Mitigating the Risks of New Product Introductions

An Assessment of Alternative Merchandise Testing Methods

White Paper

November 2012

Introduction

New product introduction is a cornerstone for growth throughout the retail industry. International leaders such as H&M and Zara have demonstrated that speed-to-market for new products can translate into a growing customer base and strong financial performance. New category expansions, such as the entry into home furnishings planned by both of these retailers, create new growth opportunities and solidify their positions as fast-fashion leaders. Other retailers are focused on growth from entering new markets. Japanese retailer Uniqlo, for example, must understand how to appeal to the U.S. customer in order to sustain the anticipated growth in this new marketplace.

Retailers face significant challenges when entering new product categories or markets, or even when introducing new products within existing categories:

- How will a new product perform?
- How well will a brand translate into a new product category?
- How well will the current product assortment sell in a new market?

According to a recent Gartner study of 87 apparel manufacturers, wholesalers and retailers published in Apparel Magazine on June 1, 2012:

> *"New products and expansion into new categories remain the leading drivers of revenue growth. It is no surprise that a significant number of new product launches are deemed failures.*
>
> *Top reasons for failures include product cost issues, being late to market or missing demand, and experiencing inventory shortages that lead to supply disruptions and missed market and trend opportunities."*

In order to mitigate the risks associated with uncertain demand for new products, many retailers turn to product testing. The goal of such testing is to predict future demand of new products and fashion items with little or no sales history, in order to manage optimal inventory investments. Furthermore, product testing can help retailers better understand how to appeal to new customer segments. Just as importantly, early product testing can inform when not to invest in a new item, category or customer segment.

This white paper presents several alternative methods for testing new products prior to introduction and evaluates the efficacy of each method in addressing retailers' objectives.

"One of the most difficult challenges merchants face revolves around product introductions. How will an item that is new to the retailer's assortment or new to the market perform?"

– Gartner, Cool Vendors in Retail, 2012 (April 11, 2012)

The Challenge

Product testing is a complex component of a retailer's strategy, balancing the desire to make informed inventory investments on new products with the need to control operational expenses. A test must be meticulously executed to deliver meaningful, actionable results. It must reach the right customers to forecast future demand with accuracy. It must be fast in order to deliver timely results which support the retailer's speed-to-market objectives. The test should also allow for testing of any product, at any time, targeting any potential customer.

Furthermore, these testing requirements should be achieved while controlling costs and minimizing the distraction to the primary business: that of selling the current in-store product assortment. Testing can deliver significant competitive advantages, however, it should not come at the expense of current business needs.

Traditional Methods

A look into the current state of merchandise testing among retailers reveals several different methodologies, with varying levels of investment and operational support required, and varying levels of accuracy.

No Merchandise Testing

Many retailers trust the judgment of merchants and buyers in predicting new product performance and do not test new merchandise. Other retailers forego testing because of increasing pressure to minimize costs and maximize SKU productivity; they believe creating test product SKUs will detract from the goal of maximizing ROI on production SKUs.

While a "No Test" policy may reduce development, merchandising and operational costs, the true costs of not testing—under-investing in winners while over-investing in losers— can be extremely high. As Wharton School researchers Fisher and Rajaram describe:

> *"[When] the decision to buy is not based on actual sales of the product, but merely on the subjective judgment of merchandisers and buyers about how well it will sell…[we] have found that these subjective forecasts have an error of 50% or more. As a result, retailers frequently buy too little of some fashion products, resulting in lost sales and profit margin, and too much of other products, resulting in excess supply that must be marked down in prices at the end of the season."*[1]

[1] "Accurate Retail Testing of Fashion Merchandise: Methodology and Application." Marshall Fisher, Kumar Rajaram. Marketing Science, Vol. 19, No. 3, 2000, p. 267.

In-store Merchandise Testing in a Cross-Section of Stores

Typically, an in-store merchandise testing program consists of selecting a subset of stores of average store sales volume and displacing current merchandise with test merchandise for several weeks to gauge customer reaction to the test product. Once the sales results are considered significant enough to read results, the test product is removed from the stores (or sent to markdown) and is replaced by the regularly planned product. Usually, tests begin alongside a new merchandise flow, when customers are most likely to be in stores shopping for full-price items. This can yield selling information for the test items, but results in loss of full-price selling of the regular product that has been displaced.

