



Case Study: Catalog Cover and Insert Test

Optimizing 15 marketing-mix elements in one scientific test

Introduction

Testing is an important part of most catalog marketing programs. List, product, price, offer, and creative elements are frequently tested in catalog and Internet campaigns. But a large quantity of test cells does not always lead to a high quality of market insights.

As a leading consumer-products cataloger discovered, the right approach can make a big difference. With one scientific test—testing catalog cover elements, blow-ins, bind-ins, and other inserts—the cataloger:

- **Tested 15 marketing-mix elements simultaneously in one drop**
 - Split-run tests would have required a sample size of 5.1 million more names or completely missed 6 out of 7 significant effects

- **Quantified 7 significant elements and one important interaction**
 - Four new ideas increased response rate
 - Three new ideas hurt
 - Two elements worked together to increase response even further

- **Increased response rate 18% after implementing the best catalog cover and inserts**

#1: Creative Freedom

A senior vice president of marketing at one of the leading consumer catalog firms was intrigued by the benefits of scientific testing. As a large cataloger with mature marketing programs, the company had to work harder every year to keep sales growing. Split-run testing was useful, but a large sample size was taken up by testing, results were sometimes difficult to compare, and few test cells beat the control.

After getting some guidance, the marketing SVP planned a “scientific” catalog test. They focused on the smaller, more frequent catalog supplement which offered greater flexibility to test more versions with lower production costs. The marketing team and test guide began by reviewing current controls, past tests, and competitive catalogs. They then brainstormed new ideas to increase response rate (below).

Brainstorm List - 48 new ideas to increase catalog response

- | | |
|-------------------------------|---|
| More color | Special offer: discount/free gift/free shipping |
| Testimonials | Page 2 message from president - image/sell |
| Smaller book | Cover product: high-unit/value-priced/image |
| Free shipping | Better up-sell of accessories and products |
| Larger bind-in | Low-price display: red-line vs. compare at |
| Discount level | E-mail or postcard before catalog drops |
| Size of catalog | Messages on cover - number and copy |
| Personalization | Discount (% vs. \$ off first purchase) |
| Sticker on cover | Tear-off special offer page on cover |
| Table of contents | Emphasize special deals (starburst) |
| Pricing - high/low | Logo on cover: size and placement |
| Stronger cover offer | Paper stock - standard vs. textured |
| Less copy on cover | URL and 800 number on all pages |
| Hard-sell order form | Cover: real sticker vs. pre-printed |
| Frequency of contact | More dramatic visuals on bind-in |
| Multiple-order discount | Cover - one vs. multiple products |
| Additional bind-in card | Loss-leaders throughout catalog |
| "New selection" message | Expiration date for special offer |
| More products, less copy | Larger book for best customers |
| 2 different cover concepts | Copy - lifestyle vs. direct offer |
| Real dollar bill as a blow-in | Person with product on cover |
| Brighter colors throughout | Outer envelope over catalog |
| Interactive device on cover | Stronger copy - more direct |
| Custom covers by segment | Free gift - cheap/expensive |

#2: Scientific Discipline

The next step was to concentrate the power of their ideas. Since production costs are often an issue with catalog testing, the marketing executive decided to limit the test to cover and insert elements. With guidance, the team narrowed the list to 15 elements (below), focusing on bold, independent, and actionable changes.

<u>Test Elements</u>	<u>Control</u>	<u>New Idea</u>
A Cover paper weight	Current	Heavy
B Cover concept	Best-sellers	Image
C Logo position	Top	Bottom
D Dot on cover	Dot	Text only
E Dot message	New Spring styles	Save up to 35%
F Number of messages on cover	Three	One
G Personalize cover message	No	Yes
H Phone number on cover	No	Yes
I Page numbers for cover products	No	Yes
J Inside front cover	Control	Table of contents
K Back cover	Control	Personalized message
L Free gift blow-in card	Control gift	New gift
M Retail coupon bind-in	10% off	30% off
N Credit offer bind-in	No	Yes
O Sale items blow-in brochure	No	Yes

Density and Intensity

Since the number of variables has no impact on sample size (a unique feature of scientific tests), the team wanted to test as many cover and insert elements as they could. With greater “density” of test elements, they knew they would learn more, find new ways to increase response, and be able to analyze interactions between elements—where the impact of a variable may change depending on how other variables are set.

