

The Impact Of Different Scale Anchors On Responses To The Verbal Probability Scale

Originally published in the Canadian Journal of Marketing Research, Vol 20, No. 1 2002.

Keywords: Juster Scale, Verbal Probability Scale, Purchase Intentions scales, verbal anchors

John Dawes & Erica Riebe

Marketing Science Centre, University of South Australia

GPO Box 2471, Adelaide, 5001

Telephone: +61 8 8302 0592 Fax: + 61 8 8302 012

Email firstname.lastname@marketingscience.info

John Dawes and Erica Riebe are Senior Research Associates at the Marketing Science Centre.

John Dawes has a PhD from the University of South Australia. His research interests include service quality and satisfaction, pricing, and market research methodology. Erica Riebe's research interests include the use of the Juster & Verbal Probability scales to predict future behaviour, and the dynamics of brand switching. Erica is currently undertaking a PhD at the University of South Australia.

* we use the keyword 'purchase intentions' but note that asking about intentions is not the same thing as asking about likelihood of purchase. Many so-called purchase intentions scales actually use wording that pertains to a purchase probability.

The Impact Of Different Scale Anchors On Responses To The Verbal Probability Scale

Abstract

This paper examines how the mention of different scale "anchors" impacts on responses to probabilistic questions about future behaviour, administered via telephone. Three groups of respondents were each administered a different form of a probability scale. All forms were derived from the Verbal Probability Scale, and thus extend the work of Brennan (1994, 1995). In this paper, the three probability scale versions were: (1) where no mid point or scale anchors were mentioned other than 0 and 10; (2) where 5 was mentioned as a scale mid-point as well as 0 and 10; and (3) where 3 and 7 were used as anchors as well as 0 and 10. The results indicate that there was no statistically significant effect on the mean probability scores derived from the different versions of the probability scale. However, the version of the scale where 5 was mentioned as a mid-point, resulted in a spike in responses for that response category. This suggests that using the 5 as a scale anchor / mid point is undesirable as it distorts the distribution of responses. Thomas Juster, the original developer of the Juster scale, considered that such distortion is likely to result in less accurate predictions (Juster, 1960, Juster, 1966).

Introduction

In this paper we consider the effect of different scale point anchors on responses to probabilistic scales. Probabilistic scales have been shown to be more accurate predictors of future behaviour than attitudinal or intentions scales (Day et al., 1991), and as such are the focus of this paper. A large body of literature considers other scales that have been used for the purposes of predicting future behaviour, such as attitudinal and intentions measures, however, we focus here on probabilistic measures.

The Verbal Probability Scale is an eleven-point, zero-to-ten scale that asks respondents to allocate probabilities to their future behaviour. The scale is typically verbally anchored at the points zero, three, seven and ten. The Verbal Probability Scale is an adaptation of the Juster Scale (Juster, 1960, Juster, 1966), designed for use via telephone. Both the Verbal Probability Scale and the Juster Scale have been used in numerous studies in endeavours to predict future customer behaviour. Such studies have examined the accuracy of the scale in terms of its ability to forecast aggregate level buying behaviour such as penetration level and buying frequency (e.g. Juster, 1969, Brennan and Esslemont, 1994, Seymour et al., 1994); as well as studies at the individual level such as identifying customers who are more likely to defect from a particular service provider (e.g. Danenberg and Sharp, 1996a, Danenberg and Sharp, 1996b).

The use of the Verbal Probability Scale via telephone raises some specific issues for researchers, as it is recognised that administering scale-type questions via telephone can be problematic.

Telephone surveys differ from those that use self-completion questionnaires or those filled out by an interviewer face-to-face. For self completion or face-to-face methods that use scales, the respondent can often *see* the scale, either on the questionnaire itself or on a “show card” which

has the range of possible responses. In a telephone survey, the respondent cannot see the scale, and respondents have limited ability to recall response categories (Dillman, 1978). This is one of the appeals of the Verbal Probability Scale, in that the notion of providing a rating "out of ten" is presumably an easier task for respondents, than asking them to use unwieldy categorisations such as "very unlikely, unlikely," and so on. However, given that in telephone interviews, respondents must be guided solely by the verbal instructions of the interviewer, more information may result in better quality responses, than less information. This is a rationale for using additional instructions such as scale anchors. The question is - what effect do they have on the data gathered?

