ALERT!

MARKETING RESEARCH ASSOCIATION

SECOND QUARTER 2013 · VOL. 53 · NO. 3

AMELIA SHOWALTER, KEYNOTE SPEAKER,
2013 MRA INSIGHTS & STRATEGIES CONFERENCE
DIRECTOR OF DIGITAL ANALYTICS, 2012 OBAMA CAMPAIGN

THE WORLD OF BIG DATA

10 WAYS TO OPTIMIZE YOUR SOCIAL RESUME

WHEN A MAN LOVES A WOMAN…IN SOCIAL MEDIA DATA

HTML5 AND THE FUTURE OF DIGITAL ADVERTISING

How Sliders Bias Survey Data

By Ron Sellers

One of the catchphrases popular in marketing research over the last few years has been "respondent engagement." In the traditional methodologies of telephone and mail, "respondent engagement" was usually more the avoidance of negatives than the presence of positives – in other words, don't design a particularly dull, difficult, or confusing questionnaire, and make sure the interviewers sound upbeat rather than bored senseless when they read the questions.

With the advent of on-screen methodologies such as mobile MR and online surveys, far more options are now available for respondent engagement, e.g. gamification, drag-and-drop, cartoon icons that "lead" people through questionnaires, etc. The concept is that respondents are bombarded with research opportunities and lengthy, boring questionnaires, so anything that does a better job at capturing their interest will likely result in a higher response rate (less potential for response bias), as well as interested respondents who are really thinking through their answers rather than just clicking on things to finish their task as quickly as possible.

But how much work has been done in the industry to understand what impact, if any, these new approaches to respondent engagement have on the actual data we're collecting?

One of the approaches to respondent engagement that is relatively common in on-screen surveys is the use of sliders. Rather than a series of radio buttons that respondents click on to register their answers, sliders have a button respondents grab with their mouse and move over a pre-determined scale to record their answers. The thought process is that sliders give people a different way of interacting with the questionnaire that breaks up the routine of clicking buttons over and over. In addition, sliders can be used with graphics (a thermometer where the mercury "rises" as the slider is moved up the scale) that are visually engaging, and some researchers believe it's more natural for respondents to use a slider when the scale is particularly large (1 – 100), rather than typing in a specific number.

There's only one problem with sliders : they have a strong potential to bias the answers people give.

In the new report *How Sliders Bias Survey Data*, Grey Matter Research (Phoenix, AZ) compares the data from sliders to traditional radio buttons for the same questions. The main difference between the two approaches is that radio buttons have no on-screen "starting point." All the buttons on the scale are equally available to people; it takes just as much effort to move the mouse over a "2" and click on it as it does to do the same thing for a "7."

In the new report How Sliders Bias Survey Data, Grey Matter Research (Phoenix, AZ) compares the data from sliders to traditional radio buttons for the same questions. The main difference between the two approaches is that radio buttons have no on-screen "starting point."

Sliders, on the other hand, must have a starting point for the little button that people grab with the mouse and move up or down the scale. It's this starting point that appears to cause the problems.

Grey Matter surveyed 1,700 people using a demographically representative sample from a national online access panel, in English and Spanish. The survey was fairly lengthy (at least for us – it was much shorter than some that are used today), which led us to think about ways to break it up a bit and do what we could to encourage higher levels of respondent engagement. On two questions, this included using sliders rather than radio buttons for grids.

One question used a five-point scale, asking respondents whether certain things they recently experienced had increased or decreased their interest in various types of organizations. The scale was *strongly increased*, *somewhat increased*, *had no impact*, *somewhat decreased*, and *strongly decreased*. This question had five different choice sets.

The second question gave respondents two options, and asked which one they preferred and how strongly. Each option was at an opposite end of the slider, and the scale was *prefer A strongly*, *prefer A*, *prefer A somewhat*, *no preference*, *prefer B somewhat*, *prefer B*, and *prefer B strongly*.

At the outset, we suspected that the starting point of the sliders might bring some minor bias to the data, so we arranged for the sliders to be randomly started at the low end of the scale, at the midpoint, or at the high end of the scale. Each respondent had a consistent starting point, e.g. respondent 1 always saw the sliders start at the midpoint; respondent 2 always saw the sliders start at the low point, etc.

In addition, we forced movement of the sliders in order to continue. In other words, someone who saw the slider starting at "1" could not simply hit "next" and move on – they had to move it off "1" and then back on it again for that answer to register. This avoided allowing respondents to take the easy way out by simply leaving the slider where it was instead of actively registering an answer.

After 496 completes, we found something very disturbing : the data was being biased according to the starting point of the sliders When respondents had their sliders starting at the high end of the scale, they were more likely to chose an answer at that end of the scale. When their sliders started at the midpoint, they were more likely to choose a midpoint answer. It was at the low end of the scale that we saw the strongest impact – people who had their slider start at the

low end of the scale were dramatically more likely to give an answer at that end of the scale.

Could this be some other type of bias ? First, we checked the distribution of the 496 responses. They were almost exactly divided into thirds. Next, we checked the demographic distribution. The demographics of the respondents who had been randomly divided into the three different starting points were statistically identical. The random distribution had worked exactly as it was supposed to. We were left with the conclusion that the only difference among the three randomly distributed groups of respondents was the starting position of the sliders.

What impact did that really have ? It was substantial. On the five choice sets using the five-point scale, people who saw their slider start at the low end of the scale were 71 percent more likely to chose "1" as their answer than were those who saw their slider start at the midpoint, and 38 percent more likely to do so than were those who saw it start at the high end of the scale.

