10 Best Practices to Improve Your Concept and Product Tests

Dr. Bruce Isaacson, President of MMR Strategy Group
Debbie Lesnick, Senior Vice President and Head of Research, MMR Strategy Group
For many companies, the stakes involved in efforts to roll out new products and services have never been higher. In this era of sluggish economies and nervous consumers, new products and services can provide the engine to grow revenues by increasing market share, providing access to new customers, and opening up entirely new markets.

The problem is that while managers’ interest in new products and new services remains high, the odds of success for new products and services remain low. In the marketplace, new products and services often face crowded retail shelves, price pressure from online channels, and skeptical customers who can be difficult to reach and harder to motivate. Also, uncertain markets have caused some companies to reduce development budgets, reducing the resources available to bring new offerings to the marketplace.

In short, developing and launching new products and services is a difficult and important activity.

We believe that testing of products and concepts, when done properly, can dramatically improve the odds of a successful launch for a new product or service. Testing provides measures and diagnostics to focus development resources on the opportunities of greatest potential.

The article draws upon our firm’s experience, which spans more than 35 years of testing, to cut through the jargon and explain how concept tests and product tests work when done right. We’ll also identify some of the pitfalls and traps that often plague testing efforts, and highlight some best practices for testing methods that support the product lifecycle by incorporating the voice of the customer in development efforts.
Why Test at All?

Let’s start with the basics. Why test at all? There are many companies that launch very successful products and services without any customer testing. For example, the legendary Apple CEO, Steve Jobs, was well-known for trusting his intuition and avoiding market research or consumer testing, particularly for innovative and first-generation products.

There is a point here, particularly for radical innovations. Could consumers have told researchers developing the iPod how much they would value a portable menu-driven memory drive to carry their personal music collection? To take a different example, before the microwave oven was popular, could consumers have accurately rated the value of an item that could heat up a frozen lunch in a few minutes?

There are two problems with this line of thinking. First is that most of us do not have the vision of Steve Jobs. Second, the vast majority of development projects involve incremental improvements, such as new flavors, versions or models of existing products, rather than radically new innovations and inventions. With these types of projects, testing provides important and powerful feedback.

The goal of concept and product testing is a simple one: to determine which concepts and products can generate enough interest among consumers to merit further development and which do not deserve any more of your company’s development and marketing dollars. In other words, testing keeps development resources focused on the highest potential opportunities.
Testing typically occurs in three development stages. Those stages include:

- **Concept testing** evaluates new concepts before the product is developed. Concept tests involve ideas for products, rather than actual products or prototypes.

- **Product testing** evaluates new products and new product prototypes. It can be conducted in a laboratory or in-home but does not involve an actual marketplace.

- **In-market testing** evaluates new products and services in markets or channels, where they are offered for sale to consumers to measure consumer response.

As one proceeds through the funnel, the stakes get higher. The rest of this paper focuses on concept tests and product tests.
Concept Testing

Concept testing is conducted early in the development cycle, when there are many potential concepts still being considered, and before significant funds have been spent on any one product or service concept.

Concept testing is one of the most standardized types of research; it uses what researchers call “experimental design” to evaluate a series of concepts under consideration. By keeping the questionnaire and the sample design the same for all cells, and changing only the concept being evaluated, it’s easy to measure the differences in ratings between concepts.

Who should be interviewed in a concept test?

Concept testing should be conducted among the widest segment of the appropriate population. Many grocery products have national distribution and are sold through all major channels, so a test might be conducted among any grocery shopper who would have access to this product. Many grocery products are tested among household grocery shoppers, whose ages, gender and geography are set to match the overall population.
If a product has a specific target, the concept test should be conducted among that group. For example, cosmetics concepts would be tested among female teens and adults, while diaper concepts would use a sample of parents of young children who are responsible for diaper purchases.

