

Experimenting Outside the Box: Using SAS/QC® for Modern Applications of Experimental Design

Sharad Prabhu, Randall Tobias, Tonya Balan
SAS Institute Inc., Cary, NC

Abstract

Statistical design and analysis of experiments has revolutionized science and industry, from agriculture to industrial quality improvement. But the principles of experimental design go far beyond these traditional applications. Researchers in areas as diverse as direct marketing, computer simulation, and Web site design are also using statistical experiments to study and optimize their endeavors. For example, you can optimize Web site traffic by conducting an experiment on different choices for headline texts, product images, or button text. Or you can study customer behavior by designing a direct marketing campaign with different offers and prices. In this presentation we demonstrate specialized features of SAS/QC software that enable you to perform these kinds of experiments. Our aim is not only to show how SAS® software can enable advances in the particular applications we cover, but also to inspire you to think outside of your own box by looking for new applications of experimental design.

Experiments Come Naturally!

Experimentation is the most natural of human responses to the task of acquiring information. In nearly every new endeavor, an initial discovery and pioneering phase is followed by a phase of experimentation—trying out various changes in the methodology or technique to see which ones make for the best result. Teachers experiment with teaching methods, store managers experiment with ways to schedule their employees, and cavemen probably experimented with the best sources of stone tools! Strangely enough, modern enterprises often need to relearn this skill that came so naturally to the Neanderthal man. As Peters and Waterman note in *In Search of Excellence*,

... most big institutions have forgotten how to test and learn. They seem to prefer analysis and debate to trying something out, and they are paralyzed by fear of failure, however small. (Peters and Waterman, 1982)

That observation is key: to experiment productively is to embrace the possibility of fruitful failure, failure you can learn from—in a sense, successful failure.

Statistical design of experiments is critical for succeeding through planned failure. Early in the last century, scientists such as R. A. Fisher and R. C. Bose made the fundamental discovery that the statistical methods used to analyze an experiment *after* it is performed can also tell you how best to design the experiment *before* it is performed (Fisher 1971). Thus was the field of statistical design of experiments born, and it is difficult to overstate its subsequent impact on all aspects of science and technology. For

example, it is fair to say that the so-called “Green Revolution” in agriculture of the last century was enabled by statistically driven experimentation, taking techniques developed by Fisher and his colleagues in the 1920s and applying them to problems of famine and food production throughout the post-WWII third world (Wikipedia 2009). Likewise, in the 1950s G. E. P. Box, G. Taguchi, and many others began to apply statistical design of experiments to problems in industry, producing a similar revolution in industrial output and quality (Box, Hunter, and Hunter 2005). The pattern is clear: when scientists and engineers want to optimize a product or process, they turn to statistical design of experiments.

The New Frontier

Although traditional applications of statistical design of experiments are still being vigorously studied and improved, the dawn of the twenty-first century brings many new human endeavors that need to be optimized. In turn these bring many new applications for statistical design of experiments. Many of these applications, as indeed with many other aspects of modern life, have to do with computers and the Internet. For example, the number of websites has increased a hundredfold and more in the past decade (Netcraft 2009). Every one of these sites can be optimized to deliver its information as efficiently as possible. As another example, computers are often used to simulate physical experiments that are difficult, costly, or perhaps even impossible to perform. The design for such computer experiments shares some features with design for the corresponding physical processes, but there are also new wrinkles. A third example of a new arena for experimentation, and thus for statistical design of experiments, is marketing, where the interest is in targeting the message to customers as accurately as possible.

The next three sections discuss specific examples of statistical design of experiments in these areas, demonstrating specialized features of SAS/QC software that enable you to get the most out of such novel studies. The goal of this paper is not only to show how SAS software works for these particular examples, but more importantly to inspire you to think outside of your own box in looking for new applications of experimentation and thus for statistical experimental design.

Example 1: Website Experiment

A multinational company has just completed a major redesign of its website and wants to perform experimental testing before a big launch event. They want to do a simultaneous launch in three countries with the website localized to those countries. The experiment is to be conducted over nine weeks, with three different designs tested per week. The restrictions are that only one country, one endorsement (celebrity), and one product can be used per week. Besides these three restrictions, all other testing variables can vary.

