Exploring Conversational Interviewing in the American Time Use Survey

Behavior Coding Study Report

10.08.12

Stephanie Denton, ATUS

Jennifer Edgar, Scott Fricker, Polly Phipps, OSMR

Executive Summary

In the American Time Use Survey (ATUS), interviewers use a set of scripted open-ended questions to walk respondents chronologically through the prior 24-hour day, collecting activities and details about each activity reported. Interviewers are trained to engage the respondent in a conversational and flexible way to gather necessary information to complete the time diary. These conversational interviewing techniques, such as using a variety of strategies to assist with recall, anchoring questions, and using active listening are thought to put the respondent at ease and provide interviewers with the freedom to collect data in the best possible way.

While the 24-hour recall diary is a standard way to collect time use data, recalling activities that occurred on a previous day is a challenging task. Conversational interviewing is hypothesized to improve respondent understanding of questions and concepts as interviewer and respondent collaborate and converse on meaning. In the ATUS, conversational interviewing is also thought to improve recall by allowing interviewers to ask open-ended questions to assist respondents in reconstructing their day in an order and way that is meaningful to them rather than following a set script and sequence.

This paper summarizes the results of an analysis of 104 behavior-coded transcripts of ATUS interviews. For this initial analysis, we utilize basic descriptive statistics and univariate comparisons. Four research questions about the use of conversational interviewing in the ATUS recall diary are addressed. The conclusions that follow should be viewed as preliminary and tentative. Additional analyses that take advantage of the relationships between variables and the interactions between interviewers and respondents are planned and are expected to provide additional insight into the impact of conversational interviewing in the ATUS survey.

> Q1. Are interviewers using conversational interviewing, scripted interviewing, or some hybrid combination?
>
> a. Although ATUS is designed to be conversational, there is not much evidence that it actually is. The majority of the interview consists of activity and time retrieval questions, and a majority of those questions are asked exactly as scripted.

Q2. What conversational interviewing techniques are being used?

a. Although a majority of retrieval questions are scripted, unscripted anchors were often added to tailor the question to the respondent. For example, instead of saying "What did you do next?" an interviewer anchors the question using the prior activity to ask "What did you do after you finished eating?" Interviewers also used confirmation and feedback to verify respondents' answers or to provide clarification about the survey task, as well as non-scripted probes.

Q3. Is there variation in the use of conversational interviewing by interviewer experience?

a. Conversational interviewing techniques did not vary much by interviewer experience or coach status.

Q4. Is there evidence that conversational interviewing is associated with the quality and quantity of activities reported by the respondent?

a. There is limited evidence that conversational interviewing techniques are associated with data quality, although this analysis is a preliminary step and additional data quality measures need to be explored.

Background

ATUS

The American Time Use Survey (ATUS) is an ongoing survey, collected almost every day of the year using computer-assisted telephone interviewing (CATI). A time diary is the core of the ATUS interview. During the time diary, interviewers use a 24-hour recall diary, which is a set of scripted open-ended questions in tandem with conversational interviewing techniques, to walk respondents chronologically through the day, collecting information about time spent in activities beginning at 4:00 am on the previous day up until 4:00 am of the interview day. For each activity reported, the interviewer asks how long the activity lasted. For most activities, the interviewer also asks who was in the room or accompanied the respondent during the activity and where the activity took place. ATUS respondents are interviewed only one time about how they spent their time on the previous day. Demographic information—including sex, race, age, educational attainment, occupation, income, marital status, and the presence of children in the household— is available for each respondent.[1]

Conversational Interviewing

The ATUS was the first Census Bureau survey to use conversational interviewing to collect data and is one of the most conversational interviews in the federal government statistical system. Interviewers are trained to use a set of scripted, open-ended questions in combination with conversational interviewing techniques to walk respondents through the 24-hour reference period. By training interviewers to use conversational interviewing techniques, such as using a variety of strategies to assist with recall, anchoring questions, and using active listening, the goal is to put the respondent at ease and collect high quality data.

While the 24-hour recall diary is a common way to collect time use data, recalling activities on a previous day is a challenging task. Conversational interviewing is hypothesized to improve respondent understanding of questions and concepts as interviewers and respondents collaborate and converse on meaning. In the ATUS, conversational interviewing is also thought to improve recall by allowing interviewers to ask questions in multiple ways to ask about activities or time, and to use information the respondent has already provided to anchor future questions. Additionally, this type of interviewing may allow respondents to reconstruct their day in an order and way that is meaningful to them rather than following a set script and sequence, also thought to improve recall and data quality.

---

[1] Some of the demographic data are collected as part of the ATUS interview, while other items are collected during the Current Population Survey interview (from which the ATUS sample is drawn). Some of the CPS demographic items are verified and updated by ATUS interviewers.

Conversational Interviewing in the ATUS

Standardized wording and question reading is commonplace in surveys. The rationale for asking respondents identical questions and probes includes reducing response error that might occur with wording changes or context, and controlling survey length, ultimately minimizing costs (Belli et al., 2001). Alternatively, conversational interviewing, with origins in communication theories, suggests that shared understanding is more likely to be achieved if individuals use unscripted conversation to collaborate on meaning (Conrad and Schober, 2000). In contrast to standardized survey questions, Conrad and Schober point out that allowing interviewers to assist respondents by providing answers and deviating from scripts in ways that do not overly influence response may help assure that respondents understand the questions as intended by survey designers.

The ATUS diary appears to be an instrument very well suited to the use of conversational interviewing. ATUS diary respondents are asked to participate in an ambiguous task – no level of detail about the day is specified; interviewers simply state that individuals will be asked to provide the activities of their previous day, along with the times, who they were with, and where they were beginning and ending at 4 AM. To carry out this task, general scripted questions are provided for interviewers (what did you do next? how long did that take?). The ambiguous nature of the task may require interviewers and respondents to collaborate on meaning. Some specific uses of conversational interviewing in the ATUS include: probing for the level of detail required by the instrument, confirming activities or activity lengths, redirecting respondents who get off task and provide too much or too little information, and facilitating respondent recall by moving backward or forward in the day or encouraging respondents to visualize their activities.

Related to time diaries, event history calendars require respondent recall of past events and times, and these survey instruments commonly use conversational interviewing. Belli and colleagues (2004) describe data quality advantages of flexible interviewing with event history calendars as twofold: (1) the method encourages respondent use of retrieval cues available that are associated with autobiographical memory structure; and (2) similar to Conrad and Schober (2000), the method may resolve uncertainties in verbal exchanges between a respondent and interviewer in searching for a shared meaning.

Behavior Coding

Behavior coding is a technique to systematically code interviewer and respondent behavior and interactions during the survey interview. It is used to assess how interviewers ask questions and provide clarification and feedback to respondents, how respondents interpret questions and communicate their answers, and how interviewers and respondents interact during the survey task. Behavior coding has been used to evaluate data accuracy and quality, identify difficult survey questions, tasks, and concepts, and detect interviewing practices that are associated with data quality.

Use of Behavior Coding to Measure Conversational Interviewing

In a major study comparing interviewing methods of the event history calendar using conversational interviewing and standardized questions, Belli and colleagues (2001) utilized behavior coding to investigate respondent-interviewer interactions and evaluate data quality. Their behavior coding scheme captured interviewer and respondent behaviors that are typically associated with standardized interviewing rules (e.g., whether interviewers are reading questions as written, probing in a non-directive way, and using appropriate feedback), as well as indicators of respondent difficulties (e.g., responses that do not sufficiently answer the question, qualified or estimated responses, requests for clarification, etc.). In addition, the authors developed a behavior coding scheme that integrated retrieval cues and other conversational processes that may be associated with flexible interviewing.

Given that ATUS interviewers are provided scripted questions as well as encouraged to use conversational interviewing, we developed a behavior coding scheme that accommodates standardized rules, possible retrieval cues, and conversational processes. Our goals are complementary to those of Belli and colleagues: to use the behavior coding method to understand interviewer and respondent interaction, to ascertain what respondent behaviors and interviewer techniques are exhibited in the time diary, and to gain insight into data quality.

Research Questions

In this paper, we utilize the behavior coding method to determine if ATUS interviewers are using techniques associated with conversational interviewing, and if so, which techniques. We also explore variation in interviewing technique by interviewer experience, and investigate whether conversational interviewing is associated with the quality and quantity of the activities reported by the respondent.

Method

Recordings

For this study, 108 ATUS interviews were recorded over the course of three weeks in August and September of 2008. Only the time-diary portion of the ATUS interview was recorded. Interviewers asked for consent prior to recording. A taping station was set up at the CATI center, and the 36 interviewers on the ATUS staff at the time were instructed to record three consecutive interviews at the station during their shift and, if necessary, to finish the task on their next assigned shift (see Attachment 1). The dates and times of the taped interviews indicated that interviewers followed procedures.

Comparison of Study and ATUS Samples

Table 1 shows the demographics of the study sample compared to the full 2008 ATUS sample.
No significant differences were found between the study sample and the 2008 sample (alpha
level = .05) for sex, age, race, education level, employment status, and marital status.. Thus, the
study sample can be viewed as demographically representative of the full ATUS sample.

Table 1. Sample Comparisons

	2008 ATUS Sample %	**Study Sample*** %
Sex		
Male	44.3	43.3
Female	55.7	56.7
Age		
15-24	11.3	13.5
25-34	16.6	13.5
35-44	21.4	15.4
45-54	18.5	21.2
55-64	14.9	15.4
65 and older	17.3	21.2
Race		
White only	80.8	81.7
Black only	13.9	12.5
Alaskan Native only	0.8	1.0
Asian only	3.0	1.9
Hawaiian/Pacific Islander only	0.2	0.0
More than one race	1.3	2.9
Education Level		
LT High School	16.5	15.5
High School graduate, no college	26.1	18.5
Some college or associates degree	26.5	28.2
Bachelor's degree	19.7	27.2
Post-graduate degree	11.2	10.7
Employment Status		
Not in Labor Force	34.0	38.8
Employed Full-Time	52.2	47.6
Employed Part-Time	11.0	9.7
Unemployed	2.8	3.9
Marital Status		
Married	50.0	46.1
Widowed	9.4	15.4
Divorced	13.6	10.6
Separated	3.1	1.0
Never married	24.0	26.9

* In some cases, demographic information was missing; we have not included the missing cases in the denominator.

<u>Transcription</u>
Due to the complexity of the diary interview, the research team decided to code from transcripts rather than the audio recordings. Transcription coding facilitates a more thorough analysis of question and answer sequences, more systematic checks of coder performance, and allows for the use of a more complex coding scheme (Ongena and Dijkstra, 2006). The interview recordings were transcribed verbatim, with the transcriber noting all utterances made by the interviewer or respondent. This included unfinished questions or statements, false sentence starts, repeated phrases, words, and conversations that seems unrelated to the diary interview. Language or terms were transcribed unchanged. The only exceptions were exclusions of background noise (such as typing, music, children crying). In addition, to protect respondents' confidentiality, Personally Identifiable Information was excluded in the final typed transcription, replaced with a generic description.

The transcriptions were done by four Census interviewers with previous transcribing experience. The selected interviewers had interviewing experience, but did not have ATUS interviewing experience. These interviewers were selected to prevent ATUS interviewers from transcribing their own interviews and introducing possible bias. All transcribers were trained using a transcription manual and completed training and evaluation exercises before transcribing the ATUS tapes (see Attachment 2).

The transcription process took place over three waves. During each wave, a subset of tapes was transcribed. The research team reviewed a random subset of the transcriptions against the original tapes for accuracy and provided feedback on any problems they identified before the Census interviewers began transcribing the next subset of audio tapes. One transcriber was removed from the study during the first wave due to poor quality transcriptions.

