

>appendix

Test Markets

This section examines traditional and emerging designs for test marketing, including the characteristics of six test market types and the strengths and weaknesses of each type.

A *test market* is a controlled experiment conducted in a carefully chosen marketplace (e.g., website, store, town, or other geographic location) to measure marketplace response and predict sales or profitability of a product. The objective of a market test study is to help marketing managers introduce new products or services, add products to existing lines, identify concepts with potential, or re-launch enhanced versions of established brands. By testing the viability of a product, managers reduce the risks of failure. Complex experimental designs are often required to meet the controlled experimental conditions of test markets. They also are used in other research in which control of extraneous variables is essential.

The successful introduction of new products is critical to a firm's financial success. Failures not only create significant losses for companies but also hurt the brand and company reputation. According to ACNielsen, the failure rate for new products approaches 70 percent.¹ Estimates from other sources vary between 40 and 90 percent depending on whether the products are in consumer or industrial markets. Product failure may be attributable to many factors, especially inadequate research. Test-marketed products, typically evaluated in consumer industries, enjoy a significantly higher success rate because managers can reduce their decision risk through reality testing. They gauge the effectiveness of pricing, packaging, promotions, distribution channels, dealer response, advertising copy, media usage patterns, and other aspects of the marketing mix. Test markets also help managers evaluate improved versions of existing products and services.

Test Market Selection

There are several criteria to consider when selecting test market locations. As we mentioned earlier, one of the primary advantages of a carefully conducted experiment is external validity or the ability to generalize to (and across) times, settings, or persons. The location and characteristics of participants should be *representative* of the market in which the product will compete. This requires

consideration of the product's target competitive environment, market size, patterns of media coverage, distribution channels, product usage, population size, housing, income, lifestyle attributes, age, and ethnic characteristics. Not even "typical" all-American cities are ideal for all market tests. Kimberly-Clark's Depend and Poise brand products for bladder control could not be adequately tested in a college town. Cities that are overtested create problems for market selection because savvy participants' prior experiences cause them to respond atypically.

Multiple locations are often required for optimal demographic balance. Sales may vary by region, necessitating test sites that have characteristics equivalent to those of the targeted national market. Several locations may also be required for experimental and control groups.

Media coverage and *isolation* are additional criteria for locating the test. Although the test location may not be able to duplicate precisely a national media plan, it should adequately represent the planned promotion through print and broadcast coverage. Large metropolitan areas produce media spillover that may contaminate the test area. Advertising is wasted as the media alerts distributors, retailers, and consumers in adjacent areas about the product. Competitors are warned more quickly about testing activities and the test loses its competitive advantage. In 2002, Dairy Queen (DQ) Corp., which has 5,700 stores throughout the world, began testing electronic irradiated burgers at the Hutchinson and Spicer locations in Minnesota. No quick-service restaurant chains provide irradiated burgers, although McDonald's and Burger King also researched this option. DQ originally focused information about the test at the store level rather than with local media. When the *Minneapolis Star Tribune* ran a story about the test, DQ had to inform all Minnesota store operators about the article, although all operators had known about the planned test. The article created awareness for anti-irradiation activists and the potential for demonstrations—an unplanned consequence of the test market.² Although relatively isolated communities are more desirable because their remoteness aids controlling critical promotional features of the test, in this instance media spillover and unintended consequences of unplanned media coverage became a concern.

The *control of distribution* affects test locations and types of test markets. Cooperation from distributors is

essential for market tests conducted by the product's manufacturer. The distributor should sell exclusively in the test market to avoid difficulties arising from out-of-market warehousing, shipping, and inventory control. When distributors in the city are either unavailable or uncooperative, a controlled test, in which the research firm manages distribution, should be considered.

Types of Test Markets

There are six major types of test markets: standard, controlled, electronic, simulated, virtual, and Web-enabled. In this section, we discuss their characteristics, advantages and disadvantages, and future uses.

