

*Journal of Research Practice*  
Volume 11, Issue 1, Article D1, 2015



*Research Design:*

## **Exploring Slider vs. Categorical Response Formats in Web-Based Surveys**

**Catherine A. Roster**

Anderson School of Management  
University of New Mexico, UNITED STATES  
[roster@unm.edu](mailto:roster@unm.edu)

**Lorenzo Lucianetti**

Department of Business Administration  
Università degli Studi “G. d’Annunzio” Chieti – Pescara, ITALY  
[llucianetti@unich.it](mailto:llucianetti@unich.it)

**Gerald Albaum**

Anderson School of Management  
University of New Mexico, UNITED STATES  
[albaum@unm.edu](mailto:albaum@unm.edu)

### **Abstract**

Web-based surveys have become a common mode of data collection for researchers in many fields, but there are many methodological questions that need to be answered. This article examines one such question—do the use of sliders to express numerical amounts and the use of the more traditional radio-button scales give the same, or different, measurements? First, we review the central debates surrounding the use of slider scales, including advantages and disadvantages. Second, we report findings from a controlled simple randomized design field experiment using a sample of business managers in Italy to compare the two response formats. Measures of topic sensitivity, topic interest, and likelihood of participation were obtained. No statistically significant differences were found between the response formats. The article concludes with suggestions for researchers who wish to use slider scales as a measurement device.

**Index Terms:** Web-based survey; response format; categorical scale; visual analog scale

**Suggested Citation:** Roster, C. A., Lucianetti, L., & Albaum, G. (2015). Exploring slider vs. categorical response formats in web-based surveys. *Journal of Research Practice*, 11(1), Article D1. Retrieved from <http://jrp.icaap.org/index.php/jrp/article/view/509/413>

## **1. Introduction**

Survey researchers in business and society in general have come to rely upon Web-based (i.e., online) surveys for data collection. Online surveys have numerous advantages over traditional data collection modes, including significant cost and time savings, greater flexibility, convenience, and anonymity for survey respondents (Couper, 2000; Couper & Miller, 2008; Miller, 2006). These advantages generally outweigh disadvantages associated with online surveys, which include low response rates (Lozar Manfreda, Bosnjak, Berzelak, Hass, & Vehovar, 2008; Vicente & Reis, 2010) and lack of respondent engagement (Couper, 2008; Downes-Le Guin, Baker, Mechling, & Ruyle, 2012).

A key advantage of online surveys over paper-and-pencil surveys is their ability to harness the Web's rich visual capabilities. Use of graphic elements can enable online researchers to create a more engaging and interactive experience for survey-takers. Examples include visual analog scales (VAS) including slider scales, where respondents drag sliders or bars to express numeric amounts, online card sorting tasks in which respondents drag and drop visual elements into one of several piles or buckets, and ranking tasks that involve a similar drag and drop action to sort objects (for examples, cf. Couper, 2008; Smith & Albaum, 2013). Slider scales are often displayed in a matrix grid featuring multiple scales or rating items, where the radio-button response format is used traditionally. There has been some discussion in the professional literature on the two response formats, with some divergent opinions emerging. In addition, some methodological studies have examined differences between VAS response formats and traditional radio buttons, with sometimes conflicting findings (Couper, 2008).

The first objective of this article is to provide a detailed review of the discussions and empirical studies to date surrounding the use of slider scales. Second, we hope to contribute to research in this area by presenting results from a field experiment in which we varied the response format between traditional radio buttons and a slider scale in a pilot study conducted with a sample of European marketing managers. We conclude with suggestions for researchers who wish to use slider response formats in online surveys.

## **2. Sliders vs. Traditional Rating Scales: Debates and Evidence**

Compared to other methodological issues regarding online surveys, there has been relatively little attention given to the growing use of slider scales as an alternative to traditional rating scale formats. Rather than clicking a radio button to respond to online survey questions, respondents click on the start button or bar of the slider scale and drag and drop it to the desired response position. Examples of a traditional radio-button scale and various visual analog scale (VAS) versions of the same scale are illustrated in Figure 1. The primary argument for utilizing sliders is that they are less repetitive and more engaging for online survey respondents than traditional radio-button style scale formats. The assumption is that a more interactive experience may reduce survey fatigue and nonresponse, and potentially, lead to higher quality data. A second argument for their use is that the data obtained from use of slider scales may be equivalent or superior to traditional Likert-style scales that employ radio buttons.

**Example A: Traditional radio button grid format**

Please rate this survey compared to other surveys you've completed with 1 being much worse and 5 being much better by clicking on the number next to each item that reflects your opinion.

	1	2	3	4	5
Survey was well written	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Time and effort required met my expectations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overall, the survey experience was good	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 1a. Examples of response format (Example A: Traditional radio-button grid format).

**Example B: Slider bar grid format**

Please rate this survey compared to other surveys you've completed with 1 being much worse and 5 being much better by moving the bar next to each item to the number that reflects your opinion.

	1	2	3	4	5
Survey was well written	<input type="range" value="1"/>				
Time and effort required met my expectations	<input type="range" value="1"/>				
Overall, the survey experience was good	<input type="range" value="1"/>				

Figure 1b. Examples of response format (Example B: Slider bar grid format).

**Example C: Graphic star grid format**

Please rate this survey compared to other surveys you've completed with 1 being much worse and 5 being much better by clicking on one of the stars below next to each item that reflects your opinion.

Survey was well written	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
Time and effort required met my expectations	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
Overall, the survey experience was good	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

Figure 1c. Examples of response format (Example C: Graphic star grid format).