While in-store product testing can be an effective way to understand the customer's reaction to a new product offering, there are several drawbacks to this method. Below we explore each of these challenges.

Issue #1: Inaccuracies
The complexity of in-store testing exposes itself to a variety of issues that can contribute to flawed test results.

First, the retailer must apply an algorithm to the store test results which extrapolates these results to the rest of the store chain. The retailer must create a test group of stores that performs to the average store volume of the chain, or else adjust the sales results to reflect the subset's deviation from the average. Even with these adjustments, the impact of geographic, demographic and psychographic attributes on sales is ignored. Further subsetting to reflect an accurate cross-section of these attributes is complex and costly. According to Fisher and Rajaram, "The decision of how many test stores to use must trade off the increased accuracy that comes from using more test stores against the cost of running the test, which is greater if more stores are used."[2]

Second, reading the test results is not a simple matter of assessing sales volume for the test merchandise. It requires an analysis of external factors that could have impacted test sales. For example, are sales of test product negatively impacted by substitutable items already on the floor, or by markdown items? Or, if testing out of season merchandise—for example, heavy outerwear in the spring—are the results going to be as meaningful if the test product looks out of place in a floor-set?

Finally, the results of the test are only as good as the execution of the test itself. Often, stores will execute the tests differently as it relates to timing of set, receipt of inventory, location in store, and overall procedures. If the product is flawed from a quality or fit perspective, or if test inventory suffers from early stock-outs due to poor sizing allocation, or if the test product is forced into a weak visual presentation due to fixture limitations, the test results can be compromised and yield an inaccurate sales forecast.

Issue #2: Limitations of Out-of-Season Testing
One of the fundamental challenges with in-store testing is that the test product must tie seamlessly to the regular in-store merchandise. For example, a retailer would not want to place a

[2] Fisher, Rajaram, 267.

swim test among sweaters and outerwear in the winter. Such a merchandising disconnect would be aesthetically confusing to the customer, and risk damaging the integrity of the brand.

Some retailers utilize a "sunbelt strategy" whereby stores in geographically warm regions such as Arizona and Florida receive a merchandise test of warm weather product (swim, shorts, tanks) that fully displaces a cold weather presentation (outerwear, sweaters). This test is conducted during winter in order to get reads on spring and summer product. While such a test does accommodate the issue of testing out of season merchandise, the results cannot be accurately extrapolated to the balance of chain. This is because the sales reflect only the behaviors of these "sunbelt" region customers, and the stores are not necessarily representative of the average store in terms of format, demographics or psychographics.

Issue #3: Limitations of Testing New Product Categories or Customer Segments
In-store testing does not allow for testing of new product categories that do not fit within the current blueprint of the store. For example, an apparel store that is exploring adding shoes would not have the requisite fixtures, floor space or sales support necessary to execute a reliable test.

Often the goal of a new product is to broaden not just the assortment, but also the customer base. However, how can a new customer segment be targeted if that customer is not currently in the store? A "girls-only" store that wants to explore the potential in the boys market, a mid-western apparel retailer looking to expand its store base to the West Coast—these are examples of logical brand extensions that would be difficult to assess by putting test product in front of the current customer base.

Issue #4: Scalability and Biased Results
Due to the cost and complexity, most retailers that run in-store testing programs test only a subset of the candidate new products. Merchants typically edit a much larger set of potential new products down to a smaller set that are tested in stores. Often retailers miss out on a potential winning product from among the styles that were eliminated by the merchants through the initial editing process.

Because of the lack of scalability of in-store testing, the initial editing process therefore introduces bias into the new product introduction cycle. Those products that are tested in stores carry an inherent bias from the merchant that performed the editing.

Issue #5: High Production and Operational Costs
Test merchandise requires the same level of attention to detail as regular merchandise: it must be executed to the same quality, fit and finishing standards as regular merchandise in order to preserve the integrity of the brand. These rigorous standards require substantial lead-times, and due to the small scale of test orders and the expedited nature of testing, the production costs associated with test items is usually significantly higher than regular inventory.

In addition to the high production costs, executing in-store testing results in high operational costs. As Fisher and Rajaram describe, "The cost of running a test is incurred from administrative costs, the need to provide extra inventory to avoid stockouts during the test, [and] possibly the cost of air-freighting merchandise to the test stores."[3] This significant financial and operational

[3] Fisher, Kumar 267.

investment is all for product that is already considered "high risk," and will be sold only for a short period of time.

