

# Is a User-Friendly Diary More Effective? Findings from a Field Test.

Eric Figueroa, Jeanette Davis, Sally Reyes-Morales, Nhien To, Lucilla Tan

U.S. Bureau of Labor Statistics

2 Massachusetts Ave., N.E. Rm. 3985, Washington, DC 20212

**KEY WORDS:** Consumer Expenditure Diary Survey, field test, form design

**Any opinions expressed in this paper are those of the authors, and do not constitute policy of the Bureau of Labor Statistics.**

## I. Introduction

Diary surveys are often used to collect information on daily activities such as consumer spending. They are particularly useful for collecting daily records of small frequently purchased items, which are normally difficult to recall<sup>1</sup>. The Consumer Expenditure Survey (CE) conducted by the U.S. Bureau of Labor Statistics (BLS) and U.S. Bureau of Census (BOC) uses a diary survey to collect data on weekly household expenditures.

Recent efforts to improve the performance of the CE Diary survey have focused on designing a more user-friendly form. Several prototype diaries were developed and refined using feedback from survey respondents, field interviewers, and program staff. Based on this feedback, CE management selected one of the designs (the Redesigned Diary) for field testing. This diary was intended to stem declining response rates and improve data quality by reducing respondent burden associated with the Diary form currently used in production (the Current Diary). The Redesigned Diary is smaller and shorter than the Current Diary, has a simpler organization, and highlights important instructions and examples.

The Redesigned Diary was tested in the field during the last four months of 2002. The primary objective of the field test was to compare response rates and data quality obtained using the Redesigned Diary with those obtained using the Current Diary. The results showed no significant difference between Diary forms in completion response rates, and only few significant differences in expenditure means and allocation rates. Allocation rates measure the proportion of expenditures requiring further

processing because they are reported with insufficient detail<sup>2</sup>.

However, the Redesigned Diary performed significantly better than the Current Diary in a majority of tests for the collection of item attribute information needed for classification<sup>3</sup>. And when considering all differences—not just those judged statistically significant—the Redesigned Diary produced higher expenditure means and lower allocation rates in a majority of expenditure categories. Finally, the (BOC) Field Representatives who worked on the field test expressed a strong preference for the Redesigned Diary.

## II. Background

### The Consumer Expenditure Survey

The CE is a nationwide survey of households conducted by the BLS and the BOC to find out how Americans spend their money. The CE provides information needed to construct weights for the Consumer Price Index, as well as data for other economic analysis. The Diary survey, one of two survey instruments in the CE, is intended to obtain out-of-pocket expenditure data on small, frequently purchased items which are normally difficult to recall. These items include detailed expenditures for food and beverages, both at home and away from home; housekeeping supplies and services; nonprescription drugs; and personal care products and services. The Diary survey is not limited to these types of expenditures, but rather, includes all expenses which the respondent incurs during the survey week. Business expenses, expenses incurred by family members while away from home overnight, and credit and installment plan payments are excluded.

<sup>2</sup> Although *allocations* are often used to account for item-nonresponse, in the diary this term refers to an expenditure that does not identify individual items at the required level of detail (e.g., a respondent reports “groceries \$150,” rather than the specific items purchased). This type of entry requires additional processing to assign the aggregate expenditure to target items.

<sup>3</sup> *Attribute information* is needed for item classification; percent of entries missing attribute information measures the portion of entries for which respondents did not provide the needed attribute information.

<sup>1</sup> Sudman, S., & Bradburn, N. (1982). *Asking questions*. San Francisco: Jossey Bass Publishers.

### *Definition of Consumer Unit (CU)*

A consumer unit is the unit of analysis and is defined as follows: (1) all members of a particular housing unit who are related by blood, marriage, adoption, or some other legal arrangement, such as foster children; (2) a person who is financially independent, whether sharing a household with others or living as a roomer in a private home, lodging house, or in permanent quarters in a hotel or motel, and (3) two or more unrelated persons living together who pool their income to make joint expenditure decisions. Students living in university-sponsored housing are included in the sample as separate CU's. CU membership is limited to the civilian, non-institutionalized population of the U.S.