Example D: Smiley Meter format (non-grid)

Please rate this survey compared to other surveys you've completed with 1 being much worse and 5 being much better by moving the bar under the smiley face to a position that reflects your opinion.

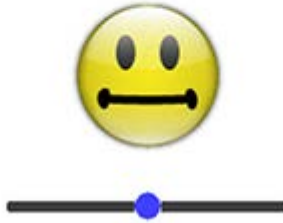


Figure 1d. Examples of response format (Example D: Smiley meter format).

Whether or not these claims are empirically supported is a matter of debate among research professionals and academic researchers. For instance, in a blog review the founder of *CheckMarket* survey software, Alexander Dobronte (2012) examined the two general arguments in support of slider scales. Dobronte concedes that while sliders may create a more pleasing experience for survey respondents, there is little scientific evidence to support the argument that data quality is better with sliders than with traditional Likert scales. Alternatively, Peter Cape (2009) of Survey Sampling International argues that slider scales are superior to traditional scales in many ways, and should be used more often. Systematic studies by academic researchers have investigated data quality in multiple ways with mixed empirical results.

The versatility of sliders may be responsible for some of the conflicting evidence, as there are many options and variations available to researchers who wish to use slider scales. Design choices include the range of scale points and whether values are discrete or continuous, the initial starting position of the slider, variations of graphics, use of labels, how many labels to include and where they appear on the slider, and so forth. Ultimately, online researchers must make multiple design choices when selecting and formatting scales, any of which can have an impact on data collection or quality (Derham, 2011). In the following sections, we review all sides of the slider debate by examining arguments and empirical support for the two central questions surrounding use of slider scales, namely engagement and data quality. After presenting results from our own field experiment with European marketing managers, we synthesize best practices and factors online survey researchers should consider before using slider scales.

## 2.1. Sliders as an Engagement Device

The advanced graphics technology supported by most Web browsers has greatly enriched the capabilities of Web surveys. Web surveys can incorporate not only slider scales, but also pictures, colors, and interactive components such as progress bars. The general assumption of researchers is that these attractive and interesting design elements will make Web surveys more fun to complete. Puleston (2011), for instance, lists the “the fun factor” as one of the main advantages of slider scales. Slider bars can easily be used instead of

traditional radio buttons in repetitive, boring matrix grids, which have long been associated with high nonresponse and drop-out rates (Couper, Tourangeau, Conrad, & Zhang, 2012). Vicente and Reis (2010, pp. 260-262) in their article discussing use of questionnaire design to fight nonresponse bias in Web surveys agree that the use of graphically enhanced response formats can make surveys more attractive and engaging, but warn they may increase break-offs if respondents become frustrated by increased complexity, time commitments, or if respondents encounter software/hardware compatibility problems. For these reasons, they advise visual enhancements be used sparingly in Web surveys.

Only a handful of studies have directly assessed respondents' satisfaction with sliders compared to traditional response formats. These studies provide mixed support for their use as an engagement device, but reveal some interesting insights. Stanley and Jenkins (2007) conducted a study with established Internet panelists from the UK in which respondents were asked to evaluate their survey experience after completing either a graphic image-based survey or a traditionally-formatted survey in terms of usability, engagement, and enjoyment. In their field experiment, respondents received either a traditional radio-button response format survey or a survey employing slider scales. Engagement scores (i.e., subject interesting and question style enjoyable) from respondents who received the slider scale survey were higher than those from respondents who received the standard radio-button survey and the difference was statistically significant, especially from adults in the 25 to 34 years age group. Time to complete the survey was longer for respondents in the graphic format survey as opposed to the traditional format survey, but the difference was not statistically significant.

Based on open-ended text responses to questions at the end of the survey, Stanley and Jenkins attribute the additional time needed to complete the slider scales to increased time respondents spent reading instructions about how to use and interpret the scale. Open-ended comments revealed that respondents found the graphic scales to be "slightly more complicated than other surveys" (Stanley & Jenkins, 2007, p. 87) and they valued the explanatory guidance provided in the instructions to the question. Over two-thirds (73%) of respondents presented with the graphic version stated they spent at least some time reviewing the guidelines. When correlated with educational background data, the authors found that respondents with some university/college credentials were more likely to refer to the scale instructions.

The educational level of respondents has been cited as a factor by other investigators who have compared VAS versus traditional scale formats. Funke, Reips, and Thomas (2014), in a survey of health-related products that compared slider scales and radio-button categorical scales in both horizontal and vertical orientation, found that slider scales led to higher break-off rates and a substantially higher response time. Since problems with slider scales were prevalent in respondents with less than average education, these researchers suggest that the slider scale format is more challenging in terms of the cognitive load it creates.

Sikkel, Steenbergen, and Gras (2014) conducted a two-wave Internet survey field experiment with Dutch marketing research panelists that compared "clicking" (i.e., traditional radio buttons) versus a variety of "dragging" scale response formats (i.e., sliders

and other drag and drop versions) and found statistically significant differences in the panelists' ratings of their survey experience across the two waves. Panelists who received the dragging version the first time rated their experience as significantly more pleasant, interesting, and the topics more important to them than did the panelists who received the clicking version, but they also rated the survey as more time-consuming. These results, however, were statistically reversed in the second wave, in which panelists received either the same or the opposite version of the survey. Panelists who received the dragging version for a second time rated their experience statistically significantly less pleasant and less interesting than they did in the first wave. Their ratings were also statistically significantly lower than those from the respondents who received the dragging version for the first time in the second wave after completing the clicking version in the first wave. These results suggest a novelty effect may drive respondents' positive reactions to graphically enhanced scale formats, one that dissipates after repeated exposure if it increases time needed to complete the survey.

