are randomly generated in proportion to the number of listed households in each block. This method guarantees coverage of every assigned phone number regardless of whether that number is directory listed, purposely unlisted, or too new to be listed. After selection, the numbers are compared against business directories and matching numbers are purged. The RDD sample yielded 1,500 interviews.

Pre-screened Sample

To supplement the RDD interviews, an additional 2,000 interviews were completed from minority households identified in the PDS Demographic Tracking Omnibus Survey. This short 10 minute demographic survey asked a number of questions about individual demographics and household composition. Sample for the original Demographic Tracking survey was drawn using the same RDD sampling as described above.

[27] For this survey, boomers are defined as those born between 1946 and 1964.

Questionnaire Development and Testing

The questionnaire was developed by PSRA International in collaboration with AARP in 2002 and readministered in 2003 with slight modifications. The 2003 questionnaire was pretested with a small number of respondents from an RDD sample. Pretest interviews were monitored by PSRAI and AARP 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. The questionnaire was translated into Spanish by Princeton Data Source. All interviews, both English and Spanish, were conducted using a fully-programmed CATI instrument.

Contact Procedures

Interviews were conducted during the period April 3 through June 8, 2003. As many as ten attempts were made to contact a person at every sampled telephone number. Sample was released for interviewing in replicates, which are representative subsamples of the larger sample. Using replicates to control the release of sample ensures that complete call procedures are followed for the entire sample and that the geographic distribution of numbers called is appropriate.

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.

Weighting and Analysis

Weighting is generally used in survey analysis to adjust for sample design effects and to compensate for patterns of non-response that might bias results. The weighting for this project was accomplished in two stages; a first stage to adjust demographic distortions due to non-response and age screening, and a second stage to put the various racial/ethnic groups back into their proper proportions and to make the final weighted n equal to the final unweighted n for the total sample.

First Stage: Demographic Adjustment

In the first weighting stage, the demographic composition of each racial/ethnic subsample was weighted to match national parameters for sex by age, sex by education, age by education, marital status and census region. These parameters came from a special analysis of the March 2002 Current Population Survey (CPS) that included all telephone households in the continental United States.

This stage of weighting 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*. This weighting adjusts for non-response that is related to particular demographic characteristics of the sample and ensures that the demographic characteristics of each race/ethnic group's sample closely approximate the demographic characteristics of that group nationwide. Weights from this stage were *trimmed* to prevent individual interviews from having too much influence on the final results. Table 1 compares weighted sample distributions to population parameters.

TABLE 1: TOTAL SAMPLE DEMOGRAPHICS

Gender	Parameter	Weighted
Male	47.8	46.4
Female	52.2	53.6
Age		
18–38	39.1	37.8
39–49	23.0	23.9
50–57	13.3	13.8
58+	24.5	24.6
Education		
Less than HS	16.1	14.6
HS graduate	35.9	36.5
Some college	23.1	23.1
College graduate	24.9	25.8
Marital Status		
Married	57.5	58.5
Widowed/ Separated/ Divorced	18.6	17.6
Never Married	23.9	23.9
Race/Ethnicity		
White, not Hispanic	72.9	72.9
Black, not Hispanic	10.8	10.8
Hispanic	11.4	11.4
Other, not Hispanic	4.9	4.9
Region		
Northeast	19.9	19.4
Midwest	23.1	23.3
South	35.1	35.7
West	21.9	21.6

Second Stage: Final Adjustment to Racial/Ethnic Groups

The second and final stage of weighting adjusted the racial/ethnic distribution of the entire sample to match the distribution of the U.S. adult population. The final weight is the product of this adjustment factor and the weight after the first stage. Table 2 outlines the computation of the final weighting adjustment factor. Cases where the respondent refused to give their race were given an adjustment of 1.00. A final adjustment was made to make weighted and unweighted total sample counts equal. Tables 3 through 5 compare weighted and parameter distributions for major racial and ethnic subgroups.

