ATTACHMENT J: VSP Review

Complete Text of Dr. Don Dillman's Review of Visitor Services Project Questionnaires

October 1, 2007

To: Margaret Littlejohn

cc: John Tarnai

From: Don Dillman

Subject: Review of Visitor Services Project (VSP) Surveys

The Request

It is my understanding that you would like for me to review two surveys, "The Independence National Historical Park Visitor Survey," and the "Agate Fossil Beds National Monument Visitor Study," which are representative of the kinds of surveys now being done.

I understand that the general concern that I should contemplate is the extent to which these surveys are burdening respondents in an excessive way. Some of the factors to be considered are:

1   Length.
2   Ease of understanding by respondent.
3   Booklet layout, e.g. white space, navigation through the booklet, use of maps, skip instructions, etc.
4   Question formatting including level of complexity, ordering of questions, and use of multi-part questions.
5   Instructions, both for the survey as a whole and for individual questions.
6   Can these surveys be completed in 20 minutes?

Some Relevant Background

I think you are aware that I have reviewed different aspects of the VSP surveys and the procedures involved in their use previously. The first effort was with Dana Dolsen and Gary Machlis. We wrote a paper that was published in the Journal of Official Statistics (1995) that described the procedures used for implementing these surveys since their beginning in 1988. I was impressed then with the unusual delivery mechanism and extraordinarily high response rates which the surveys obtained.

In about 1999, I was asked by Dr. Machlis to analyze factors affecting response rates to the surveys done prior to 2000, which I did and reported to him in a technical report authored by myself and Lisa Carley-Baxter (2001). Thus, I have some familiarity with the history of the surveys.

After this request was received, I asked Margaret Littlejohn for updated information on the performance of surveys in recent years, and have also used it to form some observations about overall performance between the 1988 instigation of the visitor surveys and their use through 2006.

The questions that I have been asked to address in this memorandum can be informed by results of these various analyses. I will refer to them in appropriate places in the discussion that follows.

Effects of Increased Length on Response Rates

In 1988 the four initial surveys achieved a response rate of 86%, which immediately dropped to 78% in 1989, stayed near that amount through 1997, and 74% for the last10 years. This stability in response rates, which is shown in Figure 1, is remarkable. Excluding the initial year of surveying (1988) that is a drop of only four percentage points in an 18 year period when response rates on many surveys, and in particular telephone surveys, have been dropping by a much greater amount.

It is also apparent that the number of questions being asked has increased significantly in this 18 year period. However, it is also significant that the number of pages, another indicator of length, has increased by only one sheet of paper or four additional questionnaire pages. Considerable research suggests that greater length of surveys is a factor causing response rates to decline, but is less clear on whether people react to number of questions or to number of pages. It is my understanding that before 1992 all questionnaires were 12 pages long (3 sheets of paper folded) and after 1999 all questionnaires were 16 pages, with some mixing in between. Thus, it is important to recognize that the VSP has been able to maintain response rates fairly well, despite the change in this measure of questionnaire length (Figure 2).

However, it is also important to recognize that the procedures used by VSP have also changed, and over the years they have added more contacts and replacement questionnaires. Whereas in the beginning they only used delivery of a questionnaire and a follow-up thank-you picture postcard, they later started using one replacement questionnaire and then two of them. The analysis completed in 2001 showed, for example, that without the second replacement in1999 the final response would have been about 64% instead of 72%, and that without either replacement the response rate would have been slightly under 50% (Dillman and Carley-Baxter, 2001).

These data can be interpreted in two quite different ways. First, from a best practice perspective, I have believed for many years (since before publication of my 1978 book on Mail and Telephone Surveys) that one should use multiple follow-ups in order to assure high quality responses are achieved. Thus, I think one could argue that the VSP has simply started using best practices and the addition of these follow-up procedures is what enabled them to keep their response rates high. In many respects, it can be argued that the response rates now being obtained to this voluntary survey of the general public is remarkably high compared to nearly all other voluntary general public surveys I have observed. An average response rate of 72% in 2006 is unusually high, if not simply remarkable.