Lastly, slider scales may or may not be accompanied by emoticons, which are pictorial images, including faces or other icons, used to represent variables being measured (see Example D in Figure 1d). Derham (2011), in a series of online survey field experiments about banking services, compared multiple scale formats. The experiments indicated that respondents preferred a traditional click style response scale with category labels over the moveable emoticon response format. However, the emoticon format was preferred second over the traditional click format featuring a non-categorical numerical scale format. When asked about their enjoyment and intentions to respond to future surveys employing the emoticon format, respondents reported the emoticon scale format was "cute" (Derham, 2011, p. 23) but did not influence their intentions to complete similarly-formatted surveys in the future. Respondents also reported that the emoticon scales were more difficult to respond to and that their answers may not have reflected their "true" responses due to confusion about how to use the scale. Derham (2011) concludes that Web surveys should not employ sliders with emoticons, as their visual appeal can be offset by difficulties faced by respondents, which include having to give more thought to responses, and lack of understanding from respondents about how to enter their true opinions, including "no opinion" or "can't say" options. The findings from Derham's study echoes cautions voiced by previous researchers regarding the confusion that can result from incorporating visual icons in survey response formats (Couper, Conrad, & Tourangeau, 2007; Couper, Tourangeau, & Kenyon, 2004).

No clear picture emerges from the little research that has directly examined respondents' enjoyment, interest, level of engagement, and intentions to complete future surveys that employ slider scales or other similar VAS formats as opposed to traditional scale formats. The take-away seems to be that, at least initially, these alternative formats can be more engaging and fun for respondents. However, the novelty effect may wane if VAS formats lengthen the time it takes respondents to complete the survey. Currently, many Internet survey respondents in the US and elsewhere are members of consumer access panels who agree to participate in surveys regularly in order to earn rewards and other direct compensation for their participation (Brick, 2011). Keeping panelists engaged is of central concern to the viability of Web surveys (Couper, 2000). Use of alternative response

formats that employ VAS, such as slider scales, may be a good way for researchers to make surveys more interesting, but not if their use compromises data quality.

## **2.2. Sliders and Data Quality**

A growing body of slider scale research has examined data quality issues, in particular, whether response data from sliders is at least equivalent to data obtained from traditional scale formats. Ganassali (2008, p. 27) has proposed a conceptual framework for examining the impact of questionnaire features on quality of responses. Her framework includes two input sources: (a) questionnaire features, including length, illustration, question wording, interactivity, and response formats, and (b) survey context, including topic and nature of invitation. These two inputs impact respondents' comprehension, meaning, retrieval, and judgment processes. The output or result is quality of responses, assessed by response rate, drop-out rate, completeness, and depth and variety of responses, with data quality impacting and impacted by respondent satisfaction. Discussion and research involving data quality produced by slider scales mostly center on these indices of response quality.

Here again, results are mixed. Some studies report little or no differences in data obtained from sliders as opposed to traditional scale types. Couper, Tourangeau, Conrad, and Singer (2006) explored the utility of VAS in a Web survey, comparing it to radio-button input and numeric entry in a text box on a series of bipolar questions eliciting views on genetic versus environmental causes of various behaviors. The response distributions for the VAS did not differ statistically from those using the other scale types, but the VAS had higher rates of missing data and longer completion times. Bayer and Thomas (2004) used Java applets to create sliders in an experiment that compared vertical and horizontal sliders to various formats of radio-button scales (end-anchored vs. fully anchored) and a numeric box entry version. No advantage or disadvantage was found in sliders, from a validity perspective. Validity coefficients were high and comparable to other scale formats. Lastly, slider scales are being used for online surveys conducted via smartphones. Buskirk and Andrus (2014) conducted a randomized experiment to compare mode effects of an online survey completed by computer and by smartphone. Results for slider structured questions showed no statistically significant difference across survey mode.

Other studies have shown the researcher's decision about where to place the starting point of a slider (e.g., low, high, or in the middle of the scale) can bias responses. Unlike traditional radio-button scales, a respondent must grab the slider and move it to a desired point on the scale to register his or her response. If the respondent fails to move the slider, no response is recorded, which leads to missing data. In the Bayer and Thomas (2004) study, the default position of all sliders was located at the middle of the 7-point scale. The authors found that this caused average values for the slider scales to be higher than for the non-slider scales. A similar bias associated with where the starting point is located on slider scales was reported by Sellers (2013) in a field experiment with a U.S. national online access panel that compared traditional radio-button scales to sliders. This study, conducted by Grey Matter Research, found that sliders initially positioned at the high point of the scale produced higher scores. Sliders positioned at the midpoint of the scale increased midpoint responses, and interestingly sliders positioned at the low end of the

scale resulted in higher scores, compared to scores of respondents who were presented with a traditional radio-button scale format.