### *Diary Survey Instruments*

Two paper-and-pencil questionnaires are currently used to collect Diary data. The first is the Record of Daily Expenses, the actual Diary form. This is designed as a self-reporting form on which respondents record a detailed description of all expenses for two consecutive 1-week periods. Data collected each week are considered independently. The diary is divided by day of purchase and by broad classifications of goods and services—a breakdown designed to aid the respondent when recording daily purchases. In current production, the major classifications are:

- Food Away from Home
- Food for Home Consumption<sup>4</sup>
- Clothing, Shoes, and Jewelry
- All Other Purchases and Expenses

Each classification is further divided into numerous sub-categories. The items reported within these categories are subsequently coded by the BOC so that BLS can aggregate individual purchases for representation in the Consumer Price Index and for presentation in statistical tables.

The second questionnaire is the Household Characteristics Questionnaire. This form is used to record information pertaining to age, sex, race, marital status, and family composition, as well as information on the work experience and earnings of each CU member. This socioeconomic information is used by BLS to classify the CU for publication of statistical tables and for economic analysis.

<sup>4</sup> Includes food and beverages purchased as gifts.

### *Redesigning the Diary Form*

The Redesigned Diary was designed to be more user-friendly than the Current Diary. The BLS and BOC began developing the Redesigned Diary in 2000. They sought to design a form that was easier to understand, less complicated to navigate, and less time-consuming to complete than the Current Diary. In addition, they wanted the new form to be more attractive and appealing. The thought was that if a more user-friendly diary could be created, this would lead to higher response rates and better quality data.

A team was chartered to create a user-friendly diary. In order to begin, the team had to gain a better understanding of respondent burden. By analyzing these burden issues, they could then design a diary that would minimize respondent burden and therefore be easier for a respondent to fill out. Some respondent burden issues that the team aimed to reduce include diary length, bulky diary size, dense instructions, and the time consuming task of having to subcategorize every expense.

After two rounds of cognitive testing of several prototype diaries, a new Diary form, with many of the user-friendly features recommended by test participants, was selected for field testing<sup>5</sup>. The recommended user-friendly features of the Redesigned Diary include:

- *Decreasing size:* The Redesigned Diary is smaller (8 ½ x 11), has fewer pages (44 pages), and is in a portrait format.
- *Simplifying the layout:* The Redesigned Diary has a simpler organization; each day is reduced to only 4 pages and broken down into five major expenditure categories with the food related expenditures on the first 2 page spread and the non-food related expenditures on the second 2 page spread. It also does not require the respondent to classify their purchases under any further subcategories. As a result, the cues are removed from the recording pages and placed on a flap, which provides a cleaner look that is less intimidating and makes the task seem easier.

<sup>5</sup> Davis, J., Stinson, L., To, N. (2002). What Does it Really Mean to Be User-Friendly When Designing an Expenditure Diary? Paper presented at the annual meeting of the American Association of Public Opinion Research.

- *Clarifying instructions, rules, and definitions:* The Redesigned Diary gives clear instructions in multiple ways.
  1. The General Instructions section is compact and concise. It clearly explains and illustrates with concrete examples.
  2. An outline provides an overview of all the major categories at the beginning of the diary so that respondents will have a mental map of what lies ahead.
  3. The Frequently Asked Questions (FAQs) answer common questions asked about the diary keeping task.
  4. The Examples of Expenses lists different types of expenses and shows the section under which the expense should be recorded.

The FAQs and Examples are printed on flaps that are easy to access. The flaps can also be used as bookmarks to help the respondents keep their place.

- *Detailing the example pages:* The Redesigned Diary increases the variety of the examples, focuses on difficult cases, and highlights important data entry instructions and examples using color, white space, bolded text, and superimposed balloons.
- *Using checkboxes to facilitate the recording task:* The Redesigned Diary allows the respondent to mark a checkbox to classify expenditures instead of requiring the respondents to subcategorize or choose a code.
- *Making the diary look current and appealing, but maintain a professional and official quality:* The Redesigned Diary uses color and photos to cue respondents and to make the diary more attractive.

### III. The 2002 Field Test

#### Sample Design

To assess the performance of the Redesigned Diary, a field test was conducted from September through December 2002. In addition to the Redesigned form, a Computer Assisted Personal Interview (CAPI) component was also tested. This was used to administer the Household Characteristics Questionnaire, replacing the paper-and-pencil version used in current production.