At the same time, the need to use more follow-ups suggests that the greater length of the surveys may be having a significant negative effect on response rates. I wish to emphasize "may" because I believe it can also be argued that the general culture of surveying has changed so that response rates are simply lower across the board than they were in the past.

Ease of Understanding By Respondent

I am routinely exposed to numerous surveys each week. From the standpoint of general clarity and communication the VSP surveys are well done, compared to most other surveys I review.

Similar procedures to these have been used for many years. This has probably given the VSP a sense of when questions work and do not work, so that I believe there may be a corrective mechanism at work to eliminate survey formats that produce poor results. I cannot say this with certainty, because I have not examined item nonresponse rates. If one wishes to conduct such an analysis I think that would provide a more informed understanding of whether problems may exist, and I would suggest that at some time in the future such an analysis be done.

In recent years I have done a great deal of visual design and layout research, and there are some things I would do differently in these surveys (which I will describe below). However, overall I want to say that I think the questions and general format are ones that are used by many surveyors and probably work reasonably well.

Double Response Requests

One of the practices used in these surveys (e.g. question 1 a and b) of the Independence survey is to ask people to answer in a left hand column about a prior visit, and then to ask them to answer on the right of the item with regard to their likely behavior on future visits. I do not think it is a good practice to ask people to respond to the same listing items twice. I have observed in some surveys that the item nonresponse to the second question tends to be higher than for other items to which respondents are only asked to respond once to each statement or phrase rather than twice. Thus, I would encourage this practice to be reviewed with regard to possible negative consequences. I cannot say with certainty that it happens here, because my observations were not experimental, but I have observed the differential item non-response enough to be concerned about it.

Directions for Branching Instructions

The VSP uses a somewhat different branching pattern method than I do. This is a topic of research on which I have published experimental results (e.g. Redline, Dillman, Dajani and Scaggs, 2003). The VSP practice differs from the convention I have found works well, e.g. on question 4 a b, they provide a wide separation between the yes and no categories and direct people from these separate locations to go different directions. I think the method they use may work, and there are visual theory reasons that it could work, which I will not go into detail about here. The point that seems important to make here is that their method may not only conserve space, but be effective. I sense it has probably been developed over a period of years, and may

reflect experience with finding out which other methods did and did not work.

Also, I should note that I thought the branching on question 14 of the Independence survey was particularly creative, and is one that I might consider trying at some time in the future.

Check-All-That-Apply Question Formats

The questionnaires I was asked to review contain a number of check-all-that-apply questions. I believe this format has two negative consequences. The first is to bias answers towards the earlier categories and the second is to get fewer responses marked. The second problem was reported in a test of 16 questions reported in an article by Smyth, Dillman, Christian and Stern (2006). All 16 comparisons of check-all vs. use of a yes/no format for each item produced resulted in significantly fewer items being selected in the check-all format. In addition, responses in the bottom half were less likely to be marked. Further, proper wording of the yes/no format can prevent the occurrence of a high item nonresponse rate.

Another article forthcoming in Public Opinion Quarterly (Smyth, Dillman and Christian, In Press), shows that the yes/no format compares more closely to answers obtained on telephone surveys. This research is quite recent, and I suspect it will result in more and more organizations abandoning the check-all format. This is a topic worth considering for the present surveys, but would have the effect of increasing the time required for answering each item.

General Observation

The visual techniques that I now use for constructing mail questionnaires have advanced beyond those in use here. For example, I prefer using colored backgrounds with white spaces to highlight answers locations. I use reverse print numbers to provide a stronger navigational path.

However, I also believe strongly that other simpler and less expensive formats may obtain data that are of excellent quality. It also seems important to note that I do not think that visual appearance of questionnaires is a major determinant of response rates—number of follow-ups and contact strategies are much more important.