Their guide through the testing process also pushed them to maintain “intensity” in the difference between test levels. They worked diligently to push the limits of their current knowledge and test bold changes at the edges of their comfort zone. This also led to the selection of two levels for each test element. By testing a bold change against their control, the team could prove which elements were most important before spending the resources to test numerous levels.

Sample Size

In scientific testing, the number of test elements has no effect on sample size, so the team could use the same sample size for this 15-variable test as they would use for a simple A/B split. With an average response rate of 4%, they calculated that a total of 150,000 names—divided equally among all test cells—was sufficient. This would give them a 50-50 chance of seeing any element that changed response by 5% or more (increasing response rate from 4% to 4.2%) and an 80% chance of seeing any effect that was 7% or larger.

As the team debated sample size, they came to realize that a larger test mailing had important statistical benefits and no real downside. Once the presses were set up for each cover, the cost of producing additional copies was minimal. Plus the team expected that most test versions would perform as good or better than the control. Statistically, a larger sample size would reduce “experimental error,” so more test elements could be significant if they allocated a larger portion of their list to the test. Ultimately, they decided to use 600,000 names for this one 15-element test. (Interestingly, if the team tested all 15 elements as separate one-variable tests, they would have needed 5.7 million names for equal statistical confidence!)

Step #3: Statistical Power

A scientific test design is based on highly-sophisticated statistics, but is straightforward to execute when someone else can take care of the technical details. In this case, the test expert developed the statistical test design and gave the marketing team “recipe sheets” to follow in creating each test cell (or “test recipe”).

Test recipes in a scientific test are much different than common split-run test cells. A scientific test is like a jigsaw puzzle: you have a number of pieces (i.e., recipes) but each piece gives you another bit of information about every variable in your test. The test is very efficient because every piece adds to your knowledge about every test element. Then when you combine results from all test recipes, you clearly see the impact of each element on its own and in combination with others. Instead of just one new variable in each test cell, a number of variables are changed all at once following a precisely-defined scientific algorithm.

Three test recipes for the team’s 15-element test are shown below.

3 Recipes from the 15-element Scientific Catalog Test

Test Elements	Test Recipe #1 (control)	Recipe #2	Recipe #3
A Cover paper weight	Current	Heavy	Current
B Cover concept	Best-sellers	Best-sellers	Image
C Logo position	Top	Top	Top
D Dot on cover	Dot	Dot	Dot
E Dot message	New Spring styles	Save up to 35%	Save up to 35%
F Number of messages on cover	Three	One	One
G Personalize cover message	No	Yes	No
H Phone number on cover	No	No	Yes
I Page numbers for cover products	No	Yes	Yes
J Inside front cover	Control	Table of contents	Table of contents
K Back cover	Control	Personalized msg	Control
L Free gift blow-in card	Control gift	New gift	Control gift
M Retail coupon bind-in	10% off	10% off	30% off
N Credit offer bind-in	No	No	Yes
O Sale items blow-in brochure	No	No	No

This sample of three test recipes shows a few versions of the catalog that the creative team had to produce. Test recipe #1 is simply the control. Test recipe #2 has seven elements set at the control level and eight elements changed to the new level. The changes include:

- Heavy paper stock for the catalog cover
- The dot on the cover has the message, “Save up to 35%”
- Two of the three control messages are eliminated from the cover
- The recipient’s name is added onto the cover (personalize – yes)
- Page numbers are added in a small font beside each cover product
- A table of contents is added to the inside front cover
- The back cover also has a personalized message
- The free gift blow-in card is changed to a new gift offer

After combining the correct elements for recipe #2, the creative team then changes eight elements to create recipe #3:

- Cover paper is changed back to the lighter (control) paper
- The cover products and photographs are changed to the “image” concept
- The personal message is removed from the front cover and back cover (elements G and K)
- The toll-free phone number is added onto the front cover
- The control free-gift blow-in card is used
- The retail coupon bind-in is changed to offer a 30% discount
- A credit offer bind-in card is added

The other recipes are similar—about half the elements are kept at the control level and half are changed (but a different half and half for each recipe). The recipe sheets show the marketing team the exact combinations that must be created for a statistically-valid test. Although these recipes may seem random, they actually follow a very specific statistical algorithm that results in the greatest wealth of information from the fewest test cells. Only a small fraction of all 32,768 (2^{15}) possible combinations are tested, but the selected combinations give just the right data to allow for the independent analysis of each element and key interactions.