<u>Participants</u>

Of the tapes transcribed, three were removed from the study because the coding scheme did not accommodate the situations: interview conducted in Spanish, a proxy respondent, and a respondent who reported activities for an incorrect day and then re-reported activities for the correct day. A fourth transcript was removed from the study because the audio quality prevented a useable transcription from being made.

<u>ATUS Data</u>

Each transcript can be linked to production ATUS data files. This allows for the behavior coding data to be linked with the actual activity and duration data, as well as to the ATUS survey methods data. The latter includes information such as the number of call attempts and the time of day the interview was conducted. In addition, each tape can be linked with data about the interviewers, including the number of years of interviewing experience, both on the ATUS and on other Census surveys overall, whether the interviewer is an ATUS coach, and the usual number of ATUS shifts worked per month.

Coding Scheme Development

The research team consisted of three survey methodologists from the Office of Survey Methods Research and one ATUS staff member. The team was tasked with developing a coding scheme, coding the transcripts and analyzing the data. An outside expert in cognitive psychology with behavior coding research experience was consulted at various times throughout the project.[2] Ongena and Dijkstra (2006) note two concerns that arise when researchers do the behavior coding. First, it is possible that the coding could be biased by the researchers' hypotheses about the outcomes. Second, researchers may have such an in-depth understanding of the coding scheme that it may be less reliable when used by other researchers. A more ideal setup would be to have field staff or specially trained coders do the behavior coding to avoid these disadvantages. However, this was a small study with a limited budget and thus no alternative coding arrangement was possible.

To our knowledge, this is the first to use behavior coding to assess conversational interviewing in time diaries, so the coding scheme development was an extensive process. To develop the coding scheme, the research team first identified several research questions related to ATUS, conversational interviewing and data quality. The team also reviewed other coding schemes (Belli, 1998; Belli et al., 2001; Cannell and Oksenberg, 1988; Esposito, 2004; Ongena and Dijkstra, 2006; Ongena et al., 2007; Schober, Conrad, and Fricker, 2004) to identify commonly used categories and variables. As Ongena and Dijkstra (2006) note, it is not always possible to identify all relevant behaviors in advance and thus the development of a coding scheme was an iterative process. Over the course of a year, the team met as a group about 20 times to listen to interview recordings and identify relevant respondent and interviewer behaviors to address the research questions. These initial meetings lasted anywhere from 1 hour to 7 hours, and upon completion of the 20 monitoring sessions, the team had a working coding scheme.

One of the first questions the project team addressed was how to define the unit of analysis, or turn. One option was to code at the activity level, where a change in turn occurs each time a new activity is reported. However, there could be three to four or more questions to code within one turn, or activity, when defined this way. Another option was to code at the utterance level, where a change in turn occurs each time the speaker makes an utterance. An utterance is a turn-of-talk by one speaker in the interview. It begins when a speaker starts talking and ends when the utterance naturally concludes (or is interrupted). At the end of an utterance, another speaker begins a turn-of-talk. Since one of the unique aspects of conversational interviewing in the ATUS time diary is the back-and-forth exchange between the interviewer and respondent that is designed to get activity and time information, the team opted to code at the utterance level to capture unique changes and retrieval behaviors about each reported activity. For example, the following exchange includes three utterances:

Interviewer: What did you do next?

Respondent: I went to the store. They were having a good sale.

Interviewer: How did you get to the store?

During the development of the coding scheme, the team faced the challenge of developing codes for a large and diverse set of behaviors that could be consistently assigned (across interviews and coders) and still provide a fine level of detail. Respondent behaviors proved to be the most challenging when developing codes, as they were often unpredictable and difficult to categorize. Although interviewers used a conversational interviewing style, they tended to follow a predictable question sequence, typically starting with a general retrieval question followed by probes (as necessary) to elicit more information or clarification from respondents. Respondents, on the other hand, provided a wide range of responses and other types of comments during the interview, with varying levels of relevance to the interview or research questions of interest.

Once the team had developed a working coding scheme, they conducted independent coding. During this phase, each researcher coded a single transcript and then met as a group to discuss the coding results and evaluate inter-coder reliability. The team identified which verbal behaviors could be coded reliably and those that seemed to be too subjective to code. The team also identified new behaviors and codes to add to the scheme. Overall, the team coded three cases in the second, independent coding phase. Each researcher spent about 2 hours coding a transcript per week over the course of 6 months. The team met about 18 times, with meetings lasting about 2 hours, for a total of 4 hours per week spent on coding.

In the third phase of coding scheme development, the team double-coded three transcripts, using a double-blind design where the case assignment was done by a non-team member. After coding, the researchers examined and discussed the codes they disagreed on and attempted to clarify the coding scheme to improve inter-coder reliability.

The final coding scheme identified 159 behaviors; 90 interviewer behaviors and 69 respondent behaviors. Attachment 3 provides the final coding scheme.

Coding Process

After the coding scheme was finalized, each researcher was assigned 42 transcripts to code. For reliability purposes, 18 transcripts were double coded. Reliability analysis was conducted and is described below. Each researcher went through assigned interviews sequentially. Although there was minor researcher variation in the exact coding process, generally researchers went through the transcript sequentially, assigning codes as the behaviors were observed. Researchers did ask clarification questions during the coding process, but no changes were made to the coding scheme during this time. Any questions or clarifications made were shared with all team members.

Reliability was calculated twice during the coding process. This allowed the research team to evaluate the coding consistency and how the coding scheme was working as a whole. In a few

instances, situations where the codebook was not being followed correctly by individual researchers or consistently across researchers were identified. In these situations, researchers went back and reviewed and revised their coded transcripts. In a very few cases, problems were identified with the coding scheme, and corrections were made. This happened for *qualified anchors*, respondent *non-substantive responses*, and *missed probes*. Following the changes to the coding scheme, researchers reviewed their coded transcripts and corrected items as appropriate.

Reliability

To allow reliability statistics to be calculated, researchers asked a colleague not involved in coding to randomly assign the transcripts among the researchers. Each researcher was assigned 42 transcripts. Of those 42 transcripts, twelve were double-coded; six with each of the other two researchers. This gave a total of 18 double-coded transcripts, or 17% of the total transcripts.

To ensure differential effort was not applied to coding transcripts that were and were not used for reliability, researchers were blind to which transcripts would be used for calculating reliability. Each researcher was given an assignment sheet with transcript identification numbers listed. Double-coded transcripts were dispersed throughout the list and researchers were told to follow the list in order to ensure that increased coding experience and fatigue would not systematically affect reliability results.

Reliability was calculated for all of the interviewer and respondent codes using the kappa statistic. To do this, two-by-two matrixes were created for each code, and used to identify the number of times the researchers agreed that a code should be assigned (A), the number of times researchers agreed a code should not be assigned (D), and when the researchers disagreed on whether a code should be assigned (B and C)[3]. Next, the proportion of agreement between the researchers was calculated for each code using Equation 1. The proportion of agreement that would be expected by chance was then calculated using Equation 2. Finally, kappa was calculated using Equation 3.

Equation 1
$$\frac{A+D}{A+B+C+D} = Proportion\ of\ Agreement$$

Equation 2
$$\frac{((A+B)*(A+C))+((B+D)*(C+D))}{(A+B+C+D)^2} = Chance\ Agreement$$

Equation 3
$$\frac{Proportion\ Agreement - Chance\ Agreement}{1 - Chance\ Agreement} = Kappa$$

To obtain a single measure of reliability for each code category across the three coding pairs, a weighted, average reliability value was calculated. A weighted value was used since the number of times a coding pair used a particular code may have varied. This value was calculated by

---

[3] Note, that for respondent sequentiality and non-sequentiality, the agreement was examined by activity number rather than by turn because that grouping made more sense for those categories.

summing the products of each coding pair's kappa value multiplied by the total number of times the pair had assigned the code. The sum of products was then divided by the number of times all coders, in all coding pairs, had assigned the code (see Attachment 3 for weighted Kappa scores for the code categories).

A Kappa score of <=0 indicates no agreement between raters, while a kappa of 1 indicates complete agreement. Only codes which met a minimum reliability of 0.40, which is considered moderate agreement, are reported in this paper.

Analysis

The analyses detailed in this report were performed on a dataset containing summary counts of the incidence of the individual behavior codes for each of the 104 interviews. To create this dataset, we began by merging the 86 single-coded interview files, and one randomly selected file from each of the 18 double-coded transcripts, into a single SAS data file. This data file contained over 21,000 rows of turn-by-turn data (verbatims and their associated interviewer and respondent codes) from interviews in the study sample. In order to make comparisons at the interview level, we then computed interview-level count variables for each of the 159 behavior codes, as well as variables to indicate the number of interviewer and respondent turns for each codeable ATUS activity and the total number of turns per interview. The resulting data file contained 104 rows of data, each corresponding to a single interview, and summary measures (counts, rates of occurrence) of code incidence per interview[4]. To this file we appended information about each respondent (e.g., demographics), interview (e.g., duration), and interviewer (e.g., Census experience, ATUS experience).

Results
Conversational Interviewing

In this analysis, we explored the types of conversational interviewing techniques being used in the ATUS. Because the ATUS does not follow a traditional script, we defined conversational interviewing as the use of information provided by the respondent or situation to change the question. This contrasts with a "script" of straightforward, simple questions (e.g. *What did you do next? How long did you spend…?*). For the ATUS, scripted items consist of both on screen questions as well as those learned in training. The following conversational techniques were identified for this analysis: the use of non-standard retrieval questions, 'anchors,' confirmation/feedback, digressions, and non-scripted probes.

---

[4] For the purposes of these analyses, we excluded turns in which the interviewer or respondent made no substantive utterances (i.e., turns for which the code category "Adequacy" was coded as 'xx' or 'yy' – see Attachment 3). By definition, those turns received no coding of the key substantive variables, so we left them out when constructing our incidence and turn count variables. In addition, we excluded turns in which the utterances were devoted exclusively to the collection of 'who' and 'where' information (i.e., turns for which the code category "Activity Number" was coded as '999' or '998').

<u>Retrieval Questions</u>

In addition to the three scripted questions (the first question of the diary, 'What did you do next?" and "How long did you spend [activity]?"), to collect time and activity information, interviewers are trained on a variety of questions to help respondents retrieve activity and time information. The non-scripted questions, such as those which ask the respondent to work backwards, visualize an activity, or fill in time gaps, require the interviewer to think creatively about the best way to get the desired information from the respondent or to otherwise tailor the interview to the situation. Use of these non-scripted questions shows that interviewers are taking advantage of the flexibility allowed in the ATUS interview, and being more conversational in their interviewing approach.

The distribution of the different types of retrieval questions (scripted and non-scripted) provides one indicator of how conversational an interviewer was being during an interview. An interview filled primarily with scripted retrieval questions is not conversational; the interviewer is simply reading the script from the instrument screen. An interview filled with non-scripted questions is conversational; the interviewer is changing the question and approach based on what the respondent has said. Table 2 summarizes the frequency each type of retrieval question was used in the ATUS diaries.

Activity and time retrieval questions (RQs) make up 54.9 percent of all interviewer turns[5]. On average, there were 36.4 retrieval questions in each interview, ranging from 11 to 88 (Table 2). Of all the RQs, a higher percent asked about time (54.1 percent) rather than activity (45.9 percent).

