



Achieving Breakthroughs in Non-Manufacturing Processes via Design of Experiments (DOE)

Inspirational examples* compiled by Mark J. Anderson, Principal, Stat-Ease, Inc.

**If you do not see your application here, that does not mean that it cannot be done. Press ahead and give two-level factorial design of experiments a try. If/when you make a breakthrough, let me know by emailing mark@stateease.com — I will add it to our list of successes. Thus you can ‘pay it forward’ by providing inspiration for others in your field. Mark*

Advertising, Sales & Marketing:

- Could response rate be improved by going from two- to four-color printing on a direct-mail postcard? In addition to this factor, the experimenters looked at two sizes (small vs big) and two types of stock (thin vs thick). The eight different designs ($2 \times 2 \times 2 = 8$) were sent to eight equal segments of the company’s client list. Surprisingly, the highly technical audience responded better to two colors, thus saving many thousands in needless printing costs.

“If you test factors one at a time, there’s a very low probability that you’re going to hit the right one before everybody gets sick of it and quits.” (Forbes)

(*Stat-Teaser*, Summer 1996, “New Spin on DOE from Forbes Inspires Case Study by Stat-Ease® Marketers,” Anderson. A recap of this experiment, plus further thoughts and advice based on subsequent trials with direct mailing, was provided in a follow up article by Anderson titled “A Crash Course on Sales and Marketing,”** which can be found posted at www.stateease.com/news/news0712.pdf (originally published in *Stat-Teaser*, December 2007).**

***The two-day Crash Course on Sales and Marketing workshop, which teaches how to put DOE to profitable business use, is available exclusively for on-site presentation. For details, go to www.stateease.biz. The author, Dr. Paul Selden, has uncovered dozens of examples proving that this powerful tool can yield tens of millions of dollars by identifying what appeals most to customers and by eliminating expensive traditional practices that prove ineffective.*

- Crayola conducted an experimental design to help attract people to their new Internet site via an e-mail to parents and teachers. They discovered a combination of factors that made the best script 3.5 times more effective than the worst script. (*Harvard Business Review*, October 2001, reprint R0109K, “Boost Your Marketing ROI with Experimental Design,” Almquist, Wyner.)
- A travel club wanted to increase response to direct mail. It tested 17 factors in 20 mail pieces, including: copy on envelope and cover letter, the offer, graphics, and even fonts and logos. They found that short and simple worked best, text on envelope helped, but a free offer made no difference. An extra insert actually lowered response. The DOE increased predicted response from 0.3 to 0.5%, worth \$20 to \$40 million in annual revenue. (*Multivariable Testing Methods in Marketing*, Gordon H. Bell.)

- A more modern variation on the test described above tested 19 factors on an e-mail advertisement (some people call this “spam” ☹). The marketers did all this in only 20 variations of their message sent to a total of 500,000 recipients. The biggest benefits came from setting a deadline of 3 days on their offer and including more products. Being more creative in their subject line, which the staff anticipated would be helpful, turned out to be a negative factor. Armed with this information on significant pros and cons, the e-mailers achieved a 24 percent increase in response on their next ‘drop.’ (“Beyond the A/B Split,” *Target Marketing*, October 2003, p. 111, Gordon H. Bell. For more details on this study and many, many, more, see the author’s web site: http://www.lucidview.com/case_studies_articles.htm.)
- A major web retailer’s designed experiment on an e-mail solicitation studied these factors in a half-fraction of all possible two-level combinations:
 - The subject line
 - The salutation (such as; “Greetings!” or “Hi [customer name]”)
 - A call to action (pitch to take the next step)
 - A promotional giveaway (such as entry in a \$100 drawing or a \$25 gift card)
 - Closing statement