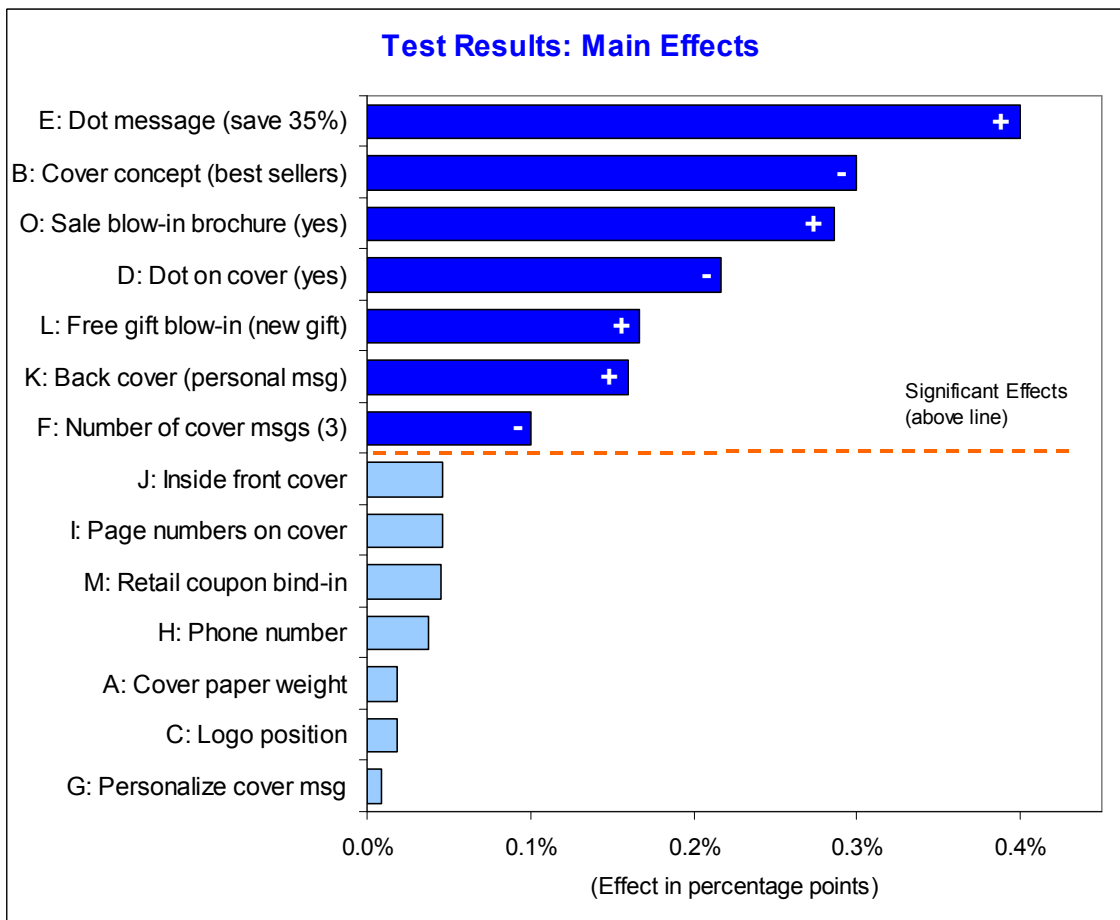
After all catalog covers and insert cards were printed, each version of the catalog was produced and keycoded by recipe number, and a random selection of names was assigned to each. All inside catalog pages remained the same since only the cover and insert cards were tested.

Following their usual response curve, the marketing team collected response and sales data for each catalog test recipe and sent all data to the test expert for analysis. After all statistical analyses were completed, the team met to discuss the results.

#4: Marketing Insights

All 15 main effects (the impact of each element on its own) were summarized in the bar chart shown below. The “main effect” of each element is the increase or decrease in response rate when the control level was changed to the “new idea,” or test level.

In this chart, the main effects are shown from largest to smallest (top to bottom). The name of the element is on the left with the optimal level in parentheses. The length of the bar shows the magnitude of the effect and the sign shows whether the “new idea” is better (+) or the control setting is better (-). The dashed “line of significance” is a measure of market variation. Every effect above the line is statistically significant. Every effect below the line can be explained as a result of natural market variation.



Main Effects

Seven main effects were significant. The four “new ideas” (+ effects) that increased response rate include elements E, O, L, and K:

E: Dot message (save 35%)

The message in the dot on the cover “Save up to 35%” was the largest main effect, increasing response by 10% (0.4 percentage points) over the control message, “New Spring styles.” The marketing team knew that both messages were strong, but this main effect showed that the savings message should have the most prominent cover position.