Issue # 6: Slows Down Product Development/Launch Cycle
As described previously, in-store testing requires extensive lead-times to design, approve, produce, ship, execute in stores and read results. This can add up to many months from test conception through execution. As a result, in-store product testing cannot effectively be used early enough in the merchandise pipeline to evaluate designs, when the goal of testing is to make an informed decision on which new products to deliver early-to-market. In-store testing is best used for products that have already been selected, where the results can inform depth of buy or magnitude of demand.

Issue # 7: Distracting from Core Business Needs
The complexity of in-store testing requires coordinated management effort across many business functions: design, production, merchandising, planning, allocation, store operations, and testing analysts. This effort often comes at the expense of core business operations and has the potential of compromising core business objectives.

Issue #8: Opportunity Costs
In order to place test merchandise in-stores, regularly planned inventory must be removed from the floor to yield the space, and as Fisher and Rajaram describe, "[this creates] an opportunity cost on the store space used for the test, because test merchandise by its nature usually sells less well, on average, than regular merchandise."[4] In-store testing creates a trade-off between higher margin planned merchandise and lower margin test product. Because the high margin product is removed from the stores, a loss of full-price sales occurs for the duration of the test. Subsequently, there is often an excess of inventory when the non-test product returns to the selling floor, which in turn calls for inventory to be cleared at steeper markdown.

In-store Testing Using Store Clusters

In 2000, Fisher and Rajaram developed a new approach to in-store merchandise testing by which they clustered stores within a chain that shared similar attributes based on sales history and store descriptors. These descriptors included urban vs. suburban location, format (mall, stand alone, etc.), typical customer ethnicity and geographic location. From each cluster they selected one test store, creating a test store group that was considered optimal. This method is significantly more involved than the traditional method of selecting stores that perform at the average store volume of the chain.

Using the clustering method, Fisher and Rajaram were able to determine a sales forecast for each cluster, which in turn could allow for a better purchasing and allocation strategy based on the performance of each cluster.

While the clustering method appears to provide a more robust forecast for product testing than using the "average" store method, most of the limitations discussed with respect to traditional store testing still apply. The operational costs of these tests are just as high, if not higher, due to the additional analysis required for store clustering. This method also requires significant time

[4] Fisher, Rajaram, 267.

to develop, produce and execute the test; it demands rigorous management attention to execution and operational details; and the results can still be flawed by inaccuracies due to product shortfalls and operational failures. Furthermore, it has the same scalability and bias limitations described previously.

E-Commerce Testing

Another merchandise testing method that some retailers are using is e-commerce testing. With this method, retailers put candidate new products on their e-commerce sites and apply marketing strategies to generate early reads, assessing how online customers are responding to the test product. This method addresses several shortfalls of in-store testing.

Faster than in-store methods
There is significant time saved by not waiting for a product to be shipped to stores and timed with a new floor-set. With e-commerce testing, the product can be posted online as soon as the merchandise is ready to sell, and is not dependent upon fixtures, floor space, or tying into other merchandise stories.

Lower opportunity costs
E-commerce testing does not require that other merchandise be removed from the site in order to sell additional product. Although highly substitutable items run the risk of customer trade-offs, particularly if they are choosing the lower margin test product over the higher margin non-test product, there is less margin risk with this method as the non-test merchandise does not lose several weeks of full price selling as it does with the in-store method.

Ability to test new product categories & customer segments
As described earlier, new category testing can be challenging or even impossible in stores due to fixture limitations. E-commerce testing is a better venue for such new category introductions, since it is exempt from the limitations faced by the store format.

Similarly, reaching new customer segments is more readily achievable through e-commerce because of the nature of online marketing which can direct traffic to the website through email messaging or search engines. In this way, customers that would not be induced to enter a brick & mortar store can be directed to the online channel to explore new products and generate reads on test product.

Although e-commerce testing does have some advantages over in-store testing, several challenges remain:

Issue #1: Inaccuracies
The greatest liability of the e-commerce testing method is that ***the online sales results cannot provide an accurate proxy for the product's performance in the balance of chain.*** The reads demonstrate how the ***online*** customer reacts to the new product; however the behavior of the online customer cannot accurately be extrapolated to the ***offline*** (store) customer in the balance of chain across all store formats, geographic regions, and demographic segments.

According to Forrester Research, 7% of total retail sales occurred online in 2011, and this is

expected to increase to 9% by 2016.[5] It is clear that online shopping will continue to represent a small fraction of total retail business.