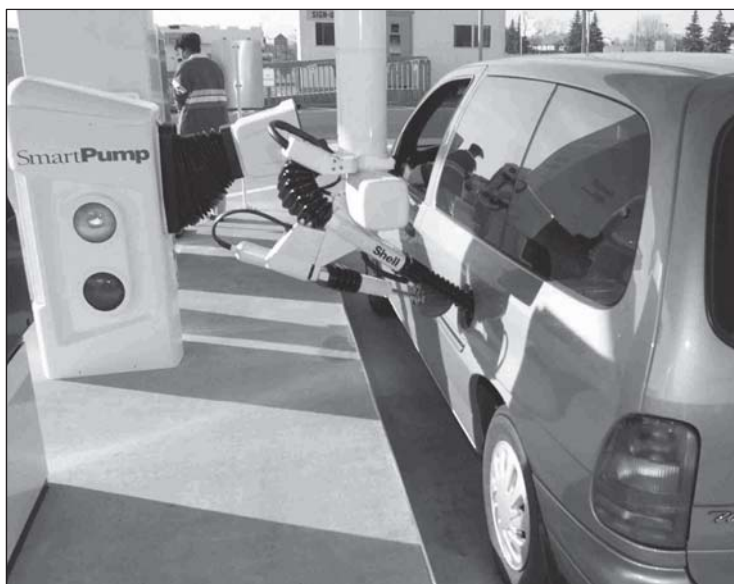
Standard Test Market The *standard test market* is a traditional test of a product and/or marketing mix variables on a limited geographic basis. It provides a real-world test for evaluating products and marketing programs on a smaller, less costly scale. The firm launching the product selects specific sales zones, test market cities, or regions that have characteristics comparable to those of the intended consumers of the product. The firm performs the test through its existing distribution channels, using the same elements as used in a national rollout. **Exhibit Ta-1** shows some U.S. cities commonly used as test markets.

Standard test markets benefit from using actual distribution channels and discovering the amount of trade support necessary to launch and sustain the product. High costs (\$1 million is typical, ranging upward to \$30 million) and long time (12 to 18 months for a go/no-go decision) are disadvantages. The loss of secrecy when the test exposes the concept to the competition further complicates the usefulness of traditional tests.

In March 2000, in an affluent suburb of Indianapolis, Shell Oil Co. test-marketed the first robotic gas pump that allows drivers to serve themselves without leaving their cars. The innovation, which uses a combination of robotics, sensors, and cameras to guide the fuel nozzle into a vehicle's gas tank, took eight years to develop. Its features allow a parent to stay with children while pumping gas and enable a driver to avoid exposure to gas fumes or the risk of spillage, static fire, or even bad weather. Unfortunately, the product requires a coded computer chip containing vehicle information that must be placed on the windshield and a special, spring-loaded gas cap, which costs \$20. The introduction could hardly have been more ill-timed. Just as gasoline prices began their upward advance and the end of winter removed the incentive for staying behind the wheel, Shell planned to charge an extra \$1 per fill-up.³

Controlled Test Markets The term *controlled test market* refers to real-time forced distribution tests conducted by a specialty research supplier that guarantees distribution of the test product through outlets in selected cities. The test locations represent a proportion of the marketer's total store sales volume. The research firm typically handles the retailer sell-in process and all distribution activities for the client during the market test. The firm offers financial incentives for distributors to obtain shelf space from nationally prominent retailers and provides merchandising, inventory, pricing, and stocking control. Using scanner-based, survey, and other data sources, the research service gathers sales, market share, and consumer demographics data, as well as information on first-year volumes.

