Best Segmentation Practices and Targeting Procedures that Provide the Most Client-Actionable Strategy

Presented by Frank Wyman, Ph.D.
Director of Advanced Analytics
M/A/R/C® Research

Presented November 7, 2005
IIR's The Market Research Event conference, San Francisco

[Slide 1 – title slide]
Welcome. Segmentation is an important aspect of marketing research. Because there are a myriad of approaches to segmentation it is easy to get sidetracked away from good, actionable segmentation results. While the exact approach used in any given study tends to be highly custom to that study, it is helpful to have a guide to the approaches that tend to work best in most cases. This presentation lists 23 separate good practices as an attempt to provide such a guide.

[Slide 2]
Let's start with a basic definition of what we're talking about here. I think of Segmentation as the dividing of a market's customers into subgroups in a way that optimizes the firm's ability to profit from the fact that customers have different needs, priorities, and economic levers. It's always important to me to remain cognizant of the fact that MR is done in order to improve the firm's profit. We're not doing this stuff for any other reason than to help out clients' bottom line profit. This leads directly to my first best practice [Slide 3]: If you keep in mind that the end goal is always to enhance the firm's profit, this can often help increase the actionability of your segmentation. This happens because segmentation involves a good number of steps, some of which call on judgment, and keeping your eye on "the money," so to say, can help guide you to making the most profit-enhancing decisions along the way to your final segmentation. At each step ask "how can these results help improve profits?"

[Slide 4]
Before jumping into the nuts and bolts of segmentation, I want to mention how important it is to have good product position information against which to compare the segmentation. In order to gain the most actionable segmentation results, you'll want to know the strengths and weaknesses of each offering in the market, both true functional differentiations as well as perceived ones. This slide shows a MARC Differentiator map, a great way to reveal a given product's differentiating features. Since positioning research is easily an entire presentation itself, I'm not going to go into any detail, but I would definitely recommend positioning research as a good compliment to any segmentation research. In fact, it really is a necessity for new products that are still forming their optimal positioning in a market. It's kind of empty knowledge to do a segmentation study and find out that such and such segment requires A, B, and C yet in the end you don't really know how well the competition delivers on A, B, and C and thus what position your offering should ultimately take. Oftentimes, segmentation and positioning research can be done in a single study. At very least, in a segmentation questionnaire you'll want to include basic items covering competitive product usage and basic satisfaction. [Slide 5]: So, our 2nd best

practice is to have timely product positioning information on hand to improve the interpretability and enhance the actionability of the segmentation.

[Slide 6] Generally speaking, there are three basic approaches to segmentation. There is any method that is not based on quantitative analytics. Often this type segmentation is done out of convenience – it may simply be the way a firm has always looked at consumers in their market and hasn't seen need to update that view. Sometimes this a priori type segmentation <u>can</u> be appropriate, as with geographic segmentation of customers served by different regional sales forces, or SIC-based B2B segments served by divisions with different technical specializations. All too often, however, a non-analytic approach is merely a weak attempt to accomplish what an analytic approach is going to do much better. When it is appropriate to segment - and let me stop there and say that some markets, the highly commoditized ones, do not segment well because price competition easily undermines the objectives of segmentation – but in the vast majority of markets where segmentation *is* appropriate, an analytic approach is typically the approach that's going lead to better profitability. Let me reiterate that this is *typically* going to be the case, but not always. There are certainly instances where a non-analytic approach to segmentation is more cost effective than an analytic one. For instance, most electric utility companies treat their large commercial customers especially well – they segment them out, if you will, in a non-analytical way and their special treatment of them is a cost-effective strategy for retaining that highly valued customer. I've done several segmentation studies for electric utility companies and always the large commercial customers are purposefully excluded from the sample of customers to be analytically segmented. The cost effective strategy for that segment is to cater to each one customer on an individual basis….no analytics needed there! In my experience, however, I'd say the majority of the time I see a non-analytic approach to segmentation going on, it's being done in lieu of what could be done better analytically. [Slide 7]... leading to my suggested best practice #3 [Slide 8] Please help me feed and clothe my children by continuing to … choose to do analytic segmentation!