For most categories, selecting the sample population is easy. But, it can get tricky with niche or high-end products. You want to cast a net wide enough to accurately represent your sample population, but not so wide as to include people who are unlikely to buy your product. For example, if you are testing a concept for a $500 espresso machine, you probably want your sample to have a higher-than-average annual household income and perhaps a likelihood to consider a high-end kitchen appliance. You must understand your target, and locate respondents who will give you relevant results.
How many concepts should one person evaluate?
Concept testing is most often done one of two ways, either as a monadic test or a sequential monadic test. In a monadic design, each respondent sees only one concept, and then evaluates it. This design provides the cleanest read because there are no other factors to influence the respondent. The disadvantage of this method is that testing multiple concepts can require a very large sample size and is therefore more costly.

In a sequential monadic design, each respondent is exposed to several concepts, one at a time, and then evaluates each. The benefit of using a sequential monadic design is that you get more evaluations of each concept with fewer respondents, typically making it less expensive.

Concepts must be presented in random order because each concept that a respondent sees may affect their evaluations of subsequent concepts. However, it’s important that each concept is shown first enough times to be able to analyze the results by those who saw it first as well as by all respondents who rated the concept.

What exactly does a concept test interview ask?
A typical questionnaire used in concept testing begins by describing the concept, usually with a headline, a basic description of the product, an overview of product benefits and uses, a listing of flavors or varieties, product size, and a picture of the product. At a very early stage of development, a concept may not include a price. Whenever possible, we favor testing with prices, because price is typically an important element of the purchase decision. The concept must be a brief description, not a full-size ad.
The concept should be well-written, easy for consumers to follow, and provide enough information to understand the product - but not too much. If your concept cannot be explained in a simple, concise description, then it is better to consider other types of research, such as qualitative research.

Some truly innovative products cannot be properly explained in a relatively simple concept statement, but almost all incremental products can. For example, the first generation iPod may have been difficult to describe in a few sentences, but subsequent generations, with features such as larger storage capacity or the ability to shuffle songs, would have been easier to explain in a concept test to consumers already familiar with the device.

The concept questionnaire typically asks consumers to rate a concept, using measures such as these:

- **Purchase interest:** How interested the respondent is in buying the product.
- **Purchase quantity:** The number of products a consumer would buy at their initial purchase of the product.
- **Purchase frequency:** How often the respondent would be likely to buy the product.
- **Value for the money:** How the respondent perceives the product’s benefits, compared to its price.
10 Best Practices to Improve Your Concept and Product Tests

- **Uniqueness:** How new or different the concept is.
- **Overall liking:** How much the respondent likes the concept.
- **Believability:** How realistic the concept sounds; how true the claims seem to be.
- **Confusion:** How uncertain the respondent is about what the concept proposes.
- **Brand fit:** How closely the concept seems to match the respondent’s idea of the brand image.
- **Purchase interest by variety:** How interested the respondent is in buying each variety or flavor of the product.

The first three measures are essential and are also used as inputs for first year sales forecasts. Some of the other questions may be omitted, such as overall liking, confusion, and brand fit, but they can provide helpful direction and alert you to potential issues, such as a high rate of confusion or weak fit with the brand.
The next section of the survey to test a product concept may include ratings on diagnostics such as the following:

- **Product characteristics**, such as whether the product would taste good, be light weight, or be of high quality.
- **Product benefits**, such as whether the product would save time, provide a healthy snack, or leave floors clean.
- **Brand characteristics**, such as whether the product is made by a trusted company or by a company that knows this product category.

The last sections of the survey would include category and brand usage questions, followed by demographic measures. These sections can later help profile consumers with the highest level of purchase interest.
How does a concept measure source of volume?

Whether you are testing a concept for a new product or one for a line extension, a source of volume estimate will help you understand what brands will contribute to your sales volume or how much of the parent brand’s share could be cannibalized by this new product.

A simple method is to ask whether this product would be an additional purchase in this category, a replacement for a current brand, or a first purchase in this category. If it will replace current purchases, a follow-up question asks which brands would be replaced.

