TESTING DIRECT MAIL, TELEPHONE, AND WEBSITE PROMOTIONAL DESIGNS: A CASE EXAMPLE

Joseph L. Orsini, Ph.D., Professor of Marketing California State University, Sacramento, CA 95819-6088 [tel] 916-278-6992
Miguel de Anquin, VP of Marketing, Premier Power Renewable Energy Inc.

ABSTRACT

The testing of direct marketing promotional designs is the marriage of traditional market tests with scientific experimental methodology. It is a powerful tool that allows the direct marketer to choose, with a known degree of statistical certainty, the most profitable mailing package or website design. This paper discusses the use of experimental designs in the testing process, and provides an example of a simple fractional factorial design, which may be used in a classroom demonstration or student assignment.

BACKGROUND

As has been pointed out in numerous venues, rapid advances in technology have been a driving factor for business at the dawn of the 21st century. In particular, with increases in computer database storage capacity and the development of the internet, marketers have had available opportunities to measure the effects of promotion on an individual level, in the form of responses to direct mail, telephone, or business website. These, in turn, led to the development of the subdiscipline of direct and interactive marketing.

Testing Promotional Designs

Basic Testing

The basic test in direct mail promotion has been the "A/B Split". In this form of test, two competing versions of the promotion are distributed to random samples of the market. The responses of each group (A or B) can then be compared, costs and profits of each computed, and the best option used in the full market promotion. The A and B version may differ in only one aspect, or ***factor***, of the promotion (e.g. a price of $7.95 vs. a price of $8.49), or may differ on many factors at once.

Testing Using Full Factorial Designs

The immediate need of the direct marketer is to develop a promotional package that works. Because there is usually not sufficient time to sequentially test all of the variables that may effect the profitability of the campaign, simultaneous testing of several promotion factors is desirable. The principles of design of experiments can be fruitfully employed to increase the efficiency of promotional tests. The application of these principles is relatively new to the field of direct marketing, but they have already proven their profitability by using more efficiently designed tests.

When more than one promotion factor is utilized in a given test, information is obtained which would be difficult or impossible to obtain even in a long series of the simple "A/B Split" tests. In addition to testing several factors at once, there is the prospect of ***interaction*** effects among the items tested (i.e. a "synergy") which should be considered. Interaction effects may lead to a case, for example, in which individual testing of lists identifies list X (e.g. moderate income prospects) as the most profitable, and individual testing of price identifies price Y (e.g. a high price) as the most profitable, but the *combination* of moderate income list X and high price Y is not the most profitable combination of list and price because of the "synergy" of those two factors. Understanding this interaction effect can be very important in determining the best overall combination of factors.

A ***factorial design*** permits the researcher to evaluate not only the main impact of each test factor, but this "synergistic" or combined effect of two or more factors. When all the possible combinations of all the factors being tested are included, the design is called a ***complete factorial design***. Unfortunately, the number of tests required increases quickly as the number of factors, and levels (or categories) of each factor increases. However, methods have been developed which substantially reduce the number of different test conditions required, while retaining important information.

Testing Using Fractional Factorial Designs

When there are many factors being simultaneously tested, it is desirable that the design will consist of fractions of the full factorial design; otherwise, there are simply too many combinations of levels of factors to carry out the experiment in reasonable time and at reasonable expense. These are called ***fractional factorial designs*** – designs in which only a portion, or fraction, of all of the possible combinations of factors are included. While these methods have not previously been widely used in marketing research,

they have a long history of application in the physical sciences, and are more recently being used in website and database testing.

This type of design works by simply not testing every possible combination of the full factorial design. While some information is indeed lost by using fractional designs, careful design will not lead to the loss of important information. What managers often consider unimportant is what statisticians call ***higher-order interactions***. These are interactions resulting from combinations of three or more factors, above and beyond the effect of each factor itself or the effects of each two-factor combination.

Experience indicates two important aspects of higher order interactions that allow us to successfully utilize fractional designs. First, in the real world the occurrence of *significant* higher-order interactions is infrequent. Second, higher-level interactions are very difficult to interpret, and for most managers do not form a very comfortable basis for action.