Table 2. Use of Retrieval Questions (RQs)

	Mean	Minimum	Maximum	Std. Deviation
Sum of all RQ	36.4	11.0	88.0	14.9
Sum of all Activity RQs	16.6	4.0	49.0	7.2
Sum of all Time RQs	19.8	7.0	49.0	8.8
Percent of Activity RQs	45.9%	27.5%	70.0%	7.8%
Percent of Time RQs	54.1%	30.0%	72.5%	7.8%

To determine if interviewers were using conversational interviewing techniques, the percent of RQs that were scripted and non-scripted were identified (Table 3). An overwhelming majority, 88.7 percent, of all RQs were scripted, with interviewers using non-scripted RQs only 11.3 percent of time. Other than the diary start question, which only occurs at the very beginning of

---

[5] Again, this excludes 'who' and 'where' turns as well as non-substantive utterances.

the diary, the two scripted RQs were used with about the same frequency; 40.2 percent of the RQs were "Activity: Sequential Forward" (what did you do next?) and 43.5 percent were "Time: Duration" (how long did you spend *activity*?). Of the non-scripted RQs, there was one that was used more than the others, "Time: Start/End" (what time did you start/stop that?) which was used in 10 percent of all the RQs. The rest of the non-scripted RQs were used very rarely.

Table 3. Scripted and Non Scripted Retrieval Questions

	Average % of all RQs	Min % of all RQs	Max % of all RQs
Scripted	*88.7%*	*38.7%*	*100%*
Activity: Diary Start	4.%	1.3%	13.3%
Activity: Sequential Forward	40.2%	20.0%	61.5%
Time: Duration	43.5%	9.7%	58.8%
Non-Scripted	*11.3%*	*0%*	*61.3%*
Activity continuity*	0.3%	0%	9.7%
Activity decomposition	0.1%	0%	3.2%
Activity Time gaps	0.4%	0%	5.9%
Next activity you remember*	0.4%	0%	11.1%
Time: Start/End	10.0%	0%	51.6%
Timing decomposition*	0.0%	0%	4.4%

*The reliability of these codes was not assessed due to their low incidence rates.

To explore the effectiveness of the scripted and non-scripted RQs, we looked at the number of RQs used for each reported activity (Table 4). On average, there were 23.1 activities reported per interview, and 36.4 RQs used, resulting in an average of 1.6 RQs per activity. There is a difference in the ratio of scripted and non-scripted RQs to the number of activities; on average there were more activities reported per each scripted RQ (1.4) than for non-scripted RQs (0.2), indicating that non-scripted RQs were less efficient in eliciting usable activity information.

Table 4. Number of Retrieval Questions per Activity

	Average per interview	Min per interview	Max per interview	St Dev
Number of Activities	23.1	8	56	9.23
Number of RQs	36.4	11	88	14.9
Ratio of Retrieval Question per activities (# of RQs/ # of activities)	1.6	0.7	2.4	0.3
Ratio of Activities to Scripted RQs	1.4	0.6	2.3	0.3
Ratio of Activities to Non-Scripted RQs	0.2	0.0	1.0	0.2

Anchors

ATUS interviewers are trained on several forms of active listening, to use respondent-provided information to help administer the diary in a way that collects high quality data while reducing the cognitive burden of the recall task. One way that interviewers often incorporate active listening is through the use of anchors. Anchors are pieces of information that the respondent has previously provided used in a retrieval question. For example, instead of just saying "What did you do next?" an interviewer might incorporate an anchor and ask "What did you do after you finished eating?" By using information the respondent already provided, the fact they were eating, in a question asking for new activity information, the interviewer has changed the question and tailored it to the respondent and the situation.

Interviewers use anchors to help respondents recall the desired activity or time information by anchoring the question on information they have already reported. Even when combined with scripted retrieval questions, anchors show an effort of the interviewer to make the interview more conversational by demonstrating active listening and modifying question wording based on respondent input. The following tables summarize the use of non-scripted anchors in the ATUS. Three infrequently used anchor codes that did not meet reliability standards are not included[6].

On average, interviewers used non-scripted anchors 14.1 times in an interview (Table 5). There was a wide range in their use, with one interviewer using only three anchors in an interview, and another using 45. Of the RQs where a non-scripted anchor could have been used, 69.1 percent of them had an anchor added. Again, this ranged considerably, between 21.4 and 94.1 percent.

---

[6] There are some anchors programmed into the ATUS interview, such as "How long did you spend {activity}?" These scripted anchors were excluded from analysis.

Table 5. Use of Non-Scripted Anchors

	Average # per interview	Min #	Max #	St Dev
Number of non-scripted anchors	14.1	3	45	6.57
% of RQs with a non-scripted anchor	69.1%	21.4%	94.1%	16.1%

Digressions

Digressions occur when an interviewer asks a question or makes a comment that is not directly related to diary completion. For example, a respondent states, "I went back inside the house because I forgot my lunch." The interviewer responds "Oh, I do that all the time!" Because interviewers are not following a script, conversational interviewing makes it easier for interviewers or respondents to stray off the course of the survey. These side conversations, though not directly related to the diary completion, may allow the interviewer to build or maintain rapport with the respondent. However, digressions may also indicate respondents are "off task and interviewers are trained to redirect these respondents.

Tables 6 and 7 show that ATUS interviewers digressed very little overall. In total, digressions occurred in about 2.4 percent of all interview turns, with an average of 1.9 digressions per interview. About half of all interviews had no digressions, while five or fewer digressions occurred in about 90 percent of interviews. There was a considerable range in the number, with five interviewers digressing more than 10 times in a single interview.

Table 6. Overview of Digressions

	Mean	Std Dev	Minimum	Maximum
Number of Digressions per interview	1.9	3.8	0	26

Table 7. Frequency of Digressions

Number of Digressions per Interview	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	54	51.92	54	51.92
1	19	18.27	73	70.19
2	7	6.73	80	76.92
3	10	9.62	90	86.54
4	2	1.92	92	88.46
5	2	1.92	94	90.38
7	3	2.88	97	93.27
8	2	1.92	99	95.19
10 or more	5	4.8	100	100

The occurrence of digressions is not distributed equally among interviewers (Table 8). Two of the 36 interviewers accounted for 32 percent of all digressions. These interviewers strayed off the course of the interview a total of 24 to 39 times over three interviews each. Further examination of these interviewers showed that both had less than one year of ATUS interviewing experience and one or more years of Census interviewing experience.

Table 8. Distribution of Digressions across interviewers

Number of Interviewers	Number of Digressions	Percent of all Digressions
7	0	0.0%
10	1	5.1%
7	2 to 4	8.7%
5	6 to 9	18.9%
5	11 to 17	35.2%
2	24 to 39	32.1%
Total	196	100.0%

Confirmation and Feedback

In a strictly standardized interview, interviewers are required to adhere to scripted probes. In the absence of instrument-scripted probes, interviewers' attempts to verify respondents' answers or to provide clarification about the survey task go beyond the standardized protocol. In practice, however, interviewers sometimes do make 'unscripted' use of information that has been provided by a respondent, either to confirm the accuracy of the information or to communicate the interviewer's understanding and use of a previous report. Although such behavior may enhance data quality (e.g., by ensuring shared understanding) and make interviews seem more conversational, one potential problem with unscripted verification probes and feedback is that their use may vary considerably across interviewers (both when and how they are used). This is because they are based on interviewers' interpretations of and inferences about previously reported information.

We examined four types of interviewer Confirmations (i.e., confirming reported duration, clock time, activity, or some mixture of the three) and four types of Feedback (i.e., task-related, focus on yesterday, focus on respondent, and operational problems – see Attachment 3). Together, these Confirmation and Feedback (CF) statements accounted for nearly a quarter (23%) of all interviewer turns. Moreover, the prevalence of CF statements was substantially higher when we restricted our analysis to only those turns in which researchers were permitted to use the CF codes (i.e., by design, researchers did not assign the four types of interviewer Confirmation codes on turns that received a Retrieval Question code).[7] As can be seen in Table 9, confirmation and feedback statements were made on nearly half (45%) of the turns that were eligible to receive a CF code.

---

[7] If an interviewer turn included a specific retrieval question as well as a statement that provided feedback to the respondent about the duration, clock time, or activity reported in their previous answer (e.g., "Ok, you said that you were walking the dog. What did you do next?"), we assigned the turn a Retrieval Question code, and treated/coded the feedback as a type of Anchor. No Confirmation/Feedback codes were assigned on these turns. As a result, for the Confirmation/Feedback statement analyses, when the statement referenced an activity, clock time, or duration (or combinations thereof), the denominator was all non-Retrieval Question turns (excluding 998/999 and Who/Where turns). For analyses of all other Confirmation/Feedback codes, the denominator was all interviewer turns (excluding 998/999 and Who/Where turns), since it was permissible to use Confirmation/Feedback codes 'e' through 'h' (e.g., task-related feedback) and Retrieval Question codes on the same turn.

Table 9. Percentage of Eligible Interviewer Turns That Were Confirmation or Feedback (CF) Statements

	Mean	**Min**	**Max**	**Std. Dev.**
Total CF codes	45.3%	2.7%	100.0%	17.2%
Confirm duration	5.4	0.0	17.1	4.6
Confirm clock time	9.1	0.0	44.4	7.1
Confirm activity	18.9	0.0	72.7	12.3
Confirm dur./clock/act. mix	8.5	0.0	30.0	7.6
Task-related feedback	2.1	0.0	12.5	2.7
Yesterday feedback	0.4	0.0	4.0	0.8
Refocus on respondent	0.1	0.0	3.3	0.5
Operational feedback (problems)	0.8	0.0	5.6	1.3

Also evident in Table 9 is the fact that interviewers were far more likely to ask respondents to confirm their reports (41.9% of eligible turns) than they were to offer feedback (3.4% of eligible turns). This result makes sense since there were many more opportunities for interviewers to try to verify something the respondent said than there were for interviewers to try to correct some perceived respondent error or to comment on instrument problems (both of which were relatively rare events).

Follow-up Probes

ATUS interviewers have a number of follow-up probes in the diary that they use to follow up and gather more information on specific topics and responses. One type of follow-up probe is automated; with this probe, a pop-up screen in the CATI instrument displays the probe question text. As an example, any activity that lasts more than two hours, with the exception of sleeping, working, or attending high school, requires the interviewer to verify the activity by asking the question: Did you stop (activity) to do anything else at that time? Another type of follow-up probe is not automated. Instead, the interviewers receive training on the question, including the specific wording. An example of this type of probe occurs when a respondent reports that they are taking a class. The interviewer is trained to follow up with: "Is that a class that you are taking mainly for a degree, for your current job, for personal interest, for volunteering, or some other reason?"

We investigated whether interviewers ask automated and training follow-up probes as scripted, or whether they modified the script in some manner, utilizing the flexibility of conversational interviewing[8]. Scripted probes are defined as asking the probe verbatim with no intervening inserts, and including all relevant concepts/main words without substantive changes. Non-scripted probes are defined as when interviewers deviate from the scripted language; this includes modifying the language so it renders a probe as incomplete conceptually (for example, if an interviewer only asks if a class is taken for a degree, and leaves out current job, personal interest, volunteering, or something else), as well as modifying the language so it generally covers the probe concepts.

Follow-up probes make up 13.6 percent of all interviewer turns. Table 10 shows that interviewers use conversational interviewing techniques when asking follow-up questions. On average, interviewers ask five non-scripted follow-up probes per interview, and less than one scripted probe per interview. Overall, non-scripted probes form the largest share of follow-up probes, accounting for approximately 85% of all follow-up probes.

Table 10. Use of Scripted and Non-scripted Follow-up Probes

	Mean	Minimum	Maximum	Std. Deviation
Sum of Scripted Probes	0.79	0	5.0	1.1
Sum of Non-Scripted Probes	5.0	0	13.0	2.9
Percent Scripted	15.3%	0	100%	21.9%
Percent Non-Scripted	84.7%	0	100%	21.9%

Differences by Interviewer Experience

We looked at each conversational interviewing technique to see whether the frequency of use varied by three types of interview experience: Census interviewing experience (less than 4 years vs. 4 years and higher); ATUS specific interviewing experience[9]; (less than 12 months vs. 12 months and higher) and ATUS Coach status (Coach vs. non-Coach). ATUS coaches are usually ATUS interviewers with more ATUS interviewing experience and higher response rates.

