

Global Segmentation – Dealing with Cross-Cultural Differences in Survey Rating Scale Usage

By John V. Colias, Ph.D.

Developing segmentation solutions that are global in scope requires dealing with cross-cultural differences in scale usage. To a greater or lesser degree, respondents of different countries or cultures:

- Tend to rate high for all questions.
- Tend to rate low for all questions.
- Bunch responses at end points of the scale.

Given cross-cultural differences in scale usage, marketing research analysts frequently develop ways to adjust survey responses, so that a particular survey response value means the same thing regardless of country of origin.

Perhaps the most sophisticated example of this approach is a Hierarchical Bayes scale usage model developed by Rossi, Gilula and Allenby (2001)¹. This scale usage model estimates mean and standard deviation adjustments for each individual respondent. Once adjustments are made to respondent attribute ratings, data across countries may be pooled together and segmentation analysis proceeds as usual.

Another approach is to avoid scale rating questions altogether. With the Maximum Difference (MaxDiff) survey task, respondents do not use a rating scale at all, but rather make choices². For example, instead of rating the importance of each attribute on a scale from 1 to 10, they select a most and least important attribute from among small subsets of the total set of attributes.

For example, in the MaxDiff survey task, the respondent might read, say, four attribute descriptions and then decide which

1 Rossi, Peter E., Zvi Gilula and Greg M. Allenby (2001) "Overcoming Scale Usage Heterogeneity: A Bayesian Hierarchical Approach," *Journal of the American Statistical Association*, 96, 20-31.

2 Cohen, Steve and Bryan Orme (2004) "What's Your Preference?" *Marketing Researach*, 16 (Summer 2004), 32-37.

one is MOST important and which is LEAST important in making category purchase decisions.

Each respondent would be presented with multiple sets of four attribute descriptions and would make their choices. The total number of sets and the number of attributes per set depends on the total number of attributes and attribute complexity.

The selection of which attributes would appear together in each set of four would be determined by an experimental design. For example, for a total of 20 attributes, we might develop an experimental design with 2 blocks of 15 sets of 4 attributes. Each respondent would be randomly assigned to a block and would select most and least important attributes in each of the 15 sets.

The most-least task offers the following key benefits:

- Since the survey task forces respondents to make a discriminating choice about which statement is the most and which the least influential to them, there is no possibility to encounter cultural scale bias (e.g., respondents of a particular culture are high-raters or low-raters for all attributes).
- The most-least survey task is less difficult for the respondent to do than a full sort and rank, but still produces a full ranking of all statements for each respondent.

The most and least choices can be analyzed using Latent Class (LC) choice modeling, which produces distinct segments of customers. Each segment would have unique attribute importance scores.

Alternatively, most and least choices can be analyzed using Hierarchical Bayes choice modeling which produces unique attribute importance scores for each individual respondent. In the Hierarchical Bayes model, segmentation solutions would be developed through application of clustering algorithms applied to the respondent-level attribute importance scores.