IDENTIFYING FEATURE IMPORTANCE:
A COMPARISON OF METHODS

White Paper Series

TRC

Consumer **Insights**.
Market **Innovations**.

Understanding what customers want is fundamental to the new product development process as well as to the process of keeping existing products fresh and relevant. To be successful in this area we need to be able to correctly identify what features are important to consumers. Feature importance can be measured using a variety of methods of differing effectiveness. In this paper we will deal with the following methods:

– Importance Scales
– Pick data
– Pairwise Comparisons
– Max-Diff

Importance Scales
This is the most popular way of measuring feature importance primarily because of its ease of use. Respondents simply indicate on a (say, 1-10) scale how important they think each feature is. The advantages are that respondents are not taxed, need to evaluate each feature only once and can evaluate each feature independently. However, these advantages have downsides too. Since not much is demanded of respondents and they are not constrained in any way, there is no incentive to prioritize importance among the various features. As a result we have the "everything-is-important" syndrome. Results often have minimal discrimination between features and there is generally a gradual decline in importance scores from the most to the least important feature.

Pick Data
In contrast to importance scales, in this method respondents are asked to pick a certain number of features from a list that are most important to them. Based on the number of people

picking each feature, percentages are calculated at the overall level indicating the importance of each feature. Obviously, respondents have to see all the features before they can pick, and hence this will not work for phone data collection. The task is harder than using an importance scale since multiple comparisons need to be made. A common question is the number of features a respondent should be asked to pick. If too many are picked, the results will resemble importance scales in having a slow decline from top to bottom in feature importance scores. If too few are picked some of the less desirable features may never be chosen, leading to two groupings of very high and very low importance. Research has shown that picking approximately a third of the total number of features presented is likely to provide the best results.

Pairwise Comparisons
In this method features are presented as pairs and respondents select the one that is of more importance to them in each pair. The task is simple enough for a child to answer and also avoids bias due to respondent scale usage tendencies. But the number of pairs to be evaluated quickly becomes very large as the number of features increases. It is possible to use designs to reduce the number of pairs to be evaluated, but the results may not be as reliable. Advanced statistical analysis can be used to correct for this, but the fundamental problem is under-utilization of the information processing capability of respondents. That is, people are capable of evaluating more than two features at a time. Hence if more features can be evaluated at a time, the total number of evaluations comes down and better

quality data can be obtained for the same effort. The next method uses this principle.

Max-Diff
Max-Diff or Maximum Difference scaling is an enhancement over pairwise comparison, where respondents are shown more than two features at a time and asked to pick the one they like best and the one they like the least. Typically 3 to 5 features are shown at a time, as this seems to provide the best information. A number of such sets of features are shown to respondents using a mathematical design such that each feature is shown an equal number of times and in equal number of positions. The results are analyzed using an advanced statistical technique called Hierarchical Bayes estimation. Final results are converted to a 0-100 scale, which is much like percentage scores that add up to 100 across all features. Research has shown that this method does a much better job of discriminating between features. As a result, managers should be in a good position to identify features that are truly important to consumers.

An Example
We conducted a web study with a split sample design where respondents evaluated the importance of twelve banking features that were important to them in opening a checking account. A third of the respondents rated the features on a 10-point importance scale, a third of the respondents picked the five most important features and a third went through a Max-Diff task. In the Max-Diff task, each respondent saw twelve sets of

four features and picked the feature they liked best and liked the least in each set. The results are presented next.

If based on research on scales we believe that Max-Diff provides the best result, then it is clear that pick data are much closer to the Max-Diff results and that the importance scale information is clearly different. The importance scale generally identifies the most and least important sets of features correctly but the ordering is not right. Further, the lack of discrimination between the importance scores makes it difficult to clearly draw a line anywhere except to identify the three least desired features. The Max-Diff results are ratio scaled and hence 20 can be interpreted as twice as important as 10. They therefore clearly show the high importance associated with Free Checking (and to some extent the pick data reflect this too). Since the Max-Diff data are available at the individual respondent level, one could segment the data to identify pockets of respondents who highly prefer certain features or combinations of features. This would not be possible with pick data, but if Max-Diff cannot be used pick data seem to provide a much better alternative to importance scales.

End Note

Telephone research may have contributed to the widespread use of importance scales, since methods that require feature comparison are almost impossible to do over the phone. Now the emergence of the web as a viable data collection tool is changing that and hence we are seeing more interest in other ways of identifying feature importance.

Features	Max-Diff		Pick Data		Importance Scale	
	Score	Rank	Chosen %	Rank	Top 3 box %	Rank
Free Checking	27.3	**1**	78	**1**	88	**1**
Online Banking	17.7	2	54	4	70	7
Balance/Fees	16.8	3	58	3	84	2
Accounts/Services	11.6	4	40	5	81	3
Branch Locations	8	5	68	2	79	5
Reputation	6.4	6	39	6	81	4
Interest Rate	6	7	33	7	64	9
Customer Service	2.3	8	25	9	73	6
Branch Hours	1.3	9	28	8	64	8
Bonus/Gift	1	**10**	14	**10**	30	**12**
Recommendation	1	**11**	5	**12**	31	**11**
Phone Hours	0.5	**12**	6	**11**	46	**10**