

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

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

Eric Figueroa, Nhien To, and Lucilla Tan are economists in the Branch of Research and Program Development, Division of Consumer Expenditure Surveys, Bureau of Labor Statistics.

Jeanette Davis is a senior economist in the Branch of Research and Program Development, Division of Consumer Expenditure Surveys, Bureau of Labor Statistics.

Sally Reyes-Morales is a mathematical statistician in the Division of Price Statistical Methods, Branch of Consumer Expenditure Surveys, Bureau of Labor Statistics.

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 recall¹. The Consumer Expenditure Survey (CE), sponsored by the U.S. Bureau of Labor Statistics (BLS) with data collected by the 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. Such a form would have a simpler recording scheme and be more attractive in appearance than the form currently used in production. Several prototype diaries were developed and refined using feedback from survey respondents, field interviewers, and program staff². 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 physically 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 from October through December of 2002³. The primary objective of this 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 statistically 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 detail⁴.

However, the Redesigned Diary performed statistically significantly better than the Current Diary in a majority of tests for the collection of item attribute information needed for classification⁵. In addition, the BOC Field Representatives who worked on the field test expressed a strong preference for the Redesigned Diary because of its more attractive layout and simpler recording scheme.

Based on the field test results, it was decided to continue research on the Redesigned Diary before implementing it in production. The focus of the research was to test modifications to the Redesigned Diary that would increase expenditure levels in *Food Away from Home*, and reporting detail in *Food for Home Consumption*.

Background

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

² Davis, J., Stinson, L., To, N. (2002). Creating a "User-Friendly" Expenditure Diary, *Consumer Expenditure Survey Anthology*, 2003, p 3.

³ A field test is designed to reproduce data collection conditions as close as possible to those that exist in the production environment.

⁴ *Allocation* is an adjustment performed on expenditure entries that do 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.

⁵ *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 (e.g., a respondent reports "peas" but does not provide attribute information on the package type—fresh, frozen, or canned).

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 their consumer units 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 Consumption⁶
- 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. Beginning in 2003, the Household Characteristics Questionnaire has been administered using Computer Assisted Personal Interviews (CAPIs).

Re-designing the Diary Form. The objective of the form redesign was to produce a more "user-friendly" form to encourage higher response rates and more accurate expenditure reporting. The BLS and BOC began developing the Redesigned Diary in 2000. Findings from focus groups were used to define features of a "user-friendly" form. These features are: a form that is easier to understand, less complicated to navigate, simpler to complete than the Current Diary, and looks more attractive. Through a series of cognitive tests of several prototype diaries designed with these user-friendly features, one was selected for testing in the field - the Redesigned Diary.

A summary of the differences in the features of the Current Diary and the Redesigned Diary follow:

- *Decreasing physical size.* The Redesigned Diary is smaller (8 ½" x 11"), has fewer pages (44), and is in a portrait format. In contrast, the Current Diary is 14"x 8" with 66 pages and is in landscape format.
- *Simplifying the layout.* The Redesigned Diary has a simpler organization than the Current Diary. In the Current Diary, each day's reporting space consists of 7 pages, broken down into broad classifications and numerous subcategories. In the Redesigned Diary, each day's reporting space is reduced to 4 pages. These are also broken down into broad classifications, but there are no subcategories, simplifying the diary task and the form's appearance.
- *Clarifying instructions and examples:* The Redesigned Diary's instructions were formatted so that topics would be easier to find.
 1. The Current Diary's instructions are evenly spread over two pages, divided into eight topical areas. These are distinguished by their titles, which compete with numerous subtitles. The Redesigned Diary's instructions are also contained on two pages, but the different topics are more easily distinguished from one another. The information is grouped into three topics, and these are graphically set apart from one another through the use of frames and large font title blocks.
 2. Frequently Asked Questions (FAQs) were added to the Redesigned Diary. These answer common questions asked about the diary keeping task and are found on an easily accessible flap on the Diary's back cover. Examples of expenditures are contained on a flap on the front cover. Both flaps can be used as bookmarks to help the respondents keep their place.

