Respondent-Generated Intervals: Do They Help in Collecting Quantitative Data?

Dominic Lusinchi, Far West Research, San Francisco, California

KEYWORDS: anchoring, item nonresponse, measurement/response error, quantitative data, recall, uncertainty, web-based survey research

Background

The respondent-generated-interval (RGI) technique was devised to collect quantitative data ("How many electrical outlets are there in your residence") in survey research. This response protocol, promoted by statisticians Press and Tanur[1], promises to have a positive impact on survey data quality, both by reducing item nonresponse and by improving the accuracy of estimators. This paper reports on an experiment, designed as part of a survey, to test the potential impact of RGI on item nonresponse. The survey, a self-administered web-based scroll-type questionnaire, measured ninety-seven open-ended quantitative response items all related to the development of an integrated circuit. Each respondent described a unique integrated circuit design that was completed or was near completion in 2002.

Research Methodology

In its orthodox format, as implemented by Press and Tanur, RGI asks respondents to provide not only a point estimate but also the lower and upper bounds of an interval that is most likely to capture the quantity of interest.

In our experiment, we used a truncated version of the RGI protocol: respondents (engineers) in the treatment group were encouraged (in the form of instructions before each question) to provide an interval. The exact wording (in red font) went: "IF you are unsure of the exact number requested (*e.g.*, gates, bytes, transistors, clock speed, etc.), please provide an interval (*e.g.*, "5-10", meaning, "between 5 and 10") in which you believe the number is most likely to fall. Thank you."

The control group received no such encouragement, and only saw the question. Therefore, for respondents in the treatment group interval generation was purely voluntary, not "hardwired" into the questionnaire itself. The respondents were provided with one field area to enter their answer. Thus, both the nature of the instructions and the response-entry format did not require respondents to provide a point estimate along with an interval but one or the other.

The respondents were randomized to the treatment and control groups based on the time they accessed the web survey page. The sample size that we ended up with, and, that we used for the purpose of this report is $n = 260$: $n = 144$ for treatment group, and $n = 116$ for the control group. The data we analyze in this paper are the raw results of the survey: *i.e.*, before the quality review process. For the purpose of this analysis we looked at 80 of the 97 quantitative variables.

Descriptive Results

Table 1 RGI Usage by Experimental Group

RGI Usage	Experimental Groups: Treatment	Control
YES	70	6
NO	74	110
Total	144	116

$\chi^2 = 58.602$, $df = 1$, $p < .01$

Seventy-six respondents (see Table 1) provided a RGI (~30%) to one or more response item. Ninety two percent (70/76) of those respondents were in the treatment group (*i.e.*, respondents who were assigned to the questionnaire in which they were encouraged to provide a RGI). Requesting respondents to provide a RGI clearly had an effect, as almost half in the treatment group made use of this response protocol at one time or another during the course of the survey. Without the RGI prompting, only 5% (6/116) voluntarily answered one or more response item with an interval: the vast majority provided a point estimate. It should be noted that a small minority of respondents provided an open interval – for lack of a better description – to one or more response items. By this I mean that they used some kind of symbol next to the number they reported: for example, "30+", ">5", or "~250". In all, one-third of all respondents provided, in at least one response item, an

[1] S. James Press is Distinguished Professor of Statistics, at the University of California, Riverside. Judith M. Tanur is Distinguished Teaching Professor of Sociology, at the State University of New York, Stony Brook.

answer other than a point estimate, and of those, one-sixth did so without prompting. We conclude that the treatment convinced approximately 43% of that group to use RGI – without the treatment these respondents would otherwise not have done so.

Next we take a look at the rate of RGI usage. We examine the prevalence of this response format from two related perspectives. First, from the user side: how often is RGI used when giving an answer? In other words we measured the number of items in which RGI users provided a response using that protocol. Second, from the response item side: how many answers come in that format? Thus, for each response item, we counted the number of RGI answers.

RGI is not a protocol that is relied on intensively by users. The modal category for the usage rate is comprised of respondents who used RGI on three to ten percent of response items they provided answers to. Nearly two-thirds of respondents only used RGI in less than 10% of response items. The median value of this distribution is 7% and the mean is 12%. The rest of the time, which means most of the time, RGI users rely on point estimates.

We then examined the 80 variables used in this analysis, to determine how often the RGI protocol was used as a response format. Ninety-one percent of response items had at least one response in RGI format. However, on average, only about 6% of answers to a response item were respondent-generated intervals (median ~3%). In other words, the vast majority of response items have few answers in RGI format.

