Understanding Consumer Response to Category Management Through Virtual Reality

STEPHEN P. NEEDEL
President
Simulation Research, Inc.

Category management is becoming a common business practice in the consumer packaged-goods industry. However, few research tools exist to help manufacturers and retailers study the impact of category management initiatives. This paper presents a virtual reality shopping simulation, known as Visionary Shopper®, which allows us to consider the sales implications of changing product assortments and layouts in the store, without the cost, timing, or difficulties of in-store testing. A series of examples of how this system has been used are presented.

INTRODUCTION

In their 1991 book The Marketing Revolution, Kevin Clancy and Robert Shulman argue that there will be a marketing revolution in which marketers will be forced to

... do new things in new ways. They will compel marketing departments to abandon myth and ignorance and consider hundreds ... of alternatives to every marketing decision to find the optimal one. And they will hold marketing executives accountable for a measurable return on the marketing investment (Clancy and Shulman, 1991).

Many consumer packaged-goods companies believe that category management is this “new way.” These companies are expending great amounts of energy and cash to be the player in the category management arena. This is often accomplished by becoming a retailer’s “category captain.” Retailers appoint companies to provide management expertise for certain categories or sectors in their stores. The CPGs get a measure of preference in the store in return for managing this business. They maintain their position by enhancing the retailer’s business, even if at the expense of their own. In this paper, we want to discuss the rise of category management as a marketing tool, the need for research techniques that help market- ers to “abandon myth and ignorance,” and provide examples of virtual reality and its ability to provide answers to many of the questions raised by category management.

The easiest place for us to begin is to show you an example of a virtual reality system. This system, known as Visionary Shopper, has been available since 1993, and it simulates an in-store shopping environment. Consumers are intercepted in shopping malls and are brought into the interview site under the guise of testing a new method of in-home shopping. They are seated in front of the large computer monitor, where the interviewer explains the purpose of the test and the operation of the computer simulation (how to move around the shelf and how to pick up, turn, read, and purchase products). Using a touch-screen, consumers can move left or right, up or down, and move closer to and farther away from the shelf. Touching a product brings it to the front of the screen, where it can be turned around to show any side. Touching an icon on the screen pulls up a high-resolution image of the package for readability. A separate icon places the product in the shopping cart. The respondent’s task is simple: shop as they would normally shop.

Respondents typically go on a number of shopping trips, allowing for multiple data points from each respondent and allowing us to introduce manipulations within respondent. The Visionary Shopper system has been validated on a number of occasions, both with respect to its ability to match normal shopping behavior (Burke et al., 1992) and its ability to predict purchasing changes based on manipulation of in-store variables (Burke, 1996). We want to show you how we’ve used this type of system to get at some of the key category management questions.

CATEGORY MANAGEMENT

Nielsen (1992) defines category management as “a process that involves managing product categories...
as business units and customizing them on a store-by-store basis to satisfy customer needs." The ideal in this process is to "identify the optimal product mix and stock each store with the specific products ... customers wish to purchase." While we've given it a new name and brought the retailer into the mix, this is not much different from the age-old definition of marketing as the process of bringing the right product to the consumer at the right price. We can ask, then, why the renewed interest in this activity. We can point to five factors that may have precipitated this trend:

1. The U.S. market is much more fragmented than in marketing's heydays of the '50s and '60s, when conformity was "in" and the population was more homogeneous.

2. Following the demographic and lifestyle diversification, media has fragmented, making delivery of marketing communications much more difficult and expensive.

3. We've recognized, after years of propaganda from the Point of Purchase Advertising Institute (POPAI) that many of the consumers' choices are made in the store. This increases the importance of shelf assortment and shelf layout.

4. U.S. retailers have woken up to the fact that they are not simply a distribution channel, but rather they need to be marketers in their own right. The globalization of everything has made the examples from European retailers available and attractive to American retailers.

5. Thinning retail margins demand methods of cost reduction and space optimization, giving rise to efforts such as Efficient Consumer Response (ECR) and Micro-Marketing.