Thus, I do not believe that these surveys present an excessive burden to respondents. Burden in my view constitutes in part being asked to respond to boring questions that do not make sense to the respondent, just as much as it involves time. I think that for the most part the questions asked here will make sense to respondents, and even be considered "interesting," inasmuch as the topic reflects an experience that may have been unique for the respondent. I also have the feeling that these questions are ones for which people feel like others will benefit from their response. I do not think that such a benefit is perceived in many surveys that people are asked to complete on other topics and personal experiences. Had I been asked in the abstract to compare these questionnaires to many other questionnaires I see on a daily basis, I would tend to rate the overall burden as low.

White Space

In general, "white space" in questionnaires is desirable (it helps separate questions from each other) and sometimes it can be manipulated to make navigation easier. By comparison, these pages are fairly full. However, I do not see something that I would quickly change to provide more of it. If the response rates were a lot lower, then that might be an avenue to pursue for improving response rates, but I am not sure that I would put much of a priority on this point. I should also add that because of the small size of the printed questionnaires used in these studies I do not think the benefit of white space would be as high as on larger-size paper formats.

Booklet Layout

Nearly the same booklet layout was used in the surveys conducted in the late 1980s as is now being used. The VSP does not use envelopes. I believe that was a good decision, and would not suggest any other format that is likely to improve response rates to these surveys (Dillman, Dolsen, and Machlis, 1995). Return envelopes are likely to get lost when people are on vacation trips.

Instructions

I think the instructions for answering questions are reasonably clear. The writing style is reasonably terse in the sense there are no wasted words in the introductions to questions. I sense the sponsors have built upon their experiences in earlier years to develop a parsimonious style that works. There is nothing obvious to me that would suggest a dramatic change in instructions would improve response.

However, I would encourage the VSP to consider switching from blank lines for response spaces to boxes, which are becoming more conventional in our computer age. That may be one of the benefits of the proposed optical scanning. I see that circles are being used in the new optical scanning format and think that will be useful. That format is also used frequently by surveys and is understood by them, so I believe that change in format will work fairly well.

Order of Questions

There are two issues here that I look at when examining questionnaire order effects. One is whether answering earlier questions bias answers to later questions. The other one is whether the questions are ordered in a way that enables the respondent to use a natural organization of the experience to recall what happened to them.

Attitude and opinion items are the main items for which I worry about earlier questions changing answers to later ones. There are few if any of these items. In addition, as I looked through the questionnaires it seems to be that items are in a logical order to the respondent, starting with before the visit, going through activities of the visit, and ending up with demographics. I would not be inclined to change the order of the items in any significant way.

As someone who designs a lot of surveys, I rely a lot on cognitive interviews to give me insight into how respondents think about the experiences they are reporting. If I were to do cognitive interviews with real respondents for these questionnaires I might be able to make specific

suggestions for reordering. However, I am too far removed from the experiences the questionnaires ask about (I have not visited most of these parks) to be able to effectively emulate a respondent's behavior.

Can These Questionnaires be Completed in 20 Minutes?

This is a difficult question to answer. People vary greatly in the speed at which they complete surveys. When I am faced with making estimates I usually ask some knowledgeable people (in this case it would need to be park visitors) to complete questionnaires and see how long it takes them to do that.

Theoretically, the current estimate of 20 minutes means that a person has about 1 ½ minutes to complete each of the 15 pages of questions. There are some open-ended questions but they are for the most part short. The number of "responses" requested on each page or items to be processed runs from about 12 to 20. Fifteen items in 90 seconds means spending about 6-7 seconds per item. However, people do not have to go look up information in order to answer virtually all of the questions. My best guess is that the burden may be a little higher than 20 minutes on average, but not by much. However, I would encourage VSP to ask people to fill out questionnaires while timing them. Items that seem complex to a survey designer and questionnaire reviewers who have not gone through the "Park" experience may be much simpler for respondents who have had that experience.

Concluding Observation

I enjoyed reviewing these questionnaires, in part because of not having seen any of them since the report I completed in 2001. Thus, it came as a pleasant surprise to see the response experience of the last several years. These surveys and the history of their use is unique, and quite revealing of how scientists have been able to maintain high response rates by adopting strategies for which there is clear literature support. Most surveyors are now rushing madly to develop mixed-mode survey strategies as a means of retaining high response rates, and it is impressive to see that such a strategy is not necessary for these VSP questionnaires. The procedure they have used for the last 20 years seems remarkably robust.