Another scale design issue that has spurred debate over the quality of data from sliders versus traditional scale formats pertains to the scale range or number of points, and whether responses are scored as discrete values (categorical or numerical) or on a continuum that could include in-between values. Proponents of slider scales claim that the ability of sliders to capture a more precise reflection of respondents' opinions than the traditional 4- to 11-point Likert scale offers an improvement in the scale's reliability and validity (Taylor, 2012). This argument, however, appears to be predicated on two assumptions. The first assumption is that sliders are an exact online replica of the traditional paper-and-pencil graphical scale proposed by Freyd (1923) in which respondents indicate their response by placing a mark on an unnumbered line format for surveys. Naturally, the online slider version of graphic scales eliminates the tedious manual measurements to record responses formerly required by researchers when such scales were presented in a paper-and-pencil survey mode. The second assumption is that an increase in scale points increases data reliability and/or validity in a linear fashion. Over the years, a considerable body of scientific evidence has demonstrated that gains in reliability and validity begin to taper off after about seven response alternatives (e.g., Lozano, Garcia-Cueto, & Muniz, 2008; see also Dobronte, 2012 for a review). Both assumptions ignore the versatility that online researchers now have in determining the format, range, and presentation of sliders to online respondents.

It is beyond the scope of this article to review the adjoining debates that circle around these basic scale format issues. Some studies have, however, directly tested the proposed superiority of slider scales as opposed to traditional radio-button scales in terms of reliability and validity. Cook, Heath, Thompson, and Thompson (2001) in a study assessing users' perceptions of university libraries, compared data obtained from a 9-point radio-button format scale to a slider scale with a continuum from 1 to 100 in a test of scale reliability between the two response formats. Reliability differences, assessed by comparing Cronbach's alpha coefficients between the different formats, were relatively small, with a slightly higher alpha recorded for the radio-button scale as opposed to the slider scale. However, alphas for both scale formats scored well within psychometric standards ( $\alpha \geq .70$ ), leading these researchers to conclude that both sliders and radio buttons are psychometrically acceptable ways to gather attitudinal data. Cook et al. (2001) did find, however, that slider scales took a longer time to complete—a difference that was statistically significant.

Cape (2009) reports findings from a random experiment with online survey respondents that compared the validity of responses obtained from traditional Likert scales to those obtained from slider scales using a unique research design. In this study, respondents were first presented with traditional Likert scales for a series of items and then given the opportunity to re-score the same items on a slider scale in which they could adjust their original responses in accordance with the range offered by the original scale. Cape found that a majority of respondents chose to re-score their original opinions when presented with slider formats for the same items. Respondents who had selected a categorical response indicating "slightly" (whether agree/disagree) re-scored it in the direction of a



more definitive agree/disagree response. Based on these findings, Cape concluded that slider scales afford a more accurate reflection of respondents' true attitudes and opinions, but he notes that data distributions were not equivalent across the different slider formats presented to respondents. Each slider design tested produced a different data distribution, especially those with visual/pictorial components. Therefore, Cape advises against the use of sliders with pictorial elements, especially when results need to be compared across waves of data collection.

One fairly consistent finding across slider scale comparative studies is that sliders take longer to complete than traditional radio-button click response formats. Husser and Fernandez (2013), in a field experiment using a computer-assisted telephone survey that varied response formats by clicking, entering text, or dragging, found that dragging formats took the longest of the three styles to complete. These authors suggest that the longer response time may indicate that slider scales require additional cognitive processing and may not be as intuitive as simply clicking a radio button.

In summary, the jury is still out regarding the overall value of using slider scales over more traditional radio-button categorical scales. While some research supports the merit of slider scales as an engagement device and as a superior measurement tool in comparison to traditional scale formats, practitioner and academic research findings appear to be mixed on both fronts. Part of the problem appears to be that comparative rating scale studies often employ apples-to-oranges comparisons. In reality, the decision to use slider scales evokes a number of questionnaire and scale design choices, not all of which are directly comparable. Puleston (2011) summarizes scale design issues that researchers who might wish to use slider scales should be aware of before using them. The issues are the following:

- (a) Starting point of slider scales
- (b) Range and labeling protocols
- (c) Slider size and appearance on devices used
- (d) Use of icons or pictures

Puleston warns that any or all of these design choices can result in significant differences between data obtained by standard scales versus sliders, but that careful and prudent use of slider questions can help overcoming boredom and rote responses from survey respondents.

### **3. A Direct Comparison Study of Sliders vs. Radio Buttons**

As described above, most studies that have compared sliders to traditional scales contain a mixture of design choices, any of which can impact data quality. Therefore, as part of an exploratory pilot study designed to assess Italian business managers' views regarding the sensitivity of several potential survey topics, we created two versions of a 5-point scale designed to capture their opinions, one that used a slider scale and one that used a traditional radio-button scale. Respondents randomly received either one or the other scale formats. Our central research question was: Do slider and categorical response formats in Web-based surveys provide similar results? We kept every element in our design exactly the same except for the response format itself, including the invitation,

scale instructions, scale anchors, and the scale values, which were constrained to discrete numerical values, no in-between values, on 5-point scales anchored by categorical labels, as illustrated in Figure 2.