Armed with sufficient motivation, retailers and manufacturers are actually working together to achieve the goals of category management. We need to recognize that this activity is still very much in its infancy as an accepted practice. More important for marketing researchers, we need to realize that category management is even less developed as an understood practice. We believe that the ECR initiative, as one of many category management programs, has received so much attention because it is a process that is easily comprehended by both sides. Retailers are using scanner data to forecast product movement and stocking needs faster, and more accurately, than ever before. Feeding this process back through the distribution channel to the manufacturer makes ECR an exercise in logistics that is relatively easy to apply. Other areas of category management are much less tractable than ECR; determining "optimal" product configurations and defining superior shelf layouts are two examples of areas where we as an industry are on less surer footing.

Product assortment
The retailer needs to provide the products that consumers want, while attempting to maximize profits from the selection offered. These two goals may be contradictory and may vary by store or homogeneous store group. In this section, we want to talk about how retailers and manufacturers determine what comprises a good product assortment.

The worst and, unfortunately, the most common method for determining which items to add or to discontinue is to use a "Brand Ranking Report." This report, which may go by many names, is a listing of sales for a fixed period of time (usually quarterly) in which each SKU is ranked ordered by its sales, revenue, profits, or, occasionally still, DPP. A retailer looking to add an item to a fully stocked category may scan the list and drop the weakest performer. A manufacturer deciding on a new flavor will often look at a list of rankordered flavors in the marketplace and add the first new flavor that is not in its portfolio. The beauty of this approach is its simplicity—it takes no thought to make the decision. The ugliness of this approach is that it is easy to make a weak decision.

There are two fallacies to the ranking approach:

- First, it assumes all products are equally available to be ranked. While this is usually true when the report is run at the individual store level, it is less common at the retail chain level and even less likely at the market level. A particular product may be very strong in the one chain where it is available, but in mixing that data with that of other chains in the market, the product becomes a small factor.

- The second fallacy is that ranking reports assume equal product desirability across all consumers. Therefore, the worst-selling products are the least desirable and should be discontinued. However, we know that all products are not equally desirable by all consumers. The classic example of this is the split between premium and budget frozen entrees, one of the few cases where household income is an excellent discriminator of purchase probability. Lower-income families buy budget frozen entrees; upper-income households go for premium entrees. A ranking report combines the two groups of consumers and the two groups of products, leading to possible errors in a delisting situation. Again, this is less of an issue at the individual store level when we can assume a homogeneous trading area.

What we need to do is understand the consumer's structuring of our product cat-
category, then match that structure to a particular store. In this way, we can provide a more optimal shelf set to the consumers in that store.

Let's use a simplified example here. Mustard is not mustard; products come in a variety of flavors (e.g., yellow, spicy, Dijon, deli), containers (e.g., jars, squeeze bottles), sizes, and prices. A stocking decision that splits the category into two segments (flavored versus yellow perhaps) and then offers all containers, sizes, and prices in each segment is not likely to be optimized. In some stores, yellow mustards will be woefully overstocked while in other stores the upscale mustards may be wasting valuable space. We might expect a hierarchy of purchase decisions where flavor is the first split. Under yellow mustards, the container may make the next big difference in consumer choice behavior, while under flavored mustard price becomes the purchase driver. (See Figure 1.) The category must be understood from the consumers’ point of view, before undertaking category management efforts.

IRI pioneered behavioral mapping from household panel data in the mid 1980s. The basis of the technique is to take some measure of cross-purchasing, such as combination purchasing, share of requirements, or probability of next purchase matrices, and submit them to cluster analysis or factor analysis. The solutions from these models can provide a spatial representation of the dimensions consumers use in selecting products and place each product on the map. This is very similar to the more common perceptual maps. Products which are proximal in the behavioral maps are similar; this may mean that they are competitive or, if the products are part of a brand, just similar. Products which are distanced are not competitive, in the sense that there is not a constant switching or co-purchasing and are unlikely to be purchased by the same cohort.