[Slide 9] There are two distinctly different approaches to quant segmentation: "interdependence" segmentation (or clustering) and "dependence" segmentation (CHAID, or Chi-Square Interaction Detection, being my favorite of these forms). While there are other methods of interdependence segmentation than clustering and other types of dependence segmentation than CHAID, clustering is by far the most widely used interdependence method and CHAID, well CHAID is simply my preferred method for dependence segmentation. In a general way I think its best to associate the word "strategy" with interdependence segmentation and the word "tactical" with dependence segmentation. If you want a "natural" view of the lay of the land to use in forming general strategy – you're asking "What do consumers in this market want and how many "camps" are there and how big is each?" – then a clustering approach is appropriate. With more specific tactical questions such as "Who specifically will buy our product and how do we reach them and trigger them to buy?" then a *de*pendence approach such as CHAID is appropriate.

[Slide 10] Thus, our fourth Best Practice becomes: for setting general strategy, pursue *inter*dependence segmentation; in forming *specific* tactics you'll usually want to pursue a *de*pendence approach to segmentation.

[Slide 11] A whole host of considerations should go into the decision of whether a given project requires an interdependence method or a dependence method or both. In general the younger a product and the more general the objective, the more one wants to start with the market overview type segmentation produced by clustering. Then, with good product positioning information, you can try to figure out where your new product fits in. After optimizing the positioning for the new product, you might then move into the more tactical fodder of dependence type segmentation. The more mature a product and/or market the more one needs to use a dependence type segmentation to identify appropriate, specific tactical opportunities for product differentiation and customer targeting.

- Is a primary objective "to know the entire market" (how many needs groups exist and how large is each)? [interdependence]
- Is a primary objective "to know how to target those who will buy my offering/product/service"? [dependence]
- How old is the market? [younger/dynamic -> interdependence; older ->dependence]
- How well are key needs in the market being met? [not well - > interdependence; well -> dependence]
- How old is your product (existing, new launch, line extension, etc.)? [younger -> interdependence + dependence; older -> dependence]
- How crowded is the market? [uncrowded –> interdependence; crowded -> dependence]
- How much room for differentiation is there? [more room for differentiation -> interdependence; less room for differentiation -> dependence]
- If an older product, how long since last interdependence segmentation? [perhaps its time to update the market-wide view of customer needs via interdependence segmentation – depends on market but should be updated around every 3 years average]

Often the case can easily be made that both approaches are needed. For instance, it might be that a product is nearing maturity and needs the type boost that can be realized through the specific customer-targeting tactics that a dependence segmentation can uncover. But at the same time the information the firm has on the entire market has become stale and outdated and needs to be revamped through an interdependence segmentation to set the proper stage for interpreting the dependence segmentation.

[Slide 12] In general, then, you need to let the parameters of your market and product offering determine which of the 2 analytic approaches to take.

[Slide 13] We don't have time to cover all the relative pros and cons of each quantitative segmentation method -- especially as it requires getting into the gory detail of all the distance and similarity measures, linkage techniques, node-splitting criteria, etc. Nonetheless it's instructive to take time to see how the methods I'm going to recommend as best practices fit into the general tool bag of quant segmentation.

