Concept Testing (And The “Uniqueness” Paradox)

Contributed by Jerry W. Thomas
Thursday, 10 February 2011

A well-designed new-product concept testing system that is overseen by experienced and knowledgeable researchers can vastly improve a companyâ€™s ability to develop successful new products or services.

The great potential of concept testing is seldom achieved in actual practice, however.

Having tested thousands of new-product concepts over the past 30 years, we have observed many methodological failures, system flaws and analytical errors.

We have also worked with corporations that are masters at developing and testing new product concepts. Based on these experiences, we have evolved some theories about what works best. The purpose of this article, then, is to suggest some guidelines and best practices to improve new product concept testing.

Research System

Concept testing works best as a research system. A research system involves standards and standardized processes and procedures at every step in the process, so that each new-product concept is tested in exactly the same way as all other concepts.

A â€œsystems approachâ€• permits comparisons to be made over time and across concepts. A systems approach allows predictive models to be developed and evolved. A systems approach helps provide the information, understanding and knowledge for sound, new-product decision-making.

Testing concepts as isolated, unconnected events (without a system and standards) does produce data, but little information and knowledge.

Research Design

Should you test new-product concepts monadically (i.e., each person sees only one concept), or test them in batches where each respondent sees multiple concepts? If you have a large number of early-stage concepts to evaluate, then you might choose to test them in batches (a dependent design) to save money. Each respondent would see and rate several concepts, and the results would help you measure the relative appeal of the concepts.

You would know that some concepts are better than others, but you would not know how good any of the concepts are. The problem is bias or distortion caused by interactions among the concepts. A really appealing concept will depress the scores for all of the other concepts in the batch. Likewise, a really bad concept will tend to raise the ratings of the other concepts. The batch approach, however, is an economical way of weeding out the weaker concepts.

Monadic testing is the recommended method for most concept testing. Interaction effects and biases are avoided. Results from one test can be compared to results from other monadic tests. A normative database can be constructed.

Action standards can be set. The data from monadic tests can be used for modeling and volumetric forecasting (i.e., forecasting year-one sales of the new product once it is introduced). Monadic testing is the superior research design, but it is more expensive on a â€œper conceptâ€• basis than batch tests.

Concept Quality

The concepts that many clients send to us for testing are poorly written and inconsistent in style, appearance and content. Some have illustrations, some do not.

Needless to say, you need to start with good new-product concepts to end up with winning concepts that might make it to market. Generally you must have a small group (no more than two or three people) that reviews, edits and approves all concepts to enforce standards and achieve consistency.

This small group will get very good at improving the concepts because they review and work on so many each year. This small groupâ€™s specialization will lead to much better concepts and it can reside inside of the corporation or inside of the research agency.

Concept Standards

Concept standards are absolutely essential. Standards help achieve consistency from concept to concept. Standards make the results from testing Concept A comparable to the results of testing Concept B. Concept standards can vary

from company to company, but should at least address the following issues:

* Print or video?
* Type and size of illustration(s)
* Content of illustration (include retail package or not?)
* Priced or unpriced?
* Degree of finish (rough or magazine-ready?)
* Style, tone, complexity, and length of copy
* Branded or unbranded?
* Font types and sizes
* And so on.

Without consistent concept standards, comparisons to historical concept tests, or across current tests, are largely meaningless. Also, the differences in concept standards across companies (where standards exist at all) mean that research firmsâ€™ normative databases are largely meaningless. Those firmsâ€™ norms are loaded with historical test results - all based on and biased by inconsistencies in concept style and execution across all of the concept tests in their databases.