If the higher-level interactions occur infrequently, and when they do they are not particularly useful, then it does not seem worthwhile to spend time and money to measure them. Managers do want information about the effects of the main factors themselves, however – the effect of price differences, the effect of different graphic backgrounds, and so on. It is also typically desirable to have information on the two-factor interactions, which occur reasonably frequently in the real world and are useful in determining promotion design.

The advantage of the fractional factorial approach, therefore, especially for designs with more than three factors, is that they do not lose much useful information, while maintaining their cost savings. The trick is to design fractional factorial experiments that unambiguously measure the basic impact of the main factors themselves, and as many of the two-way interactions as the manager wishes, while not testing the higher-order interactions.

A Direct Mail Case

Premier Power Renewable Energy Inc. provides solar power systems for residential, agricultural and commercial applications. Located in a major metropolitan area of northern California, the company has been in business for over twenty years.

One of their promotional methods is the mailing of flyers to consumers, where the company frequently performs the "A/B Split" type of testing. The factors and alternatives they were interested in evaluating for this test were as follows:

- Flyer headline: *antipathy* toward paying high electric bills vs. presentation of a *new product* (solar tiles)
- Type of graphic used: photograph of satisfied *customers* vs. photograph of the *product*
- The "next step" for an interested customer: obtaining a more detailed *brochure* vs. arranging a *site survey* for a solar tile application.

The Research Design

Rather than performing three separate "A/B Split" testing efforts for each factor, for a total of six test conditions (3 x 2 = 6), the company considered a factorial research design. A full factorial design would require eight tests: one for each unique combination of the three factors at two alternatives each (2 x 2 x 2 = 8). For example, one of the tests would be a flyer with the *antipathy* headline, with a graphic showing satisfied *customers*, and indicating how a *site visit* could be obtained as the next step in the search for the product. All other aspects of the flyer would be the same for all tests.

The advantage of a full factorial design is that it would not only test each factor, but also test the effects of the three pairs of factors in combination with each other, and the effects of all three factors combined. The disadvantage is cost: eight tests vs. six tests. Desire for appropriate statistical accuracy indicated that each test would require 2,500 mailings; therefore the three A/B Split tests would require a total of 15,000 mailings, while the full factorial design would require 20,000 mailings.

An alternate and somewhat more elaborate design was considered: a one-half factorial design using the three factors with two alternatives for each factor (described above). While not providing all the information of a full factorial design, it would provide more than the three A/B Split designs, including each of the three test factors and the combination of the three pairs of factors. Also it would be cheaper, requiring only four tests, or 10,000 mailings.

The direct testing of a 3 X 2 half factorial design produces confounded results, where each of the main effects (factors 1, 2 or 3) is confounded with one of the two way interactions, and the three way interaction is a constant However, the three 2-factor tests are full factorial 2 X 2 designs. For a more complete discussion see Montgomery 2001.

The fractional factorial design was selected and implemented. The four test flyers were prepared and

printed, and the flyers randomly distributed to the list of 10,000 potential customers and mailed (2,500 receiving each flyer). Response to each of the alternatives, through return mail or website access, was kept track of through the use of four separate flyer test codes: 101, 102, 103, and 104.

Findings

The response rates (percentages) are indicated in the table below; actual numbers are disguised for confidentiality; see the Appendix.

Test	Response
101	1.2%
102	7.2%
103	4.0%
104	2.8%

Average 3.8%

Figure 1 of the Appendix illustrates the fractional factorial design. In a full factorial design each corner of the cube would represent a test condition; e.g. Test 101 would be a flyer containing an *antipathy* headline, a graphic illustration of satisfied *customers*, and the next step of requesting a detailed *brochure*. Because a fractional design was used, only the four indicated corner combinations were tested.

Figure 2a shows a complete full factorial 2 X 2 design, analyzing headline and graphic. In effect, the three dimensional cube has been collapsed on the next step variable; similar collapsing for analysis is illustrated in Figures 3a and 4a. The results of the analysis of Figure 2a are shown in Figure 2b: there are significant main effects of each variable ($p = .05$), but no significant interaction. Figure 3b shows the results of the headline X next step analysis, indicating there are significant main effects of headline, no significant main effects of next step, but a significant interaction between the two variables. Similar results are found in Figure 4b for the analysis of graphic X next step: significant main effects of graphic, no significant main effects of next step, but a significant interaction between the two variables.

