Obtaining Purchase Probabilities via a Web Based Survey: The Juster Scale and the Verbal Probability Scale

Mathew Parackal and Mike Brennan

The purpose of this research note is to report the findings of a study that compared two purchase probability scales in a Web based survey: a pull-down form of the Juster Scale, and a written version of the Verbal Probability Scale. Although the response rate to the survey was rather low, limiting the conclusions that can be drawn from the experiments, the results do identify important issues regarding the use of these scales in Web surveys.

Keywords: purchase probabilities, Juster Scale, Verbal Probability Scale

Introduction

The unique characteristics of the World Wide Web, and the rapidly increasing importance of the Web for both communication and commerce, provide unique opportunities for Web based survey research. However, while a great deal is known about questionnaire design and layout with regard to conventional surveys, it is not certain that these same methods will work on the Web.

Conventional questionnaires are text based, even in telephone and face-to-face surveys. But, given that the Web typically utilises sophisticated graphics, colour and animation, and can use sound and video as well as text, it is not at all clear that a text based questionnaire would be suitable. Furthermore, the Web requires the use of keyboard, mouse or touchscreen to elicit responses, rather than a pen, pencil, or simple spoken response, and this has implications for the way questions and scales are presented.

One of the exciting features of the Web is that it can make use of multimedia applications, including virtual reality. An increasing number of sites are utilizing virtual reality software to allow visitors to inspect three dimensional models of products, such as cars, cameras, and medical equipment, to name a few (Urban, Hauser, Qualls & Weinberg 1997). This technology opens up possibilities for research, as it would be a relatively easy matter to manipulate product characteristics (colour and design and price, for example) to determine the best mix (see Needle 1995, 1996; Burke 1997).

To estimate demand for the variants, given that actual sales data may not be an option, an instrument such as the Juster purchase probability scale (Juster Scale) could be used. This scale, developed by Thomas Juster (Juster 1966), has been used quite successfully to estimate demand for a range of products and services, including durables, fmcgs, and even specific brands (Juster 1966; Day, Gan, Gendall & Esslemont 1991; Hamilton-Gibbs; Esslemont & McGuinness 1992; Seymour, Brennan & Esslemont 1994; Brennan, Esslemont & U 1995). Forms of the scale have been used in self-completion questionnaires (Gendall, Esslemont & Day 1991) and telephone surveys (Brennan, Esslemont & Hini 1995).

While the printed form of the Juster scale can easily be included in a Web based questionnaire, only a limited amount of information can be displayed on a computer screen at one time. Scrolling up or down pages to view the scale when it has disappeared off the screen

would be tedious, as would repeated presentations of the scale. A more compact form of the scale is needed for Web applications. One option is to present the standard scale as a pull down menu. Another is to use a different use a different form of the scale altogether.

The purpose of this research note is to report the findings of a study that compared two purchase probability scales in a Web based survey: a pull-down form of the Juster Scale, and a written version of the Verbal Probability Scale. Although the response rate to the survey was rather low, limiting the conclusions that can be drawn from the experiments, the results do identify important issues regarding the use of these scales in Web surveys.

Method

Respondents were recruited via a letter sent with a newsletter mailed by a local Internet provider to all of its clients. The letter briefly outlined the purpose of the study and encouraged people to participate. Those willing to do so were directed to a Web site. The letter informed readers that all participants would be entered into a prize draw for $200 worth of products or services from the Internet provider. The URL for the survey site was posted on the Internet provider's home page for the duration of the study. No reminder letters or emails were used, on the insistence of the Internet provider.

The questionnaire comprised three web pages. The first page required respondents to complete some demographic questions, and included a reminder about the prize draw. The second page asked respondents to indicate, using a purchase probability scale, the likelihood that they would choose each of five alternative billing options. The third page asked them to indicate, again using a purchase probability scale, the likelihood of using two services ("in-shop" or "on-site" help) at a particular price.

Three separate Web pages were needed for the questionnaire to allow respondents to be allocated to different experimental treatments. This was accomplished by using CGI (Common Gateway Interface) scripts. Respondents who completed and submitted the first page of the three-page questionnaire were assigned to one of two treatment groups, each of which was exposed to a different version of the purchase probability scale used on the second page. On submitting the second page, respondents were assigned to one of four treatment groups, each of which was exposed to a different version of the pricing options listed on the third page.

On submitting the third page, respondents were sent to the Marketing Bulletin homepage. This displayed one of four versions of a banner ad, as part of another study (see Rae & Brennan 1998). The research design is shown in Figure 1.

Instruments

The study compared two forms of purchase probability scale: The Juster Scale, and the Verbal Probability Scale.

