



How to improve your segmentations with max-diff

When conducting segmentation analysis there are several issues to consider, not the least of which are the questions used for analysis. More than almost any other factor, the questions asked will influence the quality of the segmentation results. The two main issues with regard to questions are question content and question type.

Question content refers to the subject areas covered by the questions. They could be attitudinal, behavioral, demographic, etc. While there are plenty of issues to discuss on this topic, that is not the focus of this article. We will confine ourselves to a discussion of question type.

Question type refers to how questions are asked. Various types of scales are often used to collect data for segmentation analysis. The primary purpose of all of these scales is to be able to get sufficient discrimination between respondents. If a scale is unable to dis-

criminate between respondents then it is not contributing anything useful to a segmentation analysis. Using such a scale would be no different from using a constant in the analysis.

Consider that one of the most popular types of questions asked in segmentation studies is the importance question. This could take various forms such as importance of product features, brands, decision criteria and so on. The traditional way of asking the question is to use an importance scale where each of these items is rated (perhaps on a 1-10 scale). The problem with this approach is that the respondent considers each item in isolation and, further, has no incentive to say that anything is unimportant. As a result one often sees data where many items are rated as important. More damagingly from a segmentation perspective, the questions don't sufficiently discriminate between respondents. This greatly

Editor's note: Rajan Sambandam is chief research officer at TRC, a Fort Washington, Pa., research firm. He can be reached at 215-641-2251 or at rsambandam@trchome.com. To view this article online, enter article ID 20091102 at quirks.com/articles.

reduces the usefulness of these questions in the analysis. So, what is the alternative?

Make comparisons

One could use an approach where respondents to a survey are asked to make comparisons, rather than rate each item in isolation. For example, a pairwise comparison task could be used where the respondent indicates the item in each pair that is more important. While this has the capacity to provide better discrimination between respondents, it is also more tedious, as the number of pairs to be evaluated quickly balloons. Designs can be used to pare down the number of pairs but the fundamental problem is that we are not making use of the respondents' full cognitive capacity. Comparing two items and choosing the more important one is often a very easy task. Respondents have the capacity to

snapshot

The author uses a checking-account example to show how maximum difference scaling, or max-diff, can deliver finely-tuned segmentations without subjecting respondents to an onerous number of comparison questions.

choose from more than two items at a time, and it is precisely this ability that maximum difference scaling (max-diff) exploits to give us better results.

Best and worst

Max-diff is a recent development in statistical analysis. It is a comparative method where respondents are shown sets of items and asked to pick the best and worst, or most- and least-important item, in each set. The number of items shown per set usually varies from three to five. The manner in which the items are grouped together and the order in which they appear are carefully selected through an algorithm. Data are then analyzed using hierarchical Bayes estimation to provide importance scores for all of the items used in the design. Scores appear like percentages and add up to 100. Since respondents have to make comparisons and choices, the problems mentioned with traditional importance scales are largely absent.

It therefore stands to reason that using max-diff to collect data for segmentation is likely to be more fruitful than using importance scales. In order to demonstrate this, a test was conducted and the results are reported here (Tables 1 and 2).

A split-sample design was used to identify feature importance when opening a checking account. A random half of the sample rated 12 features that were important to them in opening a checking account using a 1-10 importance scale, anchored by “not at all important” and “very important.” The other half of the sample was given a max-diff task for the same 12 features. They saw 12 sets of four items each and chose the most- and

Table 1

Features	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5
Interest rate	8	7	7	9	5
Gift	7	4	6	5	5
Accounts	10	9	8	10	5
Locations	9	9	8	9	6
Reputation	10	8	8	10	6
Fees and balance	10	10	7	10	6
Recommendation	7	5	6	5	5
Online banking	9	9	8	3	5
Branch hours	9	7	7	8	6
Phone hours	9	4	6	6	6
Customer service	9	8	7	9	6
Free checking	10	10	8	10	5

least-important feature in each set.

The importance scores from the max-diff analysis are available at the individual respondent level and hence are ready for segmentation analysis. A neural network-based segmentation technique called self-organizing maps was used to analyze both sets of data. The five-segment solution obtained with the importance scale information (rounded mean scores) is shown in Table 1. As can be seen in the table, the segments don't show much variation between them. One of them (Segment 1) has high scores on all variables, while another (Segment 5) has low scores on all variables. This is a typical pattern when the scale does not discriminate well between respondents. Other segments show only sporadic variation. Overall the results are not particularly interesting to a manager looking for differences in perception.

Results from the max-diff data-based segmentation are shown in Table 2 (rounded mean scores). It is immediately apparent that these results are very dif-

ferent from the previous results. Seven of the eight segments are very clearly defined by a single variable. Some segments (such as Segments 1 and 5) are overwhelmingly defined by one variable, while others place some amount of importance on one or two other variables also. But it is quite clear that the max-diff method has been able to clearly identify the differences in importance placed on these features by the respondents, and the segmentation analysis has been able to capitalize on the variance in the data to produce (seven) interesting segments.

Very worthwhile

Using max-diff in a Web survey is very straightforward. It is possible to use in a phone survey too, but the problem has to be limited to a few features and those features have to be defined very simply. Given the need for data discrimination in segmentation analysis, it would be very worthwhile to consider if comparative methods like max-diff can be used in your next market research project. |Q

Table 2

Features	Segment 1	Segment 2	Segment 3	Segment 4	Segment 5	Segment 6	Segment 7	Segment 8
Interest rate	2	3	2	6	1	55	2	7
Gift	0	0	0	1	0	2	0	6
Accounts	3	10	48	13	4	5	9	9
Locations	2	6	7	4	1	3	63	10
Reputation	2	3	9	37	1	3	2	5
Fees and balance	6	57	11	10	2	13	9	10
Recommendation	0	0	0	1	0	0	0	6
Online banking	4	3	9	5	85	4	6	18
Branch hours	0	1	1	1	0	1	1	6
Phone hours	0	0	0	1	0	0	0	3
Customer service	1	1	3	12	0	1	1	4
Free checking	80	15	9	9	5	13	7	17