

Sample Design

Introduction

The Princeton Affect and Time Survey, based on the American Time Use Survey (ATUS) and Day Reconstruction Method, employed a questionnaire designed to imitate key aspects of the ongoing ATUS survey, but with two primary differences. First, and most important, it incorporated a module that probed respondents' feelings about the activities reported in the activity diary section of the ATUS questionnaire. Second, because the persons interviewed for the pilot had not participated in the Bureau of Labor Statistics' Current Population Survey (as regular ATUS respondents have), it was necessary to collect demographic data which, in the ATUS survey, had been captured in the course of prior CPS interviewing.

The ATUS Feelings Module Pilot questionnaire was administered to two samples of respondents: (1) A Random Digit National Sample of persons 16 years of age and older living in the continental United States, and (2) an over-sample of retired adults recruited from Gallup's national Panel Sample. The questionnaire and administration was identical for these sub-samples, but because of their different sources and the overlap of their target populations they will be described separately in the discussions of sampling and weighting.

The questionnaire was programmed for CATI in the BLAISE software used by the Bureau of Labor Statistics for data collection on the ongoing ATUS program, in order to provide for as much consistency as possible between the data collection undertaken by Gallup for the Pilot Survey and the regular ATUS conducted by the Bureau of Labor Statistics. Gallup Interviewers used Computers For Marketing Corporation's call-scheduling/sample management software, because of its familiarity to Gallup interviewers when faced with a wide variety of household contact situations.

Random Digit National Sample: Telephone Sampling and Within-household Selection

The Random Digit Samples used to select respondents for the ATUS Pilot survey was purchased from Survey Sampling, using their "Random A" method, an EPSEM (Equal Probability Selection Method) procedure which ensures that every possible telephone number – including business and protected numbers – in any working block with at least one directory-listed telephone number has an equal probability of selection. These numbers were screened against a national business database to eliminate business numbers from the sample to be used for residential calling.

National Retired Over-sample Drawn from Gallup Panel

The Gallup Panel uses random digit dial (RDD) methodology to recruit its members—the same type of sample as was used for the RDD segment of the ATUS Pilot, described above. No incentives are given to panel members, who reside in all 50 states. Those who elect to join the panel commit to the completion of two to three surveys per month, with the typical survey lasting 10 to 15 minutes.

In order to obtain an over-sample of retired Americans, we selected and screened a random sample of Gallup Panel members who answered that they were retired in their original recruitment interview. During this recruitment interview, Gallup had collected demographic background information on these respondents, including a variety of variables that were appended to the data file forwarded to the ATUS Pilot's principal investigators, as well as information used in weighting the final dataset

Calling Design & Data Collection

The sampling design involved a random assignment of each sample record to a day of the week, in order to ensure that the survey data would be as representative as possible of time use patterns across all 7 days of the week. Respondents were contacted on the day of the week to which they were assigned, and asked about their experiences on the previous day. Once the time diary section of the interview was begun, respondents could request a later call to complete the remainder of survey (including the feelings module), but the balance of the survey was required to be completed within one day of the beginning of the time diary section.. Toward the end of the field period, to improve the response rate and the speed of completion of the fieldwork, the restriction of interviewing to pre-assigned days was relaxed for the remaining sample records for which the diary section was not begun (including records for which no contact had been made, or no respondent selected within the household.

The entire sample of telephone numbers was divided into random sub-samples (replicates). Replicates were released sequentially as interviewers complete the calling design on previously released replicates, a device that preserves the randomness of the sample while permitting flexible completed interview targeting and the maximization of response rates.

Call scheduling for the RDD sample involved an initial daytime call placed to each number, to allow the early identification and exclusion of numbers associated with businesses (which are incompletely identified by the business database matching described above). An appointment set with a respondent would pre-empt the system's automatic call scheduling, but if an interviewer did not set a specific call-back time, four additional calls would be placed, two of them between 5:00 pm and 7:00 pm, and two of them between 7:00 pm and 9:00 pm on weekdays or between 12:00pm and 9:00pm on Saturday or Sunday. These procedures are designed to improve the likelihood of an interviewer making a household contact. In order to improve the response rate, toward the end of the field period the 5-call limitation was lifted from numbers that had not been resolved, extending the calling design to at least 6 calls. By the end of data collection,

21% of the interviews had been completed on a sixth call or higher, and more than 60 interviews were completed only after 10 or more calls had been placed.

Interviews were conducted in English and Spanish, according to the selected respondents' preference

Once contact was made within a sampled household, Gallup interviewers used the “most recent birthday” selection technique to designate a selected respondent among all household members 15 years of age or older.