Previous Research

“Mid Points” are often used as additional scale anchors. Previous research on the effects of scale mid-points is both minimal, and provides mixed results. Worcester & Burns (1975) concluded that use of a scale without a mid-point “pushes more respondents to the positive end of the scale”, but this conclusion was based on the comparison between responses to Likert scale scores derived from respondents making marks on blank lines. Worcester & Burns (1975) did not report on any differences in the proportion of respondents that gave a particular response according to the presence or absence of a mid-point on the Likert scales.

Three other studies have examined how the inclusion of a mid-point alters the proportion of positive or negative responses obtained. Spagna (1984) found that respondents allocated fewer positive responses and more negative responses when a mid-point was included in a scale. Garland (1991) reported quite different results - respondents allocated fewer positive responses and more negative responses when a mid-point was not included in a scale. Dawes (2001) found

that the use of a scale mid point resulted in greater use of low responses (ie responses below the mean level) for customer satisfaction data. So the available evidence on the effect of scale mid-points is rather inconclusive. However, of these four studies, three used either mail or face-to-face survey methods for data collection, leaving only Dawes (2001) supplying evidence as to telephone methods. Furthermore, we could not find any previous study that has looked at the issue of scale mid-points or anchors for the Verbal Probability Scale. Since such scales are very often used to try to predict future behaviour, it is important to assess how well different methods of administration affect the responses. Hence the rationale for this study.

Method

A survey of companies was conducted in 2000 as part of a commercial market research project. In total, 332 firms were surveyed. The topic of the research was the likely future usage of a government organisation that provides assistance and funding to firms. Two types of firms were surveyed, being (a) firms that were previously identified as possible users of the government organisation, and (b) firms that were previously identified as possible referral points for other firms who might use the government organisation. 191 referral firms and 141 user firms were surveyed.

As part of this survey, three versions of the Verbal Probability Scale were used. The first version used no verbal anchors apart from zero and ten. The second version used the scale mid-point, 5, as an additional verbal anchor. The third version used the verbal anchors of 3 and 7. The descriptive statistics from the study are shown in Table 1. We see that there is little difference in response profile according to scale type. However, there is a small difference in mean response level according to the type of firm, with user firms having a higher average probability of using

the service than the average probability of referrers actually referring a client to the service (2.7 compared to 2.3 respectively).

Table 1: Descriptive Statistics

Results for user firms	Scale type	Mean	SD	N
	Only zero and ten as anchors	2.6	3.3	86
	Zero, five and ten as anchors	2.7	2.9	24
	Zero, three, seven and ten as anchors	3.1	3.5	30
	<i>Average</i>	2.7	3.2	
Results for Referral firms	Scale type	Mean	SD	N
	Only zero and ten as anchors	2.1	2.9	79
	Zero, five and ten as anchors	2.5	2.8	63
	Zero, three, seven and ten as anchors	2.2	2.9	46
	<i>Average</i>	2.3	2.8	

In order to ascertain whether the different scale anchors had any statistically significant effect on responses, we used a one way ANOVA. This was done for each of the two types of firms. The results are shown in Table 2 and Table 3. They show that there was no statistically significant difference between the response levels for the three scale types – for either user firms or for referral firms.

