COMMENT ON COHEN

JAY MAGIDSON,
STATISTICAL INNOVATIONS INC.

Overall, I found Steve's presentation to be clear, concise, and well organized. More important is the substance of his primary message, that segmentation can be improved upon, substantially in many cases, through the use of the new methods of maximum difference (max-diff) scaling and latent class (LC) modeling.

Steve presents strong arguments in favor of LC over traditional ad-hoc clustering techniques such as K-Means, by listing many weaknesses in K-means that are remedied by LC. The fact that only 58% of the cases classified by K-means into one of 6 segments remain in these segments (the rest being reassigned to one of the other 6 clusters) when simply repeating the analysis following a reverse ordering of the cases, is a striking illustration of the inherent inconsistency of K-means clustering.

To Steve's excellent list of problems with the traditional approach to clustering, I would add that use of Euclidean distance to measure closeness between cases works *only* if within each segment, all variables have equal variances. This is an unrealistic limitation which has been shown to result in high rates of misclassification even when all variables are standardized to Z-scores prior to the analysis (Magidson and Vermunt, 2002a, 2002b). LC models on the other hand, do not make such assumptions. Moreover, LC segmentations are invariant to linear transformations made to one or more variables.

The maximum difference approach to scaling, introduced originally by Jordan Louviere, is an important contribution that may well alter the way that conjoint data is collected in the future. In Steve's example, he shows that 5 of the 6 pair-wise comparisons are captured by just 2 selections – best and worst choices. In my own research I am finding that the choice of best and worst in discrete choice studies provides extremely powerful information. When used with LC, max-diff outperforms by a significant margin the equally parsimonious design involving the first and second choice.

To obtain the final segments, Steve pointed out that he first estimated individual coefficients using HB, and then subjected these coefficients to a LC analysis. A better approach would be to perform a simultaneous (1-step) analysis using LC rather than using this *tandem* approach. While individual coefficients would then be unnecessary to obtain the segments, should such be desired for other reasons, they could still be attained – directly from LC.

The trick to obtaining individual coefficients for each case with LC is to weight the segment-level coefficients by the posterior membership probabilities obtained for each case. This approach makes use of the fact that HB may be viewed as a parametric and LC as a nonparametric method for random effects modeling (Vermunt and Magidson, 2003). In the HB case the random effects are assumed to be continuous, while for LC they are assumed to be discrete. In my current research with maximum difference scaling, I am finding that use of individual coefficients obtained from LC produce higher hit rates than those produced by HB.

Overall, Steve's paper makes an important contribution to the field. It is for good reasons that his presentation was voted best at the conference.

REFERENCES:

Magidson, J. and J.K. Vermunt (2002a) "Latent class models for clustering: A comparison with K-means" in *Canadian Journal of Marketing Research*, Vol. 20, pp. 37-44.

Magidson, J. and J.K. Vermunt (2002b) "Latent class modeling as a probabilistic extension of K-means clustering", in *Quirk's Marketing Research Review*, March issue.

Vermunt J.K. and Magidson, J. (2003, forthcoming) "Random Effects Modeling", in *Sage Encyclopedia of Social Science Research Methods*, Sage.