Let me quickly address the two oddballs on this slide: Q-factor analysis and Neural networks. Q-factor analysis is factor analysis applied to cases (that is respondents) rather than to variables as is the usual application of factor analysis. Its use in segmentation is to group folks who have the same "relative *pattern*" of responses though not necessarily the same LEVEL. So the rank order of needs will be the same among Q-factor segment members but some will have much stronger absolute levels of need than others. Keeping in mind Best Practice #1 and the goal of enhanced profitability, note that a person with a strong need for something is quite likely to spend much differently than a person with a weak need for that same thing – and they do not belong in the same segment! Q-factor analysis is a poor choice for an interdependence method. Neural networks, though a really cool technique and highly applicable in the medical field, is of little use in marketing research, especially in deriving segments. This is because it is the quintessential "black box" – while it indeed models and predicts a dependent variable as does, say, linear regression, it does so without revealing the actual mechanisms at work between the set of drivers and the dependent variable….in the end you are left not knowing *why* any given person belongs to a particular segment – just that whatever the response dynamics of that person they are similar to others in that segment. While the approach sometimes has an appropriate use in scoring huge customer databases, it is not a method to use to derive segments. [Slide 14] I submit that using Q factor analysis or neural networks for segmentation … [Slide 15] … is <u>not</u> a best practice!

[Slide 16] Instead, I recommend as what works best in most instances are k-means clustering and CHAID. K-means is the simplest approach to clustering and in my experience typically mirrors reality more closely than other clustering approaches such as hierarchical clustering or latent class clustering. Now, I rarely if ever rely *solely* on k-means in conducting cluster analysis. Although *<u>final</u>* segments are derived via k-means, nearly always I answer the big question of "How many clusters?" using hierarchical and/or latent class clustering. But in the end, the final interdependence segments in my work are usually derived via a k-means clustering. As for CHAID, I recommend it for dependence segmentation because among the treeing methods it tends to be the most general and thus most applicable. For instance, some treeing methods only allow binary splitting at each node. CHAID allows any number of splits; furthermore CHAID works great with all types of data, including missing data.

… [Slide 17] … Generally speaking, then, I recommend the use of k-means and CHAID as best practices.

[Slide 18] Let's talk a little bit more about what one gets from each of the two major segmentation types…
With interdependence clustering segmentation we find out what the existing "camps" of needs among (potential) consumers are. How many "natural camps" there are. How each is defined with regards to their needs/beliefs/behaviors. What the size of each segment is. We can profile these cluster segments and possibly identify the key markers of each segment, … if any exist. Now this underscores a key distinction between cluster segments and CHAID segments -- in cluster analysis there is no guarantee that the resulting clusters will differ by demographics, media use, or any other "marker" type variable that allows them to be targeted….in CHAID the "marker" variables are actually the very definition ***of*** the segments since those variables are the drivers of the model. This is what makes CHAID a generally more actionable type segmentation than clustering.

With dependence segmentation we find out who will buy the firm's product and who will not. We find out how many groups there are with distinctly different propensities toward buying the firm's product and what the key drivers (again, the "marker" type variables) are that define the segments. It is through such techniques as CHAID that we are best able to answer the important question of how to best reach those most likely to buy the firm's product.

Now, in terms of what makes a "good target segment" notice that in clustering the answer is **indirectly** pieced together by careful examination of the juxtaposition of segment needs vis-à-vis the features and positionings of the market products including, of course, the subject firm's product. On the other hand, in CHAID the answer is **directly** yielded *by* the methodology itself. A "good target segment" has a high propensity to purchase the firm's product and is actually identified *because of* that desirable property. (Of course a good target must also be sizeable enough to pursue with the scarce resources of the firm.) I am not attempting to make the case that CHAID is a better segmentation approach – only that the two approaches differ significantly in the type information they yield.