Analysis of the results of the "three variables" trial (headline, graphic, and next step) illustrates the benefit of performing a factorial design in direct market testing If the typical "A/B Split" test was performed for different headlines, and the background for both test versions was the *site visit* alternative of next step (Fig 3b), then there would have been no significant difference between the two headlines; a similar result would have been found for a test of the two types of graphics (Fig. 4b).

However, if the background for both tests was the *brochure* alternative of the next step variable, rather than a *site visit*, then there would be significant differences in the headlines used (Fig 3b), and significant differences in the type of graphics used (Fig 4b). By simultaneously testing the next step variable in conjunction with headline and graphic, the interactive effects between next step and headline, and between next step and graphic, were discovered. In other words, there were synergistic effects of the combination of next step with both headline and graphic. The *site visit* had a suppressive effect on the attractive alternatives of both headline and graphic, yet the difference in the next step alternatives (*site visit*/*brochure*) did not, by itself, appear significant. These results would have also been discovered in a full factorial design. However, a full factorial design would have required sending out a mailing of 20,000 brochures, while the fractional factorial design used in this test required sending only 10,000 brochures, even less than the three A/B Split tests. This substantial savings can be applied to future testing.

Conclusions

While making conclusions based on one test must always be approached with caution, the following may be inferred. First, the type of headline does seem to have an impact on the response rate. Second, given that this is a relatively new product class on the market, an ad showing the product appears to provide more useful information than seeing apparently satisfied customers. Third, and perhaps most important, the market appears to want the "low risk" next step of seeing a more detailed *brochure* on the product rather than having a salesman call (*site visit*).

However, further analysis on the purchase rate for each next step group would appear to be warranted. It could be, for example, that the purchase rate for those wanting a *site visit* is much higher than those wanting a *brochure*. Thus, it could be more **profitable** to use the *site visit* alternative in the promotion even though the initial response level is not as high as the *brochure* alternative.

(References available upon request)

Appendix - *Design and Findings*

Figure 1

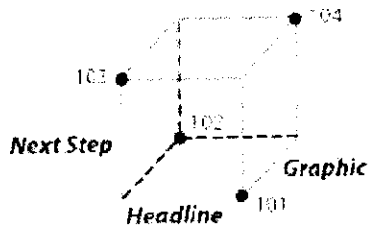

Figure 2a ***Headline & Graphic***

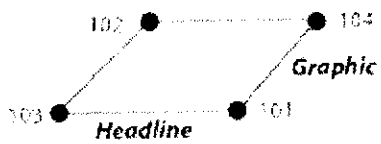

Figure 2b

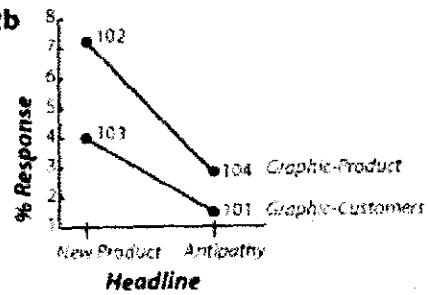

Figure 3a ***Headline & Next Step***

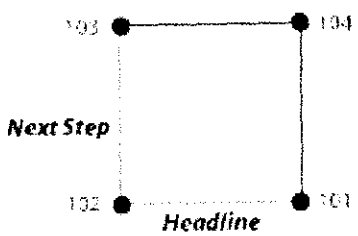

Figure 3b

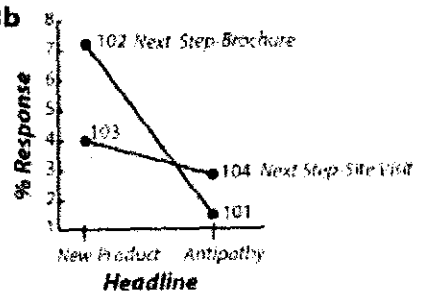

Figure 4a ***Graphic & Next Step***

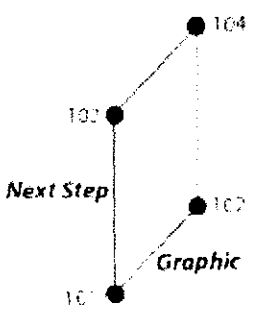

Figure 4b