Thank you for the opportunity to review these surveys.

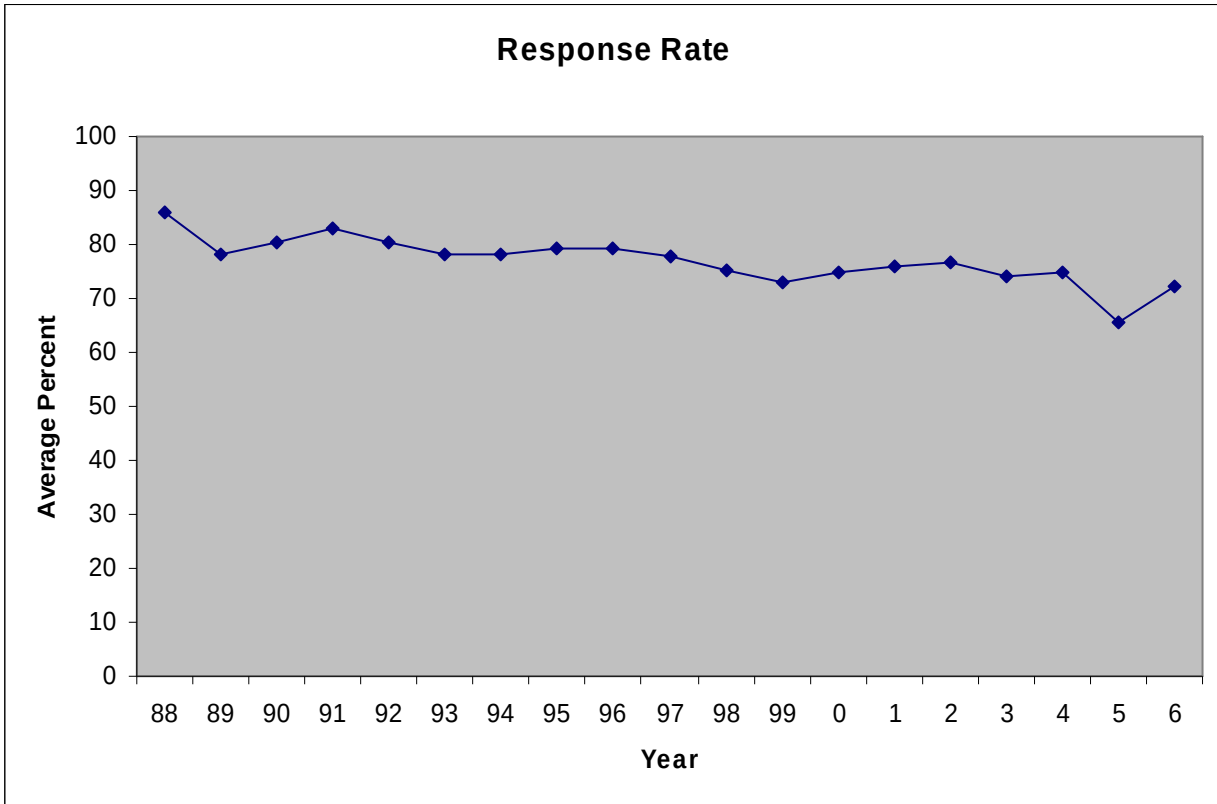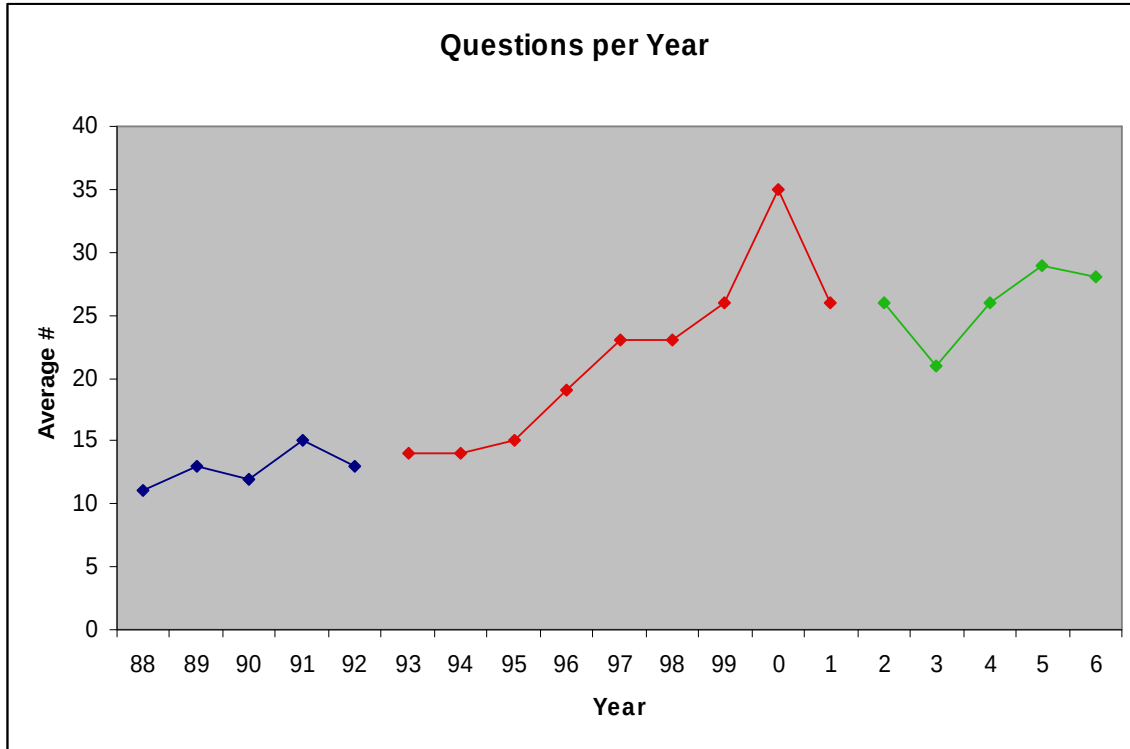Figure 1: Response rates for all Visitor Park Surveys, by year, 1988-2006.