While we believe that these behavioral mapping approaches represent a significant leap in research technology, there are problems associated with their use arising from the quality of the data that serves as input into the models. These problems include:

- All of the algorithms for analyzing cross-purchasing assume complete data from the respondent. However, many of our data-collection systems do not provide this complete record. It is common knowledge that neither is true. A single missed purchase can dramatically affect an analysis for infrequently purchased categories. A consistent bias in missed purchases (e.g., one member of the household picks up the products only he or she uses and doesn’t scan) can be even worse.
- Data from time-period-based brand switching/source of volume analyses, where we compare household purchases between two equal time periods, have a share-allocation method for determining brand switching. These techniques tend to inflate category expansion beyond the level shown in scanner or shipment data. Also, the results are highly dependent on how the brand stub is specified in the analysis. Any technique that requires brand stubs to be mutually exclusive and all inclusive, yet produces different answers depending on the level of detail in the stub’s specification, has a problem; it is impossible to determine whether stub specification, and hence the analysis, is correct.
- Most algorithms based on panel data consider the household as the unit of analysis. Categories such as toothpaste, shampoo, cereal, and snacks are individual-oriented. We can be misled into thinking that more switching is going on when we analyze any category at a level higher than the purchase decision or end-user.
- The algorithms do not consider, and generally cannot consider, the internal and external causal factors that drive brand switching and brand loyalty. None of the collection techniques can identify the purchase intention: Did the consumer intend to switch brands or was it driven by external factors, such as out of stocks, coupon availability, pricing, or promotion activity? The inability to know the consumer’s intent and the lack of data permitting us to relate the causal activity to the switch is a fundamental shortcoming of switching methodologies in this context.

We have altered our basic technique to collect data that is not subject to these problems. On any section of any shopping trip we choose we can create an out-of-stock situation. The consumer picks up the product, chooses to purchase it by touching their shopping cart icon, but instead of putting the product in the shop-

![Figure 1 Mustard Purchasing Decision Tree](image-url)
ping cart, a sign comes up that says, "Sorry, product temporarily unavailable. Please select another product or touch the exit sign if you do not want to substitute another." The sign goes away and there is an empty space where the product used to reside. In this way, we get a measure of true purchase intent, the substitute for that specific product (if any), and all in a controlled retail environment.

The example in Figure 2 is taken from a study we conducted in another country, where the client had conducted research suggesting that there was a usage-occasion segmentation of the market. They believed that consumers would buy particular cookies and biscuits for a particular usage occasion (breakfast, dessert, snack, for company, and so forth). What we showed from our analysis is that when it came to substituting for products, brand was much more important than the type of cookie. Consumers remain relatively brand loyal and are more likely to switch within brand rather than between brands when faced with the substitution choice.

Once they understood the brand-loyalty aspect, they looked within brand to see if there was a reasonable internal structure. They found no pattern to the switching, not surprising to them given the proliferation of products they offered and the lack of product differentiation. The client used this data to reduce the number of SKUs on the shelf, saving a variety of costs in the process, with little loss of brand sales.

**SKU reduction and optimization**

Two of the natural outcomes of understanding the market structure for a product category is to look at a shelf set that either reduces the number of SKUs or optimizes the product selection. A research project conducted for the Food Marketing Institute showed two simple but important findings: variety is good, duplication is bad (FMI, 1993). With this in mind, one manufacturer looked at reducing the number of SKUs, with the goal of reducing category clutter and duplication while maintaining its share position. The manufacturer determined that reducing SKUs from 125 by 30 percent led to their share declining by 1.4 points. Reducing the shelf by 40 percent of the SKUs led to a share loss of .67 points. While they hoped to gain share, this minor loss was more than offset by the cost savings involved in the SKU reduction process.

In another case, our client, which manufactures frozen novelties, wanted to strengthen its position in convenience stores, which typically load heavily with local brands. Our client tested four shelf sets (see Figure 3):

- a current set which was overloaded with local dairy products
- the client shelf overstocked the store with our client's products
- the national set with only one or two local products and the rest nationally known brands (our client's products included)
- the premium set, which overloaded our client and competitors' premium products

Results showed that all three new shelf options generate more retail dollar sales compared to the current shelf set. Moreover, the premium shelf set generates higher dollar sales than the other three sets. Our client benefits most when their own products (CLNT) dominate the shelf. However, they also perform better in the national and premium shelf sets relative to control.

Our recommendation was to go with the national shelf set as the easiest and


least self-serving option to sell in to retailers.

Shelf layout

The layout of the products on the shelf is an issue that is not often subject to rigorous research. This is due in part to the high cost of conducting controlled store tests and partly due to the lack of influence many manufacturers have over the planogram. However, we have seen a number of times where making shopping easier for consumers through a rearrangement of the shelf set has a measurable advantage for manufacturers and also, at times, for the retailer.