There was no evidence that interviewer experience was related to the use of the conversational interviewing techniques involving non-scripted RQs, anchors, or digressions. For each of these techniques, no significant differences were found by ATUS experience, Census experience, or Coach status (p<.05).

---

[8] The content of the probe (e.g. travel, education, sleep) was not distinguished in this study.
[9] ATUS interviewers also work on other, non-conversational surveys, and so their experience on those surveys may be impacting these results.

There was evidence that interviewer experience was related to the use of follow-up probes and to confirmation and feedback statements. ATUS coaches were significantly more likely to use non-scripted probes than non-coaches ($p < .001$), and were less likely to use confirmation statements than non-coaches ($p<0.5$).

<u>Impact on Data Quality</u>

We next looked at the impact of these conversational interviewing techniques on data quality. We identified a few, basic variables with which to measure initial data quality. The first data quality variable we identified was the percentage of respondent turns coded as inadequate (respondent did not report an activity, duration, or time that enabled the interviewer to code the response). A high percentage of inadequate respondent turns indicate that the respondent may not understand the questions and also contributes to longer interview lengths, both which negatively impact the respondent's interview experience.

The next data quality variable we identified was the percentage of respondent activities that could not be coded by ATUS coders. In the ATUS, all activities are assigned an activity code from which estimates can be made. For example, walking may be for exercise or travel. If the respondent does not provide a sufficiently detailed response (and the interviewer fails to follow up to get enough detail), the activity is assigned a data code which cannot be categorized into a detailed activity.

Finally, we identified the number of reported activities as a sign of data quality. A high number of reported activities indicates respondents who are engaged and reporting not only the obvious activities that take up a large portion of their day (e.g. sleeping and working), but also the more mundane, short activities that occur in between longer activities.

There was no evidence that the use of non-scripted RQs or non-scripted anchors, use of digressions or scripted or non-scripted follow-up questions were related to the data quality measures examined ($p<.05$). There was a significant association, however, between interviewers' use of confirmation/feedback statements and the percent of inadequate respondent turns. The percent of inadequate respondent turns was significantly reduced as the percent of CF statements per interview increased ($p<0.5$).

Discussion

The purpose of this report was to document the objectives, analytic methods, and preliminary findings of a small-scale, behavior-coding study of ATUS interviews. This study was motivated by contrasting two approaches to survey administration – a strictly standardized method in which interviewers are required to read questions exactly as worded and use a limited set of scripted probes in well-defined situations, and a more conversationally flexible approach in which

interviewers are permitted (even encouraged) to tailor the assistance they provide respondents on the basis of their individualized assessments of the interview situation. We noted that the ATUS interview uses a hybrid of these two approaches, but that it is generally considered to be one of the most conversational interviews in the federal statistical system. One of the aims of this study was to systematically investigate the extent to which conversational interviewing techniques were actually being used by ATUS interviewers, and if so, whether their use varied by interviewer experience, and whether they had any demonstrable impact on ATUS data quality.

Before summarizing the main study findings, it should be emphasized that our behavior coding analysis of the 104 ATUS transcripts generated an extremely rich dataset. The majority of the team's work to date was spent on the important iterative development processes underpinning the coding scheme, ensuring the reliable coding of transcripts, and the creation of the project data files. The results contained in this report reflect the team's initial effort to provide a general overview of some key attributes of the interview interactions and to explore basic descriptive statistics and univariate comparisons. Additional analyses of these data should (and we hope will) be conducted to examine multivariate associations among the data elements, expanded treatment of data quality, and closer inspection of respondent factors. With this caveat in mind, we next briefly highlight the main findings from the present analyses, with special attention to those findings that may have implications for interviewer training and/or program office interventions.

How conversational is ATUS? Although it is viewed as 'conversational,' our results provide only limited evidence that it actually is in practice. The bulk of the interviews we analyzed were made up of retrieval questions, and the majority of retrieval questions were asked as scripted. Most interviewers proceeded through the instrument in a straightforward fashion, linearly walking respondents through the day, without much observable need to utilize some of the more conversationally-oriented recall aides (e.g., visualization, leveraging gaps, working backward). The occurrence of interview digressions in these interviews also was generally very low. Moreover, when interviewers did digress or use unscripted retrieval questions, those behaviors were unrelated to our three ATUS data quality indicators. Interviewers did exhibit more conversational behaviors in three areas: the relatively high incidence of unscripted anchors, unscripted follow-up probes, and confirmation and feedback statements. However, only confirmation and feedback statements appear to be related to gains in data quality (and in only one of our three data quality metrics). Together, these preliminary analyses suggest that many of the conversational interviewing techniques that ATUS interviewers are taught may not be needed to get respondents through the diary or to improve data quality. This leaves as an open question the efficacy of interviewer training protocols to the extent that they continue to emphasize conversational approaches to collecting diary information.

Somewhat surprisingly, years of ATUS interviewing and overall census interviewing did not have much of an impact on the types of behaviors evidenced in these interviews. ATUS coaches were more likely than other interviewers to use non-scripted follow-up probes and less likely to

ask respondents to confirm a previous report. Both of these behaviors may reflect coaches desire to streamline and shorten the interview (i.e., by paraphrasing probes rather than reading them verbatim, and accepting respondents' answers without additional corroboration). But, the broader consistency of interviewer behaviors despite experiential differences may demonstrate the inherent sway of the step-by-step temporal flow of the time diary, or perhaps a lack of recognition by even seasoned ATUS interviewers of the value of conversational techniques.

Study Limitations and Next Steps

This study is an exploratory look at a small, non-random sample of ATUS interviews. Although it provided much insight into the extent of conversational behaviors in these ATUS interviews, there are some limitations to this exploratory study to consider when drawing conclusions.

The coding scheme used in this study drives the analysis and conclusions that can be drawn. There was no existing diary coding protocol, so development of the coding scheme was done using a set of preliminary research questions, and was relatively ambitious. The large number of codes, and the subtle distinctions between some of them, resulted in several codes that were not reliability coded and so could not be analyzed. A 'master' coder or more straightforward coding scheme would likely improve coding reliability and increase the codes that could be reported. The complexity of the coding scheme also required some data processing and cleaning, and it is likely that not all coding inconsistencies have been identified, documented, or corrected.

This analysis excluded 'who' and 'where' questions, as they were thought to be secondary to the activity and time retrieval questions. The extent to which these questions differ in their conversational nature limits the conclusions from this report.

We identified a few simple measures of data quality (percent of inadequate respondent turns, percent of uncodeable respondent activities, and number of reported activities), but suspect there may be others which might shed further light on the impact of conversational interviewing on data quality. Although there are some variables thought to be indicators of data quality, we believe there is no direct measure of actual data quality, and so only to the degree our indicators accurately capture that construct are our conclusions valid.

The analyses in this study are preliminary and generally univariate. Additional analysis, taking advantage of the relationships between variables and the interactions between interviewers and respondents are planned and expected to provide additional insight into the impact of conversational interviewing in the ATUS survey. Specifically, we plan to examine:

- Additional interviewer behaviors
  o Create a scale of conversational interviewing (e.g., a highly conversational interview) and rate each interview. Use that variable in multivariate analysis with respondent and interviewer variables.

- o Differentiate between the kind of probe used (e.g., travel, education, leisure) and see if there are differences in frequency of probes, completeness of probes, and number of missed probes.
- Respondent analysis
  - o Are there some respondent characteristics (e.g., old, unemployed) where conversational interviewing is used more often/is more effective?
  - o Are there some respondent behaviors (such as initial confusion) that are related to conversational interviewing techniques being used?
- Additional data quality measures
- Multivariate analysis
- Sequential analysis (e.g., interviewer did X, then respondent did Y)
  - o Points in the interview when conversational interviewing is more/less used
  - o Are nonscripted questions usually asked after scripted questions fail to elicit a usable activity?
- Do interviews with higher percents/numbers of behaviors associated with conversational interviewing take longer than those with lower numbers?
- How often do interviewers perform rapport building behavior?
- How often are interviewers leading?
- Does the interviewer use precode terminology to redefine respondent activities?
- Do interviewers ask probes without communicating the full intent of the probe?

References

Belli, R. F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, 6, 383-406.

Belli, R.F., Shay, W. L., & Stafford, F. P. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, 65, 45-74.

Belli, R. F., Lee, E.H., Stafford, F. P., & Chou, C.H. (2004). Calendar and question-list survey methods: Association between interviewer behaviors and data quality. *Journal of Official Statistics*, 20, 185-218.

Cannell, C. and Oksenberg, L. (1988). Observation of Behavior in Telephone Interview. In *Telephone Survey Methodology*, R.M. Groves, P.P. Biemer, L.E. Lyberg, J.T. Massey, W.L. Nicolls, II, and J. Waksberg (eds). New York: Wiley, 475–495.

Conrad, F. G. and Schober, M.F. (2000). Clarifying Question Meaning in a Household Telephone Survey. *Public Opinion Quarterly* 61: 1-28.

Esposito, James L. (2004). Iterative, Multiple-Method Questionnaire Evaluation Research: A Case Study. *Journal of Official Statistics*, 20, 143-183.

Schober, M.F., Conrad, F.G., and Fricker, S.S. (2004). Misunderstanding Standardized Language in Research Interviews. *Applied Cognitive Psychology*, 18:169-188.

Ongena, Y.P. and Dijkstra, W. (2006). Methods of Behavior Coding of Survey Interviews. *Journal of Official Statistics*, 22, Vol. 3, 419-451.

Ongena, Y.P. Bautista, R., Kruse, Y, Perez Berestycki, J. & Vitense, J. (2007) The Development of a Multivariate Coding Scheme for Evaluation of Event History Calendar and Standardized Interviews. Paper presented at the 62nd Conference of the American Association for Public Opinion Research, Anaheim, California, May.

<u>**Attachment 1: Interviewer Taping Instructions**</u>

INTERVIEWER TAPING PROCEDURES

TAPING INSTRUCTIONS FOR THE TELE-RECORDER 500

One of the methods that BLS uses to evaluate questionnaires and survey questions is behavior coding. This involves analyzing the interaction between the interviewer and the respondent and coding the way the questions are asked and answered. Using the behavior coding method to assess the ATUS diary questions will help us understand the ease and difficulty of the task that faces respondents as they recall their "yesterday" activities and times. To conduct this evaluation, we need to tape record some of the interviews. You should be aware that the purpose of the taping is to evaluate the diary questions and interview procedures, not the interviewers.

<u>Tape Recorder Settings</u>

1. Set the SPEED button to Standard.

2. Set the VOX / NORMAL button to Normal.

3. Set the MONITOR button to Off.

4. Set the H/M/L sensitivity button on the bottom of the recorder to H.

5. Plug one end of the adapter into the back of the tape recorder marked DC 6V. Plug the other end into an outlet.

PLEASE, PLEASE ensure that all settings on the recorder are correct before you begin recording. Do a test recording of your own voice and replay the tape and listen to it to ensure that all recorder settings are correct.

<u>Preparing & Using the Cassette Tape</u>

1. You will be using 90-minute tapes. Please use one tape for each interview. Each side of the tape is 45 minutes long. That should cover most diary interviews, but if an interview

requires more time, use Side B to continue the interview. If you do this, be sure to fast forward the tape to the end before beginning an interview on Side B (press the FAST-F button).

2. Insert the cassette tape into the Tele-Recorder 500 by pushing the STOP/EJECT button and inserting the tape in the space provided under the lid. Press the lid back down. The full reel of tape should appear on the left. Run the tape forward to clear the tape header before you start an interview. You will ask respondents permission to tape the interview and this must be included on the tape (see below). If you don't run the tape forward, it may not get taped.

