

The American Time Use Survey: Test of Coding Procedures and Lexicon

Sid Schneider, Jennifer Crafts, and David Cantor

Westat

Tina Shelley, Diane Herz, and Lisa Schwartz

Bureau of Labor Statistics

Abstract

The American Time Use Survey (ATUS) telephone interviews are scheduled to begin January 2003. Respondents will report all of their activities over a 24-hour period. Once these data are collected, coders will assign a three-tiered code to each activity, based upon a lexicon and coding rules developed by the Bureau of Labor Statistics (BLS). Our poster presents the results of a test of the ATUS coding procedures, including the lexicon, the documentation, and the Blaise software.

For the test, nine coders received four hours of training. Then, for the next four days, they coded respondents' activities. Periodically while they worked, they filled out questionnaires and rating forms and participated in focus groups. Four of the coders were individually videotaped while coding in a usability laboratory and while participating in think-aloud interviews.

Our poster presents the accuracy and productivity of the coders, the coders' improvement over time, and the sources of the coders' confusion and errors. The poster presents the coders' opinions of the lexicon, software, documentation, and coding rules, and the manner in which these opinions changed as the coders gained experience. Our poster also covers potential enhancements to the lexicon and the software that might increase coder accuracy, consistency, and efficiency.

Background

A time-use survey collects data on the amount of time people spend in various activities, such as paid work, childcare, volunteering, commuting, and socializing. About 50 countries have fielded, or will soon field, time use surveys. The Bureau of Labor Statistics is currently developing the American Time Use Survey (ATUS). One purpose of the ATUS is to value unpaid work and measure its contribution to the national economy (Stinson, 2000; Herz, Schwartz, & Shelley, 2001). Researchers will be able to use ATUS data to compare time-use patterns by such characteristics as industry, occupation, marital status, and age of children. Researchers will also be able to contrast time use by Americans with similar data from other countries.

ATUS interviewers will contact respondents and ask them to recall their previous day's activities and estimate the duration of each activity. Coders will then code each activity within these 24-hour intervals into aggregate categories of time-use by assigning a six-digit code.

The ATUS coding system or "lexicon" has a three-tier structure. The first tier consists of 19 major activity categories. Under each of these 19 activity categories is a second tier of subcategories. Under each second tier are the specific, third tier activities; many of these activities have examples listed with them to help coders make coding decisions.

Westat conducted a test of the ATUS coding system, to evaluate the usability of the software that the coders will use, to study the manner in which coder training could be enhanced, to evaluate the ATUS coding lexicon and coding rules, and to suggest ways in which the coding process could be improved.

Method

Nine experienced coders with varying demographic characteristics were recruited for this test, which had the following components:

- All nine coders participated in a one-half day training session presented by BLS staff.
- The coders then coded activities over a period of four days. They worked with some activities that were collected in actual test time use surveys and some activities which were invented for this evaluation. The coders completed a questionnaire after each interview to indicate the activities that were most difficult to code and the steps they took to make their coding decisions.
- Four coders were individually videotaped while coding in Westat's usability laboratory. They also participated in "think-aloud" interviews, articulating their thoughts while they coded.

The coders filled out questionnaires about their attitudes about the lexicon and the software at three points: after training, after 4 hours of coding, and after the four days of coding. The coders'

attitudes and opinions were also assessed in a debriefing discussion after two days of coding, and in two focus groups after the four days of coding.

Results

The main findings of this evaluation were as follows:

- Coding accuracy and speed varied considerably among the coders. On average, the nine coders were correct on all three tiers for 70 percent of the activities that they coded.
- Coders who were relatively fast tended also to be relatively accurate. Coders who were relatively slow tended also to be relatively inaccurate.
- Both accuracy and speed tended to increase as coders gained experience.
- Coders were most accurate when coding commonly-occurring activities, such as those in the Tier 1 categories of Personal care; Eating and drinking; Socializing, relaxing, and leisure; and Household activities.
- Coders were least accurate when coding activities within the Tier 1 category of Traveling and when coding infrequently-occurring activities (such as those in Tier 1 categories Arts and entertainment, Volunteer activities, and Child and adult care services [paid]).
- Coders' attitude data indicated that:
 - As the coders gained experience, they thought that the coding task became progressively easier. Their self-rated speed, accuracy, and confidence improved over time.
 - Coders had generally positive attitudes toward the overall usability of the ATUS coding software, both after training and after becoming more experienced. They thought that the software was easy to use, and allowed them to navigate easily among the tiers and to conduct searches.
 - The coders found the categories of the lexicon to be logical and comprehensible.

Conclusions and Recommendations

Screen Design. Although coders were generally satisfied with the screen layout and content, they had several suggestions:

- Making the lexicon and coding rules available on the screen on command, with the capability of conducting searches of these materials.
- Allowing experienced coders to adjust and customize features of the screen set-up (e.g., position of the panes, the text colors, and the text fonts).