[Slide 19] As far as data considerations for segmentation, the following hold in general for both interdependence and dependence approaches. Sample size should be larger rather than smaller. Segments as small as 5% can often promise high profitability once they are weighted according to their spend levels. But generally, one would never want to make inferences from a subsample any smaller than around n=30. Thus, one would have had to start with an overall sample of 600+ in order to ensure that a segment as small as 5% could be reliably described. To the degree you desire higher reliability and/or think promising segments of smaller size than even 5% exist, then you will require higher sample size than 600. To the degree you are able to live with less reliability and/or feel that segment sizes need to be substantially larger than 5% then you might be ok with as few as 400. When deciding on sample size for a segmentation study you must of course weigh in the type audience under study and any required incentives as well as the potential return on research for the firm given its product's competitive differentiation and the size of the market. For instance, in segmenting physicians, note that in some specialties 600 might be impossible given the limited number of the specialty – also with honoraria approaching $200 the research costs increase dramatically. Physician segmentation research is more typically done with 300-400 responses. But, very generally, in most markets, segmentation analysis will require AT A MINIMUM 400-600 respondents. More typically, segmentation studies are based on 1000 to 2000 cases. And for true testing of the stability of some models such as CHAID it is best to have some so-called "holdout sample" – so the analysis might be conducted on 1000 cases but another 500 are also used to apply the model to ascertain how well results hold up, bringing the total sample N to 1500 even though the actual modeling is done on only 1000.

Random sample. Ha ha – like *this* ever happens in survey research! Well, the idea is to at least get a *representative* sample and not to oversample any particular subgroups -- we're not so much interested in inter-group comparisons here as we are in correctly describing the groups, including correctly estimating the segment sizes and so we want to be as close as possible to true representativeness of market composition. The usual quoting mechanisms to ensure proper representativeness are appropriate.

[Slide 20] So, in segmentation we want large, random samples.

[Slide 21] The questionnaire should include lots of demographics and media and channel use items. To be fully actionable, we'll want to be able to know how to best reach our segments. These should go toward the end of the questionnaire. You can hardly ever have too many of these type variables. Also, whether or not positioning research is also being conducted at nearly the same time, it's a good idea to field a few of those type questions. What competitive products have been used and how do they rate on key attributes or at least overall. If a full positioning study is being conducted concurrently then only a few of these for validation and/or calibration purposes is necessary – if not then more detail on this important aspect of the market is desirable.

[Slide 22] As a best practice, include lots of demographics as well as media and channel use items in your questionnaire. Also, some market positioning type items are usually a good idea.

[Slide 23] For clustering type segmentation in particular, you will want to include an exhaustive "needs" battery of items. While for the most part it is the "needs" of a market's consumers that we typically want to be the basis for clustering segmentation, depending on the market, variations on this are things such behaviors, opinions, or beliefs. At any rate, it is always best to gain this exhaustive list through fresh qualitative study, keeping in mind things can change rapidly in most of today's markets. The scale for these items should be short. I like to use a 1-to-5 agree-disagree scale. It cannot be too short, however. The variables in clustering are assumed to be ratio or interval level … and so a 5-point scale is about as small as you want to go.

The biggest bane to clustering segmentation – yea to all of market research possibly – is what I call differential scale use, also known as response style bias, which happens when two respondents truly feel the same way and truly want to respond the same way but use the scale differently. You like chocolate the same amount that I like chocolate but you respond 5 whereas I respond 4. This phenomenon causes different segments to artificially arise when in fact no differences exist. A short scale usually is the first step in avoiding this type bias. It also helps to place anchors, labels, over each point on the scale to help the respondent know exactly what each response means. Furthermore, these labels should be clearly distinct so as to further help the respondent know what each point means and how it is different from each other point. For instance, a 4 being "somewhat" and a 3 being "moderately" is not very distinct. I like to use the following scale: 1=totally disagree, 2 = somewhat disagree, 3=neutral, 4= somewhat agree, and 5 = totally agree. On this scale it is also helpful for each item to be extreme….it is much clearer to interpret the difference between a somewhat agree and a totally agree response to an item such as "I absolutely love chocolate" than it is to do so for an item such as "I like chocolate." This battery is the heart and soul of interdependence segmentation. It is not atypical for this battery to contain 50 or more items. An added benefit of a short 5-point scale is it helps alleviate respondent burden on these long batteries. Also, it's not a bad idea to randomize the order of the items -- to reduce potential order bias. Also note that all respondents should answer all items - no skip patterns, since those missing values really screw up cluster analysis. Finally, note that the basis variables for interdependence segmentation do not always arise form **stated** responses as this slide presumes. While stated responses are used more often than not, more and more these days respondent-level conjoint and discrete choice utilities are used as the measures of consumer needs in clustering; … conjoint utilities are *excellent* basis data for clustering.