TABLE 2: FINAL ADJUSTMENT FACTOR

Race/Ethnicity	% after First-Stage Weight	Population Parameter	Adjustment
White, not Hispanic	22.64	72.9	3.221
Black, not Hispanic	30.00	10.8	0.360
Hispanic	45.90	11.4	0.248
Other, not Hispanic	1.49	4.9	3.295

TABLE 3: AFRICAN AMERICAN SAMPLE DEMOGRAPHICS

Gender	Parameter	Weighted
Male	44.0	42.7
Female	56.0	57.3
Age		
18–38	44.6	42.5
39–49	24.0	24.9
50–57	12.6	13.1
58+	18.9	19.5
Education		
Less than HS	20.8	19.2
HS graduate	37.3	38.0
Some college	26.1	26.2
College graduate	15.8	16.6
Marital status		
Married	38.5	39.6
Widowed/ Separated/ Divorced	23.2	25.9
Never Married	38.3	34.5
Race/Ethnicity		
White, not Hispanic	NA	NA
Black, not Hispanic	NA	NA
Hispanic	NA	NA
Other, not Hispanic	NA	NA
Region		
Northeast	17.9	18.5
Midwest	18.3	18.6
South	55.2	53.9
West	8.6	9.0

TABLE 4: HISPANIC SAMPLE DEMOGRAPHICS

Gender	Parameter	Weighted
Male	50.4	46.9
Female	49.6	53.1
Age		
18–38	57.7	53.2
39–49	20.3	22.4
50–57	9.0	9.9
58+	13.1	14.5
Education		
Less than HS	41.5	36.3
HS graduate	31.1	33.3
Some college	17.5	19.4
College graduate	9.9	11.0
Marital status		
Married	55.0	56.9
Widowed/ Separated/ Divorced	14.0	14.8
Never Married	31.0	28.3
Race/Ethnicity		
White, not Hispanic	NA	NA
Black, not Hispanic	NA	NA
Hispanic	NA	NA
Other, not Hispanic	NA	NA
Region		
Northeast	14.3	13.6
Midwest	7.6	9.4
South	34.2	35.8
West	43.9	41.2

TABLE 5: WHITE/OTHER RACE SAMPLE DEMOGRAPHICS

Gender	Parameter	Weighted
Male	48.0	46.7
Female	52.0	53.3
Age		
18-38	35.7	34.8
39-49	23.3	24.0
50-57	14.1	14.4
58+	27.0	26.8
Education		
Less than HS	11.7	10.7
HS graduate	36.4	36.8
Some college	23.5	23.2
College graduate	28.4	29.2
		0.5
Marital status		
Married	60.6	61.3
Widowed/ Separated/ Divorced	18.6	16.8
Never Married	20.8	21.8
		0.8
Race/Ethnicity		
White, not Hispanic	NA	NA
Black, not Hispanic	NA	NA
Hispanic	NA	NA
Other, not Hispanic	NA	NA
Region		
Northeast	21.0	20.4
Midwest	26.1	26.1
South	32.4	33.1
West	20.5	20.4

Effects of Sample Design on Statistical Inference

Specialized sampling designs and post-data collection statistical adjustments require analysis procedures that reflect departures from simple random sampling. PSRAI 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 a disproportional sample design and systematic non-response. PSRAI calculates the composite design effect for a sample of size n, with each case having a weight, w_i as:

$$deff = \frac{n\sum_{i=1}^{n} w_i^{\,2}}{\left(\sum_{i=1}^{n} w_i\right)^2}$$

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 ($\sqrt{deff}$). Thus, the formula for computing the 95% confidence interval around a percentage is:

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

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

The formula for computing the 95 percent confidence interval around the difference between two percentages, p_1 and p_2, of sizes n_1 and n_2, is:

$$(\hat{p}_1 - \hat{p}_2) \pm 1.96 \sqrt{\frac{deff_1\, \hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{deff_2\, \hat{p}_2(1-\hat{p}_2)}{n_2}}$$

where $\hat{p}_1$ is the estimate of p_1, $\hat{p}_2$ is the estimate of p_2, and $deff_1$ and $deff_2$ are the design effects for each group.