We conducted a study in the frozen foods section of the store which, at the time, was most commonly stocked by segment. For example, all single-serve entrees were put together on the shelf, followed by family-style entrees, kids’ entrees, pot pies, and so forth. Our client had reason to believe that if they could stock by brand, the consumer would find shopping easier and would spend more at the shelf. We created two different layouts for the shelf in addition to the current set. In the first, brands were blocked together and then segments within the brand were blocked. In the second set, brands were still blocked, but key competitive items between our client and their main competitor were adjacent on the edges of the shelf. Our client had both a price and size advantage: this layout, which we called the “competitive” set, should perform better than the other two sets. As you can see in Figure 4, this was indeed the case.

Our client’s brand dollar sales increased by 16 percent over the control condition when they blocked by brand and put their key products head-to-head with their main competitor. The retailer also won in this scenario: blocking increased dollar sales by 18 percent and 16 percent when adjacencies were also considered. The results of this test were sold in to retailers across the country with many sections now looking like our test condition shelves.

SECTOR MANAGEMENT

Occasionally we test the larger issues in category management. In this final example, a manufacturer of home cleaning products utilized a supplier’s proprietary statistical model to optimize the entire household cleaning aisle. Each of the 31 separate business segments was “optimized,” as well as the two main categories (laundry and household), and then the entire aisle. The results of the optimization model indicated changes to a number of the segments in terms of product selection, product location, and number of facings per product. Being concerned with strictly model-derived results, the client used Visionary Shopper to see whether the model-driven recommendations translated into actual sales changes.

The results of this between-groups test design show that optimizing via this scanner-data-based model does not produce a consumer-friendly outcome. Consumers spent fewer dollars and bought fewer products with the optimized shelf set than they might otherwise do. Consumers purchased in fewer segments with the optimized shelf set (9.3 versus 10.2 segments out of 31 possible segments). For these 31 segments, 14 had higher unit sales with the current shelf set, 6 had higher sales with the optimal shelf set, and 11 showed no difference. (See Figure 5.)
The search for category management models based on scanner data is one of the more important and worthwhile research endeavors we can undertake.

Advantages of virtual reality
There are a number of reasons why virtual reality would be preferred over central location or controlled store tests:

- **Speed**—We often complete a test, from the start of the design phase to presentation of actual results, in four to six weeks. Controlled store tests may take this long just to set up, never mind running them for six months to a year.
- **Control**—One of the problems with other research tools is that experimental control is difficult to maintain. In our system, what the last person sees on the shelf is the same as what the first person saw. We control the retail environment, rather than fall subject to it.
- **Flexibility**—We can create almost any type of shelf layout with virtual reality and can easily make alterations to the shelf set. There are significant difficulties in rearranging entire store sectors for testing purposes.
- **Efficiency**—Using the computer system makes many operational factors quite simple. This lets us charge much less for tests than one might spend in controlled store testing. We also make more efficient use of our respondents. By exposing a single respondent to multiple shelf sets, we can either reduce the sample size required or test more options than might otherwise be considered.

**SUMMARY**

The research industry needs to pick up the pace in development of research tools for category management issues. The search for category management models based on scanner data is one of the more important and worthwhile research endeavors we can undertake. However, caution is required when we evaluate these models. The state of market research today does not yet permit the types of scanner data analyses that can predict consumer behavior based on changing the assortment or configuration of products in a store. Custom research is required to verify that the suggestions from the models will actually produce the intended effects. Systems such as Visionary Shopper can create reasonable simulations of these retail environments and let us make inferences regarding the effects of category management driven changes. This can be accomplished quickly, cost effectively, and reliably, providing researchers with a
powerful tool for understanding the consumer responses to changes in the retail environment.

REFERENCES


Index of Advertisers

American Marketing Association (http://www.ama.org) ................................................................. page 6
The Burke Institute (http://www.BurkeInstitute.com) ................................................................. Cover 4
Decision Analyst Inc. (http://www.decisionanalyst.com) ............................................................... page 1
The Market Research Society ........................................................................................................ page 46
The R&D Initiative .......................................................................................................................... Cover 2
rsc The Quality Measurement Company (http://www.rscquality.com) ............................................... Cover 3