The Juster Scale

The Juster Scale (Juster 1966) is an eleven-point scale from 0 to 10. Each point on the scale has a numerical and written description attached to it (see Figure 2).

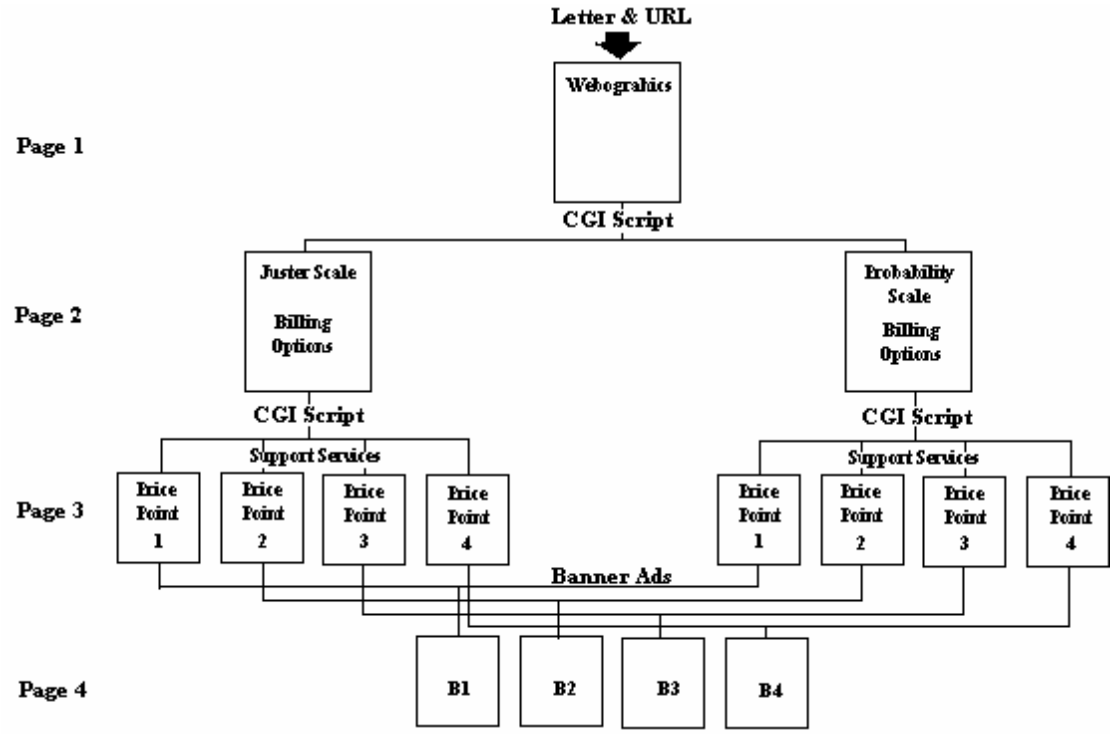


Figure 1. Experimental Design

The Juster Scale was presented as a drop down menu after each probability question. Respondents clicked on an arrow on the right hand side of an empty answer box to cause the menu to drop down, then clicked on the option they wished to select from the scale.

The first purchase probability question concerned the five billing options. These were described in the questionnaire at the top of the page. The following directions for using the scale were provided:

> "We would like to know what the chances are of you choosing **each** of these five options. For each option, please select an answer from the pull-down scale provided.
>
> **If you are certain, or practically certain that you would choose the option then you should choose the answer '10'. If you think there is no chance or almost no chance of choosing the option, the best answer would be '0'. If you are uncertain about the chances, choose an answer as close to '0' or '10' as you think it should be**."

Respondents were then asked questions of the form:

"Taking everything into account, what are the chances that you would select **OPTION 1?"**

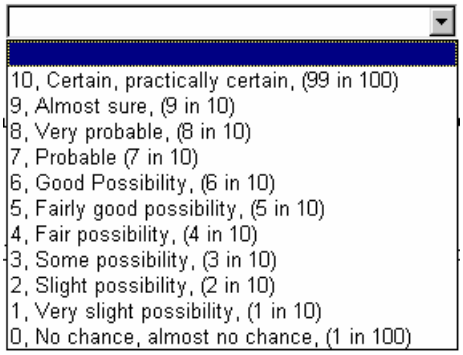


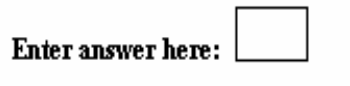
Figure 2. The Juster Scale presented as a drop down menu

Verbal Probability Scale

The Verbal Probability Scale is an eleven-point scale with values ranging from 0 to 10. It has no graphical form; it is a text version of the spoken Verbal Probability Scale (Brennan, Esslemont & Hini, 1995) developed for use in telephone surveys. Respondents were simply given the following direction:

> "We would like to know what the chances are of you choosing **each** of these five options. For each option, please give an answer between '0' and '10'.
>
> **If you are certain, or practically certain** ….<as for Juster Scale>

The respondents indicated their probability by typing their answer (a numeral between 0 and 10), into the space provided after each probability question.