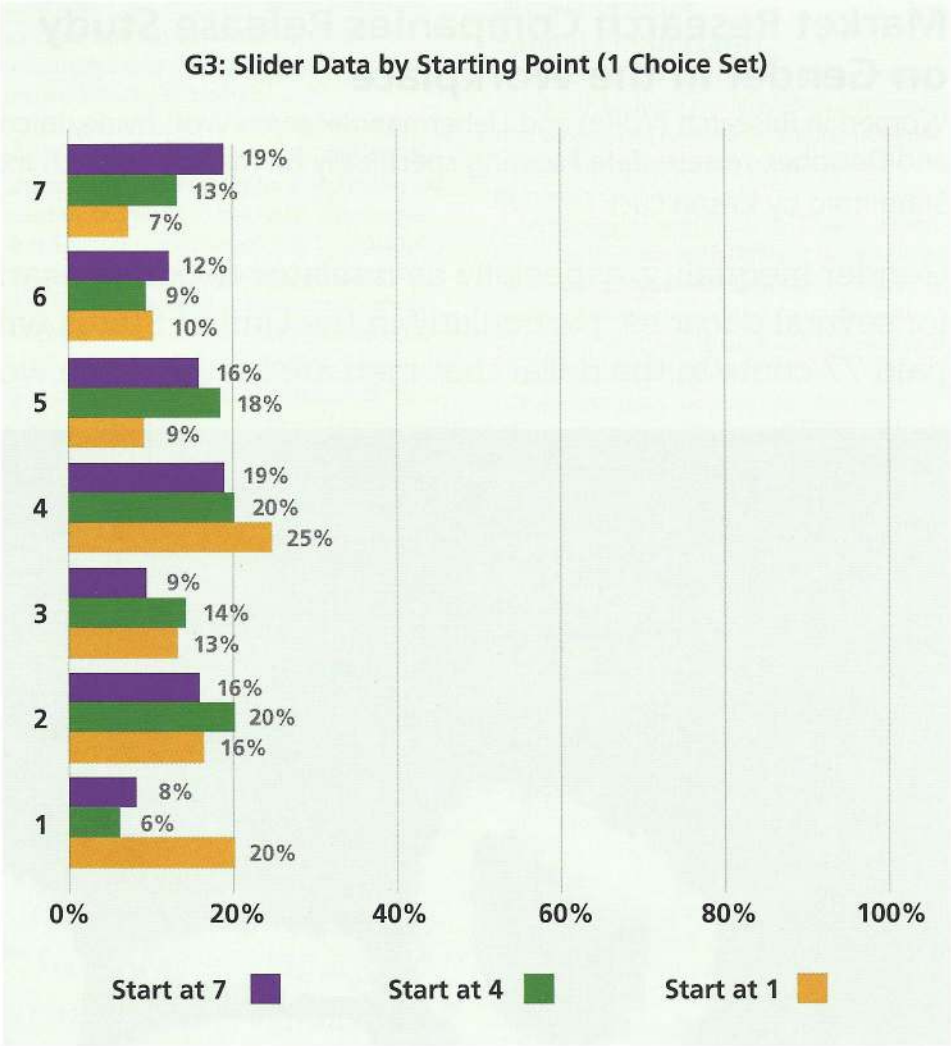
On the nine choice sets employing the seven-point scale, respondents who saw their slider start at the low end of the scale were 64 percent more likely to select a response of "1" than were those who saw it start at the high end, and a whopping 109 percent more likely to do so than were those who saw their slider start at the midpoint.

An example of the differences according to slider starting position can be seen in the graph, which gives data for one of the nine choice sets with the seven-point scale. The effect was not present to the same levels in every single one of the choice sets, but it was consistent that responses differed according to the starting point of the sliders.

One of the keys of looking at our results was comparing data from each of the slider starting points to data from the other starting points, rather than just the combined slider data compared to data from traditional radio buttons. Because we randomized the starting points, the bias which was very real for individual respondents was somewhat balanced out by the randomization, leading to combined data which looked fairly similar to the data we got once we switched to radio buttons for the remaining 1,200 respondents.

(As a side point, this really emphasizes the importance of looking at data from a variety of angles. Just because something looks right doesn't mean it is.)

Since the overall results for sliders and radio buttons were similar, doesn't this mean sliders and radio buttons are comparable ? No. Averaging together three



wrong figures in order to come up with a right one doesn't remove the problem that the individual figures being averaged together are wrong. In addition, since there was even more bias at the low end of the scale than elsewhere, the combined slider data still showed a stronger bias toward the bottom end of the scale being chosen than did the radio buttons. Finally, randomization in our study was fairly simple because the largest scale we used with sliders was seven points. Exactly how would slider starting points be sufficiently randomized on a scale of 0 – 100, or on a scale where there is no actual midpoint?

Think of it this way, are you willing to tell your client that their new concept could be twice as likely to get ratings of "not at all appealing" by potential customers because you wanted to heighten respondent engagement by employing sliders in the survey programming ?

Although Grey Matter Research only tested sliders, similar questions must be raised about other attempts at respondent engagement. Are we biasing our results by how respondents are able to answer the questions and by what they see or experience on screen?

The broader question is; why do we need overt attempts at respondent engagement to begin with ? The best way to engage respondents is not to browbeat them with dozens of survey invitations each week, to respect their time by not surveying them about irrelevant topics, and not to create overly long, boring questionnaires with repetitive questions and lengthy grids that make respondents mentally check out of the survey. If the length, topic, and design of a questionnaire are done well, there should be little need for specialized attempts at respondent engagement – particularly if those attempts at engagement actually change the data we are getting.

Ron Sellers is president of Grey Matter Research and can be reached at 602-684-6294 or ron@greymatterresearch.com. For a free copy of the full report *How Sliders Bias Survey Data*, contact Grey Matter Research.