Additionally, a 2000 study published by two California University researchers highlights the **differences between online and offline shoppers.** Online consumers tend to be focused and goal-oriented, describing their objective as "buying" rather than "shopping". The paper cites a Jupiter Communications study which found that 77% of shoppers go online with a specific purpose in mind. Nielsen confirms this, indicating that the average e-commerce web visit is approximately 10 minutes. In contrast, offline (store) shoppers are more impulsive and experiential.[6] In addition, Lab42, a market research firm, in an August 2012 study found that **45% of online shoppers had bought something online that they would not buy in stores.**[7]

A 2005 study published in the Journal of Electronic Commerce Research provides an even more detailed analysis of the differences between online and offline shoppers. Online shoppers tend to be "younger, more educated and wealthier and to have a more 'wired lifestyle,' but also more time-constrained than non-Internet shoppers." The study also highlights the differences in the types of products that are purchased online versus offline.[8]

Another challenge with online testing accuracy occurs when testing out-of-season merchandise. Some retailers may view e-commerce testing as beneficial in this situation; for example, a test of swimwear in the winter months could be arranged online without creating an "out-of-place" in-store presentation. However, utilizing e-commerce sales to predict core customer behavior would be misleading when considering who is purchasing the out-of-season merchandise. Sales of swim in winter months would appeal to customers in warm-weather regions or perhaps a niche demographic looking for resort wear. Neither customer base could accurately be used to predict core customer behavior.

Issue #2: Operational costs remain high

Although there are economies to e-commerce testing relative to in-store testing (eliminating product distribution to stores and in-store set-up costs), this method still requires that there is enough product on-hand to fulfill the customer demand. Insufficient inventory will result in stock-outs which will yield poor results and, furthermore, will disappoint customers. Too much inventory will result in low margin product that can only be sold through this channel at steep markdowns.

Issue #3: Distracting from core business

As with the in-store method, all business functions must be engaged throughout the product development and manufacturing phases of the test merchandise. While there are fewer functions

[5] **U.S. Online Retail Forecast, 2011 to 2016,** February 2012, Sucharita Mulpuru, Forrester Research

[6] **Consumer Motivations for Online Shopping,** 2000, Mary Wolfinbarger, California State University Long Beach, Mary Gilly, University of California Irvine

[7] **Shopping and Shipping – A Look at How People Shop Online** – Lab42, Aug 2012.

[8] **A Multi-attribute Analysis of Preferences for Online and Offline Shopping; Differences Across Products, Consumers and Shopping Stages**, Journal of Electronic Commerce Research, VOL 6, NO.4, 2005, A. Levin (Northern Kentucky University), I. Levin & J. Weller (University of Iowa)

operationally involved with the test execution, the e-commerce team is now called into play through this process to execute a controlled test environment that produces usable results.

Issue #4: Slow response time
Although we noted that e-commerce testing can save several weeks of waiting for product to arrive and placed in stores, this method still requires months in the product pipeline. This is critical time lost that can put a retailer behind the curve.

An Alternative Solution Needed—Fast, Accurate and Scalable

Retailers are looking for a product testing solution that is accurate, fast, and cost-effective. They need solutions for evaluating merchandise early in the product development cycle, without the need to manufacture store-ready product, without planning inventory strategies, and without the distraction to their core businesses.

The First Insight solution helps retailers execute product testing with unprecedented speed and accuracy, without the cost and operational burden created by traditional methods. First Insight is easy to integrate into any retailer's current operational structure, and can be used at any point in the product development cycle. There are no constraints to what can be tested: no floor space restrictions, no store presentations to complement, and no seasonality limitations.

First Insight—A Better Testing Solution

First Insight is the leading testing solution that helps retailers and brands maximize profitability of new product introductions— identifying the right products at the right inventory levels and prices— with an eye to controlling the operational costs and minimizing management effort on testing.

Gartner, Inc., the world's leading industry analysis and advisory firm, recently wrote a case study on David's Bridal's use of the First Insight solution. Compared to David's Bridal's in-store testing process, **First Insight improved David's Bridal's forecast accuracy by over 20% and removed approximately 35% from the product launch cycle time by taking a 3 month store testing process down to approximately 1-2 weeks.**

The Gartner study states: *"With this application,* ***the retailer gathered more detailed information about customer preferences than it could after six weeks of in-store testing****. This empowered merchants to make more informed and timely business decisions."*[9]

Below we explore each of the ways that the First Insight solution has turned the limitations of in-store and e-commerce testing into a win for its customers.