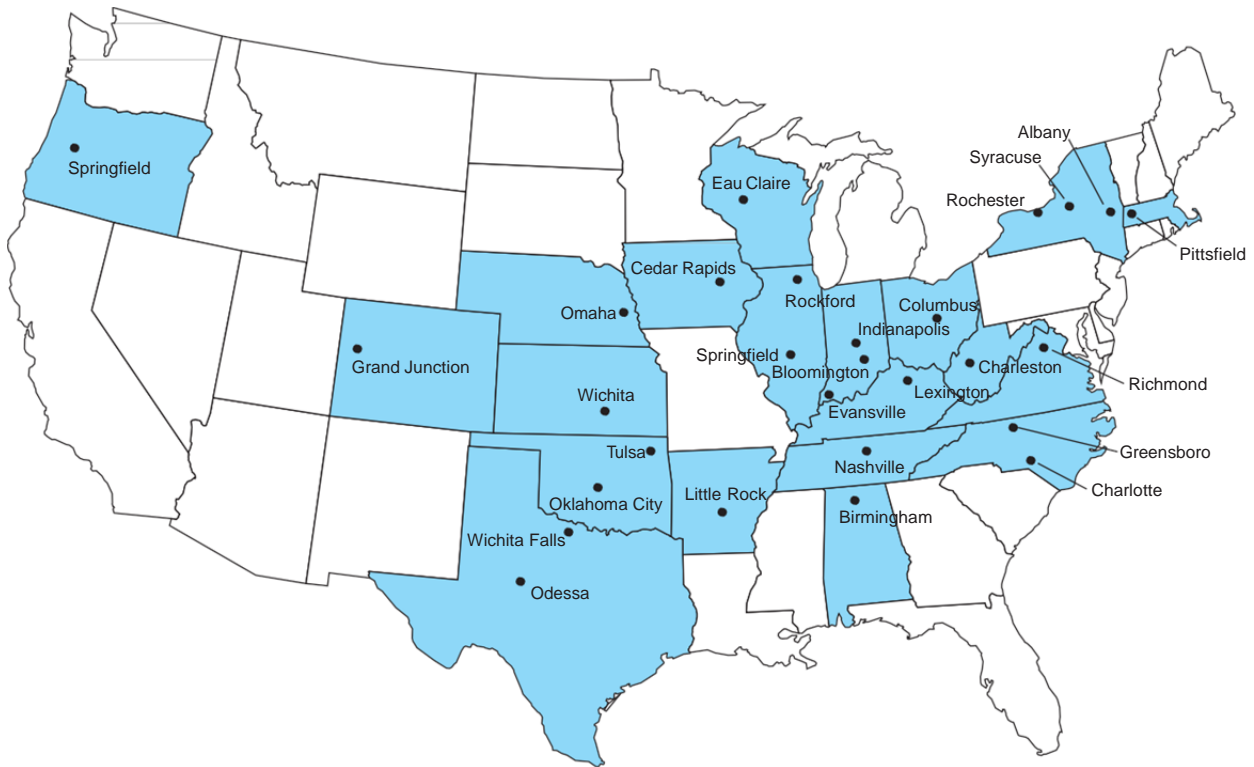
Companies such as ACNielsen Market Decisions and Information Resources Inc. give consumer packaged-goods



The SmartPump is a robotic gas pump that dispenses fuel without the customer ever getting out of the car. Customers pay an additional \$1 for the service.

www.shell.com

>Exhibit Ta-1 Test Market Cities



Source: Acxiom Corporation, a database services company, released its first “Mirror on America” May 24, 2004, ranking America’s top 150 Metropolitan Statistical Areas (MSAs) on overall consumer test market characteristics. “Which American City Provides the Best Consumer Test Market?” http://www.acxiom.com/default.aspx?ID=2521&Country_Code=USA. Also see <http://www.bizjournals.com/phoenix/stories/2000/11/20/daily5.html> and <http://celebrity-network.net/trc/business.htm>.

(CPG) manufacturers the ability to evaluate sales potential while reducing the risks of new or relaunched products prior to a national rollout. Market Decisions, for example, has over 25 small to medium-size test markets available nationwide. Typically, consumers experience all the elements associated with the first-year marketing plan, including media advertising and consumer and trade promotions. Manufacturers with a substantial commitment to a national rollout also have the opportunity to “fast-track” products during a condensed time period (three to six months) before launch.⁴

Controlled test markets cost less than traditional ones (although they may reach several million dollars per year). They reduce the likelihood of competitor monitoring and provide a streamlined distribution function through the sponsoring research firm. Their drawbacks include the number of markets evaluated, the use of incentives—which distort trade cost estimates—and the evaluation of advertising.