Selection of Behavioral Episodes from the Time Use Diary for Feelings Module Pilot

The primary purpose of the ATUS Pilot survey was to test a set of questions asking about how respondents felt during activities that they reported in the time use diary. Because the time use diaries contained an entire day's worth of activities—too many to ask even a short battery of questions about all of them—a set of three activities were selected.

The selection algorithm eliminated two types of activities from consideration: sleeping and grooming. The remainder of the day was divided into 15 minute intervals, and this time was selected with equal probability. The result of this procedure was that each activity had a probability of selection proportional to the length of time that the respondent reported having engaged in it. Within the period of time the respondent reported engaging in that activity a specific quarter-hour period was selected, and the respondent was asked to rate his/her feelings during that specific 15 minute period.

For the second selection, the entire period of time spent on the activity selected first was excluded from selection—not just the fifteen minute period specified for purposes of the feelings battery, but the entire block of time spent on that activity—and another selection was made from the remaining activities, again with probability proportional to time that the respondent reported having spent on each of the remaining qualified activities. As with the first selection, a specific 15-minute period was selected, and the feelings module was focused on that period.

Finally, a third activity was selected, in the same way—without replacement and with probability proportional to time spent on the remaining activities—and the feelings battery was asked for a third time, focusing on a specific 15-minute period randomly selected within the period of time spent on that activity according to the respondent's reporting in the time diary.

Calculation of Response rates

Gallup used their own call scheduling software and their own disposition codes, an elaborate set of 48 distinct categories that interviewers have been trained to distinguish. Gallup has mapped these dispositions to the dispositions used in the calculation of AAPOR response rates, and based on the AAPOR categorization, Gallup is able to

calculate response rates that correspond to the Standard Definitions document available on the AAPOR website.

A detailed description of the AAPOR disposition categories can be obtained on the AAPOR web site [http://www.aapor.org/pdfs/standarddefs_4.pdf]. The AAPOR Component Rates used in the calculation of response rates and associated rates are as follows (AAPOR numerical designations are in parentheses):

- I** = Complete interview (1.1)
- P** = Partial interview (1.2)
- R** = Refusal and break-off (2.10)
- NC** = Non-contact (2.20)
- O** = Other (2.30)
- UH** = Unknown if household/occupied HU (3.10)
- UO** = Unknown, other (3.20)
- e** = Estimated proportion of cases of unknown eligibility that are eligible

The table presented directly below lists the detailed Gallup disposition codes in detail and shows how they are mapped to the AAPOR code categories for purposes of calculating AAPOR Rates:

Gallup Disposition Code	AAPOR Disposition Code	AAPOR Response Rate Component Category Mapping	Description
001	1.10	I	Completed interview
003	2.33	O	Language barrier screening complete
003	3.20	UO	Language barrier screening incomplete
004	4.50	NE	800 phone number (consumer non-target)
005	4.32	NE	Bad/Disconnected phone number
006	4.50	NE	Non-target (incorrect respondent type (bus./res.) or wrong number)
007	4.20	NE	Fax or data line
008	2.32	O	Hearing barrier screening complete
008	3.20	UO	Hearing barrier screening incomplete
009	2.32	O	Illness (respondent is ill) screening complete
009	3.20	UO	Illness (respondent is ill) screening incomplete
010	2.31	O	Death of respondent screening complete
010	3.20	UO	Death of respondent screening incomplete
011	2.32	O	Death in family, inappropriate to continue, screening complete
011	3.20	UO	Death in family, inappropriate to continue, screening incomplete
012	2.25	NC	Respondent is unavailable during study field period, screening complete
012	3.20	UO	Respondent is unavailable during study field period, screening incomplete
014	2.32	O	Respondent is incapable of completing a legitimate survey, screening complete
014	3.20	UO	Respondent is incapable of completing a legitimate survey, screening incomplete
015	2.11	R	Corporate refusal screening complete
015	3.20	UO	Corporate refusal screening incomplete