The survey's *margin of error* is the largest 95% confidence interval for any estimated

proportion based on the total sample-one around 50%. For example, the margin of error for the total sample is ±2.7%, this means that in 95 out every 100 samples using the same methodology, estimated proportions based on the entire sample will be no more than 2.7 percentage points away from their true values in the population. It is important to remember that sampling fluctuations are only one possible source of error in survey estimates. Other sources, such as measurement, may contribute additional error of greater or lesser magnitude. Table 6 shows design effects, sample sizes and margins of error for the total sample and selected target subgroups.

TABLE 6: SAMPLE SIZES, DESIGN EFFECTS AND MARGINS OF ERROR

	Number of Interviews	Design Effect	Margin of Error
Total	3500	2.69	2.7%
Age			
18–38	736	2.10	5.2%
39–49	1075	2.37	4.6%
50–57	941	2.11	4.6%
58+	748	2.41	5.6%
Gender			
Male	1281	2.59	4.4%
Female	2219	2.58	3.3%
Education			
LT HS graduate	492	2.89	7.5%
HS graduate	1246	2.53	4.4%
Some college	849	3.18	6.0%
College graduate	900	2.40	5.1%
Marital Status			
Married	1850	2.30	3.5%
Widowed/ Separated/Divorced	909	3.38	6.0%
Never married	721	2.97	6.3%
Race/Ethnicity			
White, not Hispanic	1113	1.42	3.5%
Black, not Hispanic	1190	1.44	3.4%
Hispanic	1132	1.81	3.9%
Region			
Northeast	537	2.42	6.6%
Midwest	642	2.27	5.8%
South	1534	2.61	4.0%
West	787	3.22	6.3%

Response Rate

Table 7 reports the disposition of all sampled telephone numbers from the RDD sample. The *response rate* estimates the fraction of all eligible respondents in the sample that were ultimately interviewed. At PSRA International it is calculated by taking the product of three component rates:[28]

- Contact rate—the proportion of working numbers where a request for interview was made—of 79.9 percent[29]
- Cooperation rate—the proportion of contacted numbers where a consent for interview was initially obtained, versus those refused—of 49.3 percent
- Completion rate—the proportion of initially cooperating and eligible interviews that were completed—of 86.5 percent

Thus the response rate for the RDD portion of this survey is 34.0 percent.

Table 8 reports the disposition of all sampled telephone numbers from the pre-screened African-American and Hispanic samples. These samples yielded post-screening response rates of 44 percent for the African-American oversample and 50 percent for the Hispanic oversample. Figures in these tables represent rates after the pre-screening of eligible households. The response rate for the original screening interview was approximately 33 percent.

TABLE 7: RDD SAMPLE DISPOSITION

Total numbers dialed	12635	
Business	1088	
Computer/Fax	800	
Other not-working	2034	
Additional projected NW	792	
Working numbers	7921	62.7%
No Answer	197	
Busy	67	
Answering machine	799	
Callbacks	253	
Other non-contacts	276	
Contacted numbers	6329	79.9%
Initial refusals	1129	
Second refusals	2081	
Cooperating numbers	3119	49.3%
No adult in HH	57	
SO–Age quota	1249	
Language barrier	78	
Eligible numbers	1735	55.6%
Interrupted	235	
Completes	1500	86.5%
Response rate		34.0%

[28] PSRAI's disposition codes and rate formulas are consistent with standards of the American Association for Public Opinion Research.

[29] We assume that 75 percent of cases that result in a constant disposition of "No answer" over 10 or more attempts are actually not working numbers.