Results

Scale Comparisons

In the first task, respondents were asked to give the probability of choosing each of the five billing options.

The results highlight a difficulty with this type of question. Implicit in the technique is the expectation that respondents will assign probabilities across the five options in such a way that the sum of the probabilities will equal 1, but this clearly did not happen. In reality, a respondent would have to adopt only one of the five options. But in this task, the respondents appear to have treated the options independently, that is, as if the other options did not exist when the probability of adopting a particular option was given. As a consequence, the sum of the probabilities often far exceeded 1. To estimate the proportion of the sample that would adopt each option, the probabilities had to be weighted.

To weight the probabilities, the probability for each option was divided by the sum of the probabilities across all five options. This was done separately for each respondent, before the

mean probability for each option, across all respondents, was computed. It is these mean probabilities that provide the weighted estimates of the purchase rates for the five options. Both the unweighted and weighted probabilities for the five options are reported in Table 1.

Table 1. Comparison of the unweighted and weighted mean probabilities for the two purchase probability scales

	Juster Scale		Verbal Probability Scale	
	Unweighted	**Weighted**	**Unweighted**	**Weighted**
Option 1	0.80	0.23	0.24	0.10
Option 2	0.80	0.23	0.24	0.10
Option 3	0.46	0.14	0.64	0.40
Option 4	0.71	0.19	0.32	0.20
Option 5	0.75	0.20	0.35	0.20
	n = 84		n = 117	

While the weighting makes little difference to the rank order of the options, it does affect the interpretation of the results. For the Juster Scale data in particular, the unweighted probabilities suggest strong demand for four of the five options, whereas the weighted data suggests a much lower demand. Unfortunately, it is not possible to say which level is the most accurate, but the need to weight the data at all is cause for concern.

Since both the treatments estimated the adoption rate of the five billing options in the same population, the estimates produced by the two scales should be the same. However, the estimates produced by the two scales are very different, suggesting that the scales are not equivalent. This is in marked contrast to the findings of Brennan, Esslemont & Hini (1995), who developed the Verbal Probability Scale. Although the instructions to respondents in the original studies were spoken rather than written, the wording is identical. Brennan, Esslemont & Hini (1995) reported only minimal differences in purchases predictions using the Juster Scale (mailed out to respondents) and the Verbal Probability Scale (used in a telephone interview). The present result is therefore unexpected.

An examination of the probability distributions for the two scales confirms the conclusion that respondents are using the two scales differently. The mode for the Juster Scale for each option is typically 2, with only a small proportion of respondents choosing 0, whereas the mode for each option for the Verbal Probability Scale is 0.

Since the Internet provider did not introduce the billing options, it is not possible to say which scale provided the more accurate predictions. Until this validation is done, predictions undertaken using either of the scale, using the procedures used in this study, should proceed with caution.

Price - Demand Estimation

In order to construct simple demand curves, respondents were asked to give the probability of purchasing two services. A different price was used with each of four groups, so the estimates are independent. The results, shown in Tables 2 and 3, again show major differences in the estimates produced by each scale. For both services, the Juster Scale gave much higher probabilities than the Verbal Probability Scale, at each price.

Table 2. Comparison of the mean purchase probabilities for the "in-shop" service

In-shop service	Juster Scale			Verbal Probability Scale		
	Mean	SE	n	Mean	SE	N
@ $50 per hour	0.78	0.07	18	0.54	0.09	16
@ $70 per hour	0.82	0.05	18	0.30	0.10	10
@ $90 per hour	0.92	0.04	14	0.30	0.10	10
@ $110 per hour	0.96	0.02	10	0.30	0.10	6

Table 3. Comparison of the mean purchase probabilities for the "on-site" service

On-site service	Juster Scale			Verbal Probability Scale		
	Mean	SE	n	Mean	SE	N
@ $60 per hour	0.79	0.09	15	0.54	0.09	16
@ $80 per hour	0.83	0.03	15	0.40	0.10	9
@ $100 per hour	0.86	0.07	10	0.30	0.10	9
@ $120 per hour	0.93	0.02	9	0.70	0.30	5

By plotting the mean probability against the price levels, the demand schedules for the services can be estimated. Curiously, the demand schedules for the two services are in the opposite directions for the two treatments, and the demand schedules produced from the Juster Scale estimates are the inverse of what one would expect; they indicate an increase demand as the cost of the service increases. However, because of the small sample sizes, this result could be spurious and discarded. What is of note is that the level of demand estimated for each option is consistently higher for the Juster Scale than for the Verbal Probability Scale.