The company would like to test variables such as font size (small, medium, large), font type (Roman, Arial, Helvetica), sound (none, low, full), color (black, red, blue), background (flowers, stars, mountains), celebrity endorsement (Britney, Josh, Lindsay), and product (economy, special, deluxe).

A full-factorial design in eight factors of three levels would be 6,561 runs. Instead of running such a large experiment, you could consider an orthogonal array in 18, 27, or 36 runs that exist for the same specification. Usually you would choose the 27-run design, completely randomize the run order, and perform the experiments. However, the restrictions mentioned previously have repercussions for the critical question of how the experiment is to be randomized.

Randomization neutralizes the effects of systematic biases that might be involved in implementing the design and provides a basis for the assumptions underlying the analysis. But, in this example you can test only one combination of Country, Endorsement, and Product per week, and you are limited to performing only three designs per week over a nine-week period. This restriction on randomization induces a split-plot structure in which weeks are whole plots, the three restricted variables Country, Endorsement, and Product are the whole-plot factors, and the other five variables are subplot factors.

A new feature in SAS/QC 9.2 is the ability to construct full- and fractional-factorial split-plot designs with the FACTEX procedure. The following program constructs a resolution III design for this experiment, with the BLOCKS statement specifying the Week effect, and the new statement UNITEFFECT specifying that Country, Endorsement, and Product should be constant for the week and that Size, Color, Background, Font, and Sound should vary within the week.

```
proc factex;
  factors Country Endorsement Product Size Color
          Background Font Sound / nlev=3;
  model r=3;
  size design=27;
  blocks units=(Week=9);
  uniteffect Week /
    whole=(Country Endorsement Product)
    sub =(Size Color Background Font Sound)
    ;
  output out=WebDesign randomize
    Country      cvals=( "USA"      "Japan"    "India"    )
    Size         cvals=( "Small"    "Medium"  "Large"   )
    Color        cvals=( "Black"    "Red"     "Blue"    )
    Font         cvals=( "Roman"    "Ariel"   "Helvetica")
    Background   cvals=( "Flowers"  "Stars"   "Mountains")
    Product      cvals=( "Economy"  "Special" "Deluxe"  )
    Sound        cvals=( "None"     "Low"     "Full"    )
    Endorsement  cvals=( "Britney"  "Josh"    "Lindsay" )
    ;
run;
```

Table 1 shows the resulting split-plot design with each Week assigned a fixed set of Country-Endorsement-Product combinations. The experiments are run over nine weeks with three experiments per week.

Run	Week	Country	Endorsement	Product	Size	Color	Background	Font	Sound
1	1	USA	Britney	Economy	Small	Black	Flowers	Roman	None
2	1	USA	Britney	Economy	Large	Blue	Mountains	Helvetica	Full
3	1	USA	Britney	Economy	Medium	Red	Stars	Ariel	Low
4	2	India	Lindsay	Deluxe	Large	Blue	Mountains	Ariel	Low
5	2	India	Lindsay	Deluxe	Medium	Red	Stars	Romar	None
6	2	India	Lindsay	Deluxe	Small	Black	Flowers	Helvetica	Full
7	3	Japar	Josh	Special	Medium	Red	Stars	Helvetica	Full
8	3	Japar	Josh	Special	Small	Black	Flowers	Ariel	Low
9	3	Japar	Josh	Special	Large	Blue	Mountains	Romar	None
10	4	India	Josh	Economy	Large	Red	Flowers	Helvetica	Low
11	4	India	Josh	Economy	Medium	Black	Mountains	Ariel	None
12	4	India	Josh	Economy	Small	Blue	Stars	Romar	Full
13	5	Japar	Britney	Deluxe	Medium	Black	Mountains	Romar	Full
14	5	Japar	Britney	Deluxe	Small	Blue	Stars	Helvetica	Low
15	5	Japar	Britney	Deluxe	Large	Red	Flowers	Ariel	None
16	6	USA	Lindsay	Special	Small	Blue	Stars	Ariel	None
17	6	USA	Lindsay	Special	Large	Red	Flowers	Romar	Full
18	6	USA	Lindsay	Special	Medium	Black	Mountains	Helvetica	Low
19	7	Japar	Lindsay	Economy	Medium	Blue	Flowers	Ariel	Full
20	7	Japar	Lindsay	Economy	Small	Red	Mountains	Romar	Low
21	7	Japar	Lindsay	Economy	Large	Black	Stars	Helvetica	None
22	8	USA	Josh	Deluxe	Small	Red	Mountains	Helvetica	None
23	8	USA	Josh	Deluxe	Large	Black	Stars	Ariel	Full
24	8	USA	Josh	Deluxe	Medium	Blue	Flowers	Romar	Low
25	9	India	Britney	Special	Large	Black	Stars	Romar	Low
26	9	India	Britney	Special	Medium	Blue	Flowers	Helvetica	None
27	9	India	Britney	Special	Small	Red	Mountains	Ariel	Full

