

How To Estimate Questionnaire Administration Time Before Pretesting: An Interactive Spreadsheet Approach

Sandra H Berry*

Tags: survey practice

Survey Practice

Vol. 2, Issue 3, 2009

How To Estimate Questionnaire Administration Time Before Pretesting: An Interactive Spreadsheet Approach

Survey researchers are often faced with the problem of managing an instrument design process within the framework of a budgeted time constraint for administering the final instrument. While survey managers worry about costs and respondent burden, researchers want to make sure that every relevant construct and potentially important covariate is included, preferably with multi-item measures. The problem is: How can you estimate how long a draft questionnaire will take to administer without actually pretesting it? Or, if some pretesting has occurred, how can you predict the effect of additions or deletions on administration time without more pretesting? Faced with this problem for several large studies, I developed an approach to address the problems of managing the length of an instrument using a simple spreadsheet. Here's how it works.

GETTING STARTED

You can begin with an outline and/or a list of specific candidate measures or items or an actual draft instrument. If the instrument is long, it's useful to divide the draft instrument into modules and sections within modules. Modules define basic subject areas, for example, demographics, economic variables, or health related quality of life. Sections within modules define distinct subsections, for example, current sources of income within an economics module.

Once your instrument is broken down into modules and possibly sections, you will want to count how many questions are within a module or section. Do this by giving each question a raw item value. The rule we use is to count anything that could be asked of respondents as an item and to count each potential data field as a single item. Of course, this does not include checkpoints where the interviewer or computer refers to previously asked question. Using this system, the raw item count for a "circle one" question would be one (see Fig. 1,

* Institution: RAND Corporation

questions 1 and 3), while a “circle all that apply” would have a raw item count for as many items as there are possible response categories (see Fig. 1, question 2). If the question asks for an amount of time and then a unit (weeks, months, years, etc.), the raw item count is two (see Fig. 1 question 4). If the measures are well established, the item count is usually documented, for example, the health related quality of life measure called the SF-36 has an item count of 36.

What was your total income in the past year? (Income can come from salaries, wages, social security, welfare, dividends, interest, and any other income.)

TOTAL INCOME IN PAST YEAR:\$

RAW ITEMS = 1

2. What were the sources of your income in the past year?

CIRCLE ALL THAT APPLY

Salary or wages 1
 Loan(s) 2
 Investment dividends 3
 Welfare 4

RAW ITEMS = 4

3. At this time, are you:

CIRCLE ONE

Working full time 1
 Working part-time 2
 On sick leave 3
 Not working..... 4

RAW ITEMS = 1

4. ASK IF CURRENTLY WORKING: How long have you worked at your current job?

UNIT: WEEKS
 MONTHS
 YEARS

RAW ITEMS = 2

FIGURE 1 HOW TO COUNT ITEMS

To start your spreadsheet, list the module numbers in column A and a short title for each module in column B., as shown in Figure 2. Then, total the raw

item values in each section or module and put the counts in column C. This is usually quite close to the maximum number of items that could be asked of any (very unlucky) respondent although you may need to refine it as described below, depending on the structure of your instrument.

FIGURE 2

A	B	C	D
	Module Name	Raw Item Count	Shortest Path
1	Work status measures	10	1
2	Insurance plan description	4	1
3	Satisfaction with hospital stay	6	1
	TOTAL ITEMS	20	3
	EST ITEMS PER MINUTE	4	
	ESTIMATED ADMIN TIME IN MINUTES	5	0.75

Next, go through each module, paying attention to skip patterns, and take the shortest possible path through the items – the path a respondent would take who skipped as many items as possible. Enter the raw item totals for each module in column D. Individually total columns 3 and 4. You have now established the maximum (column C) and minimum (column D) number of items that could be answered.

Now that you have an idea of the number of items in the modules, you are ready to obtain a preliminary time estimate. For a CAPI survey I recommend using a rule of thumb of about 4 items asked per minute. If you divide the individual column C and D raw item totals by 4, you will have theoretical maximum and minimum times for the draft instrument. This rate works well for an interview that includes short scales with a common set of response categories as well as wordier factual recall or opinion items. If most of your items are of one kind or another, you might adjust the rate up or down, from 3 to 6 items per minute. To consider refinements to the standard items per minute rate, it is helpful to put the rate into a reference cell and use it as a “constant” in formulas. This allows you to quickly see the impact of different assumptions about the rate. A simple spreadsheet like this one is very useful for developing a ballpark estimate of administration time at an early stage. It can be computed for a partial instrument and additional modules can be added later. It can also be used to plan an instrument and allocate approximate numbers of items to each section of the instrument, though these typically need to be adjusted later.