⁶ Includes food and beverages purchased as gifts.

3. Compared to the Current Diary, 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 more checkboxes to facilitate the recording task.* In contrast to the Current Diary, the Redesigned Diary increases the use of checkboxes. These allow respondents to more easily classify expenditures.
 - *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. The Current Diary is printed in black and green on white paper, and has no photos.

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 CAPI version of the Household Characteristics Questionnaire was tested. This replaced the paper-and-pencil version of the questionnaire used in production⁷.

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 regions⁸. 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 renters in the Test sample. Because these characteristics affect expenditure levels, the disparities weakened the Control output’s usefulness for comparisons with the Test output.

In anticipation that the Control sample would not be large enough to provide meaningful estimates, a Production sample was 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 differences in the demographics between the Test and Control samples, the authors chose to 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 alternate hypothesis states that one diary is more effective than the other.

The more effective 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 completed⁹.

⁷ After further refinement, the CAPI version was introduced into production in 2003.

⁸ 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.

⁹ *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.

2. *Higher mean dollar expenditures per CU in the two food expenditure categories - Food Away from Home, and Food for Home Consumption*¹⁰.

These two criteria were selected, respectively, because of concern over the declining response rates in the CE, and the importance of the Diary as the major source for data on food expenditures. It would also be desirable if a diary produced higher mean expenditures in the two non-food expenditure categories, produced relative expenditure shares¹¹ consistent with the pattern in current production data, and had 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.

In addition to the quantitative analyses of the field test data, two additional analyses were undertaken to evaluate the diary:

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 regions¹². A total of 47 Control Diaries and 81 Redesigned Diaries from the months of September and October were reviewed.
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.”

To obtain these random groups for statistical analysis for the 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. Each statistic of interest (such as mean expenditure, response rate, relative importance) was computed separately for each replicate, as well as for 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 rth 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:

¹⁰ Includes food and beverages purchased as gifts.

¹¹ The relative share of each of the four expenditure classifications is the percent each constitutes of total expenditures.

¹² The geographic regions are Northeast, Midwest, South, and West.

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

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.

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 may have been due to the more stringent placement dates enforced by CAPI for the Redesigned Diary.

Expenditure means. In the Redesigned Diary, expenditures were significantly lower for *Food Away from Home* but significantly higher for *Clothing, Shoes, and Jewelry*. In terms of expenditure shares, the percentage of total expenditures spent on each component, only *Food Away from Home* was significantly lower in the Redesigned Diary. These results may be due to new titles¹³ in the Redesigned Diary for *Food Away from Home* and *Food for Home Consumption*. Because of the difference in titles, respondents using the Redesigned Diary may have thought they should omit from the *Food Away from Home* section some expenditures which respondents using the Current Diary thought should be included.

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 rates of missing attributes in the Redesigned Diary compared to the Current Diary. This may largely be due to the effectiveness of additional checkboxes. One test (*Package type*) showed significantly lower results in the Current Diary, and one test (*Age*) showed no difference between the diaries.

Content analyses. Based on the Diaries manually reviewed, it was not apparent that one type of diary has consistently higher error rates than the other diary (see Table 2).

FR Debriefings:

- *Survey of BOC Field Representatives (FRs) who administered the field test.* This 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.

¹³ 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*.

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 expenditures 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.

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 the titles used in the diaries 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 *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.

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.

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

	Test (CAPI and Redesigned Diary)	Production (Current Diary)	Significant difference
Response rates (percent)			
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	****
Mean expenditures(dollars)			
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	
Allocation rates (percent of expenditures from allocated items)			
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	
Missing attributes (percent of entries missing attribute information)			
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 (in percent)	Current Diary (in percent)
Error rate of illegibles (cannot read due to handwriting)		
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 %
Error rate of unintelligibles (can read but cannot tell what it means)		
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 %
Error rate of missing description fields		
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%
Error rate of missing total cost fields		
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 %
Error rate of missing alcohol check mark (when alcohol described or cost given)		
Food away from home	0.0 %	3.4 %