Figure 2: Average number of questions asked each year, 1988-2006. (Blue lines from 1988-1992 represent 12-page questionnaires, red lines from 1993-2001 represent a mixture of 12- and 16-page questionnaires, and green lines from 2002-2006 represent 16-page questionnaires only.)

References

Dillman, Don A. 1978. Mail and Telephone Surveys: The Total Design Method. New York: Wiley-Interscience. 375 pp.

Dillman, Don A. and Lisa R. Carley-Baxter. 2001. "Structural Determinants of Mail Survey Response Rates Over a 12-Year Period, 1988-1999." 2000 Proceedings of American Statistical Association Survey Methods Section, Alexandria, VA.

Dillman, Don A., Dana E. Dolsen, and Gary E. Machlis. 1995. "Increasing Response to Personally-Delivered Mail-Back Questionnaires by Combining Foot-in-the-Door and Social Exchange Methods." Journal of Official Statistics 11(2): 129-139.

Redline, Cleo D., Don A. Dillman, Araf Dajani, and Mary Ann Scaggs. 2003.   "Improving Navigational Performance in U.S. Census 2000 By Altering the Visual Languages of Branching Instructions." Journal of Official Statistics 19(4): 403-420.

Smyth, Jolene D., Dillman, Don A. and Leah Melani Christian. In Press. Chapter 27. "Context effects in Internet surveys." In Reips, Ulf-Dietrich, et al. (eds.) Handbook of Internet Research, Oxford University Press. Pp.427-443.

Smyth, Jolene D., Don A. Dillman, Leah Melani Christian and Michael J. Stern. 2006. "Effects of Using Visual Design Principles to Group Response Options in Web Surveys." International Journal of Internet Science, 1(1): 5-15.