Table 2 Results for User firms

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. Of F
Between groups	20.2	2	2.4	.23	.79
Within groups	1449	137	10.5		
total	1454	139			

Table 3 Results for Referral firms

Source of Variation	Sum of Squares	DF	Mean Square	F	Sig. Of F
Between groups	3.7	2	1.8	.22	.80
Within groups	1507	185	8.1		
total	1511	187			

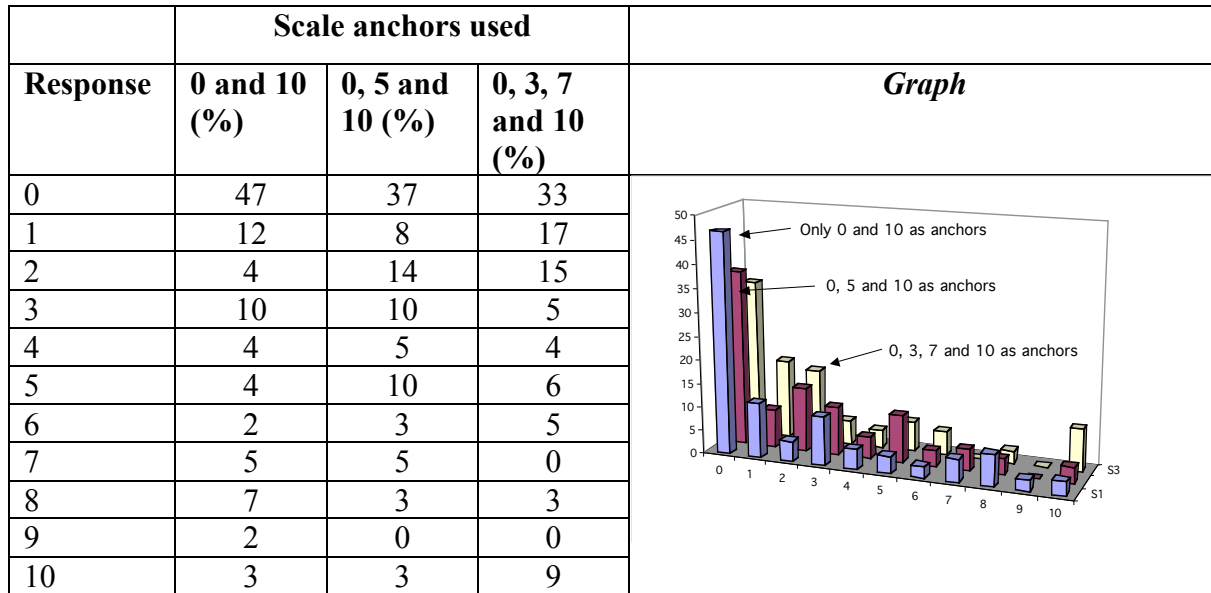
Given this result, and for purposes of illustration, we graphed the distribution of responses¹ to the three versions of the questionnaire, as shown in Figure 1. We see a spike for the response category 5 where that figure was used as a scale mid-point. This is consistent with the finding of Juster (1960) in his original work with the scale as administered in "pen and paper" format. Surprisingly, there were no corresponding spikes at 3 or 7 when those figures were used as anchors. There is also a spike for the response category 10 when the scale anchors 0, 3, 7 and 10 were used. This is likely to be simply due to chance, as there was no corresponding spike at either 3 or 7, the other non-zero anchors used for that version of the scale.

Therefore, it appears that the use of the response category 5 as a scale anchor did not have an effect on the mean average level of responses, but it did have an effect on the distribution of responses. More specifically, it appears to create an exception to the apparent "inverse J" shape

¹ Note that a two-way ANOVA (firm x scale type) showed no interaction effects, so a simple graph of responses according to scale type for the whole sample is valid.

of the data. This suggests that it may be preferable not to use the 5 mid point as an anchor, but rather, either just use 0 and 10 as anchors, or 0, 3, 7, and 10 as anchors.

Figure 1: Response distributions by type of scale anchors used



Conclusions

The different scale anchors that we used did not appear to have any association with the mean average score, but the shape of the distribution was altered slightly when the scale version with 5 as the mid point was used. This suggests that the Verbal Probability Scale should probably be used without mentioning the 5 as a mid-point to respondents.