The field test design included both “Test” and “Control” samples. These were assigned the Redesigned Diary and the Current Diary, respectively. Both samples used the CAPI Household Characteristics Questionnaire. To create the samples, the BOC selected 1,800 households from a previously unused supplemental sample. These sample units were drawn from 9 of the 12 Census regions<sup>6</sup>. The “Test” sample of 1,200 households received the Redesigned Diary, and the “Control” sample of 600 households received the Current Diary.

As the field test proceeded, significant demographic differences were found between the Test and Control samples. The largest significant differences were found in the proportions of owners and renters. In the Test sample, these proportions were close to those found in the general population. In the Control sample, there were more renters and fewer owners than found in the general population. In addition, renters in the Control sample had significantly lower incomes than in the Test sample. Because these characteristics affect expenditure levels, the disparities weakened the Control output’s usefulness for comparisons with the Test output.

A contingency plan was agreed on prior to the test. If the Control sample was not large enough to provide meaningful estimates, a Production sample would be selected for comparison with the Test sample. The Production sample was drawn from concurrent production data restricted to the regions, metropolitan statistical areas, and sample frames used to draw the field test sample. This yielded a sample of 2,703 households.

Given problems with the Control sample’s demographics, the authors chose to follow the contingency plan and focus on comparisons between the Test and Production samples. Although the production data had been collected without the CAPI component, its demographic consistency with the Test sample was thought to make it a better comparison.

#### Measures of Effectiveness

Our research goal is to compare the effectiveness of the Redesigned Diary and the Current Diary. Our null hypothesis states they are equally effective. Our

---

<sup>6</sup> The nine Census regional offices which participated in the field test are: Atlanta, Boston, Charlotte, Chicago, Dallas, Denver, Detroit, Philadelphia, and Seattle; excluded were New York, Los Angeles, and Kansas City

alternate hypothesis states that one diary is more effective than the other.

To be judged more effective, the Redesigned Diary or the Current Diary must have:

1. *higher* completion response rates than the Current Diary. Completion response rates measure the percent of all eligible diaries successfully placed and completed<sup>7</sup>.
2. *higher* mean dollar expenditures per CU in the two food expenditure categories - *Food Away from Home*, and *Food for Home Consumption*<sup>8</sup>.

It would also be desirable if a diary produced higher mean expenditures in the two non-food expenditure categories, produced relative expenditure shares<sup>9</sup> consistent with the pattern in current production data, and had lower produced lower percentages of entries missing attribute information. However, it is sufficient for one diary to be judged more effective than the other if it met the two criteria mentioned.

The diary test sponsors were interested in other insights to the effectiveness of the diaries, in addition to the quantitative analyses of the field test data. Towards that end, several additional tasks were undertaken:

1. **Content analysis of the Redesigned and Current Diaries:** The objective of content analysis is to compare the overall 'quality' of entries in the diaries – whether entries were recorded properly and clearly, and relevant checkboxes marked. Ten percent of diaries were randomly selected for content analysis, ensuring coverage in the following 3 areas: single and multiperson CUs, diaries from Weeks 1 and 2, and diaries from all geographic regions<sup>10</sup>. A total of 47 Control Diaries and 81 Redesigned Diaries from the months of September and October were reviewed for content analysis.

<sup>7</sup> Eligible housing units are those in the designated sample, less housing vacancies, housing units under construction, housing units with temporary residents, destroyed or abandoned housing, and units converted to nonresidential use.

<sup>8</sup> Includes food and beverages purchased as gifts.

<sup>9</sup> The relative share of each of the four expenditure classifications (i.e., how much money is spent in each classification vs. total expenditures) should be consistent with other data sources. For example, a variety of sources clearly show that Americans are eating more meals away from home, so one would expect this expenditure share to be increasing.

<sup>10</sup> The geographic regions are Northeast, Midwest, South, and West.

2. **Debriefing of field representatives (FRs):** FRs who participated in the field test were given an opportunity to share their impressions and reactions. In December 2002, a debriefing questionnaire was sent to all FRs who participated in the field test. The response rate for this questionnaire was 86 percent. In January 2003, 17 FRs representing the 9 Census regional offices participated in a one-day debriefing.

### Determining Significant Differences

Statistical tests were performed to measure significant differences in the output of the Redesigned and the Production Diary. For the Redesigned Diary field test, variances were calculated using the method of "random groups."