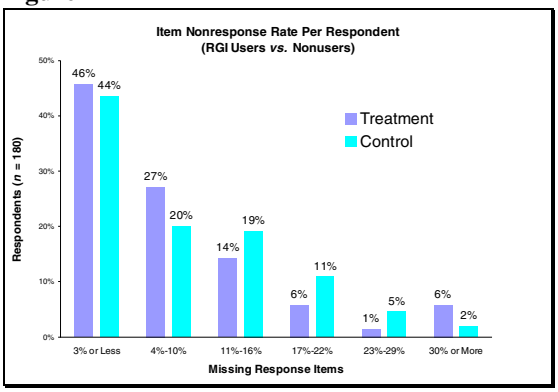
Experimental Results

The issue we wished to test experimentally was item nonresponse: does RGI (as we implemented it) help in reducing the rate of item nonresponse in a self-administered survey?

Again, we looked at item nonresponse from two different but related vantage points. First, we examined how respondents from both experimental groups responded: do RGI respondents from the treatment group have lower rates of nonresponse than nonusers in the control group? Second, we looked at the response items themselves: *i.e.*, do question items in the two experimental groups differ in terms of nonresponse.

Our research hypothesis was this: Do RGI respondents in the treatment group have a lower rate of nonresponse than nonusers in the control group? If this were so, we would expect the probability distribution of the control group to be above that of the treatment group. The null hypothesis is that the distributions of both groups are identical. As Figure 1 demonstrates, we cannot reject the null hypothesis, and we are led to conclude that no difference can be detected between the two groups of respondents (based on rank-sum test for large samples, $z = -.755$, $p > .05$, one-tail).

Figure 1



Clearly, RGI users do not appear to answer any more often than nonusers.

To analyze per item response rate, we selected the response items in which RGI was used: this left $n = 73$ response items to examine. For each of these we calculated a nonresponse score (number of nonrespondents/number of eligible respondents): each response item in the treatment group is paired with the same item in the control group. We then compared each response item as it was answered by the treatment to the same item in the control group.

We hypothesized that the rate of nonresponse would be lower among response items in the treatment group of RGI users than in the control group of nonusers. To evaluate this hypothesis, we used a signed-rank test (large sample), and a one-tailed rejection region. If the difference is in the direction that we expect, the smaller of the sums of the like-signed ranks will be the sum of the ranks from those items in the treatment group (positive ranks).

Nearly two-thirds of items in the control group have a rate of nonresponse rate of 8% or less, against only 39% of response items in the treatment group. In only 19 of the 73 items (see Table 2, next page) was nonresponse lower in the treatment group than in the control group. On the other hand, in 53 items of the 73 we examined, the response rate was higher in the control group than in the treatment group. This evidence goes against what was anticipated by our research hypothesis.

Clearly, the null hypothesis of no difference between the response items from the two experimental groups is not tenable: in fact, the results are in the opposite direction of what we originally hypothesized. We predicted a higher response rate in items from the treatment group, but we observed that very result from the control group. Table 2 details the differences between the set of response items from the two experimental groups. Both measures of central tendency show a higher rate of items nonresponse in the treatment group.

Table 2 Nonresponse Rate per Item by Experimental Group

Experimental Group	Items with Lower Nonresponse*	Statistics	
		Mean	Median
Treatment (RGI users)	19 (26%)	20.0%	13%
Control (RGI nonuser)	53 (73%)	15.6%	6%

z = -3.625 (based on 19 positive ranks: RGI nonusers > RGI users)

* There was one response item where nonresponse was the same for both groups.

Thus, these results complement what we learned earlier: we see no evidence that the RGI protocol, as implemented, helps in decreasing the nonresponse rate.

The data from this experiment, using two related measures to evaluate nonresponse, lead to one conclusion: the RGI protocol did not reduce the nonresponse rate in this survey. It is important to emphasize that this conclusion is only valid within the confines of this experiment and the methodology of this survey, *i.e.*, a self-administered survey instrument in which respondents in the treatment group were simply encouraged to use the RGI procedure. In other words, this hypothesis (*i.e.*, that RGI may help to increase item response rates) still needs to be tested experimentally within different modes (*e.g.*, telephone, face-to-face), and, perhaps, within the self-administered mode but implemented differently. Clearly, the way RGI was implemented in this experiment does not appear to be effective in reducing nonresponse.

Discussion

Press and Tanur conjectured that item *non*respondents might have been item *respondents* had they been given the RGI protocol as an alternative to the traditional request for a point estimate. The data from our experiment do not support this hypothesis: respondents will answer a question item or refrain to do so, regardless of RGI.

What our results show is that, left to their own devices, which is the essence of a self-administered survey, most respondents, including RGI users themselves, will rely on point estimates as their preferred response format.