While we found only a minor effect on the distribution of responses in this case, this is not to say that different scale anchors - particularly mentioning the 5 mid point as an anchor, will not have more of an effect in other studies. One reason is that this study produced mean probabilistic

scores of only approximately 2.5 (ie, an aggregate 25% probability of performing the action that was asked about in the survey). It is possible that when the average reported probability of the behaviour in question is low, scale anchors such as 5 will have less of an effect. This is because fewer respondents will be giving responses close to that scale point (ie say 4 or 6). It is reasonable to assume that if the inclusion of scale anchors were to alter responses, that they are more likely to alter responses that are closest to that scale anchor, as was found in Spagna (1984). So if the average probability score is, say 5 or 6, then the use of a scale anchor at the mid point or close to the mid point, will attract more responses than would otherwise be the case. In those instances, the use of the 5 as an anchor point is probably even less desirable, and market research interviewers using this scale should either use only zero and ten as anchors, or possibly zero, three, seven and ten if more guidance to respondents is thought to be necessary.

References

- Brennan, M. and D. Esslemont, 1994. The Accuracy of the Juster Scale for Predicting Purchase Rates of Branded, Fast-Moving Consumer Goods. *Marketing Bulletin* (5), 47-53.
- Brennan, M., D. Esslemont and D. Hini, 1995. A Test of Three Methods for Estimating Levels of Purchase. In: *Academy of Marketing Science - World Marketing Congress vol. VII-I*. Melbourne:
- Brennan, M., D. Hini and D. Esslemont, 1994. Obtaining Purchase Probability Data Via Telephone Surveys: A Preliminary Test of Two Techniques. *Marketing Bulletin* (5), 64-70.
- Danenberg, N. and B. Sharp, 1996a. Measuring Loyalty in Subscription Markets Using Probabilistic Estimates of Switching Behaviour. In: *Australia New Zealand Marketing Educators Conference vol. 1*. Auckland, NZ: Department of Marketing, University of Auckland.
- Danenberg, N. and B. Sharp, 1996b. Testing Probabilistic Measures of Behaviour as Measures of Customer Loyalty. In: *Southern Marketing - Theory and Applications, proceedings of the Australian Marketing Educators' Conference vol. 2*. Adelaide: Marketing Science Centre, University of South Australia.
- Dawes, J.G., 2001. The Impact of Mentioning a Scale Mid-Point in Administering a Customer Satisfaction Questionnaire Via Telephone. *Australasian Journal of Market Research* 9 (No. 1, January), 11-18.
- Day, D., B. Gan, P. Gendall and D. Esslemont, 1991. Predicting Purchase Behaviour. *Marketing Bulletin* 2, 18-30.
- Dillman, D., 1978. *Mail and Telephone Surveys*, New York: John Wiley.
- Garland, R., 1991. The Mid-Point on a Rating Scale: Is It Desirable? *Marketing Bulletin* 2, 66-70.
- Juster, F.T., 1960. Prediction and Consumer Buying Intentions. *American Economic Review* 50, 604 -622.
- Juster, F.T., 1966. *Consumer Buying Intentions and Purchase Probability: An Experiment in Survey Design*, National Bureau of Economic Research, Columbia University Press.
- Juster, F.T., 1969 *Consumer Anticipations and Models of Durable Goods Demand*. In: Mincer, J. (ed. *Economic Forecasts and Expectations - Analyses of Forecasting Behaviour and Performance*, New York: National Bureau of Economic Research, 167-233.
- Seymour, P., M. Brennan and D. Esslemont, 1994. Predicting Purchase Quantities: Further Investigation of the Juster Scale. *Marketing Bulletin* (5), 21-36.
- Spagna, G.J., 1984. Questionnaires: Which Approach Do You Use? *Journal of Advertising Research* 24 (No. 1, February/March), 67-70.
- Worcester, R.M. and T.R. Burns, 1975. A Statistical Examination of the Relative Precision of Verbal Scales. *Journal of the Market Research Society* 17 (No. 3), 181-197.