3. After each taped interview is complete, it is important to record the following information on the cassette label:
   a. 8 digit case I.D. number
   b. Interviewer Bond I.D. number
   c. Interview date (mm/dd/yy)
   d. Interview time of day (EST and AM/PM)

Taping the Interviews

There will be interviewing stations equipped for taping. Each interviewer is responsible for taping 3 diary interviews. The following guidelines for taping are suggested:

- If possible, tape the 3 interviews in one or two shifts;
- An interviewer should begin their shift at the taping station and continue there until 3 taped interviews have been completed or their shift is over.

Obtaining the Respondent's Permission to Tape

We MUST have the respondent's permission to tape record the interview and that permission must be on the tape. In order to do this, you will need to ask for permission to tape twice, once before turning on the tape recorder and a second time while the tape is running. Without the respondent's permission to tape the interview on the tape, we will not be able to use the tape for research purposes.

Taping will begin after the labor force questions, and before you begin the diary. After the labor force questions, interviewers at stations equipped for taping will ask the respondent's permission to tape using the following script, which will be provided in a job aid and posted at the interviewing station:

At the >ACTIVITY< item, stop and ask for permission to tape.

Read the following:

"To help us improve the diary questions, we would like your permission to tape record the remaining portion of this interview.   Is that OK with you?"

YES   →   Turn on tape recorder; read**, "Thank you. I have turned on the recorder. Please just say 'yes' again to indicate on the tape that I have your permission to record this interview."** This records the permission.

NO   →   DO NOT turn on tape recorder.

After the taped interview, label the cassette tape:

- Write 8 digit case I.D.
- Write Interviewer Bond I.D.
- Write interview date (mm/dd/yy)
- Interview time of day (EST and AM/PM)

Turning in the Tapes

At the end of your shift, give your taped interviews to the shift supervisor. The shift supervisor will record the information from the cassette label on the tape login sheet so that we can keep track of how many tapes have been completed.

THANK YOU IN ADVANCE FOR YOUR COMMITMENT TO ATUS!

Attachment 2: Transcription Manual and Evaluation Exercises

The following resources were used in developing the transcription instructions:

2005. <u>Transcription Training Manual.</u> Bureau of Sociological Research, University of Nebraska-Lincoln.

Garbarski, Dana, Jennifer Dykema, Nora Cate Schaeffer, and Douglas Maynard. 2006. Manual for Transcribing Interviewer-Respondent Interation in Surveys of Older Adults. University of Wisconsin-Madison.

McLellan, Eleanor, Kathleen M. MacQueen, and Judith L. Neidig. 2003. "Beyond the Qualitative Interview: Data Preparation and Transcription." Field Methods 15: 63-84.

National Center for Social Research. Guidelines for the transcription of verbatim tapes.

Contents

ATUS Audio-Tape Transcription Instructions

1. Introduction

The American Time Use Survey (ATUS) is a survey sponsored by the Bureau of Labor Statistics (BLS) that measures the amount of time people spend doing various activities, such as paid work, childcare, volunteering and socializing. A team of BLS researchers is interested in looking at the data quality of the ATUS time-diary data. Using the time-diary transcripts you prepare, researchers will undertake "behavior coding," which is the systematic coding of the interactions between an interviewer and a respondent. Behavior coding is used to evaluate survey tasks and questions, and to measure respondent and interviewer behaviors. Using this technique, researchers will explore issues such as what questions, tasks, and concepts are difficult for respondents and interviewers, and how the interaction between respondents and interviewers affects data quality. The overall goal of the project is to develop recommendations for instrument and interviewer training modifications that will make it easier for respondents to provide and interviewers to collect high quality data.

The ATUS interview is collected via computer-assisted telephone interviewing at the Jeffersonville Data Collection Center. The behavior coding project focuses on the "time diary," which is the main section of the ATUS interview. During the time diary a pre-selected household member reports about his or her activities during the 24-hour period on the day before the interview. Interviewers use a set of scripted open-ended questions along with conversational or flexible interviewing techniques to walk respondents through the 24-hour day, collecting information about time spent in activities beginning at 4:00 am on the previous day up until 4:00 am of the interview day. For each activity reported by the respondent, the interviewer asks how long the activity took, who was in the room with the respondent or who accompanied them, and where each activity took place. After completing this sequence of questions on an activity, interviewers prompt respondents for the next activity by asking them what they did next. If you would like to find out more about the ATUS interview and how the data are used, please visit the website at http://www.bls.gov/tus/.

Your work will consist of transcribing 108 audio tapes of the 24-hour time diary section of the ATUS interviews. The diary section of the interview usually takes about 15 minutes per respondent, but it can range from about five to 45 minutes. The ATUS time diary includes structured questions, but the interviewer is allowed freedom in how he/she asks the questions, and it is very important that you transcribe everything as it is said exactly on the tapes.

Since the transcripts will be behavior coded by the researchers, it is important that each member of the transcription team is transcribing in the same manner. Throughout the process, researchers will review transcriptions and provide feedback to make sure team members are consistent with each

other. TCCO will coordinate the review process with the researchers, and supervisors will relay the information to transcribers. In addition, as problems or questions arise about the transcriptions, please report them to the supervisors, and TCCO will coordinate with the researchers to resolve problems and answer questions.

2. Confidentiality

The ATUS tapes are covered under Title 13. Please keep everything you transcribe confidential. If you cannot understand a part of the tape you are transcribing, you may ask another ATUS transcriber or supervisor to help you decipher what is being said, but you must keep all other information about the case confidential.

3. Checking out the audio tapes

Tapes need to be checked out from the project supervisor.

Each tape is labeled with a Tape Number, a Case Identification Number, and a Coded Interviewer Number.

Tapes should be kept in Tape Number order and transcribed in Tape Number order.

On the tape label, please write your transcriber number so we know who is working on each tape.

The project supervisor will have a <u>Tape Checklist</u>. The checklist includes a column with all of the Tape Numbers and the Case ID Numbers. There are also columns for Transcriber Identification Number and the date. Before you begin to transcribe a tape, a supervisor will write your Transcriber ID number next to the tape number on the Checklist. If there is already a Transcriber ID number next to a tape number, move to the next tape. This will help prevent having the same tape transcribed twice. The Date column on the checklist is the End Date (when you have finished a tape). The supervisors will enter the date on the checklist only when you have finished transcribing a tape.

There is also a Tape Log with the tapes. This log has columns for the Tape Number, your Transcriber ID Number, the date and time the tape was checked out, and the date and time the tape was checked back in. *Each time* you get a tape from the supervisor, the supervisors will log it on the Tape Log form and will complete all of the information required on the log. Each time you return a tape to the supervisor, the supervisor will log that time as well. The tapes will be kept in numerical order according to the Tape Number.

When you have finished transcribing a tape, the supervisors will put it in the "Completed" box, in numerical order according to the Tape Number.

You are responsible for the tapes you use. You will be responsible for lost tapes. Unless you are transcribing, tapes need to remain with the supervisor. Please return the tapes to the supervisor each time you take a break and leave your workstation, and when you are finished with your shift for the day.

## 4. Instructions for using the RR-830 cassette transcriber

You will be transcribing the ATUS diary audio tapes using the Panasonic Cassette Transcriber (Model RR-830) and Panasonic Foot Controller (Model RP-2692). The purpose of this section is to provide you with instructions for using these transcription devices. Figure 1 (page 10) shows the locations of the Transcriber and Foot Controller controls.

To Playback:

    1.   Connect the Transcriber's AC power cord to a grounded power outlet.

    2.   Insert the connection cord of the Foot Controller into the Foot Controller Jack on the side of the Transcriber.
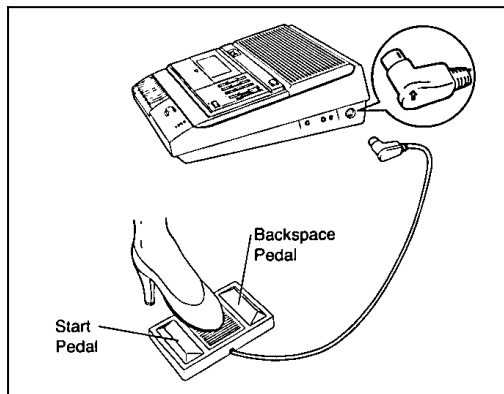

    3.   Place the ATUS audio cassette tape into the Transcriber's cassette holder with Side A facing up.

    4.   Set the Variable Speech Control (VSC) switch (item 20, Figure 1) to "OUT."

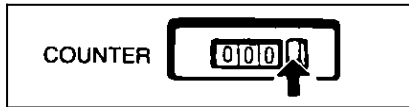   5.   Press the Power Switch. The Power Indicator will light up
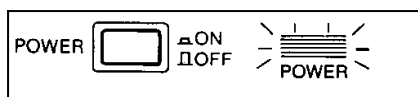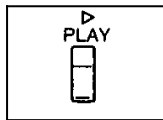
Playback, continued:

6.   Be sure that the tape is rewound to the beginning of Side A. Press the Rewind/Review
        button until the tape is fully rewound, then press the Stop button.
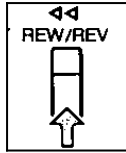
7.   Reset the Tape Counter.

8.   Press the Playback Button.

9.   Step on the Start Pedal (right side) of the Foot Controller and playback will begin. Hold
        down the pedal and playback will continue. The Tape Running Indicator will light up.

10.   Adjust the Volume and Tone Controls, as necessary.
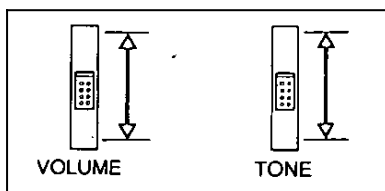
11.   Remove foot from the Start Pedal and playback will stop.

12.   Step on the Backspace Pedal (left side) and the tape will rewind. Remove foot from the
        Backspace Pedal and rewind will stop.

Auto Back Space Function

When using the Foot Controller, it is possible to have Auto Back Space when the Start Pedal is released. This function automatically rewinds the tape from 1 – 10 seconds, which allows you to re-play the most recent segments of the interview without having to manually rewind the tape. You can control the amount of Auto Back Space by adjusting the "Back Space Control" (item 19, Figure 1).

Important: This unit is put into pause mode when the Foot Controller pedals are released. Do not leave the machine in this mode for long periods of time. Be sure to press the Stop button and Power Switch to turn off the unit.

Variable Speech Control (VSC) Playback

All of the ATUS interview tapes for this project should have been recorded in a standardized manner to ensure a consistent playback speed. However, there may be instances in which you find that a tape's playback speed needs adjustment. The VSC feature on the Transcriber allows you to make these adjustments. In the event that your tape plays back at too high (or low) a speed, follow these instructions:

1.   Insert the ATUS audio tape into the Transcriber's cassette holder, Side A facing up.

2.   Set the Variable Speech Control (VSC) switch to "IN."

3.   Press the Playback button.

4.   Step on the Start Pedal (right side) of the Foot Controller and playback will begin. Hold down the pedal and playback will continue. The Tape Running Indicator will light up.