In designing an interdependence segmentation:

[Slide 24] Precede it with fresh qualitative work.

[Slide 25] Reduce differential scale use bias by designing a “needs” battery based on a short (1-5 agree-disagree) scale with clearly differentiated anchors on every point and items worded in the “extreme” sense.

[Slide 26] Consider using conjoint or discrete-choice utilities as the basis of clustering.

[Slide 27] When using stated responses, the biggest potential problem for clustering is differential scale use. You should always test for this bias first thing when data comes out of the field. Luckily the presence of this bias is easily detected and also there are ways to correct it a bit on the back end. To detect it simply run a k-means cluster solution for 2 or 3 segments using the entire battery as basis variables. If the 2 or 3 profiles are strongly correlated in *pattern* and different only in general *level of response*, then a differential scale use problem exists in the data and needs to be fixed

[Slide 28] In this slide there is no correlation between the profile patterns -- no differential scale use bias exists. The general levels of need for the two segments are much more “equal” than are the two patterns of need.

[Slide 29] Note, however, that *this* chart gives strong evidence that differential-scale-use bias exists in the data. The patterns are perfectly correlated but the general levels of need are far apart. Respondents who indeed felt the same simply used the scale differently.

[Slide 30] A definite “best practice” is to always test for differential scale use bias before clustering.

[Slide 31] If differential scale use bias is present, you will need to either fully or semi ipsatize data. Semi-ipsatizing the data re-centers each respondent’s “needs” responses around his or her particular mean response across the items. Full ipsatization, which is rarely done, not only recenters the data but also redistributes each respondent’s data so that all have the same degree of dispersion around their mean response. The big downside to ipsatizing (besides remembering how to spell it!) is that it gets rid of ALL differences in response level. So that if truly there are some segments that generally just have a lower need, globally across all attributes, than another segment, that distinction is forever lost. It is better to avoid differential scale use bias via good design up front and following best practice #11 than it is to have to fix it with ipsatization. Generally speaking outlierish responses are not terribly disruptive to cluster analysis (and are rare, anyhow, given the small scale). Missing data is a big problem for cluster analysis and needs to be fixed. I offer a few possibilities here. Finally, if the basis variables being used in the clustering are from different scales, then the variables will all need to be standardized to some common metric, for instance z scores, to make the resulting distances meaningful. As a parting shot, note that all the problems of this slide are nearly completely avoided by using conjoint or choice utilities as the basis of the cluster analysis.

[Slide 32] If differential scale use bias exists, fix it with ipsatization!

[Slide 33] When clustering data, outliers are OK but missing data is *not*.

[Slide 34] Because "needs batteries" tend to be quite long with 50 or more items, I nearly always precede the actual cluster analysis with a factor analysis of the original items. If this isn't done certain dimensions *artificially* get inflated in terms of their "importance" in defining cluster segments simply because of the large number of items representing the same dimension. Factor analysis also adds value to the study by revealing how many salient dimensions are at work in the market – are there really only 3 or 4 dimensions along which consumer deliberate in choosing product, or are there more like 7 or 8 … or what? Also, what attributes do consumers tend to see as one and the same – how do items group together in factors? In this example, we see that peanutty and crunchy are seen really as just a single dimension in this market category, as are sweet and rich.

[Slide 35] Thus, a best practice is to cluster only unique dimensions; factor analyze the original items to get these unique dimensions.

[Slide 36] In conducting the actual cluster analysis I like to use the following approach. First, determine the answer to the big question "How many segments?" I offer several approaches to this here. Not mentioned but always important is client input at this stage, keeping in mind that most brands cannot manage more than 7 to 10 segments. Once the number of segments to pursue is decided – or a small range of solutions – then conduct k-means type cluster analysis. I've included a couple of technical "good practices" on this slide.