Consumer packaged goods are consumer goods packaged by manufacturers and not sold unpackaged (in bulk) at the retail level (e.g., food, drinks, personal care products)

Electronic Test Markets

An *electronic test market* is a test system that combines store distribution services, consumer scanner panels, and household-level media delivery in specifically designated markets. Retailers and cable TV operators have cooperative arrangements with the research firm in these markets. Electronic test markets, previously used with consumer packaged-goods brands, have the capability to measure marketing mix variables that drive trial and repeat purchases by demographic segment for both CPG and non-CPG brands. Information Resources Inc. (IRI), for example, offers a service called BehaviorScan, which is also known as a *split-cable test* or *single-source test*, that combines scanner-based consumer panels with sophisticated broadcasting systems. IRI uses a combination of Designated Market Area-level cut-ins on broadcast networks and local cable cut-ins to assess the effect of the advertising that the household panel views. IRI and ACNielsen collect supermarket, drugstore, and mass merchandiser scanner data used in such systems. The BehaviorScan service makes use of these data with respondents who are then exposed to different commercials with various advertising weights.⁵

IRI's TV system operates as a within-market TV advertising testing service. The five BehaviorScan markets are Eau Claire, Wisconsin; Cedar Rapids, Iowa; Midland, Texas; Pittsfield, Massachusetts; and Grand Junction, Colorado. As small markets, with populations of 75,000 to 215,000, they provide lower marketing support costs than other test markets and offer appropriate experimental controls over the test conditions. Although several thousand households may be used, by assigning every local cable subscriber a cell, the service can indiscernibly deliver different TV commercials to each cell and evaluate the effect of the advertising on the panelists' purchasing behavior. For a control, nonpanelist households in the cable cell are interviewed by telephone.

BehaviorScan tracks the actual purchases of a household panel through bar-coded products at the point of purchase. Participants show their identification card at a participating store and are also asked to "report purchases from non-participating retailers, including mass merchandisers and supercenters, by using a handheld scanner at home."⁶ Computer programs link the household's purchases with television viewing data to get a refined estimate (± 10 percent) of the product's national sales potential in the first year. Consider the observation of a Frito-Lay senior vice president:

BehaviorScan is a critical component of Frito-Lay's go-to-market strategy for a couple of reasons. First, it gives us absolutely the most accurate read on the sales potential of a new product, and a well-rounded view of consumer response to all elements of the marketing mix. Second, BehaviorScan TV ad testing enables us to significantly increase our return on our advertising investment.⁷

The advantages of electronic test markets are apparent from the quality of strategic information provided but suffer from an artifact of their identification card data collection strategy: participants may not be representative.

Simulated Test Markets A *simulated test market (STM)* occurs in a laboratory research setting designed to simulate a traditional shopping environment using a sample of the product's consumers. STMs do not occur in the marketplace but are often considered a pre-test before a full-scale market test. STMs are designed to determine consumer response to product initiatives in a compressed time period. A computer model, containing assumptions of how the new product would sell, is augmented with data provided by the participants in the simulation.

STMs have common characteristics: (1) Consumers are interviewed to ensure that they meet product usage and demographic criteria; (2) they visit a research facility where they are exposed to the test product and may be shown commercials or print advertisements for target and competitive products; (3) they shop in a simulated store environment (often resembling a supermarket aisle); (4) those not purchasing the product are offered free samples;

(5) follow-up information is collected to assess product reactions and to estimate repurchase intentions; and (6) researchers combine the completed computer model with consumer reactions in order to forecast the likely trial purchase rates, sales volume, and adoption behavior prior to market entry.

When in-store variations are used, research suppliers select three to five cities representing the market where the product will be launched. They choose a mall with a high frequency of targeted consumers. In the mall, a simulated store in a vacant facility is stocked with products from the test category. Intercept interviews qualify participants for a 15-minute test during which participants view an assortment of print or television advertisements and are asked to recall salient features. Measures of new product awareness are obtained. With "dollars" provided by the research firm, participants may purchase the test product or any of the competing products. Advertising awareness, packaging, and adoption are assessed with a computer model, as in the laboratory setting. Purchasers may be offered additional opportunities to buy the product at a reduced price in the future.

