



QUIRK'S

Marketing Research Review

The insidious top-box and its effects on measuring line share

For companies offering more than one product in a category, market share gain and line share gain are synonymous. Self-evident as this may be, we still see widespread use of marketing research tools that either disregard or ineffectively deal with line share and product interaction. This suggests that many brands are not realizing their full potential, and many growth opportunities are being squandered.

To the extent that growing line share is a business objective, perhaps the most insidious and most dangerous research tool is the “top box” measure of purchase intent. In top-box, respondents view products and check boxes or circle numbers on scales that typically range from “definitely will buy” to “definitely will not buy.”

Top-box is insidious because everything about it seems logical and intuitive - what could possibly be wrong with asking people if they will buy your product? Top-box is dangerous because it easily misleads the marketer with data that may be statistically significant but answers an irrelevant business question. Similarly, TURF and product gap analyses identify product opportunities that many

marketers presume will minimize cannibalization. In fact, this is not always the case. Let's examine each of these approaches in turn - and then explore an alternative that better addresses the central business issue.

Suppose for a minute that consumers had unlimited budgets and an infinite appetite for products. If this were the case, releasing new products would always increase sales and category size. People would buy anything and everything they wanted, and top-box measures of desirability (purchase intent) would reliably estimate relative product potential. “I probably would buy” would mean just that! Products with higher purchase intent scores would sell more than products with lower purchase intent scores and line share would increase ad infinitum.

Limited by budgets

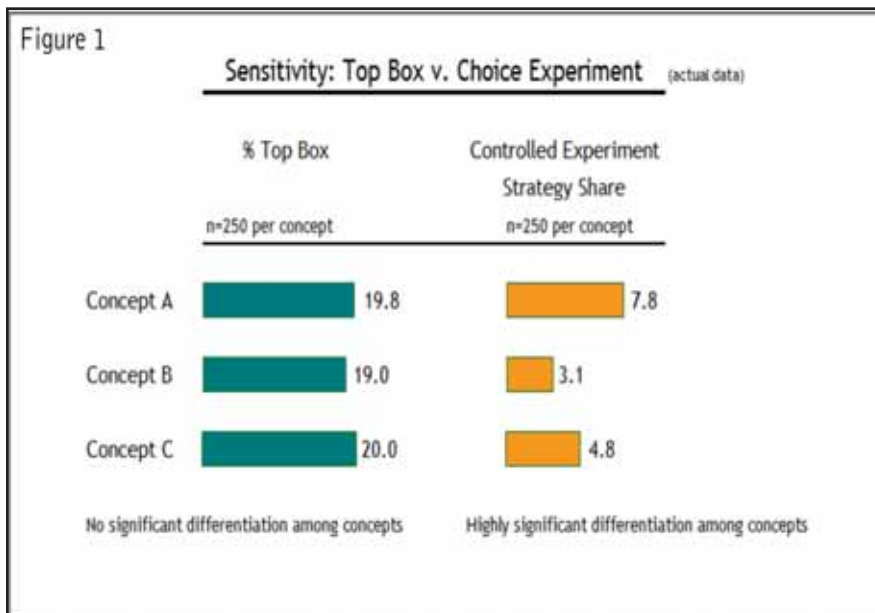
In our world, however, people are limited by budgets and by how much they are willing to consume. Choosing more of one product typically means choosing less of another. Whether we are modifying existing products or releasing

Editor's note: Bob Gerstley is senior vice president of New York research firm Eric Marder Associates Inc. He can be reached at 212-986-2130 or at bg@ericmarder.com.

new ones, the central business question is, “How will the changes we introduce affect choices people make within the category?”

Suppose you conduct a top-box test on two line extensions: “Grape” and “Lemon.” Grape scores higher than Lemon, indicating that people are more likely to buy it. Will Grape increase your line share the most? Suppose that Grape earns much of its share from one or more of your other products while Lemon earns most of its share from competitors. The top-box information is technically correct (Grape is more desirable than Lemon) but also misleading (Lemon is the better business decision).

Moreover, application of statistical formulas can make top-box data seem robust despite its lack of relevance. Sometimes top-box will produce the right answer, but only by luck. When it produces the wrong answer, you usually won't know it - and



your revenue and profit will be lower than they could be.

Inflated results

Another fundamental weakness of top-box is evidenced by how results are interpreted. It is common knowledge that top-box consistently delivers inflated results. To compensate for this, marketers apply norms to adjust the data. Restating this in accurate but less-friendly terms, top-box is a weak measurement that is known to deliver inaccurate results that must immediately be corrected. The typical correction involves application of factors based on historical data which may or may not be relevant to your product and the current market.