To get a more accurate measurement of cannibalization, a chip allocation exercise can measure which brands are likely to lose share when your product is launched. There are a number of ways to use chip allocation, and the following example will use 11 chips.

Before they see the concept, respondents are shown a list of current brands in the category and asked to allocate their next 11 purchases across these brands. Someone who is very brand loyal may give all 11 points to the one brand they use, while others may spread out their expected purchases among a number of brands. If a respondent indicates that they would not buy any brands in this category, they skip the exercise.

After exposure to the concept, the chip allocation exercise is repeated, this time including the concept in the brand list. The analysis of the pre- and post-exposure chip allocation responses shows what brands lose “chips,” or potential purchases, to the new brand.

When planning a line extension for a product already on the market, this type of source of volume estimate can help you understand how much of the parent brand’s share could be cannibalized by the new product.
How can a concept test measure which ideas will succeed?

Over time, many companies develop standards or “norms” to evaluate the performance of concepts. Norms are specific to an industry or category, and allow a company to evaluate what types of results should be expected from a concept test.

If past concept tests have included some concepts that have been successful in the marketplace, and others that have not been successful, a new concept under consideration can be compared against past tests to see where it falls. For example, many companies compare tests to the top 25% of the scores for each key measure. If the concept’s ratings fall in the top quartile, that’s an indicator that development of this idea should continue.

To develop norms, all concepts in a category must be tested among the same types of respondents. For example, most grocery concepts are tested among a sample of primary grocery shoppers, diaper concepts would be tested by moms with young children, and cosmetics concepts among female teens and adults.
How else can a concept test focus development efforts?

The concept testing format can also help focus development efforts in other ways, such as refining product positioning or improving packaging.

In a **positioning test**, different groups of respondents evaluate different ways that a product will be positioned. For example, one group may evaluate a product positioned on health benefits, while others evaluate the same product but positioned on taste or convenience. The questions described previously can identify which positioning generates the highest level of purchase interest, or which appeals more strongly to target customers.

In a **packaging test**, several versions of a product package can be tested. In each version, a different element is changed, such as the name, the font, colors, descriptions, package graphics, or other elements. By comparing across different packages, the test indicates which are most appealing to consumers.

Concept testing is the basis of product development. A well-designed and consistent program will help you focus your product development budget on ideas with the most potential for success.
Product Testing

Like concept testing, product testing measures respondents’ interest in buying a new product or service and their ratings of specific product characteristics. However, whereas concept testing tests only the idea, product testing tests the actual product or service, or a prototype version.

What are some of the types of product testing?

Product tests can be used to address a wide range of new business issues. Some of the more common types of product tests include the following:

- **Concept-to-product tests**, in which respondents who are open to buying a concept are asked to try the product that the concept described.
  - At an early stage of development, testing different varieties, flavors, or formulations can help determine which are most appealing to consumers and should be developed further.
  - At a later stage of development, testing may assess how well a product performs compared to what may be expected from the concept.

- **Reformulation studies**, which test the current product versus a reformulation. A reformulation study can be done when a key ingredient is changed to reduce costs or improve quality. By surveying current and prospective customers, the study can test whether the new version performs as well as the current product, and whether core users are alienated by the changes.
• **Discrimination tests**, which measure whether consumers can differentiate one product from another. This type of study tests whether a new or reformulated version is indistinguishable from the current product. For example, a company may use a certain type of orange in its orange juice. If this type of orange becomes unavailable or too expensive, the manufacturer may want to substitute a different type of orange. This research can be conducted with a “triangle test” in which a respondent tries three samples, including two samples of one formulation and one sample of another formulation. If the percentage of respondents correctly noticing a difference between the products is low, the change in ingredients can be made with low risk.
Where is the product test conducted?

Because product testing usually requires that the respondent touch or taste a product, it is often conducted in person. Possible locations include either a central location, such as a mall-based interviewing facility, or in consumers’ homes.

