USABLE SEGMENTATIONS

The Statisticians are Running the Asylum and it's Getting Scary Out There!

by Hilary DeCamp

Segmentation is an absolutely great business tool. It recognizes the fact that not all consumers want the same things or respond to the same business and marketing approaches. A successful segmentation provides a framework for strategy or tactics and allows a company to focus their resources where it will generate the strongest returns.

As a business strategist who is also a marketing scientist, I am alarmed when I hear clients complain that their segmentation "didn't work" and then doubt the value of this critical form of research.

What's wrong? Why are their segmentations failing?

I usually say that segmentations fail because they are either "unused" or they were "unusable".

- Segmentations are *unused* when there is insufficient buy-in from internal stakeholders, lack of consensus on business objectives, and/or inability of end users to understand how they should apply

the segments. With careful planning and partnership, this source of failure can be readily avoided…as long as the segmentation itself is a sound one.

- Way, way too often, segmentations fail because they are *unusable*. It is with increasing frequency that we are hired to reanalyze and re-report segmentation studies that were considered just not sufficiently useful. This occurs when the segmentation solutions are strategically misdefined, statistically unsound, or not easily recoverable in follow-up research.

But how do you, as a client, protect yourself against a statistically unsound or unrecoverable solution? By understanding what's required for a **usable segmentation**. Usable segmentation results require approaches and inputs that:

- Provide clarity around the types of decisions to be made
- Provide understanding of individual respondents
- Capture the differing views that consumers are trying to express using the provided scales (and not be misled by scale usage differences)
- Can be reproduced in "short forms" so as to enable further investigation of segments and to monitor progress of initiatives among the target.

Surprisingly, one of the reasons unusable solutions are created is statisticians using some pretty advanced and powerful tools. You may think, "Wait, how can that make sense? Isn't using advanced tools a good thing?" The answer is yes only if they make the segmentation a better strategic device. And sometimes these new tools don't work as well as they claim to or don't necessarily apply in the situations where they are now getting used. My two current pet peeves are MaxDiff segmentation inputs and Latent Clustering procedures. Let me tell you why.

You may think, "Wait, how can that make sense? Isn't using advanced tools a good thing?"

MAX DIFF

Maximum Difference Scaling (MaxDiff) uses forced choices and Hierarchical Bayes modeling to produce interval-scaled importance ratings that have two advantages:

- Increased discrimination between items (you don't get to say "everything's important")
- Elimination of scale usage differences (especially helpful in cross-cultural research)

It is a great measurement tool that should not be used as the source of segmentation inputs. Sound segmentation inputs need to be measured at the individual level and use a method that can be readily reproduced

in "short forms" applied during follow-up research…MaxDiff does neither.

1. MaxDiff looks like it provided individual level measurement when in reality it does not. While the analysis provides respondent-specific scores on every item, these scores are merely *estimates* — modeled results generated by a combination of the respondent's answers and information "borrowed" from other respondents. Segmentation is supposed to group individuals based on how they differ from others…you can't do that reliably when their scores are largely extrapolated from the answers of others.

2. There is no way to reproduce the MaxDiff importance scores in a short-form classification algorithm. Only vague approximations can be generated. A variety of methods are used to do this, and only some of them can be tested for accuracy. It's just as well since the accuracy rates are so low they would scare you. As a result, short forms generated off MaxDiff segmentation schemes tend to be both lengthy and inaccurate.

LATENT CLUSTERING

Latent Clustering is a powerful clustering tool that better accommodates mixing metrics from different scale types (e.g., yes/no, 5-pt scales, and continuous measures like age or income) than Euclidean distance models, such as k-means or hierarchical clustering, do. This makes it an ideal candidate for a segmentation based on objectively measured behavioral or observational data (e.g., geo-demographic traits, prescription records, purchase histories, and so on).

It is usually a mistake to apply this technique to the attitudinal data on which most segmentations are based.

However, it is usually a mistake to apply this technique to the attitudinal data on which most segmentations are based. To get *usable* segmentation results, the inputs need to reflect the differing views that consumers are trying to express and not be distorted by scale usage differences in how they express those views.

1. Consumer segmentations are generally done on survey data and respondents have the unfortunate tendency to use scales in slightly different ways from each other (see benefits of MaxDiff). The reason this is a problem in Latent Clustering is that frequently the model tends to form segments based on how people use the scale (e.g., high raters or middle raters) rather than what people were trying to tell us on the scale. Transformations (e.g., within-respondent-standardization) that are an effective solution to this issue in Euclidean distance models do not prevent Latent Clustering from generating these meaningless groups, that can even score well on statistical measures of distinctiveness.

2. The ability to mix metrics generates the temptation to throw in the kitchen sink and segment on virtually the entire survey (attitudes, needs, behaviors, demographics and even brand usage!). I recently reviewed an 8-segment solution where only five of the segments were primarily differentiated by their attitudes and needs, while two were primarily defined by most-often usage of a particular brand, and the last was differentiated by geographic region. Setting aside for a moment the questionable strategic value of segmenting directly on these latter metrics, this type of Frankenstein solution (wherein different group definitions are driven by substantively different sub-sets of the inputs) is conceptually unsound and difficult to understand and use.

As previously noted, MaxDiff and Latent Clustering are each powerful tools when used appropriately. But, when used to generate consumer market segments (particularly by inexpert analysts lacking the experience or perspective to recognize a conceptually flawed solution), they can be counterproductive. When used together, you have a genuine recipe for disaster.

As a market segmentation buyer, remember you are looking for the right lens to look at the market that will best inform strategic and/or tactical decisions. Please don't be blinded by the shiny techniques that a statistician flashes in front of you. And if that person claims that they can choose the best solution based on a statistic, run screaming for the hills.