Many marketers have become so accustomed to norms that they don't view them as a red flag. Do we apply norms when we look at our speedometer to figure out how fast we are going? Do we apply norms to the thermometer reading when we want to know the temperature? If the criterion we measure is relevant to the business issue and the measurement instrument is valid, norms should not be necessary at all. Put simply, top-box requires norms because it doesn't measure the right thing. Top-box measures desirability, not choice. Overstatement results from the fact that people desire many things but only buy a subset of what they desire.

The very real problem that top-box doesn't measure what businesses

are most interested in is compounded by its general lack of sensitivity. Because it offers few ways to express desirability (definitely will buy, probably will buy, etc.), top-box has difficulty differentiating among similar concepts. Since one of marketing's primary goals is identifying which products will sell the most, this is by no means a minor problem. Top-box is useful for preventing disaster, but its ability to maximize success is limited (Figure 1).

Some marketers recognize that top-box has limitations, but believe that TURF analysis can be used in conjunction with it to minimize cannibalization and identify products likely to grow line share. In TURF, top-box or other desirability ratings are collected on multiple products and then examined for overlap. It is presumed that product sets in which the largest number of respondents have rated something highly will maximize line share.

TURF, however, is simply another way of analyzing desirability data. All the aforementioned problems with top-box (or similar scales) remain. Purchase intent scores in a TURF study still lack sensitivity, and the measurement still overstates potential because it captures desirability, not choice. Common sense dictates that if the primary data on which a model is built is weak, strategic conclusions are likely to be weak as well. This does not mean that TURF has no application.

Products that are undesirable will sink low enough in a top-box measurement that they will usually be ruled out. Further, "big" interactions can be detected even when measurements lack sensitivity. What does this mean for a business? TURF can help the marketer avoid big mistakes and, in the rare instances that they occur, readily identify groundbreaking opportunities. TURF is less capable of effectively discriminating among the more typical strategy enhancements that move businesses forward incrementally.

Explore cannibalization

Another commonly-used approach involves the integration of top-box data with responses to add-on questions designed to explore cannibalization. As we've already seen, top-box addresses the central business issue tangentially at best. Additional responses to questions about uniqueness, substitutability or originality provide only anecdotal information since respondents are not making choices among real products as they do in the real world. Researchers are then left with two weak data sets that answer different questions and must somehow be glued together with analytical assumptions. It is no surprise that many marketers have found this solution to be only marginally effective at predicting cannibalization and hence, overall impact on the business.

Product gap analysis is also commonly used to uncover significant holes in a category. Typically, this involves mapping competing products on several dimensions and looking for uncharted territory. The logic in favor of product gap analysis goes something like this: "If there are no other products like it in the category and if I don't make anything like it, it's likely to grow my line." This logic is flawed. Regardless of its uniqueness, any product filling a gap in the market will necessarily appeal to some cluster of people. Is it reasonable to assume that these people will not share tastes for other products? What if they share a taste for another product made by your company? Digging just below the surface it becomes immediately clear that identifying market gaps (potential product opportunities) and

Figure 2

Illustrative Output from a Choice Experiment

Strategy Shares (Simulated Market Shares)

| | Current Products | Add Line Extension A | Add Line Extension B |
|-------------------------|------------------|----------------------|----------------------|
| Your New Line Extension | N/A | 6.3 | 5.7 |
| Your Current Product | 23.6 | 18.2 | 22.1 |
| Your Line Share | 23.6 | 24.5 | 27.8 |
| Competitor A | 30.7 | 30.2 | 28.4 |
| Competitor B | 24.3 | 23.9 | 22.7 |
| Competitor C | 21.4 | 21.4 | 21.1 |
| Competitor Total | 76.4 | 75.5 | 72.2 |
| Overall Total | 100% | 100% | 100% |

Line Extension A is more desirable than B (6.3 versus 5.7 share points) but when examined in product line context, Line Extension B is the better business decision (27.8 versus 24.5 line share points)

quantifying cannibalization are two very different things. Gap analysis may identify a product that will grow your line share - but then, it may not. As with top-box, if you happen upon the best answer, it's probably by chance.

We should also be wary of pre-post tests when it comes to optimizing line share. Pre-post tests have been falling out of favor but still warrant discussion. In the pre-post test, respondents are presented with a competitive product array and asked to choose products. Following this, they are shown a different product array in which a new or changed product has been introduced and are asked to reallocate their choices. Implicitly, the researcher has said, "Here is something new...I expect you to do something different." Eager to please, many respondents will do exactly what you've asked of them, delivering exaggerated results and setting unrealistic expectations.

Competitive context

The good news is that science can be implemented in marketing through choice experiments. In such experiments, choices are elicited in a competitive context. Note the important difference between measuring and modeling choice. By definition, models are weaker than measurements. In fact, models are generally designed to simulate measurements.