Trigram Search Function. Although the coders reported that the search function was easy to use, the search too often yielded conspicuously off-target results or no results at all. In these circumstances, the software offered no guidance for alternative actions, which confused novice users.

The following changes might improve the search feature:

- Populating the database with many more examples of activities. Examples could be added by using the results from ATUS interviews conducted with hundreds of respondents, or from another country's Time Use Survey.
- Adding a thesaurus that would automatically perform searches based on synonyms.
- Training coders when and how to take the context of an activity into account when conducting trigram searches.
- Adding a feature that conducts the search automatically and alerts the coder when high-confidence hits are available.

Documentation. A dictionary accessible on screen would help coders look up word meanings as needed. Coders would also benefit from having the coding rules presented in the form of flow diagrams (e.g., series of branching questions). Such diagrams would reduce the need for coders to remember the coding rules.

Navigation. Coders found the software easily navigable, but had some suggestions:

- Providing a feature to allow coders to examine the examples associated with Tier 3 activities at any time; permitting automatic entry of the 6-digit code when the coders locate a satisfactory example.
- Making examples accessible and visible at the Tier 2 level as well as at the Tier 3 level.
- In situations where coders decide to change the Tier 1 code they assigned, blanking out any codes they've already entered for Tiers 2 and 3.

Training. The group practice exercise in the training session was very well received, and coders felt that more of that type of practice was needed. By enhancing and lengthening the training program, coder speed, accuracy, and confidence may reach higher levels. Coder comments suggested the following:

- Coders need feedback during training -- from an instructor or the software itself -- so that they can learn when they have been accurate and inaccurate, and gradually sharpen their coding skills. In this test, feedback was deliberately omitted from the training; feedback will be important, however, when coders train for the ATUS.
- Coders need to be trained in the use of standard strategies for handling common difficult-to-code activities.

- Coders need refresher training, e.g., by reviewing a DVD with video of the training session.

Lexicon. Software “coding wizards” might help guide coders as they work with difficult-to-code activities. Wizards could help with coding problematic areas such as activities related to travel, vehicles, and communication. The wizards would provide a series of branching questions leading to the accurate codes.

Reducing the Error Rate. The coding software was not designed with features specifically intended to prevent the coders from making errors. Eight common sources of coder errors were identified:

- 1) Coders sometimes attended to extraneous information in the activity descriptions that they should have ignored. To prevent these errors, interviewers should be trained to record only key information, regardless of wording that the respondents use to describe their activities. Coder training should include methods for identifying the relevant information in ambiguously worded activities.
- 2) In instances where interviewers had recorded concurrent activities as one activity, coders sometimes had difficulty identifying the primary activity. To prevent this error, coders should be trained in when and how to use the context of an activity to assign the 6-digit code.
- 3) In instances where respondents had described sequential activities as a single activity, coders sometimes attended to the wrong activity. To prevent this error, interviewers should be trained not to code sequential activities as single activities. Coders need to be trained in the methods for identifying the salient activity.
- 4) Coders had difficulty understanding and applying the rules for coding activities related to traveling. To prevent this error, coders should have a diagram or set of branching questions that lead to the accurate code. Training should focus more on traveling activities. Also, a software wizard could be developed to assist coders with travel-related activities.
- 5) Coders had difficulty when coding very trivial activities. To prevent these errors, coders could be trained not to code activities of less than 5 minutes duration. Alternatively, rules for trivial activities could be developed.
- 6) Coders were not familiar with some of the words and phrases interviewers used to record activities. To prevent these errors, the coders could be provided dictionaries -- both on-screen and hard copy. Also, interviewers should be trained to avoid slang.

- 7) In instances where the verb used to describe an activity was inconsistent with the purpose of that activity (e.g., "talked to my children" was usually coded as child care, not socialization), coders sometimes made incorrect coding choices. To avoid these errors, coder training should focus on identifying sequences of activities that have a specific purpose, and to code such sets according to their purpose.
- 8) Coders sometimes found several codes that seemed to apply equally well to an activity. To avoid these errors, training should cover these situations. The database should contain examples so that the trigram search yields unequivocal results.

The results also suggested that coder accuracy could be improved with these steps:

- Reduce the number of categories in all three tiers. Merge little-used categories into other categories, unless doing so would compromise the goals of the ATUS.
- Select only the most accurate coders. Recruit and train more coders than needed, then select the most accurate and speedy coders.
- Code each interview more than once, using different coders, and then compare the results to identify errors for correction. Have experts resolve any discrepancies. This technique is often used by social scientists when studying subjective judgments.

References

Herz, D.E., Schwartz, L. K. & Shelley, K. J. (2001). *Coding Activities in the American Time Use Survey*. U.S. Department of Labor, Bureau of Labor Statistics.

Stinson, L. (2000). *Fourth Digit Reporting: A Data Collection Problem; Not a Coding Scheme Problem*. Internal memorandum, the Bureau of Labor Statistics July 12, 2000.