The 16 e-mail layouts were randomly sent to potential customers and the number that responded was recorded. Among the most coveted group of customers, the response rate was 75% greater for the best combination. (From proceedings of ASQ’s 2009 World Conference on Quality and Improvement for the “Designed Experiments in Service Quality Applications” talk by Lou Johnson and Gordon Bell.)
- A 2x2 factorial design investigated the effect of ad size versus product category on consumer recall. The effect was positive, but the benefits did not outweigh cost. (*Journal of Advertising*, Volume XXIV, Number 4, pp. 1-12, “Ad Size as an Indicator of Perceived Advertising Costs and Effort: The Effects on Memory and Perceptions,” Homer.)
- A factorial design revealed that brand memory of split TV ads improves with identical spots if presented after a short interval. (*Journal of Advertising*, V24, #3, Fall, 1995, pp. 13-23, “Enhancing the Efficacy of Split Thirty-Second Television Commercials,” Singh, Linville, Sukhdial.)
- A response surface design investigated three factors (commercial length, repetition, and delay before recall) on the average number of products recalled by twenty subjects viewing television commercials. (*Journal of Marketing Research*, 25, 1988?, pp. 72-88, “Recognition versus Recall as Measures of Television Commercial Forgetting.”)
- A major shoe retailer used MVT to simultaneously test sales techniques, advertising, separation by product color, discounts, and display configurations. They found a combination that pushed sneakers sales up 33%. (*Forbes*, March 11, 1996, pp. 114-118, “The New Mantra: MVT.”)
- An example for a candy company looks at 7 marketing factors in 8 experiments. The factors include ad medium, packaging design, size of candy bar and use of free samples. This article provides details on how to apply two-level factorial design to marketing. (*Journal of Marketing Research*, August, 1973, pp. 270-276, “Fractional Factorial Experimental Designs in Marketing Research,” Holland and Cravens.)
- Three groups of buyers, each with varying degrees of experience, evaluated saleability of a retail product. Specifically, the study tested 8 factors relating to a misses’ blouse.

These included fiber type, cut, color, country of origin, price and promotion. (*Journal of Retailing*, V62, Spring, 1986, pp. 41-63, "Retail Buyers' Saleability Judgements.")

- A supermarket used a two-level fractional factorial design to test retail price, newspaper advertising, display space and display location. The experiment produced quantifiable effects on sales and profits. (*Journal of Marketing Research*, August, 1974, pp. 286-294, "The Effects of Merchandising and Temporary Promotional Activities on the Sale of Fruits and Vegetables in Supermarkets," Curhan.)
- A study showed that the choice of high-cost brands is enhanced when the consumer is given little time to choose. This interaction could only be revealed via multifactor testing. (*Marketing Letters*, 6:4, 1995, pp. 287-295, "The Effect of Time Pressure on the Choice Between Brands That Differ in Quality, Price, and Product Features," Nowlis.)
- Carpet cleaners were evaluated with a fractional factorial design that included package designs (3), brands (3), pricing (3 levels) and Good Housekeeping seal (yes or no). Only 18 out of the 108 possible combinations were tested. The study produced a recommendation that was not actually tested. (*Harvard Business Review*, July-August 1975, pp. 107-117, "New Way to Measure Consumers' Judgments", Green and Wind.)
- A sales team wanted to improve its success rate using percent of successful closures to measure performance. They did a DOE on the following factors: attire (suit or casual), number of salespeople (one or two), presentation (high pressure or low), brochure (old or new). (From talk by Rip Stauffer on "Six Sigma in a Non-Manufacturing Environment" at 49th Annual Minnesota Quality Conference, 2002.)

Survey:

- The US Bureau of Census ran a factorial design that involved a pre-notice letter, census form and reminder card that resulted in an overall mail-back completion rate of 63.4%—substantially higher than expected (normally only 50% on average). Sending a pre-notice letter created the biggest gain in response. (*Proceedings of 1993 Bureau of Census 9th Annual Research Conference*, pp. 37-48, "Reminder Postcards Affect Mailback Response Rates for Census Questionnaires", Clark, et al.)

Billing:

- Southwestern Bell did a two-level factorial to design a better telephone bill. They varied more than a dozen factors, including: color, font, shading, alignment, and orientation. The new layout garnered a 78% preference rating versus 48% for the old bill. Savings of \$2 million in postage will result from the more efficient bill. (Michael Berry, Southwestern Bell, Austin, Texas.)
- Berry and his colleagues discovered that over 50% of billing problems came from only 18% of bills—those paid via local retail establishments offering collection services. They did a two-level factorial to investigate whether or not to introduce scanners, ten-key-pads, training videos, etc. The more expensive proposals did nothing to reduce fraction defects. Just a few simple procedural changes resulted in almost a ten-fold decrease in problems.
- A large company reduced their receivables from 200 days to only 44 days, generating a large cash flow in the process. They studied 4 factors: billing with the shipment or on a separate invoice, automation, follow-up by letter or telephone, contract or in-house billing service. They ran only 8 of the combinations—a half-fraction. Two of the factors were highly significant. (*Experimental Design*, Frigon, Mathews, J. Wiley, 1997, p. 266.)

Business (in general):

- Professor Davenport provides a general template for putting business ideas to the test along with many citations of actual applications:
 - How will newly developed burgers fare at CKE restaurants such as Hardee's and Carl's Jr.?
 - Do lobster tanks increase seafood sales at Food Lion supermarkets?
 - Does a Kmart with a Sears store inside sell more?
 - Do eBay e-commerce users bid higher when allowed to pay with credit cards?
 - How many loose checks should Wells-Fargo bank accept in an ATM?
 - Do Subway promotions of low-fat food increase restaurant sales?
 - Does competition actually increase sales of Famous Footware shoes at their retail stores located in shopping malls?
 - By staying open longer, does Toronto-Dominion bank get more deposits?
 - Which promotional offers stimulate checking accounts at PNC Bank?