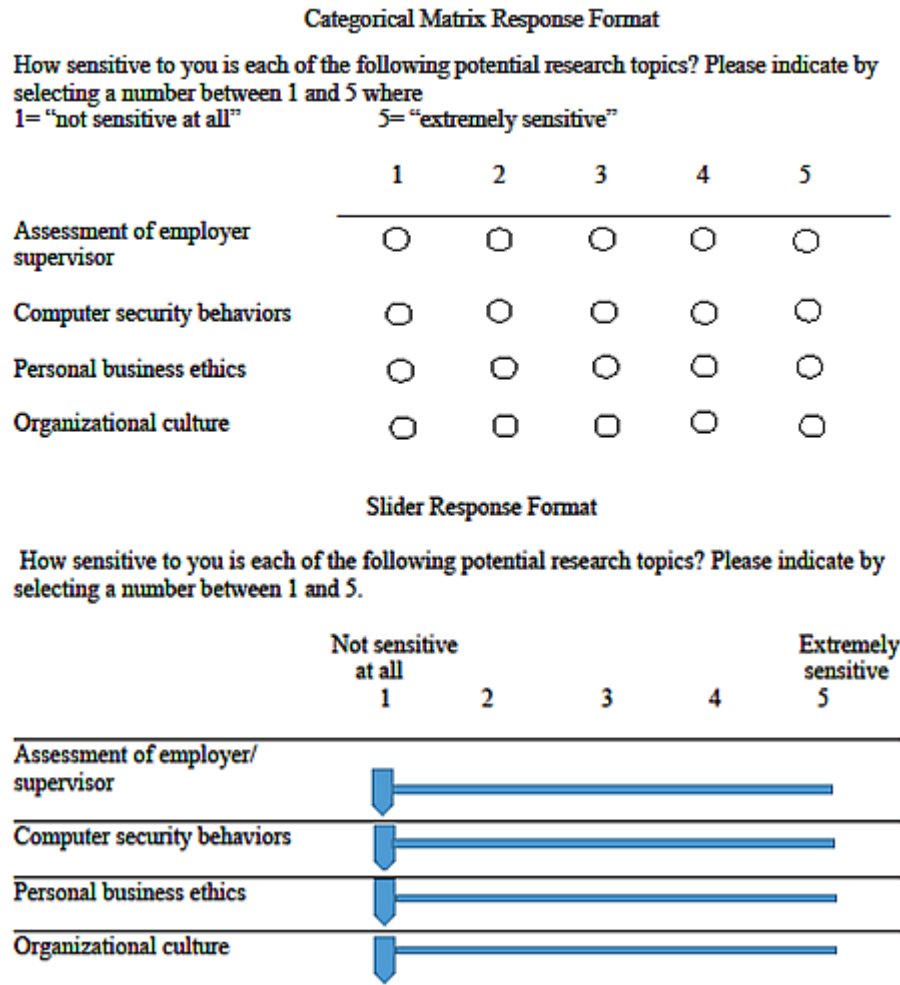


Figure 2. Illustration of response formats used.

### 3.1. Methodology

A completely randomized experimental design was used to obtain data from managers in 120 companies in Italy. The companies were randomly selected from Amadeus-Bureau Van Dijk, a database of public and private firms which includes Italian firms and multinational firms (<https://aida.bvdinfo.com/>). The treatment variable was response format (radio-button scale versus sliders). The objective our study was to examine managers’ response behaviors. Respondents were randomly assigned to a treatment group. The measurement instrument asked three sets of questions about potential survey topics—sensitivity, personally important, and likelihood of participation—as shown in Exhibit 1. For the sensitivity and interest questions, the same 12 topics were listed. For likelihood of

participation, five of these topics, chosen randomly, were listed. All topics were relevant to business managers. All scales were presented in a 5-point numerical format. The sensitivity scales ranged from *Not at all sensitive* (1) to *Extremely sensitive* (5), the personally important scale ranged from *Not at all important* (1) to *Extremely important* (5), and the likelihood scale ranged from *Very unlikely* (1) to *Very likely* (5).

Exhibit 1. *Survey Questions Instructions*

We define sensitivity of a research topic as “a topic which possesses a substantial threat to those involved as it may be perceived as intrusive and could raise fears about potential repercussions/consequences of disclosing the information requested.”

**Sensitivity (12 topics)**

How sensitive to you is each of the following potential research topics? Please indicate by selecting a number between 1 and 5, where:

1 = *Not at all sensitive* ----- 5 = *Extremely sensitive*

**Personally Important (12 topics)**

How personally important to you is each of the following potential research topic? Please indicate by selecting a number between 1 and 5, where:

1 = *Not at all important* ----- 5 = *Extremely important*

**Likelihood of Participation (5 topics)**

If you received a request to participate in a survey with one of the following topics, what is the likelihood that you would participate? Please indicate by selecting a number between 1 and 5, where:

1 = *Very unlikely* ----- 5 = *Very likely*

Two treatments were used: slider scale response and radio-button response. The two formats are illustrated in Figure 2. For the slider treatment, the slider was placed at the scale point 1 (lowest). No graphic elements were introduced to assure comparability with the radio-button format.

The invitation to the survey appears in Exhibit 2. The overall sample consisted of business managers, all of whom were sales and marketing managers—people who are considered hard to reach, but important for business research.

Exhibit 2. Survey Invitation Letter

**INTERNATIONAL RESEARCH PROJECT**

**“Sensitivity to research topics in business administration”**

Dear [Name],

The [university names] are conducting an international research project to study the sensitivity level of respondents towards certain business research topics. We would very much appreciate your participation in this pilot study. The questionnaire is one page and takes about 3 minutes to complete.

To access to the survey, please, click on the following link:

[URL link]

- A copy of the research findings will be sent to all respondents.
- All of your answers will be kept strictly confidential. No individual or company will be identified. Only summary data and aggregate results from multiple firms will be published.
- You have the right to skip any question you choose not to answer. However, we ask that you answer all questions. If you are not sure of an answer to a question, please provide your best estimate.

We will be happy to answer to any question or concerns you may have. Please, write to the following email address: **surveysentiveness@gmail.com**

Please accept my thanks for your time and consideration.

Sincerely,

### **3.2. Demographic Characteristics of the Sample**

The slider scale was given to 58 managers (of whom 35 responded) and the radio-button scale was given to 62 managers (of whom 39 responded). The demographic characteristics of the two samples are shown in Table 1. In many ways the two sample groups are quite similar, indicating they have come from the same population. The only characteristic that is different is Education (the difference is statistically significant at  $p < .10$  level).