Table 1. Split-Plot Design for Website Experiment

For an example of designing a multistep fractional-factorial split-plot that uses PROC FACTEX and how JMP® users can use the JMP scripting language (JSL) to take advantage of these advanced SAS design capabilities, see Ramirez and Weisz (2009).

Example 2: Marketing Testing

Traditionally, marketers who gauge customer preferences use a grab bag of ad hoc techniques based on expert opinion, past experience, and “We’ve always done it that way.” The most common tool is “one-factor-at-a-time” testing, also called A/B testing or split-run testing. These methods work well enough to find obvious trends in small studies, but they are not sophisticated enough for large studies with subtle effects and interactions. Industry analysts are now looking to apply techniques for statistical design of experiments to make their market research more sensitive, more accurate, and more economical.

Bell, Ledolter, and Swersey (2006) present a case study from an advertising testing application which uses statistically designed experimentation. A bank wants to test the effectiveness of their direct mail program for credit card offers. Nineteen factors are identified that are thought to influence a customer’s decision whether to respond to an advertised product. Some of the factors control how the mailing

envelope and the accompanying letter are designed; other factors include the value of a free gift, additional inserts, interest rate, and so on. The complete list of factors is shown in Table 2.

Factor Name	Low Level	High Level	Factor Label
TEASER	General	Specific	Product-specific Offer
RADDRESS	Blind	Visible	Return Address
STAMP	Yes	No	Ink-Stamp on Envelope
POSTAGE	Prepaid	Stamp	Postage
GRAPHIC	Yes	No	Additional Graphic
PGRAPHIC	Small	Large	Price Graphic
STICKER	Yes	No	Sticker
CUSTOM	No	Yes	Personalize Letter
MESSAGE	Targeted	Generic	Copy Message
HEADLINE	H1	H2	Letter Headline
LAYOUT	Standard	Creative	Benefits Layout
PS	Control	New	Letter Postscript
SIGN	Manager	Executive	Signature
SELECT	Many	Few	Product Selection
GIFTVAL	High	Low	Free Gift Value
REPLY	Control	New	Reply Envelope
BUCKSLIP	Product	Gift	Buckslip Info
BUCKSLP2	No	Yes	Second Buckslip
INTEREST	Low	High	Interest Rate

Table 2. Factors for Marketing Testing Case Study

The experimenters are interested mainly in first-stage screening, identifying potentially active factors to be investigated further in depth. They choose a 20-run Plackett-Burman design, in which main effects are aliased with two-factor interactions. This is a saturated design with no degrees of freedom available to test the effect estimates. The design listing is shown Table 3 (some column names have been edited to fit in the table):