Central location tests are used when all of the products in the study must be prepared or used in the same way. The presence of a professional interviewer provides control over the product’s preparation or usage, such as ensuring that respondents in a medical product test view educational materials, or making sure that a particular food is tested at the correct temperature.

Some products are tested more effectively through an in-home test. For example, frozen foods and beverage mixes may be tried at home, in order to learn about the preparation experience at home, or to get the opinions of other household members. In-home tests also allow respondents to test products over multiple days and in multiple ways. For example, a laundry detergent can be used on different types of clothes, a cleaning product can be tested on different surfaces, or a shampoo or body lotion can be tried out over several days of use.

A home-use test is better than a central location test whenever there are concerns about satiation or aftertaste. A respondent may have a very positive initial reaction to a sample of a rich or sweet product tried in a central location, but the at-home test may reveal that the product is too cloying or filling. Similarly, the heat of a sample of spicy food may be very appealing at first, but its lingering effects or aftertaste may negatively impact perceptions. These effects can be measured more successfully in a home-use test than in a central location test.
Respondents for home-use tests are often recruited online, and then mailed product to try at home. This allows in-home tests to offer better geographic dispersion than central location tests. Respondents for in-home tests can be recruited from any town or city, while central location tests are limited to locations with interviewing facilities. The use of online recruiting for in-home test can also make it easier to locate hard to find respondents because they can be sourced from anywhere, regardless of location.

**Who should be interviewed in a product test?**

Product tests are usually conducted among consumers who have interest in or experience with a product category. For example, if a company wished to test a new formulation of strawberry ice cream, it would make sense to test the ice cream among consumers who eat ice cream and like the flavor strawberry. A moisturizing conditioner is best tested among consumers who use moisturizing conditioners.

It is easy to find users for mature categories that have been around for a long time, but more difficult when a product category is new or a product is very different from other offerings on the market. In that case, potential respondents for a product test might be shown a brief concept description first, with only those interested in the concept invited to participate in the product test.

The respondent base for an improved or reformulated product should include category or brand users, who can indicate whether the new product is at least as good as their current product. If the New Coke example taught us anything, it’s that you want to avoid alienating your core users when changing an existing product.
What does an interview for a product test ask?

For a product test, the questionnaire typically includes questions that can be categorized as hedonic ratings or diagnostic ratings.

Hedonic ratings measure the subjective, sensory experience of the product. For example, a product test for a frozen pizza might ask respondents to evaluate how much they like the pizza overall, or on specific characteristics, such as flavor, aroma, or texture. Hedonic questions are often measured with a 6-point or 9-point rating scale.

By contrast, diagnostic ratings measure specific attributes of the product. Each main component of the product might be dissected into smaller elements, each of which can be evaluated on more detailed criteria. For example, respondents might be asked to rate the frozen pizza’s crust, sauce, cheese, meat, and vegetables overall, and then the same respondents might rate the crust on specific dimensions such as color, crispness, and thickness.

Diagnostic ratings often use Just About Right (JAR) scales. The midpoint of the scale is “Just right” and there are scale points to describe whether there is too much or too little of the element on the scale. Here’s an example of a scale measuring saltiness:

```
1 2 3 4 5
Much too salty Somewhat too salty Just about right Not quite salty enough Not at all salty enough
```

There are certain “rules of thumb” in analyzing the results of a JAR scale. Typically, we look to have a hurdle with 70% of respondents rating any attribute as “just about right”, and the remainder with a relatively even split between “too much” and “too little”.

www.mmrstrategy.com
During the stages of product development, a new product will go through both concept and product testing. When this occurs, key measures from the concept test can be re-measured during the product test. Comparing these two measures can provide very useful diagnostic information:

- **When the concept test is higher:** In this case, the product does not measure up to the expectations set in the concept. In other words, respondents like the idea of the product more than they like the actual product. This may point to the need for additional development and improvement before the product goes to market.