016	2.11	R	Second refusal screening complete
016	3.20	UO	Second refusal screening incomplete
016	3.20	UO	Second refusal at the introduction
017	3.20	UO	Screener DK/RF
019	4.80	NE	Respondent determined quota fill
020	4.70	NE	Screen failure
021	2.35	O	Grid person selected but questionnaire change prevents continuation, screening complete
021	3.20	UO	Grid person selected but questionnaire change prevents continuation, screening incomplete
023	2.30	O	Corporate referral screening complete
023	3.90	UO	Corporate referral screening incomplete
024	3.15	UO	Answering Service - respondent contact blocked
026	2.11	R	Respondent irritated - never call again, screening complete
026	3.20	UO	Respondent irritated - never call again, screening incomplete
027	2.11	R	Respondent refused due to cell phone contact screening complete
027	3.20	UO	Respondent refused due to cell phone contact screening incomplete
101	3.13	UH	No Answer
102	3.17	UH	Busy
103	3.17	UH	Busy
104	2.41	R	Specifically timed callback screening complete
104	3.20	UO	Specifically timed callback screening incomplete
105	2.42	R	Non-specific timed callback screening complete
105	3.20	UO	Non-specific timed callback screening incomplete
107	3.20	UO	Answering Machine
162	3.20	UO	Soft-refusal at Introduction
172	2.11	R	Soft-refusal beyond introduction - suspended, screening complete
172	3.20	UO	Soft-refusal beyond introduction - suspended, screening incomplete
182	3.17	UH	Busy
188	4.80	NE	Respondent type quota fill
191-197	2.33	O	Language barrier screening complete
191-197	3.20	UO	Language barrier screening incomplete
198	2.11	R	Hard refusal screening complete
198	3.20	UO	Hard refusal screening incomplete
198	3.20	UO	Hard refusal at the introduction

AAPOR Rate definitions and results for the RDD and Panel Samples

The calculation of a variety of AAPOR rates, including the overall AAPOR Response Rate (RR3), are defined below. To the right of each operational definition are the calculated values for the RDD and Panel Samples; all rates are expressed as proportions. Note that the panel sample rates are calculated and reported for comparative purposes but, as they are calculated identically to those for the RDD segment of the Pilot, do not take account of the fact that the Gallup Panel respondents represent a “second-generation RDD sample”—that they were recruited and agreed to become a part of the Gallup Panel prior to their inclusion in the ATUS retired over-sample:

	RDD	Panel
AAPOR eligibility rate (e): Proportion of Unknown Eligibility cases estimated to be eligible $(O + NC + R + I) / (O + NC + R + I + NE)$.207	.924

AAPOR contact rate: $(I + R + O) / (I + R + O + NC + e(UH + UO))$.497	.649
AAPOR refusal rate: $R / (I + R + NC + O + e(UH + UO))$.124	.058
AAPOR cooperation rate: $I / (I + R + O)$.751	.911
AAPOR response rate (RR3): $I / (I + R + NC + O + e(UH + UO))$.373	.592

Gallup delivered 3,982 cases from the RDD Study and 1999 cases from the Gallup Panel Over-sample study. Although more interviews than these were actually completed, the link between the BLAISE software used for data collection (and housing the survey data) and the CFMC software used for telephone sample management and call scheduling (and housing the sample records) was not possible to establish with certainty in all cases. Cases suffering from such imperfect linkage were not delivered.

Weighting adjustments

Weighting factors were calculated for both the Random and Gallup Panel Retired Over-samples. Two types of factors were calculated—the first to correct for probability of selection biases inherent in two-stage samples where households are first selected by generating random telephone numbers (and where households with multiple voice lines have a higher selection probability than others)—and the second to correct for probability of selection biases that favor the selection of individuals in households with smaller numbers of qualified persons over individuals in larger households. These separate weighting factors are combined in a “first stage weight” that corrects for both types of selection bias.

In addition to correcting for known sources of bias, additional post stratification weighting adjustments were calculated, and used in conjunction with the first stage weights to correct for demographic discrepancies that could have either random or systematic sources. Demographic targets were derived from the most recent Current Population Survey (CPS) for both the Retired and non-Retired populations, and the completed interview samples were divided into 3 segments: RDD Non-Retired, RDD Retired, and Panel Retired. Before addressing demographic variables, each of these samples was adjusted using second-stage weights to equalize the representation of day of the week.

Using the previously described weights as starting weights, additional ratio weighting factors were applied to adjust the following demographic variables, in order that the final weighted distributions of the data conformed to the relevant distributions derived from the CPS:

Region (4 Census Regions)

Gender

Age (Non-Retired: 15-19; 20-34; 35-44; 45-54; 55-99. Retired: 40-59; 60-69; 70-79; 80+)

Race (White versus Non-White)

Finally, weights were calculated to adjust the relative proportions of the three sample groups in order to allow them to be analyzed and projected to the total target populations either separately or combined.

Overall, the weighting procedures are designed to maximize the projectability of the data by correcting for known biases and controlling other sources of variation that could distort the representativeness of analytic results based on the weighted data.