O: Sale blow-in brochure (yes)

The marketing team developed a small multi-page insert for special sale items. As a separate blow-in, the sale insert could be created at the last minute and altered for different market segments. As they expected, the brochure increased response—and surprisingly average order size as well—far beyond its cost, increasing response 7.2%.

L: Free gift blow-in (new gift)

A gift with purchase was always a big response driver, so the team tested a new gift that was lower-cost, but looked more impressive. The new gift increased response 4.2% without hurting average order size.

K: Back cover (personal message)

A personal message with the recipient’s name on the back cover of the catalog (printed next to the mailing address) increased response about 4%. This surprised some of the team—that personalization on the front cover (G) had no impact, but a personal message on the back was helpful. This also confirmed the benefits of a cover test including changes to both front and back.

The three negative effects—where the control level was better—include elements B, D, and F:

B: Cover concept (best sellers)

When best-selling products were on the cover, response was about 7.5% higher (0.3 percentage points) than with image products. The team would not always have the same products on the cover, but this effect showed that a selection of the best-sellers should be on the cover, rather than higher-priced image products with stylish photos.

D: Dot on cover (yes)

The team knew that a dot (or “dot-whack”) increased response, but they had never tested it alone to quantify its impact. Increasing response 5.4%, the team confirmed the value of a dot on the catalog cover. Discussing these results, they decided that a future test should include different alternatives: printed dot vs. sticker, circle vs. star, standard color vs. fluorescent, etc.

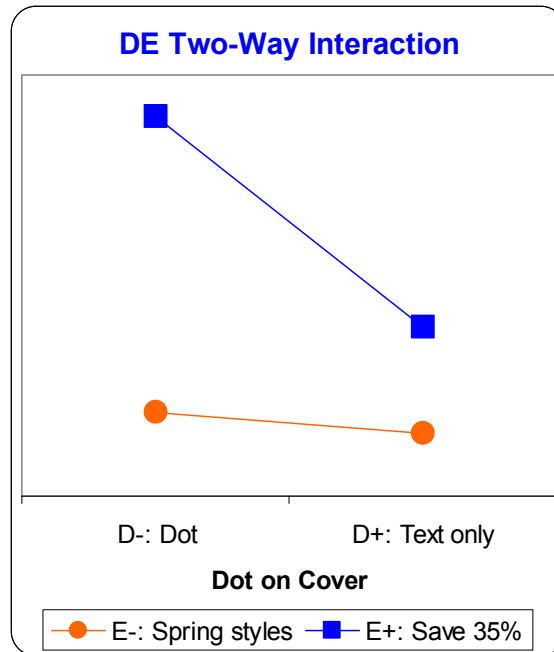
F: Number of cover messages (3)

Fewer messages on the front cover hurt response. Three cover messages increased response about 2.5% over one cover message. This helped lay to rest the complaint of “too much copy on the cover.”

The effects that were not significant were also valuable. The cover paper weight (element A) had no impact, so the team could keep using the lower-cost paper. Since the phone number and personalization (H and G) had no impact, the team decided to keep them off of the front cover. Although the higher-value retail coupon (M) had no significant impact on catalog response, the marketing team could not accurately measure the change in retail sales, so they wanted to reassess it in a future test.

Interactions

After analyzing all 15 main effects, the test guide analyzed potential interactions. In this case, one two-way interaction was significant: between the dot on the front cover (D) and the dot message (E).



This interaction plot shows how the impact of an element changes depending on how another one is set. The chart supports the main effects: “Save 35%” (blue line on top) is the better message and a dot (blue and orange points at left) is always better than no dot. However, the two-way interaction shows that the combination of the dot and “Save 35%” message is much greater than the sum of the two main effects. In other words, “Save 35%” has an even greater impact in the dot than as a text message in the same location on the cover. Alternatively—moving from right to left on each line—the dot has a greater impact with the “Save 35%” message than with the “New Spring styles” message.

#5: Profit

When the marketing team created the optimal catalog and tested it against the control in the next drop, response rate jumped over 18% .

The marketing SVP was impressed how much the team learned after only one test. The team tested about five-times more variables than usual, uncovered an interaction impossible to find with split-run tests, and increased response over 18% by simultaneously optimizing 15 marketing-mix elements. With a motivated team, the executive continued using scientific testing in catalog campaigns and spread the techniques to their Internet and retail programs.

© LucidView 2005. All rights reserved.