5.   Adjust the Tape Speed Control for desired speed from 1 to 2 times tape speed. Set the VSC Level Control to the desired pitch. It is recommended to set both selectors to the same level.
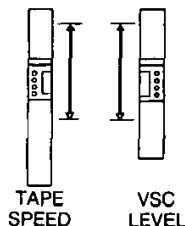

TAPE
SPEED

VSC
LEVEL

Figure 1. Location of Controls for the RR-830 Cassette Transcriber

LOCATION OF CONTROLS

(Back)

Foot Controller
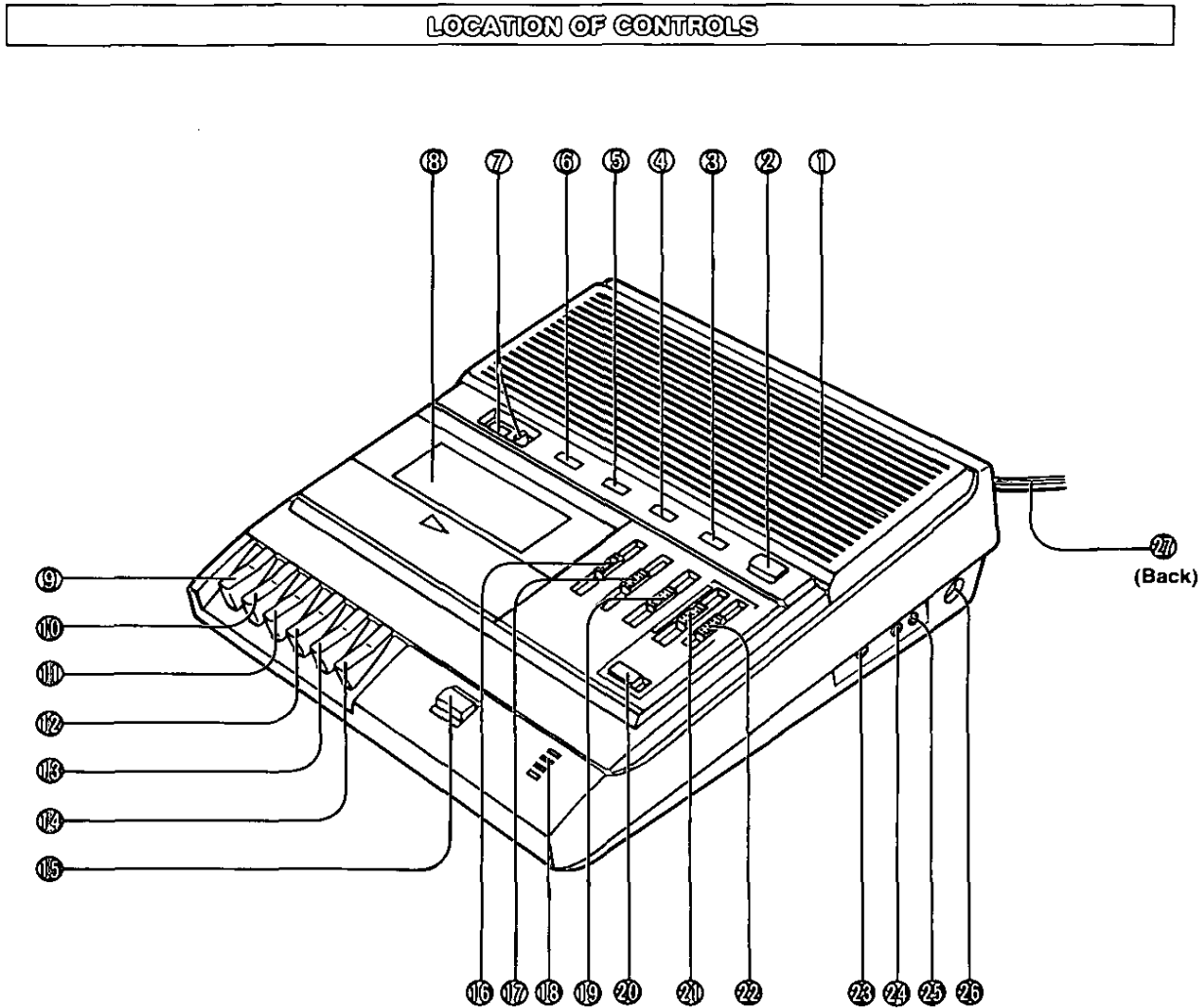
① Speaker
② Power Switch (POWER)
③ Power Indicator (POWER)
④ Recording Indicator (REC)
⑤ Quick Erase Indicator (QUICK ERASE)
⑥ Tape Running Indicator (TAPE RUN)
⑦ Tape Counter and Reset Button (COUNTER)
⑧ Cassette Holder
⑨ Record Button (O RECORD)
⑩ Stop Button (□ STOP)
⑪ Playback Button (▷ PLAY)
⑫ Rewind/Review Button (◁◁ REW/REV)
⑬ Fast Forward/Cue Button (▷▷ FF/CUE)

⑭ Eject Button (△ EJECT)
⑮ Pause Button (‖ PAUSE)
⑯ Volume Control (VOLUME)
⑰ Tone Control (TONE)
⑱ Built-in Microphone (MIC)
⑲ Auto Back Space Control (BACK SPACE CONTROL)
⑳ VSC (Variable Speech Control) Switch (VSC)
㉑ Tape Speed Control (TAPE SPEED)
㉒ VSC Level Control (VSC LEVEL)
㉓ Monitor Jack (MONITOR (8 Ω))
㉔ External Microphone Jack (MIC)
㉕ Remote Jack (REMOTE)
㉖ Foot Controller Jack (REMOTE/REVIEW)
㉗ AC Power Cord
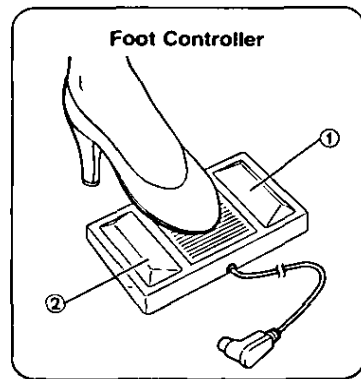
Foot Controller

① Back Space Pedal (BACK SPACE)
② Start Pedal (START)

35

5. Naming and saving cases

Each case needs to be transcribed into a separate Microsoft Word file. Open Microsoft Word by clicking on the appropriate icon located in your NAL Window. The Microsoft Word icon should look like the example below.

Microsoft
Word 2000

Once you finish transcribing the tape recording in Microsoft Word, you will need to save each transcription to your **"Completed Transcriptions"** folder. **The full directory is** Q:\Surveys\ATUS\ Transcription Project at TTC\Interviewer Access\Transcriber First and Last Name\Completed Transcriptions.

All transcription files should be named using the same format. ATUS should be first, followed by the Tape Number from the tape label, then an underscore and the Case ID Number. For example, if you have the very first ATUS tape it would be labeled as: ATUS1_07080674TU. Microsoft word will automatically add the .doc at the end of the file name.

Follow these steps to save the completed Microsoft Word transcription to the correct folder:

1. On the **File** menu, click **Save As**.
2. In the **File name** box, enter the file name to save the file as a Microsoft Word document.

Be sure to save each Microsoft Word case file separately. Supervisors will move your completed files to Q:\Surveys\ATUS\ATUS Taping\Transcription Project at TTC\Transcribed Cases.

6. Transcription

Overview

During the transcription process you will record exactly what the interviewer and respondent say and separate their talk into "utterances." An utterance is a turn-of-talk by one speaker in the interview. It begins when a speaker starts talking and ends when the utterance naturally concludes (or is interrupted). At the end of an utterance, another speaker begins a turn-of-talk.

A. Transcript formatting

Use the following formatting rules when transcribing a case. We include a formatted Microsoft word file template separately.

1. Use Microsoft Word
2. Arial 10 point face-font
3. One inch top, bottom, right, and left margins
4. All text begins on the left-hand margin (no indent)
5. Entire document is left justified
6. Use one hard return after you have transcribed the utterances of one speaker and before you begin transcription for the next speaker. When a speaker talks for some time without interruption, sentences will wrap onto multiple lines.
7. Numbers can be typed using the numerical symbol (2) or can be spelled out (two), as you wish
8. There is no need to capitalize text except I and R, but you can if you wish
9. Do not use quotation marks
10. There is no need to add periods at the conclusion of relatively short utterances (less than 10 words or so). If a speaker's utterance seems to be a complete sentence add a period at the end.
11. Add punctuation that seems to be part of the speaker's natural rhythm of conversation and/or that would help a reader of the transcript to interpret the meaning of the statement. Do your best to use punctuation in a way that represents the meaning of the respondent's words and communicates that meaning to the reader of the transcript

B. Microsoft Word transcription page heading

Each transcription should have a heading that follows the same format. An example is provided in Appendix 1 and can be used as a template (or use the separate template we have provided. At the top of the page in the left corner, the tape number should be on the first line, the case identification number should be on the second line, your transcriber id number should be on the third line, and the date you finish transcribing the case should be on the fourth line. On the fifth line, we need to have the final counter time from the recorder. The counter is three digits and is located on the right side of the transcriber above the tape. Before you start listening to a tape, be sure and set the counter to 000, by using the reset button on the right next to the counter. When you include the counter end time on line 5, please include all three digits, including leading zeros (e.g., 021). If you, as a transcriber, want to write any notes on the transcript for the researchers, do so at the end of the heading section. We're not sure if notes will be necessary, but wanted to provide the option, as there may be information you think is important for us to know.

Heading Example:

Tape Number:　　　　1
Case Number:　107080674TU
Transcriber Number:　1
Date Finished:　12/15/09
Counter end:　　　　021
Transcriber Notes:

C. Transcription begin and end points

<u>Begin</u> transcribing at the beginning of the tape. Usually the first conversation will be the interviewer asking the respondent to agree to the taping.

<u>Stop</u> transcribing at the beginning of the paid work question, which is the question after the end of the 24-hour diary. Sometimes, an interviewer will actually say: " this concludes the diary portion of the interview and now I have some other questions to ask. At others times, there might be some other conversation relating to the diary before the paid work question. We want you to transcribe everything up to the point of the paid work question. Depending on the respondent, the paid work question will be one of these two questions:

"Sometimes people do things that bring in money like selling crafts or babysitting. Were there any activities that you did yesterday that you were paid for or will be paid for?"

<u>Or:</u>

"We are interested in measuring the amount of time people spend working both inside and outside their usual work place. You said you were working from XX to XX. Were there any other activities that were done as part of your job?"

D. Identifying speakers

The respondent should be identified as R, and the interviewer as I. Although we are unaware of tapes with other speakers on the telephone other than the respondent and interviewer, if you come across such cases, please identify the speaker as O for "other." If there are multiple other speakers, identify as O1, O2, etc.

After identifying the speaker, add a colon followed by a single space and then begin the transcription. Use a hard return (enter) after the speaker finishes his or her utterances.

Example:

I: what did you do next?

R: i walked my daughter to school

I: how long did that take?

R: oh about 15 minutes

E. Verbatim transcription

The tapes should be transcribed verbatim. That means that <u>anything</u> you hear on the tape (what the respondent and the interviewer said) needs to be transcribed. The exceptions to this are background noises discussed later in this document.

It is important that you transcribe word for word everything exactly as it is said on the tape even if it does not make sense to you. You'll be transcribing everything that is said, including:

- Unfinished questions or statements,
- False sentence starts,
- Repeated phrases, words, statements or questions, and
- Conversation that seems unrelated to the diary interview.

Under no circumstances should language or terms be changed. Grammar should not be improved or altered in any way. Transcribe exactly what is said, even if it does not seem logical or make sense to you.

Do not correct the respondent's or interviewer's grammar or word choice. Transcribe words and grammar exactly as they are said. For example, do not change words like 'ain't' to 'am not' and write 'bout' or 'goin' instead of 'about' or 'going' if that is how the respondent or interviewer said the word.

F. Difficult to hear conversation

There may be several reasons you cannot hear or understand parts of the recording. The sound quality may occasionally drop due to problems with the equipment, the respondent may mumble or talk in soft tones at times, and background noise may interfere. When you cannot hear or understand a word or phrase well enough to transcribe it, use the word inaudible in brackets: [inaudible] (see Example 1). If you have difficulty hearing more than just a few words, use: [inaudible section] (see Example 2).

Example 1:

I: and what did you do next?

R: i watched [inaudible] on the tv

Example 2:

I: and what did you do next?

R: on thursdays I usually [inaudible section] and exercise

G. Interruptions/Speech overlap

Where the interviewer and respondent interrupt or overlap with one another, designate the overlapped words with asterisks (see Example 1). In some interruptions/overlaps, it may be difficult to understand what is being said (see Example 2).

