

# WHITE paper

## SERIES

Applying Knowledge™  
Improving Decisions

## Drop-down Boxes, Radio Buttons, or Fill-in-the-blank?

Web Survey Scale-Type Effects

*When it comes to Web surveys, the type of attribute rating scale used does matter. This is the conclusion of Burke Inc. research showing how responses to drop-down box, radio-button and fill-in-the-blank survey interfaces can differ significantly from each other.*

The question of what type of attribute rating scale to use arises especially in Web surveys in which respondents are asked to use the same scale repeatedly to rate, for example, an organization's performance on numerous product/service attributes.

Web surveys allow for the use of rating scale options that have no offline equivalent. For example, drop-down boxes allow survey designers to hide available response options until a respondent clicks on a box and a list appears. This approach takes up less room on the computer screen than other possible approaches; in particular, it makes it more feasible for a respondent to provide side-by-side ratings of an attribute's importance and the client company's performance on that item.

Concern has been expressed, however, that use of drop-down boxes introduces order bias, inducing respondents to choose ratings that appear higher up on the list presented. There is also concern that drop-down boxes cause respondents to drop out of the survey at a greater rate because of the need to click twice (once on the box, then once on the desired response) for each item.

The research industry's default approach to designing Web surveys has, therefore, been to design the Web screen to look approximately like a paper questionnaire.

In practical terms, this has usually meant using radio buttons stretched horizontally across the browser window, much like circles to be filled in might be stretched along a printed survey instrument.

### Research Method

To test whether concerns about drop-down boxes are warranted, research was conducted in conjunction with an effort to supplement Burke's Workforce Perspectives® employee normative database — a comprehensive set of evaluations regarding the working environment of North American adults. For this research, four different survey versions of the same questionnaire were programmed and respondents were randomly assigned to complete a particular survey version.

For each version, the main body of the questionnaire included 68 attributes in which respondents were asked to rate aspects of their work life, including their work groups and managers, their organizations' values and senior management, and their own job positions. Sixty-six of these items had positive wording, so that a higher rating would indicate a more favorable assessment. There were also two negatively worded (reverse-scored) items included at the end of the attribute battery. For all items, a five-point agreement scale was used in which a "5" was said to indicate that respondents "Strongly Agree" with a statement, and a "1" was defined as indicating that respondents "Strongly Disagree." Respondents also could indicate that statements were "Not Applicable" to their employment situations.

Two of the survey versions employed drop-down boxes. In one, the listing of rating

options began at the top of the scale, with the "Strongly Agree - 5" option. In the other, the list began at the bottom of the scale, with the "Strongly Disagree - 1" option. On both of these questionnaire versions, the "Not Applicable" option appeared last on the list.

For the third version, radio buttons used to collect responses were programmed to stretch horizontally across the width of the browser window, with the "Strongly Agree - 5" and the "Not Applicable" options appearing toward the right side of the computer screen. The fourth version used yet another option for rating attributes, asking respondents to fill in the blank by typing a whole number between "1" and "5" into a text box. The "Not Applicable" option used in conjunction with the fill-in-the-blank approach involved clicking on a box below the blank where the respondent would have otherwise typed in the number.

Respondents were U.S. adults belonging to an online panel who were paid a \$2 incentive for completing the survey. E-mail invitations to the survey advised them of the primary purpose of the study — to gauge people's opinions about their place of employment — but did not mention the accompanying research about Web rating scales.

### Results

For each survey version, more than 100 surveys were completed. After respondents completed a few initial screening questions, the overall drop-out rate was low (9.4%). There were no significant differences in the drop-out rate for the four survey versions. The biggest difference in drop-out rates, in fact, was between

the two different drop-down box survey versions, which should have required exactly the same effort of respondents. There was, therefore, no evidence that respondents found any one survey version more frustrating than the others.

Neither was there evidence that drop-down box response options induce respondents to provide answers that appear higher up on the response option list. When responses to the 66 positively worded attributes were aggregated, the mean scores provided by respondents of the descending drop-down box version and the ascending drop-down box version were identical to two decimal places (3.75). See Figure 1.

The mean score respondents provided on the fill-in-the-blank survey version (3.69) did not significantly differ from that provided by respondents to the drop-down box surveys, although the response pattern across the five scale points was much closer to a uniform distribution for this survey version. For the radio-button version, the mean score was significantly higher (3.93).