Table 1. *Sample Demographic Characteristics*

Characteristic	Slider Scale		Radio-Button Scale		$\chi^2$	<i>p</i>
	<i>n</i>	%	<i>n</i>	%		
<b>Gender</b>					1.684	<.20
Male	28	80.0	27	69.2		
Female	6	17.1	12	30.8		
Not specified	1	2.9				
<b>Number of Employees in Company</b>					1.858	<.87
10 or less	6	17.1	7	17.9		
11-50	7	20.0	7	17.9		
51-100	2	5.7	2	5.1		
101-200	4	11.4	2	5.1		
201-500	3	8.6	2	5.1		
More than 500	13	37.1	19	48.7		
<b>Education</b>					7.803	<.10
High school graduate	9	25.7	2	5.1		
Some college	8	22.9	8	20.5		
Associate degree	16	45.7	23	59.0		
Bachelor's or Master's degree	2	5.7	5	12.8		
PhD and other professional degrees	0	0.0	1	2.6		
<b>Corporate Experience</b>					0.357	<.84
5 years or less	12	34.3	11	28.2		
6-10 years	6	17.1	8	20.5		
More than 10 years	17	48.6	20	51.3		
<b>Type of Company</b>					0.117	<.74
Publicly traded	13	37.1	13	33.3		
Not publicly traded	22	62.9	26	66.7		

## 4. Findings

First, we examined response rates to the two survey scale versions. Response to the slider scale was submitted by 35 (out of 58) managers resulting in a response rate of 60%. For the radio-button scale, 39 (out of 62) managers responded resulting in a response rate of 63%. The difference between these two proportions is not statistically significant ( $z = 0.288, p < .78$ ).

A second index of data quality is completion time. Respondents in the slider treatment averaged 5.93 minutes (range from 1.92 to 49.85 minutes) to complete the survey while

respondents to the radio-button treatment took an average of 4.95 minutes (range from 0.15 to 2.00 minutes). The difference is not statistically significant ( $t = 0.593, p < .56$ ).

Third, we examined differences in the raw mean scores obtained from the two scale formats. Table 2 presents the results of measures of topic sensitivity and topic importance for the 12 topics of relevance to business managers. For topic sensitivity, all mean values are above the midpoint of the scales across both response formats with the radio-button scale generating higher scores for nine scale items, none of which were statistically significant. Only two topics had different sensitivity scores where the difference was statistically significant ( $p < .05$ ) with the radio-button format generating higher mean sensitivity scores.

A somewhat similar pattern emerged for topic importance. All mean values were above the midpoint of the scale with higher values recorded in the radio-button treatment for seven topics, but here again, none were statistically significant. Only three importance measures generated different mean interest scores where the difference was statistically significant ( $p < .09$ ), all of which were greater for the radio-button response format.

Table 2. Mean Values of Topic Sensitivity<sup>a</sup> and Topic Importance<sup>b</sup>

Topic	Slider Scale	Radio-Button Scale	t	p
<b>Sensitivity</b>				
Government actions affecting my company	3.706	3.692	.059	<.95
Market orientation	3.571	4.077	2.122	<.05
Government behavior regarding business	3.735	3.667	.280	<.79
Excellence in business	3.412	3.737	1.166	<.26
Assessment of employer/supervisor	3.697	3.949	1.072	<.29
Person values overall	4.061	4.308	.988	<.34
Organizational culture	3.794	3.923	.510	<.62
Competitors' behavior	3.441	4.000	2.211	<.04
Person behavior in doing my job	3.886	4.051	.618	<.55
Computer security behaviors	3.314	3.436	.459	<.65
Personal business ethics	4.000	4.077	.300	<.76
Social issues	3.543	3.513	.116	<.91
<b>Personally Important</b>				
Personal values overall	4.371	4.179	.936	<.36
Assessment of employer/supervisor	3.857	3.948	.447	<.66
Competitors' behavior	3.771	3.948	.863	<.40
Social issues	3.97	3.795	.878	<.39
Government actions affecting my company	3.657	3.384	1.080	<.29
Market orientation	3.885	4.236	1.887	<.06
Excellence in business	3.8	4.153	1.805	<.08

Personal business ethics	4.343	4.231	.612	<.55
Organizational culture	3.914	4.077	.840	<.41
Computer security behaviors	3.485	3.513	.106	<.92
Personal behavior in doing my job	4.057	4.487	2.087	<.05
Government behavior regarding business	3.628	3.513	.477	<.64

Notes.

<sup>a</sup>Scale ranged from 1 (*Not sensitive*) to 5 (*Extremely sensitive*). Sample size varied from 33 to 35.

<sup>b</sup>Scale ranged from 1 (*Not at all interesting*) to 5 (*Extremely interesting*). Sample size varied from 38 to 39.

Lastly, we compared the two response formats in terms of respondent likelihood to participate in a survey for five potential topics randomly selected from the 12 we initially presented to respondents. Results are shown in Table 3. There were no statistically significant differences ( $p > .15$ ) in respondent likelihood to participate based on response format style.

Table 3. *Mean values of likelihood of survey participation for response formats<sup>a</sup>*

Topic	Slider Scale <sup>b</sup>	Radio-Button Scale <sup>c</sup>	<i>t</i>	<i>p</i>
General behavior regarding business	3.062	2.846	.781	<.44
Assessment of employer/supervisor	3.647	3.743	.393	<.70
Personal values overall	3.818	3.692	.501	<.62
Personal business ethics	3.970	3.615	1.445	<.16
Personal behavior in doing my job	3.666	3.948	1.021	<.32

Notes.