Example 1: the respondent said 'about 3 o'clock' at the same time interviewer was saying 'was that'

I: what time *was that*?

R: *about 3 o'clock*

Example 2: the respondent said 'about 3 o'clock while the interviewer was speaking, but transcriber couldn't determine what interviewer said

I: what time *[inaudible]*?

R: *about 3 o'clock*

H. Pauses

For this study, a pause is any silence that lasts more than 5 seconds –Pauses should be marked as follows: [pause]. To determine if the pause is more than 5 seconds, use your watch or count the length by repeating to yourself: 1001, 1002, 1003, 1004, 1005.

Sometimes pauses can be longer. Anything longer than 20 seconds should be noted with [long pause]. A respondent might pause when thinking about a question. Also, an interviewer might pause, e.g., if he or she has to correct entries in or runs into difficulties with the CATI instrument

I. Questions/question marks

A question is a word or phrase that could get an answer, regardless of whether it actually does. Questions often involve the speaker raising his or her voice at the end of an utterance. If it is clear that the speaker is asking a question, include a question mark at the sentence.

J. False starts

A false start is when a speaker starts to say something and then makes another statement, basically interrupting or correcting themselves. Indicate a false start with two hyphens.

Example:

I: and what time did you leave for work?

R: i left at--no, i didn't leave then.

K. Fillers

Non-verbal cues or fillers are sounds that the respondent or interviewer make that are not words but are part of conversation, such is Uh or Um. It is important to transcribe all of a person's 'um's' or 'ah's'. We ask you to include them on the transcript as you would actual words.

Some common examples of fillers are as follows:

Uh, Um
Mmm, Hmm
Ba
Sigh
Gasp
Ah, Aw
Tsk
Oh, Uh-oh
Ooo
Huh

Do your best to represent these sounds by spelling what you hear. That is, there is no prescribed "correct spelling" for these sounds.

L. Other non-verbal cues

Respondents or interviewers may sometimes use non-verbal expressions of emotion that are directly relevant to the interview but are not expressed in words, such as laughter, crying, and using angry tones. Note these non-verbal cues in [ ] brackets. For example, [laughing], [crying], [angry tone].

Coughing, burping, audible breathing, or other similar types of sounds do not need to be noted or included in the transcription.

M. Background noise

You do not need to identify or include background noise (such as typing, music, tv, sirens, children playing or crying, others talking, door slamming, dog barking).

N. Transcription of time

Transcribe times in the exact manner the respondent gives them. For example, a respondent may say: 2 pm; 2; 2 o'clock; 2:15.

7. Transcription review and personal identifers

After you complete the transcript, scan your document to ensure that the transcript reflects what you have heard. Check to see that you have the interviewer (I) and respondent (R) labeled on each line and that you have used a hard return between each turn of talk.

Confidentiality is very important. After you finish transcribing a tape, go back through the typed transcription and replace all proper names of people or locations with codes such as spouse, daughter, school, business, city, state or address. There should be no person names, locations, or address information on the final typed transcription. PERSONALLY IDENTIFIABLE INFORMATION (PII) SHOULD NOT BE RECORDED ON THE FINAL TYPED TRANSCRIPTION.

If there is more than one person, school, business, city, state, address, etc. mentioned in the interview, code each different name or location with a different number (example: daughter1, daughter2, city1, city2, etc). Be sure to use the same code for the same name or location. (For example, if the respondent's daughter's names are Sally and Edna, code Sally as daughter1, and Edna as daughter2. Each time the respondent says Sally, it should be replaced with daughter1, and each time the respondent says Edna, it should be replaced with daughter2.)

8. Difficulties with tapes

On some tapes, the sound quality may be poor or the interviewer or respondent may be difficult to understand. Try rewinding and listening again to difficult parts in order to decipher what is being said. If this does not work, try adjusting the volume on the transcriber as well as the tone and/or speed of the tape. Sometimes another transcriber may be able to decipher words or phrases that are difficult for you to hear. You may ask another ATUS transcriber or supervisor to help you understand what is being said. In addition, a supervisor may reassign a tape, if one person has an easier time hearing it than another. While all of the tapes have been reviewed by ATUS interviewers and supervisors, there may be tapes where the sound is so poor that the tape needs to be removed from the study. If you think a tape is overall of such poor sound quality that you cannot transcribe it accurately, please contact the supervisor. We provide instructions later in the document on how to transcribe specific sections of text that are difficult to hear.

9. Questions

When you have questions, please let your supervisor know, and they will relay the questions back to Headquarters and then to BLS. We will respond promptly.

Thank you for your hard work. We appreciate it very much.

Attachment 3: Coding Scheme

Reliable Interviewer Codes

Variable	Description	Codes	Kappa
Activity Number (AN)		The activity number associated with the behavior. Refers to the first activity and/or last time reported in a line/turn	Kappa not calculated
Adequacy (A)	Adequacy of interviewer turn, based on conversational norms	Adequate: Interviewer asks question or makes statement or provides an answer in a conversationally appropriate manner, and in a way that conveys the intended meaning.	0.75
		(xx) No substantive utterance (including acknowledgements, inaudible, and partial questions and answers that are interrupted) – not a Question, Answer, or Feedback	0.80
Retrieval Questions (RQ)	Interviewer is asking a question to collect new activity or time information	Activity Sequential forward: Interviewer asks about the next activity.	.89
		Activity decomposition: Interviewer uses a major activity to set boundaries, then probes about what happens in between.	.67
		Time gaps: Interviewer uses gaps in time diary to probe for another activity.	.50
		Duration: Interviewer asks how long or how much time an activity occurred.	.89
		Start/end: Interviewer asks for the beginning or ending time of an activity.	.72
		Scripted: First ATUS question: Interviewer reads standard first ATUS question.	.90

Anchor (AC)	Interviewer explicitly states information previously discussed to a retrieval question	Activity Anchor: Previous activity: interviewer references a previously mentioned previous activity when asking for new information.	.81
		Activity Anchor: Start/end time: interviewer references a previously mentioned start or end time when asking for new information.	.89
		Activity Anchor: Duration: interviewer references a previously mentioned duration when asking for new information.	.57
		Time Anchor: Current/specified activity: interviewer references the specified activity when asking for new time information.	.88
Probe Domain (PD)	Topic of activity that prompted the probe.	Sleep	.51
		Personal Care	.40
		Education	.86
		Travel	.72
		Household	.50
		Work	1.00
		Child/Adult care	.80
		Leisure (including reading/writing, computer/email, exercise/sports)	.82
		Simultaneous Activity	.66
		Missed Eating/Drinking	1.00

Follow up Probe (FP)	Interviewer asks for more information related to one of the probe domains (or an instrument scripted probe, below) after information provided by R.	<u>Instrument scripted:</u> Interviewer seeks clarification as scripted on CAPI screen. (Two-hour activity, sleeping more than ten hours, working)	.81
		<u>Training</u>: Interviewer seeks clarification as instructed in training/manual.	.78
		<u>Unnecessary probe</u> INT asks a probe that is not required	.42
Missed Probe (MP)	Interviewer did not ask one of the probes as instructed in training.	<u>Missed travel probes</u>: INT did not ask travel probes as instructed in training. For example, the INT is first supposed to ask if the R made any stops along the way, and how they got from location A to location B.	.56
Follow-up Probe Verbatim (FP-V)	Whether the follow-up probe was asked verbatim or not	<u>Verbatim</u> (probe has no intervening inserts, all relevant concepts/main words, without substantive changes)	.50
		<u>Incomplete</u>	.57
		<u>Conceptually Complete But in Own Words</u>	.72
Follow-up Question (FQ)	Interviewer asks for more information (unrelated to probe domains) after activity or timing information provided by R.	<u>Request for clarification</u>: Interviewer seeks clarification about something the respondent has said, either collecting new information required for ATUS or to clarify existing information	.54
		<u>Request for repetition:</u> Interviewer asks the respondent to repeat what they just said.	.50

Confirmation or Feedback (CF)	Interviewer includes a statement about information the R provided in a turn to obtain confirmation or provide feedback about information already provided (Codes a-d not used with RQs)	Duration: Interviewer uses the time information that has already been provided by the respondent to either confirm the accuracy of the information or to allow the interviewer to convey to the respondent what information they are currently focusing on.	.62
		Clock time: Interviewer uses the information already provided by the respondent to confirm the accuracy by providing a clock time.	.71
		Activity: Interviewer uses the reported activity that has been provided by the respondent to either confirm the accuracy of the information or to allow the interviewer to convey to the respondent what information they are currently focusing on, or to aid in the retrieval process.	.73
		Duration/Clock/Activity: Interviewer uses some combination of the reported (duration or clock time) AND activity that has been provided by the respondent to either confirm the accuracy of the information or to allow the interviewer to convey to the respondent what information they are currently focusing on, or to aid in the retrieval process.	.66
		Task related feedback: Interviewer provides information about what answers are adequate for the diary or provides clarification about the interview process.	.49
		Yesterday feedback: Interviewer provides information (in form of a question or statement) to the respondent to refocus their responses to yesterday's activities, instead of usual or typical or current day activities.	.50
		Refocus on respondent: Interviewer provides prompts (in form of a question or statement) to the respondent to refocus their responses about the respondent's activities and not others.	1.00
		Operational Feedback: Interviewer provides some feedback about problems related to the instrument or interviewer data entry (e.g., waiting time).	.70
Directive-ness	Interviewer turn	Back-filling activity : Interviewer probes in an attempt to	.50

		fill in information <u>assumed</u> to be missing from R previous report	
(D)	includes some information not provided by the respondent, or is otherwise directive	<u>Back-filling time</u>: interviewer probes in an attempt to fill in missing time.	1.00
(D) Continued		<u>Sequential focusing</u>: Interviewer selects activities in a chronological manner from list of activities reported in double/triple-barreled report. Use this code for 1st, 2nd, etc. activities that comprise the multiple-barreled reports.	.52
Rapport building/Active Listening (RAL)	Interviewer turn includes some behavior to build rapport	<u>Digression</u>: Interviewer asks a question or makes a comment that is not directly related to diary completion.	.70
		<u>Distancing</u>: Interviewer distances themselves from the diary or survey process.	1.00
		<u>Interviewer uses precode terminology for GROOMING (not conversational).</u> Interviewer rephrases the respondent's activities using ATUS precode terminology for grooming. This can accompany other codes (e.g., activity probes).	.84
Interruption (I)	The interviewer turn includes an interruption	<u>Interrupted:</u> the interviewer was interrupted by the respondent	.95
		<u>Interrupting:</u> the interviewer interrupted the respondent	.92
Pause (P)	The interviewer pauses during their turn	Short – indication of a short pause from transcripts (i.e., 5 – 20 seconds)	.95
		Long – indication of a long pause from transcripts (i.e., GT 20 seconds)	1.00
Nonverbals (NV)	The interviewer turn includes some nonverbal behavior	<u>Laughter</u>: indication of laughter on transcript	.92

Inaudible (IA)	The interviewer turn includes some inaudible comment	Inaudible present	.97

Interviewer codes not used due to low reliability (Kappa < .40) or because codes were so rarely used that they were not picked up in the sample for reliability coding.