Of the four survey versions, the radio-button version also had both the shortest average and shortest median completion time (9 minutes, 55 seconds and 9 minutes, 17 seconds, respectively) approximately half a minute quicker than the average and median for the drop-down box versions. In addition, the radio-button version had the lowest minimum completion time.

Allowing respondents to complete a survey quickly is a mixed blessing. On the one hand, researchers do not want to unduly take respondents' time by burdening a Web survey with challenging technical tasks.

On the other hand, researchers want some assurance that respondents are taking time to read questionnaires and consider their responses, rather than just racing through online surveys to get them completed or to qualify for incentives. The lower average, median and minimum times associated with the radio-button version suggest this might be the format that most enables respondents to rush through surveys.

Indeed, survey results indicated that there was significantly less variation in individual scale usage among those responding to the radio-button version than there was among individuals responding to the other survey versions. Almost two-thirds (65.8%) of those responding to the radio-button version used the same scale point for at least 35 of the 66 positively worded items, compared to fewer than half of the respondents using any of the three other survey versions. See Figure 2.

Why would this happen? Perhaps repeating the same scale point involves less effort on the part of the respondent using a radio-button survey. With a radio-button survey, a respondent must adjust his or her computer mouse horizontally in order to change attribute ratings. To keep using the same scale point, no such horizontal adjustment is needed. On the other hand, with a drop-down box survey, the amount of work to provide the same attribute rating

would seem to be neither less nor greater than the amount of work required to provide a different attribute rating.

A practical consequence of this lack of individual variation in scale usage was that average ratings for the attributes included in the questionnaire were more closely clustered among radio-button survey respondents than they were for the other survey versions. If this study had been performed for a single organization, it might, therefore, have been more difficult distinguishing relatively high-performance attributes from lower performing items if a radio-button version had been used.

A complete presentation of findings from this study was nominated for Best Methodological Paper at the 2001 Congress of ESOMAR — The World Association of Research Professionals. Findings from the study were also presented at the 2001 Advertising Research Foundation Week of Workshops.

Figure 1: Combined Responses to All 66 Positive Attributes

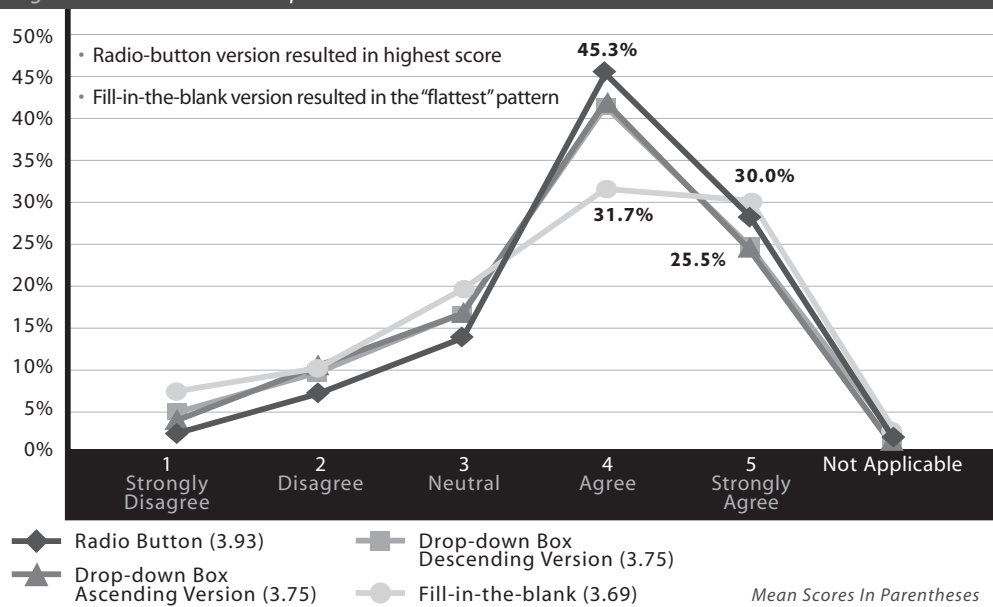


Figure 2: Variation in Scale Usage

	Average Individual Standard Deviation in Ratings	Average Maximum in Number of Same Response	Percent of Respondents Using The Same Scale Point on At Least 35 Items
Radio Button	0.74	38.0	65.8%
Fill-in-the-blank	0.82	35.2	48.3%
Drop-down Box Ascending Version	0.85	33.9	45.4%
Drop-down Box Descending Version	0.81	36.1	45.2%