Priced Versus Un-priced
Some companies favor putting a price in their concepts, arguing that the inclusion of price makes the results more representative of the real world and more predictive of in-market success. Other companies test concepts un-priced, arguing that the inclusion of price biases reactions to the concept and distracts attention from the conceptâ€™s core content. Whatâ€™s the solution? We recommend testing the concept un-priced first, and then testing it priced later in the questionnaire. The main advantage of testing the concept un-priced is that it permits a series of pricing-expectation questions (which are extremely valuable) to be added to the questionnaire. Demand curves as a function of pricing expectations can be derived, and optimal pricing levels can be determined. If a price is inserted into the concept initially, then the opportunity to accurately evaluate higher prices within the concept test is lost (unless the sample size is doubled or tripled and different prices are tested, but this is very expensive).

Branded or Unbranded
New product concepts can be tested with a brand name, or with a generic descriptive name (e.g., new orange juice). Whenever possible, we recommend testing concepts branded. A brand name conveys information; it suggests and implies things about the product itself. At times, the brand name strongly influences how consumers react to a concept, and this influence can be positive or negative. Generally, a concept test is more accurate and more predictive if the brand name is incorporated into the concept. This is certainly true of established brand names, and itâ€™s usually true even if the brand name is new and unknown.

Sampling Strategy
Two very different sampling strategies are commonly pursued in setting up a concept testing system. Perhaps the most popular is the good old random sample. That is, a sample representative of the total adult population (or whatever the sampling universe is). But, there is a flaw in this random-sample scheme. High-incidence products (i.e., products used by a high proportion of the population) will tend to generate concept scores higher than low-incidence products.

The purists will argue that the random sample is best, since it automatically reflects a conceptâ€™s true market potential. However, we have seen many strong concepts killed unawares by this sampling method -- when the problem was low incidence, not acceptance, of the product concept. A second limitation of the random sample is that it rarely provides sufficient sample size for volumetric forecasting. If a concept is tested monadically among 200 randomly selected adults, and the concept appeals to 15 percent of the respondents, then the volumetric forecast for that concept will be partially based on 30 respondents (indeed a risky foundation for forecasting).

The second sampling approach is to screen for category users (typically defined as someone who bought the product category during a time period defined by three repeat purchase cycles). So if households, on average, buy chocolate once every week, then category users would be defined as individuals who bought chocolates in the past three weeks. The main advantage of this approach is that it provides a much larger sample of category users on which to base volumetric sales forecasts.

The main disadvantage is that the new product might not fall into an established category, or that the new product might redefine the category. In either of these cases the category-user sampling method is likely to provide false signals. Whatâ€™s the best approach? If the product category is established and well defined, use the â€œcategory userâ€• sampling method, but screen a random sample to find the category users. If the product establishes a new category or redefines the category, then use the random sample approach, but bump the sample size up to 400 or 500 soÂ you have at least 150 respondents likely to express interest in buying the new product. This provides the â€œwhole marketâ€• overview of random sampling, but enough positive responders to make reasonably accurate volumetric projections.

The Questionnaire

Naturally, you will want to use or develop your own standardized questionnaire. That is, the questionnaire will be identical from concept test to concept test. You can create places in the questionnaire where customized questions or ratings could be inserted (to better tailor the questionnaire to the product category), but the core questions, and the order and wording of those questions, should be identical from test to test. The recommended questionnaire would typically include:

*Un-priced Purchase Intent: How interested are respondents in purchasing the product after exposure to the concept?
Â Â Â
* Likes: What do respondents like about the new product concept? What elements are appealing?
Â Â Â
* Dislikes: What do respondents not like about the concept? What bothers them?
Â Â Â
* Missing Information: Is the concept complete? Does it contain all of the information respondents need?

* Uniqueness: Is the new product concept unique and different, or is it similar to products already on the market?

* Image Projection: What image is the new product concept projecting? An upscale image? A healthful image? An expensive image?

* Pricing Expectations: What would consumers expect to pay for the new product?

* Priced Purchase Intent: How likely would consumers be to buy the new product at a specified price?* ?Â Â Â Purchase Volume: How often would consumers purchase the new product, if it met their expectations?

* Brand Section: What are reactions to the brand name? Does the brand name fit the product?