#	TEASER	RADD	STM	POST	GRP	PGRP	STR	CUS	MESSAGE	H	LAYOUT	PS	SIGN	SELECT	GIFT	REPLY	BCKSLIP	BLP2	INT	ORD	RES
1	Specific	Visible	Yes	Prepaid	No	Large	No	Yes	Targeted	H2	Standard	New	Manager	Many	High	Control	Gift	Yes	Low	52	1.04
2	General	Visible	No	Prepaid	Yes	Large	No	Yes	Generic	H1	Creative	Control	Executive	Many	High	Control	Product	Yes	High	38	0.76
3	Specific	Blind	No	Stamp	Yes	Small	No	Yes	Generic	H2	Standard	New	Manager	Few	High	Control	Product	No	High	42	0.84
4	Specific	Visible	Yes	Stamp	No	Small	Yes	Yes	Generic	H2	Creative	Control	Executive	Many	Low	Control	Product	No	Low	134	2.68
5	General	Visible	No	Prepaid	No	Large	Yes	No	Generic	H2	Creative	New	Manager	Few	High	New	Product	No	Low	104	2.08
6	General	Blind	No	Stamp	Yes	Large	No	No	Targeted	H2	Creative	New	Executive	Many	Low	Control	Gift	No	Low	60	1.2
7	General	Blind	Yes	Stamp	No	Small	No	Yes	Targeted	H1	Creative	New	Executive	Few	High	New	Product	Yes	Low	61	1.22
8	General	Blind	Yes	Prepaid	No	Large	Yes	Yes	Generic	H1	Standard	New	Executive	Few	Low	Control	Gift	No	High	68	1.36
9	Specific	Blind	Yes	Prepaid	Yes	Large	No	No	Generic	H2	Standard	Control	Executive	Few	Low	New	Product	Yes	Low	57	1.14
10	General	Visible	Yes	Prepaid	Yes	Small	No	Yes	Targeted	H2	Creative	Control	Manager	Few	Low	New	Gift	No	High	30	0.6
11	Specific	Blind	No	Prepaid	Yes	Small	Yes	Yes	Generic	H1	Creative	New	Manager	Many	Low	New	Gift	Yes	Low	108	2.16
12	General	Visible	Yes	Stamp	Yes	Small	Yes	No	Generic	H2	Standard	New	Executive	Many	High	New	Gift	Yes	High	39	0.78
13	Specific	Blind	No	Prepaid	No	Small	Yes	No	Targeted	H2	Creative	Control	Executive	Few	High	Control	Gift	Yes	High	40	0.8
14	Specific	Visible	Yes	Stamp	Yes	Large	Yes	No	Targeted	H1	Creative	New	Manager	Few	Low	Control	Product	Yes	High	49	0.98
15	Specific	Visible	No	Prepaid	No	Small	No	No	Targeted	H1	Standard	New	Executive	Many	Low	New	Product	No	High	37	0.74
16	Specific	Visible	No	Stamp	Yes	Large	Yes	Yes	Targeted	H1	Standard	Control	Executive	Few	High	New	Gift	No	Low	99	1.98
17	General	Visible	No	Stamp	No	Small	No	No	Generic	H1	Standard	Control	Manager	Few	Low	Control	Gift	Yes	Low	86	1.72
18	General	Blind	No	Stamp	No	Large	Yes	Yes	Targeted	H2	Standard	Control	Manager	Many	Low	New	Product	Yes	High	43	0.86
19	Specific	Blind	Yes	Stamp	No	Large	No	No	Generic	H1	Creative	Control	Manager	Many	High	New	Gift	No	High	47	0.94
20	General	Blind	Yes	Prepaid	Yes	Small	Yes	No	Targeted	H1	Standard	Control	Manager	Many	High	Control	Product	No	Low	104	2.08

Table 3. Design Listing for Marketing Testing Case Study

The experimental mailings are sent to 100,000 customers; each of the 20 different mailings in the design is sent to 5,000 customers. This sample size should be large enough to enable experimenters to detect a shift as small as 0.2% in the response rate. Note that the same experiment performed using one-factor-at-a-time method would require more than *four times* as many mailings to estimate the effects with the same statistical power as the 20-run designed experiment.

The performance of the campaign is measured by the number of orders received and the response rate. Figure 1 shows a main-effect plot from the SAS® 9.2 ADX Interface for Design of Experiments for all the factors based on the observed response rate.