This result was either due to differences in the characteristics of the respondents using each version of the scale, although they were randomly assigned to each of the eight treatment groups, or due to differences in the way people use the scales. Since the demographic composition of the respondents did not differ for the two scales, it would appear that the scales are not equivalent. Again, since we are unable to validate the results, it was not possible to establish which scale performed best. But the fact that the two scales produced such different results deserves further examination.

Discussion

While there were no complaints or queries from respondents about the probability scales or their application, there were some indications of problems with the scales. In two cases respondents using the Probability Scale gave probability scores outside the 0 to 10 range of the scale. However, in both cases their probabilities for the remaining options were within range, suggesting the errors were due to carelessness rather than a lack of understanding of the scales or task. To prevent such mistakes, it may be preferable to use a drop down menu listing the options (0 to 10), which respondents click, rather than have them type in the answer.

Of course, this does not preclude the possibility of clicking the wrong answer. Perhaps a better approach is to provide an error message if an answer is out of bounds, and perhaps allowing, or requiring, the respondent to review the completed questionnaire before submitting this. Implementing such a procedure would be a fairly straightforward programming task, and indeed, there is now available sophisticated software specifically designed to deal with this, and related, problems (e.g., Green 1998).

A second concern relates to the weighting of probability scores in those situations when the sum of the scores across a range of items should equal 1. Clawson (1971) suggests that the probability scores reflect the degree of confidence respondents have in their purchase probability decision. Weighting the data may result in data that no longer reflects this degree of confidence, and thus may affect the accuracy of the predictions. As the accuracy of the predictions has not been established, it is unclear whether weighting is a sensible step or not. However, one can argue that it is necessary at least to ensure that the responses of each respondent make an equal contribution to the estimates. Otherwise, the results would be biased in favour of respondents whose summed ratings were highest.

It is self-evident that the purchase probability task must be clearly explained to respondents, and it may be that the question wording used in this study was not clearly understood. Thus, further research into ways of explaining the task clearly are a priority for future research. An alternative method that would avoid the need to weight the data would be to use a constant sum method (e.g., Hamilton-Gibbs, Esslemont & McGuinness 1992), and this needs to be developed and tested for use over the Internet.

References

Brennan M; Esslemont D & Hini D (1995). Obtaining purchase predictions via telephone interviews. *Journal of the Market Research Society*, *37* (3), 241-250.

Brennan M; Esslemont D & U C (1995). Using the Juster Scale to estimate the demand - price relationship. *Asia-Australia Marketing Journal*, *3* (1) 27-37.

Burke RR (1997). Do you see what I see? The future of virtual shopping. *Academy of Marketing Science*, *25* (4), 352-360.

Clawson JC (1971). How useful are 90 day purchase probabilities? *Journal of Marketing*, *35* (October), 43-47.

Day D; Gan B; Gendall P & Esslemont D (1991). Predicting purchase behaviour. *Marketing Bulletin*, *2*, 18-30.

Gendall P; Esslemont D & Day D (1991). A comparison of two forms of the Juster Scale using self completion questionnaires. *Journal of the Market Research Society*, *33* (3), 257-263.

Green TM (1998). *Surveys in cyberspace: six components of wen survey design*. Paper presented at the 53rd Annual Conference of the American Association for public Opinion Research, St Louis, M O, May 14-17.

Hamilton-Gibbs D; Esslemont D & McGuinness D (1992). Predicting the demand for frequently purchased items. *Marketing Bulletin*, *3*, 18-23.

Juster FT (1966). *Consumer buying intention and purchase probablity*. National Bureau of Economic Research, Columbia University Press.

Needle SP (1995). Marrying marketing research and virtual reality: implications for consumer research. http://www.simulationresearch.com/paper1.htm.

Needle SP (1996). Virtual reality and consumer research: the future is here today. http://www.simulationresearch.com/paper2.html.

Rae N & Brennan M (1998). The relative effectiveness of sound and animation in Web banner advertisements. *Marketing Bulletin*, *9*, 76-82.

Seymour P; Brennan M & Esslemont D (1994). Predicting purchase quantities: Further investigation of the Juster Scale. *Marketing Bulletin*, *5*, 203-226.

Urban GL; Hauser JR; Qualls, WJ & Weinberg BD (1997). Information acceleration: Validation and lessons from the field. *Journal of Marketing Research*, *34* (1), 143-153.

Acknowledgement: The authors wish to acknowledge the cooperation of New Zealand Post Direct Marketing Services and, in particular, Matthew Pickering.

Mathew Parackal was a postgraduate student, and Mike Brennan is a Senior Lecturer, in the Department of Marketing, Massey University.