*"Too many 'experiments' don't prove much of anything."
(Harvard Business Review)*

(*Harvard Business Review*, February 2009, "How to Design Smart Business Experiments," Thomas B. Davenport.)

Human performance:

- The author of this listing tested four factors affecting readability of computer software displayed on an RGB projector: font size, font type, background (white vs black) and lighting level. The DOE was performed in a classroom setting. It was done in 12 runs via an irregular fractional factorial. By going to larger Arial type with white background the lights could remain on without affecting readability. (*Stat-Teaser*, Summer 1997, "New Design Makes It All Clear," Anderson.)
- A researcher studied the performance of three-man military teams in response to two-level variations in ability and motivation. (*Small Group Behavior*, 19, pp. 363-78, "Effects of Team Composition on Ranked Team Effectiveness.")
- A full two-level factorial investigated five factors (display type, orientation, crosswind, guidance, and flight path prediction) on the squared deviation from optimal landings. Thirty-two trainees, selected at random, performed the simulated trials. (*Human Factors*, 32?, 1990?, pp. 64-69, "Factors in Pilot Training and Transfer.")

Information technology (IT) and services (IS)

- This case study details the application of design of experiments (DOE) to improve performance in a client/server system. The results showed how the organization could forego the cost of replacing and upgrading their existing server, thus avoiding an expense of \$48,000. (*The Telecommunications Review*, V579, 1999, pp 73-89, "Case Study: Applying Design of Experiments to Improve System Performance," Richard F. T. Eng.)

*"Experimental design was used to by one Regional Bell Operating Company to increase directory assistance revenues. The automated directory assistance sequence and phrasing was designed to increase the probability that the majority of callers would press 1 to automatically dial the number for 30 cents. This additional revenue amounted to several million dollars per annum."
(The Telecommunications Review)*

Medical:

- St. Luke's Hospital in Kansas City tested seven factors to better educate patients on how to safely use Warfarin, an anti-blood

clotting drug which can be fatal if used improperly. They achieved a 68% improvement in patient understanding by using a standardized instruction sheet and having a pharmacist discuss the drug. (*Forbes*, March 11, 1996, pp. 114-118, "The New Mantra: MVT.")

- Several case studies and a general primer on the use of DOE for the healthcare industry are provided in *Using Designed Experiments to Shrink Health Care Costs* by M. Daniel Sloan (ASQC Quality Press, 1997).

Productivity and Quality:

- A beverage bottler looked at the effects of two different types of bottles and two workers on delivery time. (*Response Surface Methodology*, 3rd Ed., Myers, Montgomery & Anderson-Cook, J. Wiley, problem 3.3.)
- In a related experiment, the bottler investigated bottle types, shelf configurations, and coolers. They used the data to minimize stocking time.
- A candy manufacturer measured consumer acceptance of various defects in the packaging of chocolate-covered cherries. They determined the acceptable thresholds for upside down or sideways candies as well as leakers. (From Stat-Ease, Inc. files.)

Service:

- A major telecommunications provider made use of screening DOE's to reduce:
 - network outage duration time
 - service order processing time
 - response times to customers
 - and increase sales for call centers—just to name a few applications.(From talk by Harry Rever on "The Application of Large Screening Design of Experiments in the Service Industry to Improve Key Metrics of the Business" at ASQ's 2004 Six Sigma Conference.)
- The FAA studied its New York air traffic control system to minimize communication delays. They used a fractional two-level factorial to investigate 8 factors in 16 experiments on a simulation model. Factors included number and location of navigation beacons and the mix of standard versus jumbo jets. (*Statistics for Experimenters*, Box, Hunter, Hunter, J. Wiley, p. 429.)

Transactional:

- Mannington Mills had been experiencing a high rate of rework due to incomplete forms that averaged over 50% defective. Using a two-level full-factorial design that included, among other factors, the level of training, their Six Sigma quality team found a way to improve this process very significantly. (From 2nd Annual Lean Six Sigma Summit, Nov. 16-18, 2005, Dallas. Details available upon request to mark@statease.com.)

*Want to learn how to make you own breakthroughs via DOE?
Attend our Experiment Design Made Easy workshop.
Call Stat-Ease, Inc. at 612.378.9449 or visit www.statease.com
for a schedule of upcoming classes or to consider an exclusive in-house presentation.*