* Package Section (if applicable): Is the package acceptable? Does it fit the product? Does it contain all relevant information?* ?Â Â Â Category Usage: Is the respondent a light, medium or heavy user of the product category?

* Demographics: Age, income, gender, ethnicity, and education.

Naturally, every company likes to design a slightly different questionnaire. But the key is to keep it the same over a long period of time so that results of concept tests this year can be compared to tests next year and the years after.

Uniqueness Paradox
Few marketers really understand the role or importance of uniqueness. The more unique a concept is, the greater its chances of success in the marketplace. A highly unique product will tend to enjoy a monopoly since nothing else is quite like it. However, the more unique a concept is, the lower its â€œpurchase intentâ€• scores will tend to be (we call this the â€œuniqueness paradoxâ€•). That is, uniqueness and purchase intent scores tend to be negatively correlated (i.e., the relationship is inverse). We see many highly unique concepts killed by companies every year because purchase intent scores are marginal. Therefore, in modeling the effects of uniqueness, we recommend that you give extra â€œpointsâ€• to the highly unique concepts (to offset the worldâ€™s bias against the new and the different).

Normative Data
The very best â€œnormative dataâ€• is the data that you accumulate from your own concept tests over time. You know exactly how the questions were asked, how the sample was defined, what the concept standards were, how the data were tabulated, and what happened when the new product went to market. If you rely on the normative databases of research companies, you know none of these things. Moreover, research companiesâ€™ databases contain many failed concepts, which means that the research companyâ€™s norms are generally too low. If your new-product concept beats the research firmâ€™s norms, that doesnâ€™t necessarily mean it will succeed in the marketplace. Build your own norms over time. You will know exactly what you have, and then learn how to use and interpret your own norms. The normative databaseâ€™s greatest value, however, is the foundation it provides for building predictive models.

Volumetric Modeling
So you complete a new-product concept test, and you have reams of cross-tabulation tables on your desk as well as historical normative data. What do the numbers mean? How do you interpret them? We firmly believe that some type of volumetric model is essential for correct interpretation of concept test results. This volumetric model should take into account the purchase intent, the productâ€™s uniqueness (or lack thereof), purchase frequency, pricing information, and category characteristics.

We continually see companies killing new product concepts that score below a certain threshold on purchase intent, for example. If those same companies looked at the volumetric (or sales) potential, however, they might come to a different conclusion. A product with a low score on purchase intent, for example, might score high on uniqueness, might command a high price, and might be purchased frequently. When all of these factors are taken into account, it is evident that this product has high sales potential, despite the low purchase intent score. Itâ€™s best to devise a model to forecast the new productâ€™s sales potential, since thatâ€™s the only accurate way to interpret concept test results.

The Product

Concept testing evaluates only the â€œideaâ€• or the mental impression of a new product. The respondent is asked to imagine the new product and have faith that it performs as promised. At some point in the new product development process, the actual product must be developed and tested among potential users. This product testing is absolutely essential to success, and successive rounds of testing are recommended to optimize the product itself. Volumetric forecasts are much more accurate if product testing results are added to the concept testing scores.

Masters of the Future

The companies that are most successful in consistently developing winning new products tend to be the ones that have mastered the mechanics, systems, art, and analytics of concept testing. The best-managed consumer goods companies screen and test hundreds of new product concepts per year. Roughly one-in-10 concepts will be good enough to warrant investments in product development (or R&D) to create the product that fulfills the promise of the concept. And roughly one-in-five of this group will eventually be deemed worthy of taking to market. Itâ€™s a numbers game, and the odds are against you. However, for companies willing to play the game, take the risks and make the investments, the reward is a rising river of profitability.

Â #####

By

Jerry W. Thomas

(Jerry W. Thomas is president and chief executive of Decision Analyst Inc. He may be reached by email at jthomas@decisionanalyst.com or by phone at 1-800-262-5974 or 1-817-640-6166. The company address is 604 Avenue H East Arlington, Texas 76011.)