<sup>a</sup>Scale ranged from 1 (*Very unlikely*) to 5 (*Very likely*).

<sup>b</sup>Number of respondents varied from 32 to 34.

<sup>c</sup>Number of respondents is 39.

## 5. Discussion

Toepoel et al. propose that investigating design choices in Web surveys may be more important than for other modes of survey administration, simply because of the many tools available and the potential variations in how these tools are utilized by researchers (Toepoel, Das, & Van Soest, 2008, p. 988). Couper (2000) advises that more work is needed to investigate optimal response design choices for Web surveys, especially across different populations and across increasingly diverse Web-based survey platforms. Clearly, the availability of question design choices available to survey researchers online has proceeded at a faster pace than systematic studies designed to examine the impact of these new technologies on data quality.

Researchers anxious to engage survey-weary respondents are eager to utilize more interactive, visually-engaging measurement tools available in Web-based surveys. But does the use of these response formats really make Web surveys more engaging for respondents?

Overall, the empirical evidence to date suggests that respondents may enjoy surveys with slider scales; however, sliders may require more time and cognitive effort. The higher break-off rates that some studies have found suggest that sliders work best with more highly educated samples of respondents who are less likely to become confused by non-typical response formats, and who may find them more cognitively engaging (Funke, Reips, & Thomas, 2011). It is also important to keep in mind that ability of respondents to answer VAS formats may depend on the particular hardware and software configurations used by the respondents, including whether working with a mouse or touch screen. More research is needed in this area, as technological issues can contribute to confusion and frustration on the part of survey respondents (Benfield & Szlemko, 2006). Based on the findings we reviewed in this paper, we agree with researchers (e.g., Puleston, 2011; Sikkell, Steenbergen, & Gras, 2014) who have concluded that sliders can be more engaging but these should be used sparingly, keeping in mind the population to be surveyed and technological issues that may be encountered.

In this article, we articulated the two central arguments supporting slider scales and other visual analog scale (VAS) formats versus traditional radio-button scale formats: (a) sliders may be more fun and engaging for survey respondents and (b) sliders may produce comparable, or even superior data compared to traditional Likert-type radio-button scale formats. We reviewed the empirical evidence for both arguments. We then presented findings from our own exploratory field experiment, set up to compare slider and radio-button formats in a specific survey setting. We found no strong evidence in favor of the two central arguments. Comparing between slider and radio-button formats, we obtained statistically inconclusive results with regard to response rate, completion time, and data quality. We did have a small number of variables that produced a statistically significant difference between the data captured through the slider and radio-button formats, with the radio-button format yielding higher mean scores. These findings are similar to those reported by others (Couper et al., 2006; Couper, 2008, p. 124) who have directly compared slider scale formats to traditional radio-button formats.

It is important to note that the few instances in which we found statistically significant differences do not imply that these differences are also practically significant. Most differences were very small, even those that were statistically significant. This last statement is made to heed the warning by Tromovitch (2015) to ensure that statistical significance is not misrepresented as indicating practical significance.

Further research is needed to examine the effects of interactive Internet-based survey scales compared to traditional scale formats among different populations, such as “fresh” versus “trained” online panelists (Toepoel, Das, & Van Soest, 2008). Toepoel et al. report that trained panelists (i.e., those who complete surveys once a month or more) are more sensitive to the time it takes to complete surveys and are more prone to engage in satisficing behaviors than fresh panelists (i.e., new recruits with little or no experience). Whether use of sliders can counteract satisficing tendencies and encourage more thoughtful responses on the part of online panelists, while maintaining a high level of participation, is an important issue that warrants deeper empirical investigations.



Based on the empirical evidence we describe in this article, slider scales ought to include clear instructions, specifically those that inform respondents that the bar must be moved in order to register their response. This may reduce missing data, a problem commonly associated with use of sliders. We also recommend that when appropriate, slider scales incorporate a “Don’t know” response (e.g., when knowledge is required) or a “Prefer not to answer” option (e.g., for potentially sensitive information). Other formatting issues include the number of scale points offered, use of labels, and choice of comparative versus non-comparative scale designs.

It is incumbent upon survey researchers to turn their attention to the myriad of seemingly innocuous online survey design choices, such as the use of VAS or traditional scale formats, and the impact these decisions have upon respondents’ survey experience, participation, and data quality. Some things are, however, less certain—the jury is still out regarding why, how, and when to use these new measurement tools. The rapid pace of technological advances will certainly dictate and shape future discussions, and researchers must pay close attention to a rapidly changing milieu.

## References

- Bayer, L. R., & Thomas, R. K. (2004, August). *A comparison of sliding scales with other scale types in online surveys*. Paper presented at the RC33 Sixth International Conference on Social Science Methodology, August 16-20, 2004, Amsterdam, Netherlands.
- Benfield, J. A., & Szlemko, W. J. (2006). Internet-based data collection: Promises and realities. *Journal of Research Practice*, 2(2), Article D1. Retrieved from <http://jrp.icaap.org/index.php/jrp/article/view/30/51>
- Brick, J. M. (2011). The future of survey sampling. *Public Opinion Quarterly*, 75(5), 872-888.
- Buskirk, T. D., & Andrus, C. H. (2014). Making mobile browser surveys smarter: Results from a randomized experiment comparing online surveys complete via computer or smartphone. *Field Methods*, 26(4), 322-342.
- Cape, P. (2009, February). *Slider scales in online surveys*. Paper presented at the CASRO Panel Conference, February 2-3, 2009, New Orleans, LA. Retrieved from [http://www.surveysampling.com/ssi-media/Corporate/white\\_papers/SSI-Sliders-White-Pape.image](http://www.surveysampling.com/ssi-media/Corporate/white_papers/SSI-Sliders-White-Pape.image)
- Cook, C., Heath, F., Thompson, R. L., & Thompson, B. (2001). Score reliability in web- or Internet-based surveys: Unnumbered graphic rating scales versus Likert-type scales. *Educational and Psychological Measurement*, 61(4), 697-706.
- Couper, M. P. (2000). Web surveys: A review of issues and approaches. *Public Opinion Quarterly*, 64(4), 464-494.