TABLE 8: PRE-SCREENED SAMPLE DISPOSITION

	African American		Hispanic	
Total numbers dialed	5312		8390	
Business	83		211	
Computer/Fax	85		146	
Other Not-working	1224		1742	
Additional projected NW	126		61	
Working numbers	3794	71.4%	6230	74.3%
No answer	32		12	
Busy	10		9	
Answering machine	337		171	
Callbacks	214		71	
Other Non-contacts	145		476	
Contacted numbers	3056	80.5%	5491	88.1%
Initial refusals	602		442	
Second refusals	495		1412	
Cooperating numbers	1959	64.1%	3637	66.2%
No adult in HH	47		60	
SO–Age quota	575		1830	
SO–Race unqualified	120		488	
Language barrier	29		94	
Eligible numbers	1188	60.6%	1165	32.0%
Interrupted	187		166	
Completes	1001	84.3%	999	85.8%
Response rate		43.5%		50.1%

Screener

Hello, my name is _____ and I'm calling for Princeton Survey Research. We're conducting an important national opinion survey and we'd very much like to include your household. This survey is for research purposes only, we're not trying to sell you anything.

S1 Here's my first question . . . How would you rate your neighborhood as a place to live? Would you say it is excellent, good, fair or poor?

1 Excellent
2 Good
3 Fair
4 Poor
8 Don't know
9 Refused

S2 To make sure our survey includes many different kinds of families, I need to ask a few questions about who lives in your household. How many adults age 18 or older live in your household?

1 One
2 Two
3 Three
4 Four or more
5 None
98 Don't know
99 Refused

If only one adult age 18+ in HH (S2=1)

S3. May I please speak to that person?

1 Continue with current respondent—Go to Q1
2 New respondent being brought to phone—Go to INTRO2
3 New respondent not available—Schedule call back
9 Refused

If only one adult age 18+ in HH (S2=2, 3, 4)

S4. How many people between 18 and 38 now live in your household?

1 One
2 Two or more
3 None
9 Don't know/Refused

If only one adult age 18+ in HH (S2=2, 3, 4)

S5. How many people between 39 and 57 now live in your household?

1 One
2 Two or more
3 None
9 Don't know /Refused

If only one adult age 18+ in HH (S2=2, 3, 4)

S6. How many people 58 or older now live in your household?

1 One
2 Two or more
3 None
9 Don't know /Refused

If answered refused to S4, S5, S6 (S4=9 AND S5=9 AND S6=9)

CONF. Just to confirm, are you over the age of 18?

1 Yes—Go to Q1
2 No—Code as a refusal

If answered none to S4, S5, S6 (S4=3 AND S5=3 AND S6=3)

S7. May I please speak to any person over the age of 18?

1 Continue with current respondent—Go to Q1
2 New respondent being brought to phone—Go to Intro2
3 New respondent not available—Schedule call back
9 Refused—Thank and terminate

Instructions for prioritizing respondent to continue with interview

1 Has at least one 39–58 yr old (S5=1, 2) [Skip to S8] First Priority
2 Has at least one 58+ yr old (S6=1, 2) [Skip to S9] Second Priority
3 Has at least one 18–38 yr old (S4=1, 2) [Skip to S10] Third Priority

If one or more adults between ages 39 and 57 (S5=1, 2)

S8. May I please speak to the person between 39 and 57 years old?

1 Continue with current respondent—Go to Q1
2 New respondent being brought to phone—Go to Intro2
3 New respondent not available—Schedule call back
9 Refused—Thank and terminate

If one or more adults 58+ and no 39 to 57 year olds (S6=1, 2 AND S5=3)

S9. May I please speak to the person 58 or older?

1 Continue with current respondent—Go to Q1
2 New respondent being brought to phone— Go to Intro2
3 New respondent not available—Schedule call back
9 Refused—Thank and terminate

If one or more adults between ages 18 and 38 and no 39 to 57 year olds and no 58+ (S4=1, 2 AND S5=3 AND S6=3)

S10. May I please speak to the person between 18 and 38 years old?

1 Continue with current respondent—Go to Q1
2 New respondent being brought to phone—Go to Intro2
3 New respondent not available—Schedule call back
9 Refused—Thank and terminate

IF S3=2 or S7=2 or S8=2 or S9=2 or S10=2

INTRO2

Hello, my name is _____ and I'm calling for Princeton Survey Research. We're conducting an important national opinion survey and we'd very much like to include your household. This survey is for research purposes only, we're not trying to sell you anything.