For Test and Production samples, the universe of CUs was randomly assigned into 10 groups (called replicates), with each replicate containing approximately 10 percent of the universe. For each statistic of interest (such as mean expenditure, response rate, relative importance), the statistic was computed separately for each replicate, as well as the full sample.

Then the variance for the statistic is estimated by

$$Var(\bar{x}) = \frac{\sum_{r=1}^{10} (\bar{x}_r - \bar{x})^2}{10(10-1)}, \text{ where}$$

$\bar{x}$  = the full sample statistic of interest, and  
 $\bar{x}_r$  = the statistic for the r<sup>th</sup> replicate.

The standard error is estimated by

$$SE(\bar{x}) = \sqrt{Var(\bar{x})}.$$

To determine if the statistic of interest was significantly different between the Test ( $\bar{x}_{Test}$ ) and Production ( $\bar{x}_{Production}$ ) samples, z-scores (Z) that allow a statement of statistical significance were calculated using the following formula:

$$Z = \frac{[\bar{x}_{Test} - \bar{x}_{Production}]}{\sqrt{Var(\bar{x}_{Test}) + Var(\bar{x}_{Production})}}$$

where  $Var(\bar{x}_{Test})$  and  $Var(\bar{x}_{Production})$  are the variance of the Test and Production statistics, respectively.

If the absolute value of  $Z$  is greater than 2 then the difference between the statistics of interest is statistically significant.

#### IV. Findings

Based on comparisons between the Test and Production samples, the data showed:

*Response rates:* No significant difference in the response rate for completed diaries was found. (See Table 1). Compared to the Redesigned Diary, the refusal rate in the Current Diary was significantly higher. However, the Redesigned Diary also had a significantly higher rate of incomplete interviews for “other” reasons; this has been largely attributable to the more stringent placement dates enforced by CAPI for the Redesigned Diary.

*Expenditure Means:* In the Redesigned Diary, expenditures were significantly lower in *Food Away from Home* but significantly higher in *Clothing, Shoes, and Jewelry*. In terms of expenditure shares, only *Food Away from Home* was significantly lower in the Redesigned Diary.

*Allocation rates:* In the Redesigned Diary, the percentage of *Food Away from Home* expenditures coming from allocation was significantly lower than in the Current Diary. This may largely be a reflection of the effectiveness of the additional checkboxes in the Redesigned Diary. No other significant differences were found.

*Percent Missing Attributes:* Three of the five tests (*Meal type, Alcohol type, and Gender*) showed significantly lower results in the Redesigned Diary compared to the Current Diary. This may largely be due to the effectiveness of additional checkboxes. Only one test (*Package type*) showed significantly lower results in the Current Diary.

*Content Analyses:* Based on the Diaries manually reviewed, it was not apparent that one type of diary has a better overall quality of entries than the other diary (see Table 2).

*FR Debriefings:*

- *Survey of BOC Field Representatives (FRs) who administered the field test:* These showed overwhelming support for the Redesigned Diary. When asked to compare the two diaries on several criteria (Overall Impression, Ease of Administration, Ease of Respondent Use, Layout Design, Gets Complete Interview, Gets Accurate

Data), a majority of FRs consistently gave the Redesigned Diary favorable ratings and gave the Current Diary neutral or negative ratings.

- *In-person debriefing of 17 BOC FRs:* The majority of FRs felt that the format of the Redesigned Diary with fewer categories effectively reduced respondent burden. They believed that respondents were more likely to record in the diary and persevere with recording diary entries through the second week.

#### VI. Conclusions

The findings of the Diary Field Test data did not allow us to reject the null hypothesis. We conclude that both the Redesigned Diary and the Current Diary are equally effective. No significant difference was found in the test of completion response rates. Results were mixed for tests of mean expenditure in the two food categories: the Redesigned Diary had significantly lower expenditures than the Current Diary in *Food Away from Home*, and there was no significant difference between the diaries in *Food for Home Consumption*. Higher results on both tests were necessary for either diary to be judged more effective than the other.

The Redesigned Diary performed significantly better in a majority of tests of missing item attribute information. When considering all test differences—whether significant or not—the Redesigned Diary produced higher expenditure means and lower allocation rates in three of the four expenditure categories. In addition, the Field Representatives who worked on the field test expressed a strong preference for the Redesigned Diary.