Three reasons immediately come to mind for this preference for point estimates: (1) for what respondents feel is expected of them and given the amount of cognitive resources they are willing to dedicate to a survey, a point estimate is "good enough"; (2) providing an interval is not a common task, be it in our private or professional activities, we are rarely, if ever, required to give an interval for a requested quantity; (3) providing a point estimate is an easier thing to do: if we assume that RGI increases the burden on respondents, answering with a point estimate can be seen as a form of satisficing which allows respondents to provide an "acceptable" answer that requires less effort than RGI.

If participants to a self-administered survey do not use RGI as an alternative to nonresponse, then how, when and why is it used? Here are some possible hypotheses.

Nonresponse across all items in the survey stood at 18.6%. This relatively high rate was due largely to 17 items (out of 73) whose nonresponse rate was around 50%. Interestingly, these items also had a higher average rate of RGI responses than other items (19% *vs*. 4%, respectively).

This would suggest that RGIs are used in response items that have a very high degree of uncertainty (assuming nonresponse measures that); or, more precisely, whose level of uncertainty exceeds the "good enough" standard which is met by the more commonly used point estimate. However, it should be added, this is so, if respondents are instructed that self-generated intervals are a legitimate response format.

This, in turn, would indicate that RGI is unlikely to be adopted as a generalized response strategy by respondents to a self-administered survey, but will be used on a case-by-case basis, so to speak. In other words, RGI usage appears to be item dependent: if respondents believe the value requested to be unusually difficult to access, they will use RGI as the preferred format, because they believe it is better suited to the situation than a point estimate. Thus, RGI is a preferable alternative to a point estimate, given the right

circumstances, not a preferable alternative to nonresponse.

Conclusion

As this experiment clearly demonstrates, RGI is not a response format that respondents will readily adopt: roughly half of them will use it at some time during the survey, and they will use it, on the average, in only ~12% of the items they answer.

The major problem with RGI is its implementation in a self-administered survey. Web-based surveys give the data collector the ability to interact with the respondent. Certainly, one could provide fields for both the lower and upper bounds of the quantity requested, and a third for the best estimate; include pop-up warning messages when the respondent leaves two or more blank fields in a given question item, *etc.*

All of this may have the effect of increasing the proportion of respondents that will rely on RGI as a response format. But it is likely to increase the respondent's burden and task the respondent's patience.

Based on the evidence from this experiment, I do not see that RGI will have much of an impact on item nonresponse when the mode of data collection is self-administered. It might be possible to increase the proportion of respondents that use this protocol, and, concomitantly, increase the proportion of response per item that are in RGI format. No doubt we will need to weigh the risk of increased burden and page "busyness" required to implement RGI. In the end, as its very name indicates, respondents to a self-administered mode are (mostly) on their own, and there is only so much the survey researcher can do.

References

1. Beatty, Paul, Douglas Herrmann. "To Answer or Not to Answer: Decision Processes Related to Survey Item Nonresponse", in Groves, Robert M., Don A. Dillman, John L. Eltinge, Roderick J.A. Little, *Survey Nonresponse*, New York, NY, John Wiley & Sons, Inc., 2002.

2. Fowler, Floyd J., Jr. Improving Survey Questions: Design and Evaluation, Thousand Oaks, CA, Sage Publications, Inc. 1995.

3. Marquis, Kent H., S. James Press. "Cognitive Design and Bayesian Modeling of a Census Survey of Income Recall", http://www.fcsm.gov/events/papers1999.html.

4. Mathiowetz, Nancy A. "Response Error: Correlation between Estimation and Episodic Recall Tasks", *Proceedings of the Survey Research Methods Section of the American Statistical Association*, 1987

5. Press, S. James, "Respondent-Generated Intervals for Recall in Sample Surveys", (a) Oct. 2000, and (b) Jan. 2003.

6. Press, S. James, Judith M. Tanur. "Respondent-Generated Interval Estimation to Reduce Item Nonresponse", 2000.

7. Press, S. James, Judith M. Tanur, "Experimenting with Respondent-Generated Intervals", JSM August 1999.

8. Press, S. James, Kent H. Marquis. "Bayesian Estimation in a U.S. Government Survey of Income Recall Using Respondent-Generated Intervals", May 2000.

9. Tanur, Judith M., S. James Press, Diane Miller, and Diana Petitti, "A Controlled Experiment for Testing Respondent-Generated Intervals Methodology in Surveys", Proceedings of the Survey Research Methods Section of the American Statistical Association, 2000

Contact

Dominic Lusinchi
Principal consultant
Far West Research
dominic@farwestresearchh.com