STMs were widely adopted in the 1970s by global manufacturers as an alternative to standard test markets, which were considered more expensive, slower, and less protected. Although STM models continue to work somewhat well in today's mass-market world, their effectiveness will diminish in the next decade as the one-to-one marketing environment becomes more diverse. To obtain forecast accuracy at the individual level, not just trial or repeat probabilities, STMs require individualized marketing plans to estimate different promotional and advertising factors for each person.⁸

M/A/R/C Research, Inc., has what it calls its *Assessor* model with many features that address the deficiencies of previous STM forecasting models. For example, instead of a comparison of consumer reactions to historical databases, individual consumer preferences and current experiences with existing brands help to define the fit for the new product environment. A competitive context pertinent to each consumer's unique set of alternatives plays a prominent role in new product assessment. Important user segments (e.g., parent brand users, heavy users, or teenagers) are analyzed separately to capture distinct behaviors. According to M/A/R/C, the results of three different models (attitudinal preference models; a trial, repeat, depth-of-repeat model; and a behavioral decision model) are merged to reduce the influence of bias. From an accuracy standpoint, over 90 percent of the validated Assessor forecasts are within 10 percent of the actual, in-market sales volume figures.⁹ Realistically, plus or minus 10 percent represents a level of precision that many firms are not willing to accept.

STMs offer several benefits. The cost (\$50,000 to \$150,000) is one-tenth of the cost of a traditional test

>snapshot

Is Current Test Marketing Representative?

With 70 percent of new product introductions failing, the test market, a long-standing experimental research tradition in the marketing research of consumer goods, is under attack. A test market involves placing a new product in a sample of stores, usually in two or more cities, and then monitoring sales performance under different promotion, pricing, and physical placement conditions. Nonproponents claim the racial diversity of the U.S. population is not well reflected in several U.S. cities often chosen to provide crucial evidence of product viability. Cedar Rapids (Iowa), Columbus (Ohio), Little Rock (Arkansas), and Evansville (Indiana), four popular test market locations, are primarily populated by non-Hispanic white households. Yet 22.9 percent of U.S. households claimed a different ethnic background in the Census Bureau's 2000 decennial census. Information Resources, Inc.

(IRI), whose *BehaviorScan* syndicated research uses similar cities, claims household demographics are only one criterion for site choice. Valerie Skala, IRI's vice president of Analytic Product Management and Development, claims "category and brand purchase patterns, as well as representative retail development," are more important. Test marketing site choice is also determined by cost and availability of retailers willing to stock the new product. IRI and VNU's ACNielsen Market Decisions have contracts with retailers that facilitate immediate placement, eliminating the need to negotiate placement with each store's corporate office. Does the absence of comparable demographics in test markets render the test market methodology faulty?

www.nielsen.com; <http://us.infores.com/>

market, competitor exposure is minimized, time is reduced to six to eight months, and modeling allows the evaluation of many marketing mix variables. The inability to measure trade acceptance and its lack of broad-based consumer response are its drawbacks.

Virtual Test Markets A *virtual test market* uses a computer simulation and hardware to replicate the immersion of an interactive shopping experience in a three-dimensional environment. Essential to the immersion experience is the system's ability to render realistically product offerings in real time. Other features of interactive systems are the ability to explore (navigate in the virtual world) and manipulate the content in real time. In virtual test markets, the participants move through a store and display area containing the product. They handle the product by touching its image and examine it dimensionally with a rotation device to inspect labels, prices, usage instructions, and packaging. Purchases are made by placing the product in a shopping cart. Data collected include time spent by product category, frequency and time with product manipulation, and order quantity and sequence, as well as video feedback of participant behavior.

An example of a virtual environment application reveals it as an inexpensive research tool:

Goodyear conducted a study of nearly 1,000 people. . . . Each respondent took a trip through a number of different virtual tire stores stocked with a variety of brands and models. . . . Goodyear found the results of the test valuable on several fronts. First, the research revealed the extent to which shoppers in different market segments valued

the Goodyear brand over competing brands. Second, the test suggested strategies for repricing the product line.¹⁰

Virtual test markets are part of a family of virtual technology techniques dating back to the early 1990s. The term *Virtual Shopping*® was registered by Allison Research Technologies (ART) in the mid-90s.¹¹ ART's interfaces create a detailed virtual environment (supermarket, bar/tavern, convenience store, fast-food restaurant, drug-store, computer store, car dealership, and so forth) for participant interaction. Consumers use a display interface to point out what products are appealing or what they might purchase. Products, in CPG and non-CPG categories, are arrayed just as in an actual store. Data analysis includes the current range of sophisticated research techniques and simulated test market methodologies.¹² Improvements in virtual reality technology are creating opportunities for multisensory shopping. Current visual and auditory environments are being augmented with additional modes of sensory perception such as touch, taste, and smell.