[9] **"How Social Gamification Improves the Success Rate of Retail New Product Launches"** Gartner, Inc. , August 12, 2012

Benefit #1: Speed

The First Insight platform enables retailers to test new product with simply a photo or CAD of the item. By eliminating the months of product development, manufacturing and shipment and the weeks of sales required for in-store and e-commerce product testing, First Insight delivers results in as little as 48 hours.

Benefit #2: Accuracy

First Insight's solution provides an accurate forecast of new product performance for several reasons:

- Test respondents are representative of both online ***and*** offline (store) shoppers. Retailers determine the audience for each of the tests and direct the Insights accordingly, typically via email.

- Respondents and input are filtered and weighted to ensure retailers are listening to the "right" customers. Through its multi-stage algorithm, First Insight's solution identifies the individuals who have proven to be accurate predictors of new product performance and filters out the "noise" from respondents who are less predictive of actual results.

  The primary reason for the inaccuracies of most store testing methods is the fact that these methods take a sample and extrapolate the results from the sample to the entire chain. When bias and/or noise is introduced into the sample – due to promotions, store presentation, stock-outs, markdowns, weather, sales associate bias, out-of-season merchandise, etc. – the error rate becomes magnified when applied to the chain. First Insight ensures you are listening to the right consumers from the beginning, filtering out the noise so that the results are predictive. Also, there are no external variables to impact the results of a First Insight test. There is no need to adjust results on the back-end for seasonality, store attributes, or other qualitative considerations.

Benefit #3: Low Operational Costs

Because the First Insight solution requires only a photo or CAD of the test merchandise, there are no costs for approvals, manufacturing, shipping or execution in stores. There are no transactional costs associated with selling the product. All of the substantial costs associated with both in-store and e-commerce testing are removed.

Benefit #4: Minimal Distraction from Core Businesses

Because there is minimal management work required for testing on the First Insight platform, and because there is a high degree of flexibility to run tests quickly and at any time, the demands on the teams are greatly reduced. There is no concern that core business needs will be compromised for the sake of testing.

Benefit #5: No Opportunity Costs

The First Insight platform runs independently of the retailer's channels and there is no customer sales transaction required to test merchandise. Using the First Insight solution solves the issue of displacing non-test inventory and losing several weeks of full price selling, and selling lower margin test inventory at the expense of higher margin regular product.

Benefit #6: Easily test new product categories and out-of-season merchandise

First Insight is more versatile than any other traditional testing method. Retailers can test how well their brand extends to new product categories. They can also run tests to determine the

optimal product assortment for a new market (e.g. a new geography).

Using First Insight also addresses the shortfalls of in-store and e-commerce for testing out-of-season merchandise. ***Because the analysis is not dependent upon translating sales into future demand, there is no risk of listening only to a subset of customers in specific regions or demographics.*** With First Insight, you are able to reach a broad base of customers who represent your core customer base, and thus make accurate decisions based on the right customers.

Benefit #7: Easily reach new customer segments

With First Insight, you are able to target exactly the customers you want to hear from, segmenting the data by demographic, geographic and psychographic attributes. Furthermore, you can test products for a new customer base because with First Insight you are able to target new sets of individuals.

Benefit #8: Scalability; No Bias

Through online consumer engagements using gamification, input on hundreds or even thousands of new products can be obtained in two or three days. Retailers and brands can test ***all*** of the candidate new products, prior to editing by the merchant. This ensures the company does not miss a winner from among the products that never would have made it to an in-store test. One of First Insight's customers, David's Bridal, found their second best selling gown of all time by using First Insight in this way.

Conclusion

Today's highly competitive retail climate requires that merchants bring the right products to market early, and make investment decisions with as much information and as little risk as possible. Traditional testing methods are cost and resource intensive, and often not fast enough to provide an early-to-market advantage.

First Insight provides a testing solution that gives retailers the advantage to make informed decisions on new products with speed, flexibility, operational ease, and above all, accuracy.

"With this application, the retailer gathered more detailed information about customer preferences than it could after six weeks of in-store testing. This empowered merchants to make more informed and timely business decisions."

– Gartner, Inc.: "How Social Gamification Improves the Success Rate of Retail New Product Launches" (August 12, 2012)