#### VII. Further Research

The Redesigned Diary’s weaker areas merit additional research. The expenditure means in the *Food Away from Home* section were lower in the Redesigned Diary than in the Current Diary. Cognitive work is needed to determine if new titles<sup>11</sup> in the Redesigned Diary for *Food Away from Home* and *Food for Home Consumption* are confusing for respondents, possibly leading to incorrect item entry.

Additional research is also needed to develop effective cues to encourage more detailed reporting in

<sup>11</sup> In the Redesigned Diary, the *Food Away from Home* and *Food for Home Consumption* section were retitled, respectively, *Food & Drinks from Food Service Places* and *Food & Drinks from Grocery and Other Stores*.

*Food for Home Consumption, Clothing, Shoes, and Jewelry, and All Other Purchases and Expenses.* The cues should not be overwhelming or add significant amounts of respondent burden.

**Table 1: Comparison of data from the Redesigned and Current Diaries**

	Test (CAPI and Redesigne d Diary)	Production (Current Diary)	Significant difference
<b>Response rates (%)</b>			
Completed	74.5	75.2	
Eligible CUs who did not complete interview because:			
-refused	11.9	17.9	****
-not home	5.0	4.3	
- other	8.6	2.6	****
<b>Mean Expenditures(\$)</b>			
All Expenditure Categories	371	359	
Food for Home Consumption	64	64	
Food Away from Home	37	41	**
Clothing, Shoes, and Jewelry	39	33	**
All Other Purchases and Expenses	231	221	
<b>Allocation Rates (% expenditure from allocated items)</b>			
All Expenditure Categories	17.6	20.8	
Food for Home Consumption	24.3	26.3	
Food Away from Home	18.3	49.5	****
Clothing, Shoes, and Jewelry	22.2	17.5	
All Other Purchases and Expenses	15.6	16.2	
<b>Missing Attributes (% entries missing attribute information)</b>			
Package type	7.2	4.7	**
Meal type	2.8	30.3	****
Alcohol type	9.8	16.6	**
Age	17.7	21.4	
Gender	16.4	21.4	**

Source: The Consumer Expenditure Survey Redesigned Diary field test, September-December 2002

Notes: Statistical significance based on Z-score - \*\*  $2 \leq \text{abs}(Z) < 3$ , \*\*\*  $3 \leq \text{abs}(Z) < 4$ , \*\*\*\*  $\text{abs}(Z) \geq 4$

**Table 2: Content Analysis of the Redesigned and Current Diaries**

	Redesigned Diary	Current Diary
<b>Error rate of Illegibles (cannot read due to handwriting)</b>		
Food Away from Home	0.0 %	0.0 %
Food for Home Consumption	0.4 %	0.2 %
Clothing, Shoes, and Jewelry	0.0 %	0.0 %
All Other Purchases and Expenses	0.2 %	0.0 %
<b>Error rate of Unintelligibles (can read but cannot tell what it means)</b>		
Food Away from Home	0.6 %	0.0 %
Food for Home Consumption	0.9 %	5.5 %
Clothing, Shoes, and Jewelry	0.0 %	0.0 %
All Other Purchases and Expenses	0.9 %	1.8 %
<b>Error rate of Missing Description Fields</b>		
Food Away from Home	0.7 %	0.0 %
Food for Home Consumption	0.0%	0.0 %
Clothing, Shoes, and Jewelry	0.0 %	0.0 %
All Other Purchases and Expenses	0.0 %	0.0%
<b>Error rate of Missing Total Cost Fields</b>		
Food Away from Home	0.0 %	0.0 %
Food for Home Consumption	0.0 %	0.2 %
Clothing, Shoes, and Jewelry	0.0 %	0.0 %
All Other Purchases and Expenses	0.0 %	0.6 %
<b>Error rate of Missing Alcohol Check Mark (when alcohol described or cost given)</b>		
Food Away from Home	0.0 %	3.4 %

The authors would like to acknowledge the following BLS employees who contributed to this analysis: Jeff Blaha, Richard Dietz, Tammy Hagemeyer, William Mockovak, Troy Olson, Mary Lynn Schmidt, Linda Stinson, David Swanson, Clyde Tucker, Wolf Weber.