A hybrid market test that bridges virtual environments and Internet platforms begins to solve the difficult challenge of product design teams: concept selection. A traditional reliance on expensive physical prototypes may be resolved with virtual prototypes. Virtual prototypes were discovered to provide results comparable to those of physical ones, cost less to construct, and allow Web researchers to explore more concepts. In some cases, however, the computer renderings make virtual prototypes look better in virtual reality and score lower in physical reality—specially when comparisons are made with commercially available products.¹³

Web-Enabled Test Markets Manufacturers have found an efficient way to test new products, refine old ones, survey customer attitudes, and build relationships. *Web-enabled test markets* are product tests using online distribution. They are primarily used by large CPG manufacturers that seek fast, cost-effective means for estimating new product demand. Although they offer less control than traditional experimental design, Procter & Gamble test-marketed Dryel, the home dry-cleaning product, for more than three years on 150,000 households in a traditional fashion while Drugstore.com tested the online market before its launch in 1999, taking less than a week and surveying about 100 people. Procter & Gamble now conducts 40 percent of its 6,000 product tests online. The company's annual research budget is about \$140 million, but it believes that figure can be halved by shifting research projects to the Internet.¹⁴

In 2000, when P&G geared up to launch Crest Whitestrips, a home tooth-bleaching kit, its high retail price created uncertainty. After an eight-month campaign offering the strips solely through the product's dedicated website, it sold 144,000 whitening kits online. Promoting the online sale, P&G ran TV spots, placed advertisements in lifestyle magazines, and sent e-mails to customers who signed up to receive product updates (12 percent of whom subsequently made a purchase). Retailers were convinced to stock the product, even at the high price. By timing the introduction with additional print and TV ad campaigns, P&G sold nearly \$50 million worth of Crest Whitestrips kits three months later.¹⁵ P&G's success has been emulated by its competitors and represents a growing trend. General Mills, Quaker, and a number of popular start-ups have followed, launching online test-marketing projects of their own.

Reference Notes

1. Consumer Insight, Fifth Annual ACNielsen Consumer and Market Trends Report, 2001 (<http://www.acnielsenbases.com/news/New%20Product%20Introductions.pdf>).
2. <http://www.iaea.or.at/icgfi/documents/dairy-queen.pdf>.
3. Amy Forliti, "Shell Test Markets the First Robotic Gas Pump," Associated Press, March 9, 2000 (http://www.canoe.ca/TechNews0003/09_pumps.html).
4. <http://www.acnielsen.com/>.
5. http://www.infores.com/public/us/prodserv/factsheet/us_fact_btnewproduct.htm.
6. http://www.infores.com/public/us/news/us_new_102901.htm.
7. Ibid.
8. Joseph Willke, "The Future of Simulated Test Markets: The Coming Obsolescence of Current Models and the Characteristics of Models of the Future," ESOMAR Conference Proceedings, September 2002, excerpted by permission of ACNielsen (<http://www.acnielsenbases.com/news/news%20092002.html>).
9. <http://www.marcresearch.com/main.html>.
10. Raymond R. Burke, "Virtual Shopping: Breakthrough in Marketing Research," *Harvard Business Review* 74, no. 2 (March -April 1996), p. 125.
11. Reg. No. 1,881,580, February 28, 1995, Allison Hollander Corporation of Atlanta, now Allison Research Technologies.
12. <http://www.artechology.com/>.
13. Ely Dahan and V. "Seenu" Srinivasan, "The Predictive Power of Internet-Based Product Concept Testing Using Visual Depiction and Animation," *Journal of Product Innovation Management*, March 2000 (<http://faculty-gsb.stanford.edu/ssrinivasan/rp1502.pdf>).
14. John Gaffney, "How Do You Feel about a \$44 Tooth-Bleaching Kit? Procter & Gamble Discovers What the Web Is Really Good For --Test Marketing," <http://www.business2.com/articles/mag/0,1640,16977,FF.html> (October 2001 issue).
15. Ibid.; Niall McKay, "What Test-Marketing? Why E-Tailers Are Testing the Waters after Jumping In," May 2000 (<http://www.redherring.com/mag/issue78/mag-marketing-78.html>).