[Slide 37] Approach the key question of "How many segments?" from numerous angles. Once decided, finalize clusters using a k-means cluster analysis. When conducting hierarchical clustering, Euclidian distance and complete or Ward's linkage methods usually work best.

[Slide 38] Let's turn our attention to data considerations specific for dependence type segmentation, specifically CHAID. In dependence segmentation we actually have a key criterion measure, a dependent variable, for which a predictive model is built. But, unlike regression the model is not operationalized as linear coefficients for each driver. Instead, a CHAID model selects from often a large set of input variables only those that are the strongest in a predictive sense, and then decides how to divide or group the possible values of each variable, discretizing continuous variables, so as to have the strongest separation in the dependent variable. In CHAID we want a reliable and accurate measure as our dependent variable. Typically this is a measure of respondent propensity to purchase the product or service of the subject firm. One might simply use a stated purchase intention response here but more typically a stronger measure is used, a measure such as arises from conjoint simulation or MARC's Assessor measurement system. Keep in mind the dependent variable of interest isn't necessarily always propensity to purchase; other common dependence segmentations involve segmenting consumers with regards to predicting retention, ad response, targetability, and the like. Finally, the demographics, media use, and channel use items are extremely important in CHAID as they become the actual basis variables of the resulting segments.

[Slide 39] In CHAID it is important to have a good strong dependent variable measure.

[Slide 40] CHAID data requirements are easy, outliers are generally ok. Missing values are even ok, as they are seen as just another category. No real need for factor analysis or ipsatizing or standardization.

[Slide 41] In CHAID, there is no requirement of heavy data cleansing.

[Slide 42] It is usually prudent to run several different CHAID models. The skinnier models, restricted to only those drivers that the firm can definitely take action on, will be more actionable but perhaps a bit less "full" in terms of explaining all the dynamics at play. A "fuller" run might provide a better explanation of the true dynamics at play in driving propensity but it might require the softer, less actionable type variables to do so.

[Slide 43] Some technical best practices in running good CHAID include

- Defining all drivers correctly as to their level of measurement [nominal, ordinal, interval] since each type is treated differently
- Smallest child node should be set at approximately 5% of total N, parent nodes set at twice child n
- Grow trees typically to a depth of 4-5 branch levels
- Alpha set at .05, with the Bonferroni adjustment turned off.

[Slide 44] In CHAID, set the minimum size of child nodes at 5% of total N, and parent nodes at twice child node size….
…and [Slide 45] grow trees to a depth of about 4-5 branches with alpha set at .05 with no Bonferroni adjustment.

[Slide 46] A certain level of human judgment is required in any dependence segmentation – here are some common ones in CHAID:

- Merging nodes (i.e. merging segments)
- Redefining splits
- Pruning whole branches
- Eliminating some drivers

[Slide 47] In short, in CHAID it is good practice to expect a healthy dose of human-judgment overlay on the process….and thus expect a fair amount of "back and forth" interaction with the category expert and/or client.

I want to close by showing examples of what I feel are the best ways to present the segments yielded by the two approaches to quant segmentation.

[Slide 48] I like to show cluster results with profile charts such as this. It quickly reveals how many segments there are (for instance, 5 in this example) and the defining needs of each. The segment sizes are found in the labels. I usually put the attribute dimensions in descending order of differentiation across segments.

[Slide 48] The best way to show CHAID segments is usually through a table such as this, which clearly shows the definition of each segment as well as its size and most importantly its mean value for the dependent variable.

[Slide 50] This brings us to our 23rd and final best practice in segmentation research.

Thank you very much.   [Slide 51 = contact info]

Frank Wyman
Director of Advanced Analytics
M/A/R/C® Research
(864) 938-0282
frank.wyman@marcresearch.com