- **When the product test is higher:** In this case, respondents like the actual product more than they like the idea of the product as described in the concept. This suggests that the concept does not convey the features or benefits of the product well enough. Additional work on the description of the product may be helpful when it comes time for activities such as product positioning, advertising, and market strategy.

- **The concept and product test are about equal:** This suggests that the product delivers about what consumers expect based on the product test.
Are products tested branded or unbranded?

Brand matters in marketing, and the influence of brand can affect the results of a product test. The decision about whether to test branded or unbranded depends on the purpose of the test.

To reduce the impact of brand and get a clean measure of product attributes and diagnostics, it is often better to test products unbranded.

This is relatively easy to do. If the product has distinctive packaging, for an unbranded test the product can be repackaged into a plain container with no brand identification. Instead, the product might be identified for testing purposes by means of a code number or a letter.

While products are often tested unbranded, there are times when it is better to test branded products. Brand will affect purchase decisions in the marketplace, so including brand in a product test can provide better estimates of potential purchase and usage.
How many products should one person evaluate?

Like concept testing, product testing can be conducted either as a monadic test or a sequential monadic test.

In a monadic test, each respondent evaluates only one product. This design provides the cleanest read because each respondent experiences only a single product, and evaluates that product without the influence of other items tested. However, this design may be impractical if there are many prototypes to test – each respondent only evaluates one product, so testing multiple product prototypes can require a large sample size.

Sequential monadic designs commonly involve two products. In a sequential monadic test, respondents try one product, rate that product, then try a second product, and rate that second product. After trying and rating both products, respondents may be asked to compare the two products, or may be asked which product they prefer. Sequential monadic designs allow respondents to directly compare two or more products and provide a preference rating.

In many cases, the choice between monadic and sequential monadic comes down to the goals of the research. Here are some examples:

- **Evaluating a change in ingredients or design:** When development creates a less expensive formulation, service component, or product design, they may want to know whether the change is acceptable to consumers. For example, when cocoa prices increase, chocolate products become more expensive to produce. A candy company might respond by reducing the size of a chocolate bar, or reducing the amount of
chocolate in a candy bar – but want to do so without damaging consumers’ perceptions of the product. Testing the reduced-chocolate product against the current version would help the company measure the acceptance and financial benefit of reducing the amount of chocolate in each candy bar.

- **Deciding which formulation to launch**: A company may have developed several formulations of the same product variety and needs to make a decision about which formulation to use.

- **Testing against a specific competitor**: A company may wish to test the strengths and weaknesses of their products compared with a key competitor. A sequential monadic design could be used to evaluate how the brand performs versus that competitor, perhaps using diagnostics from the test to point the way for future product improvement and innovation.

- **Measuring the competitive landscape**: A company may also wish to evaluate against a series of competitors in a category, to understand which are rated most positively among users for core attributes. For example, a food company might be introducing a new breakfast pastry to the grocery market. A sequential monadic product test would let respondents sample several different breakfast pastries, and rate each one. Comparing the new product against established competitors would allow the company to determine how consumers perceive the new product against the existing marketplace, and help the company decide how to position and promote the new breakfast pastry.

Each of these situations can be evaluated with either monadic or sequential monadic testing. The right choice will depend on the budgets available and goals of the research.
Conclusions

As described, in this article, marketers and product managers today have a sophisticated and powerful array of testing methods available to evaluate new concepts and products. Those methods, used properly, can dramatically improve the odds of a successful launch for a new product or service, and can focus product development efforts on the concepts and products with the highest potential.

About MMR Strategy Group

MMR Strategy Group is a full-service market research-based consulting firm. We help our clients grow by leveraging customer insight to develop marketing and sales strategies. In order to support critical business decisions, we combine the data gathering capabilities of a research firm with the business analytics of a strategic consulting firm.

For more information about concept and product testing, please contact us at 818.464.2400 or email info@mmrstrategy.com.

© Copyright MMR Strategy Group, May, 2012. All rights reserved.