Code	Description	Notes	Kappa
Adequacy (A)	Adequacy of interviewer turn, based on conversational norms	Inadequate: Interviewer asks question or makes statement in a way that violates conversational norms	.15
Retrieval Questions (RQ)	Interviewer is asking a question to collect new activity or time information	Activity Working backwards: Interviewer asks about previous activities, to fill a gap in respondent reports.	.00
		Activity continuity: Interviewer asks whether an activity continued during a specified period.	Not in sample
		Activity visualization: Interviewer asks respondent to picture what he or she was doing either right before or after a "gap" to help fill in the gap.	Not in sample
		Next activity you remember: interviewer asks about the next activity the respondent remembers, not necessary a sequential move forward.	Not in sample
		Timing Decomposition: Interviewer uses a block of time to set boundaries, then probes about what happens in between.	Not in sample
		Mixed start/end and duration: Interviewer asks duration and start/end time.	.00
Anchor (AC)	Interviewer explicitly states information previously discussed to a	Activity Anchor: Future activity: interviewer references a previously mentioned previous future activity when asking for new information.	.00
		Time Anchor: Start/end time: interviewer references a previously mentioned start or end time when asking for	.00

	retrieval question	new time information.	
		Time Anchor: Duration: interviewer references a previously mentioned duration when asking for new time information.	.00
Probe Domain (PD)	Topic of activity that prompted the probe.	Talking	.00
Missed Probe (MP)	Interviewer did not ask one of the probes as instructed in training.	Missed '30-minute' probe	.00
		Other missed probe: INT did not ask other probe(s) as instructed in training or in the instrument, excluding missed '30-minute' and travel probes.	.09
Follow-up Question (FQ)	Interviewer asks for more information (unrelated to probe domains) after activity or timing information provided by R.	Best Guess: Interviewer **explicitly** asks R for their best guess or an estimate, OR otherwise indicates that a best guess is acceptable (a question is not necessary).	.20
Directive-ness (D)	Interviewer turn includes some information not provided by the respondent, or is otherwise directive	Skipping: Interviewer skips over activities (e.g., in a block of time reported by R)	.31
		Focusing: Interviewer selects one time from a list the respondent provided.	.00
		Leading: Interviewer includes information not previously given by respondent or discussed in interview. If a turn is both leading and backfilling, code as leading.	.34
Rapport building/Active Listening (RAL)	Interviewer turn includes some behavior to build rapport	Task related encouragement: Interviewer offers encouragement in response to respondent's efforts in an attempt to satisfy diary activity completion.	.20
		Presence of active listening: Interviewer gives feedback or makes other statement related to the respondents	.00

		preceding comment that shows they are actively listening. This DOES NOT include anchoring or confirmation/feedback.	
		Absence of active listening. Interviewer does not use active listening and asks a question that the respondent has already provided a response for.	.13
Qualified (Q)	Interviewer includes a qualification in their turn	About/approximately Interviewer includes a qualification in a question or probe by saying "around, about, approximately" etc.	.39
		Do you think/Do you estimate/Do you guess - Interviewer includes a qualification in a question or probe by saying "do you think…," "do you estimate…," or 'do you guess…"	.20
Interruption (I)	The interviewer turn includes an interruption	Overlap: there was complete overlap between the interviewer and respondent	.00
Nonverbals (NV)	The interviewer turn includes some nonverbal behavior	Crying: indication of crying on transcript	Not in sample
		Angry: indication of anger on transcript	Not in sample
		Other: add new code as other nonverbal appear in transcript	Not in sample

Reliable Respondent Codes

Variable	Description	Codes	Kappa
Activity Number (AN)		The activity number associated with the behavior. Refers to the first activity and/or last time reported in a line/turn	Kappa not calculated
Detailed Activity/Time Report (DAT)		Detailed or verbatim description of activity provided by respondent	Kappa not calculated
Adequacy (AD)	Adequacy of respondent turn, based on conversational norms	<u>Adequate substantive response</u> – Respondent reports a new activity, duration, or time that is appropriately detailed and codable, or information related to activity or time that enables the interviewer to code the response. Or the respondent asks a relevant question of the interviewer.	.64
		<u>Adequate confirmation or negation, or other adequate, task-related response that is not a new activity, duration or time</u> – Respondent provides feedback that is not providing any new activity or time information, but statement is related to the diary task	.46
		<u>Generically Inadequate</u>: Respondent provides a report that is vague, conversationally inadequate, or does not sufficiently answer the question combined with <u> ATUS inadequate</u> – Respondent provides a conversationally 'reasonable' report that nonetheless cannot be coded because some additional clarifying information is needed. This code will often be used with activities that are related to leisure, sleep and travel, if they require follow-up probes to clarify intent of activity. (See ATUS Probing Chart.)	.60
		<u>Double-barreled activity response</u>: Respondent provides 2 or more activities	.54

Adequacy (AD) *continued*	Adequacy of respondent turn, based on conversational norms *continued*	<u>Double-barreled time response</u>: Respondent provides 2 or more times	.80
		<u>Don't Know Response</u> Respondent indicates that he/she does not know the answer to the question, or states that he/she cannot give interviewer the requested answer	.76
		<u>ATUS inadequate and Double-barreled</u>: Respondent reports two or more activities and at least one of the activities is ATUS inadequate	.58
		<u>xx:</u> No substantive utterance (including acknowledgements, inaudible, partial questions, answers that are interrupted)–not a question, answer, or feedback	.60
(AD) Continued		<u>yy:</u> Respondent is conversing with someone other than the interviewer	.60
Recall strategies (RS)	Respondent provides information about their recall process	<u>Temporal anchor</u> – Respondent uses a temporal anchor as basis for time report.	.51
		<u>Think aloud reports</u> – Respondent provides a think-aloud report tied to an activity or time that *talks about the response process* without mentioning an activity- or time-based strategy.	.64
Time Reports (TD)	How respondent reports time	<u>Duration</u> – Respondent reports a duration (e.g., "30 minutes")	.82
		<u>Time</u> – Respondent provides a timing response (e.g., "… until 1:30"), or a reference to some previously reported, time-stamped activity	.67

Sequentiality (S)	Whether the reported activity moved directly forward or backwards in time or jumped to another time in the day	Sequential Forward response – Respondent provides response that moves immediately forward in time from current place in time diary	.45
		Sequential Backward response – Respondent provides a response that moves immediately backward in time from current place in time diary.	.44
		Non-sequential Backward response – Respondent reports about an activity that occurs earlier in the time diary that is out of temporal sequence and for a time period that had not previously been defined.	.41
Spontaneous digression (SD)	The respondent asks a question or makes a comment that is not directly related to diary completion.	Digression (not task related) – Respondent spontaneously provides comment/information that is not related to the diary task combined with Digression (task related) – Respondent provides a comment that is related to the diary task but is not directly responsive to the preceding question. This is NOT a straight elaboration of a preceding report.	.52
Satisficing (RSTD)	Comments or other indications that pertain to respondent level of effort	Indication of satisficing – Respondent provides a comment that indicates they are not giving their full effort	.57
		Correction of Interviewer: Respondent corrects feedback or a probe given by the interviewer. (Conditional upon interviewer feedback)	.50
Indication of task difficulty (RSTD)	Respondent comment or indication that question or survey process is difficult	Request for clarification – Respondent indicates that more information is needed about the study or to answer a question. This includes requests that a question be repeated	.86

Explicit reactions to survey/task (RSTD)	Respondent makes a comment or other reference to the specific interview question or to the overall survey task, or their ability to complete the task.	Neutral reaction	1.00
		Negative reaction	.50
Qualified (RSTD)	Respondent includes a qualification/	Qualified Activity Response: Respondent uses a phrase or expression of approximation that causes an activity to be qualified in some way, or made to seem less than certain. Uses phrases such as "about," "I think," "I guess," or "I don't know" combined with an adequate activity response.	.51
	estimation in their turn	Qualified Time Response: Respondent uses a phrase or expression of approximation that causes a time to be qualified in some way, or made to seem less certain. Uses phrases such as "about," "I think," "I guess," or "I don't know" combined with an adequate time response.	.76
Indicator of Respondent Learning (RL)	The respondent shows some sign that they have learned what the survey requires by spontaneously providing information	Spontaneous activity response: Respondent spontaneously reports a reasonable (but not necessarily ATUS adequate) activity without any probing from the interviewer	.46
		Spontaneous temporal response: Respondent spontaneously reports a reasonable time or duration for an activity without any probing from the interviewer	.77
Interruption (I)	The respondent turn includes an interruption	Respondent was interrupted by the interviewer	.81
		Respondent interrupted the interviewer	.91
		Completely overlapping speech	.40

Variable	Description	Codes	Kappa
Nonverbals (NV)	The respondent turn includes some nonverbal behavior	<u>Laughter</u>: R laughter	.84
Inaudible (IA)	The respondent turn includes some inaudible speech	<u>Inaudible present</u>	.97

Respondent codes not used due to low reliability (Kappa < .40) or because codes were so rarely used that they were not picked up in the sample for reliability coding.

Variable	Description	Codes	Kappa
Adequacy (Ad)	Adequacy of interviewer turn, based on conversational norms	<u>Generically inadequate + Double-barreled response</u> – Respondent reports two or more activities *and* at least one of the activities is generically inadequate, or two or more times plus one or more inadequate activity.	.00
		<u>Generically inadequate + Don't Know response</u>	Not in sample
		<u>Generically inadequate + Double-barreled (activity or time) + Don't Know response</u>	.00
		<u>ATUS inadequate + Don't Know response</u>	Not in sample
		<u>ATUS inadequate + Double-barreled (activity or time) + Don't Know response</u>	Not in sample

Recall strategies (RS)	Respondent provides information about their recall process	Atypical activity – Respondent mentions that the reported activity isn't what they normally do	.00
		Typical activity – Respondent describes what they typically do rather than reporting a specific event. Includes when respondent uses an anchor associated with typical activity. This can be adequate or inadequate.	.33
		Activity anchors – Respondent uses an activity anchor point to base next activity on	.25
		Visualization – Respondent describes picturing their activities	Not in sample
		Decomposition – Respondent uses a major activity to set boundaries, then recalls what happened in between	Not in sample
		Atypical time – Respondent mentions that the reported time isn't typical for their day	Not in sample
		Typical time – Respondent describes the typical time rather than reporting a specific memory	.28
		Atypical duration – Respondent mentions that the reported duration isn't typical for their day	Not in sample
		Typical duration – Respondent describes the typical duration rather than reporting a specific memory	.00
		Decomposition – Respondent uses a larger block of time to set time boundaries, then recalls how long the interim activities took	Not in sample
Time Reports (TD)	How respondent reports time	Dual/Mixed – Respondent provides both start/end time plus a duration	.34

Sequentiality (S)	Whether the reported activity moved directly forward or backwards or jumped to another time in the day	Non-sequential Forward response – Respondent reports about an activity that occurs later in the time diary that is out of temporal sequence and for a time period that had not previously been defined.	.24
Privacy Concerns (RSTD)		Privacy Concerns - Respondent gives some comment or indication that they have privacy concerns	Not in sample
Satisficing (RSTD)	Comments or other indications that pertain to respondent level of effort	Self-correction – Respondent spontaneously corrects an earlier substantive response (sign of effort).	.28
Explicit reactions to survey/task (RSTD)	Respondent makes a comment or other reference to the specific interview question or to the overall survey task, or their ability to complete the task.	Positive reaction	Not in sample

Indicator of Respondent Learning (RL)	The respondent shows some sign that they have learned what the survey requires by spontaneously providing information	Spontaneously provides any combination of activity or time	.25
		Respondent uses pre-code terminology**:** Respondent uses pre-code terminology used by the interviewers.	Not in sample
		Spontaneous with whom response**:** Respondent spontaneously provides reasonable information about whom they were with during an activity without being asked	.39
		Spontaneous where response**:** Respondent spontaneously provides reasonable information about where they were during an activity without being asked	.00
		Spontaneous who + where response	Not in sample
Pause (P)	The respondent turn includes a pause	Short –short respondent pause (i.e., 5 – 20 seconds)	Not in sample
		Long –long respondent pause (i.e., GT 20 seconds)	Not in sample
Nonverbals (NV)	The respondent turn includes some nonverbal behavior	Crying: respondent crying	Not in sample
		Angry: respondent shows indication of anger	Not in sample
		Other: Other nonverbal behavior	.00