REFINING THE ESTIMATE – “EXPECTED” NUMBER OF ITEMS

The preliminary estimate of the number of items can be refined prior to pretesting by calculating the “expected” number of items. This is the number of items that will likely be asked, taking into account the number of

respondents to whom each item will actually be administered (skip patterns) and assigning weights. This sounds difficult to do, but it can often be done with considerable accuracy. If a question is going to be asked of everyone, its weight is 1. However, many questions are only asked of a sub-group of the sample. For example, if 50% of the sample is female a series of items asked of “women only” would have a weight of 0.5 (Column E of the example below). The task is to look at each item or measure and give it a “weight” based on the proportion of respondents to whom it will be administered. You will multiply the number of items in each section by its weight (Column C * Column E below). It’s often possible to predict the probability of responding based on expectations about the sample characteristics or by making assumptions about rates for a few key parameters that can be checked from other data sources, for example, by obtaining an estimate of what fraction of the adult population in hospitalized in a year and using that as a weight for a section on satisfaction with hospital stay. It’s useful to document how you developed the weights that led to an expected number of items, since this is likely to be a key factor in explaining your estimate.

Now we can expand on our earlier example of a timing table in Figure 2. By adding up the weighted values for each item in the module, we determine the “expected” number of items in column F. Note that the expected total can be significantly lower than the raw item total. At the bottom of column F we divided the expected number of items by a rate of 4 items per minute to obtain an “expected” mean time for administration of the interview. The expected time should fall somewhere between the maximum and minimum time estimates. It’s useful to consider the maximum time as well as the expected time in thinking about acceptable respondent burden, though the expected time is most useful in relation to the budgeted administration time.

Figure 3

A	B	C	D	E	F	G
	Module Name	Raw Item Count	Shortest Path	Weight	Expected Item Count (C*E)	NOTES: note assumptions here that led you to your expected item count
1	Work status measures	10	1	0.4	4	40% of sample will be working
2	Insurance plan description	4	1	0.3	1.2	30% of sample has insurance
3	Satisfaction with hospital stay	6	1	0.2	1.2	20% of sample will have had a hospital
	TOTAL ITEMS	20	3		6.4	
	EST.ITEMS/MINUTE	4.00				
	EXPECTED ADMINISTRATION TIME IN MINUTES	5	0.75		1.6	

USING PRETEST INFORMATION TO FURTHER REFINE THE ESTIMATION

Pretest data can be used to refine the estimate, especially if the pretest sample is reasonably large and generally representative of the actual sample. Using the pretest frequencies and module-by-module time stamp data, you can obtain a count of the actual number of items asked in each module, refining the “weight” of each item to update the “expected” number of items, and the a rate of items asked per minute.

PREDICTING THE EFFECTS OF ADDITIONS AND DELETIONS

One of the most helpful uses of the spreadsheet is to model the effects of cuts or additions on the total administration time reflecting changes in the numbers of items to be asked, changes in skip patterns that affect the item weights, or the addition or removal of entire sections of a questionnaire. If you have set up your spreadsheet correctly, it will show an automatic re-calculation of the estimated administration time when you make changes. We display the spreadsheet on an overhead screen to use in meetings where changes are being discussed. It’s surprising to most people how many items must be deleted to make a meaningful change in the administration time. If the instrument administration rate is 4 items a minute, cutting 10 minutes requires eliminating 40 items that are asked of everyone, or 80 items, each of which is asked of 50 percent of the respondents. Making that point clearly, in real time, tends to focus efforts to reduce respondent burden and control costs.

HOW ACCURATE IS THIS PREDICTION?

We’ve been using this approach for about 10 years. It’s very helpful in the planning stages to give instrument designers a sense of how many questions they can afford to ask or to estimate how long it will take to administer the instrument they have planned. The approach was originally developed for CAPI surveys and the rate of about 4 items a minute worked well for that mode. We use a range of 4–6 items per minute for CATI surveys where 4 items a minute seems to work well for lower literacy or elderly populations, 5 is about right for a “general” population, and 6 is used for sections or modules that have a common stem, brief items, and a common response scale. (If the modules have very different kinds of items, it’s easy to use module specific items per minute rates by creating several reference cells and using them to make a time estimate for each module that you can sum for a total estimated administration time.) However, this is not a perfect predictor, so at the initial stages we design an instrument that’s 5–10 minutes longer than desired so we can pretest all the candidate items and then prune as needed. We are still experimenting with rates in other modes, although we expect them to be in a similar range.

CONCLUSIONS

Controlling the administration time of an instrument is the first step to a successful field effort. Using a basic spreadsheet approach, it’s possible to

develop a simple interactive model to predict the expected administration time for a draft questionnaire. Doing this at an early stage helps to guide the initial design process and sets the stage for predicting the effects of cuts or additions.

Note: I am very grateful to Shirley Nederend, Julie Brown, and other RAND SRG staff members who helped to refine this method and who provided timing information from their surveys.