- Couper, M. P. (2008). *Designing effective web surveys*. New York, NY: Cambridge University Press.
- Couper, M. P., Conrad, F. G., & Tourangeau, R. (2007). Visual context effects in web surveys. *Public Opinion Quarterly*, 71(4), 623-634.
- Couper, M. P., & Miller, P. V. (2008). Web survey methods: Introduction. *Public Opinion Quarterly*, 72(5), 831-835.
- Couper, M. P., Tourangeau, R., Conrad, F. G., & Singer, E. (2006). Evaluating the effectiveness of visual analog scales: A web experiment. *Social Science Computer Review*, 24(2), 227-245.
- Couper, M. P., Tourangeau, R., Conrad, F. G., & Zhang, C. (2012). The design of grids in web surveys. *Social Science Computer Review*, 31(3), 322-345.
- Couper, M. P., Tourangeau, R., & Kenyon, K. (2004). Picture this! Exploring visual effects in web surveys. *Public Opinion Quarterly*, 68(2), 255-266.
- Derham, P. A. J. (2011). Using preferred, understood or effective scales? How scale presentations effect online survey data collection. *Australasian Journal of Market & Social Research*, 19(2), 13-26.
- Dobronte, A. (2012, August 21). *Likert scales vs. slider Scales in commercial market research*. Retrieved June 27, 2015, from [https://www.checkmarket.com/2012/08/likert\\_v\\_sliderscales/](https://www.checkmarket.com/2012/08/likert_v_sliderscales/)
- Downes-Le Guin, T., Baker, R., Mechling, J., & Ruyle, E. (2012). Myths and realities of respondent engagement in online surveys. *International Journal of Market Research*, 54(5), 613-633.
- Freyd, M. (1923). The graphic rating scale. *Journal of Educational Psychology*, 14, 83-102.
- Funke, F., Reips, U.-D., & Thomas, R. K. (2011). Sliders for the smart: Type of rating scale on the web interacts with educational level. *Social Science Computer Review*, 29(2), 221-231.
- Ganassali, S. (2008). The influence of the design of web survey questionnaires on the quality of responses. *Survey Research Methods*, 2(1), 21-32.
- Husser, J. A., & Fernandez, K. E. (2013). To click, type, or drag? Evaluating speed of survey data input methods. *Survey Practice*, 6(2), 1-7.
- Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., & Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, 50(1), 79-104.

- Lozano, L., Garcia-Cueto, E., & Muniz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology*, 4(2), 73-79.
- Miller, J. (2006). Online marketing research. In R. Grover & M. Vriens (Eds.), *The handbook of marketing research* (pp. 110-131). Thousand Oaks, CA: Sage.
- Puleston, J. (2011, March 14). *Sliders: A user guide*. Retrieved June 27, 2015, from <http://question-science.blogspot.com/2011/02/slider-how-to-use-them.html>
- Sellers, R. (2013). How sliders bias survey data. *Alert!*, 53(3), 56-57.
- Smith, S. M., & Albaum, G. S. (2013). *Basic marketing research: Building your survey*. Provo, UT: Qualtrics Labs.
- Sikkel, D., Steenbergen, R., & Gras, S. (2014). Clicking vs. dragging: Different uses of the mouse and their implications for online surveys. *Public Opinion Quarterly*, 78, 177-190.
- Stanley, N., & Jenkins, S. (2007). Watch what I do: Using graphical input controls in web surveys. In M. Trotman, T. Burrell, L. Gerrard, K. Anderton, G. Basi, M. Couper, . . . A. Westlake (Eds.), *The challenges of a changing world* (Proceedings of the Fifth International Conference of the Association for Survey Computing, pp. 81-92). Berkeley, UK: Association for Survey Computing.
- Taylor, I. (2012, June 8). *Use slider scales for a more accurate rating*. Retrieved June 27, 2015, from <https://blog.questionpro.com/2012/06/08/use-slider-scales-for-a-more-accurate-rating/>
- Toepoel, V., Das, M., & Soest, A.V. (2008). Effects of design in web surveys: Comparing trained and fresh respondents. *Public Opinion Quarterly*, 72(5), 985-1007.
- Tromovitch, P. (2015). The lay public's misinterpretation of the meaning of 'significant': A call for simple yet significant changes in scientific reporting. *Journal of Research Practice*, 11(1), Article P1. Retrieved from <http://jrp.icaap.org/index.php/jrp/article/view/477/411>
- Vicente, P., & Reis, E. (2010). Using questionnaire design to fight nonresponse bias in web surveys. *Social Science Computer Review*, 28(2), 251-267.

---

Received 16 March 2015 | Accepted 24 June 2015 | Published 9 July 2015

Copyright © 2015 *Journal of Research Practice* and the authors