To illustrate, if you want to know the outside temperature, a model

might have you look out the window and count the number of people wearing hats, coats, gloves, sweaters, etc. The data would be plugged into a formula and the output would estimate the temperature. By contrast, a measured approach would have you simply open the window and place a thermometer outside.

Choice experiments are measured approaches that predict market behavior more accurately than modeled approaches such as conjoint analyses or discrete choice. To estimate product potential and line-share implications, choice experiments are conducted as follows:

- Divide a sample into randomly equivalent groups.
- Expose a control group to the relevant set of current, competing products including those you currently offer. Do not include your test product as the control group is your benchmark.
- Expose each test group to the identical array of current products plus one version of your test product. Do not single out test products in any way; respondents should not be able to identify the test product or the variables under scrutiny.
- Test different prices, packages, concepts, positions by systematically varying the test product from group to group - while

holding all competitive products and information constant.

- Rather than asking respondents to indicate what they like or what they think they might buy, have them allocate choices across the product array, such that any product can receive part or all of a respondent's choice.
- Hold secondary marketing variables (shelf placement, couponing, etc.) constant. Having the right products on the shelf will always be the larger contributor to product-line growth, revenue and profit. Introducing too many marketing variables into a study is a common mistake and a sure-fire way to dilute or invalidate findings!
- In each group, the test product and all competing products will earn a strategy share. If the measurement system is valid, these strategy shares will accurately reflect market share potential, holding distribution and awareness constant. Test product(s) potential and cannibalization (source-of-business) are clearly evident when comparing the control group to test groups.

Studies conducted in this way have several benefits. Compared to top-box, they more accurately differentiate among similar concepts because a more sensitive measure is used. They directly measure cannibalization because choices are elicited in a competitive context. Overall impact on the business is much easier to assess. Figure 2 shows exactly what you get from such a study. In this example, line extension A is more desirable than line extension B. In product-line context, however, line extension B clearly represents the better business decision! Source-of-business and cannibalization are unambiguous.

Readers educated in hard sciences may interpret the recommended experimental approach as an application of the scientific method - they would be correct. In the hard sciences, scientific method is generally regarded as the only reliable way to establish cause (if I do...) and effect (consumers will...).

Impact consumer choice

Marketers generally want to know how their decisions will impact consumer choice in the marketplace. If we accept the idea that measuring consumer choice is therefore more relevant than measuring desirability, and if we accept that controlled experiments - the cornerstone of the hard sciences - is a reasonable way of establishing cause and effect, a world of possibilities opens up for marketers.

An important concept in science is that controlled experiments do not care which variable is being manipulated. In a drug experiment for example, each experimental group may vary based on: the medication administered (Is the new drug more effective than the old drug?); the dosage (What is the optimal amount to administer?); the age of the patient (Who responds better?); the length a patient has suffered from an illness; etc.

By extension, choice experiments do not care which elements of the marketing mix are being manipulated. Each respondent group may be exposed to a different: concept (Which will sell the most?); price (to build a price demand curve); feature

(Which feature/flavor/color, etc., sells the most?); product claim (Which benefits sell the product best?); ad (Which induces more people to buy?); packaging; etc.

Conceptually weak

The notion that different research methods must be used to study different marketing variables is conceptually weak and severely limits the marketer's ability to make sound business decisions. There is tremendous value in the ability to make apples-to-apples comparisons of strategies across elements of the marketing mix.

In addition to using choice experiments to more effectively study one element of the marketing mix, choice experiments can be used to compare diverse strategies such as: adding a new feature versus adding a line extension; changing price versus changing packaging; bringing two versus three line extensions to market, and so on. Results estimate the market share potential for each strategy, using the same method and the same criterion.

So what are some of the reasons behind the popularity of tools that weakly (and often incorrectly) quantify cannibalization, fail to properly

differentiate among similar concepts, and use measurements that don't capture the right information?

History. Companies often do what they have always done because it is the path of least resistance.

Perceived risk. Managers sometimes focus on the risk and visibility associated with championing a new idea rather than return on investment and growth potential for the company.

Misunderstanding central business issues. Managers may think that top-box works sufficiently well because it helps generate reasonable volume projections. But estimating volume within +/-10 to 15 percent (that's a 20-30 percent range!) does not mean that you've introduced the best product, nor does it mean that you've minimized cannibalization - and the latter two are likely to be the bigger contributors to revenue and profit.

Worth re-thinking

Since one of the most powerful ways to grow a business is to gain a better understanding of product interactions and thereby increase line share, the tools used to accomplish this may be worth re-thinking. | Q