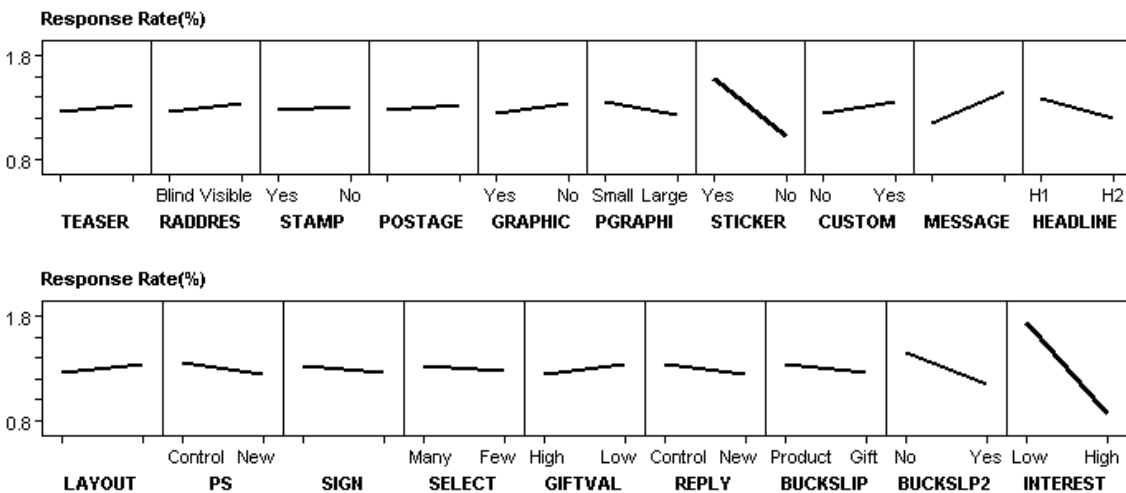


Figure 1. Main-Effect Plot

Zooming in, Figure 2 shows the five effects that stand out as active.

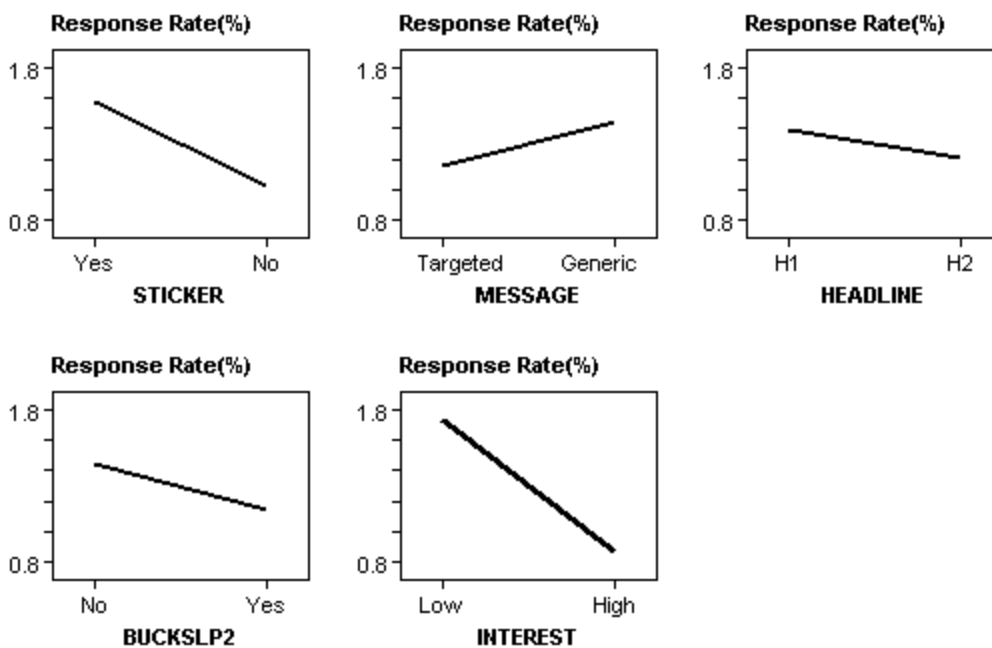


Figure 2. Active Effects

Several of these effects are in accord with expert opinion:

- Increasing the credit card interest rate reduces the response rate percentage 0.86.

- Placing an attention-getting sticker on the letter increases the response rate percentage by 0.55, easily outweighing the additional cost of the sticker.
- The improved headline wording H1 is preferred to the old style H2.

However, a couple effects are counterintuitive:

- The respondents are not swayed by the targeted message, although the creative team was certain that appealing to a customer’s interest would improve the response.
- Adding a second buckslip (a small sheet of paper to highlight product information) negatively impacts the response. Perhaps the second buckslip contains unnecessary information and obscures the main message.

It is acceptable, indeed desirable, to have an experiment reveal that some hypotheses are *not* borne out by experience. It is the creative team’s job to come up with potential improvements; if you require that all of their guesses be correct, you will get a lot fewer guesses. Experimentation encourages experts to take a risk in their notions and guesses, knowing that fruitless ideas will be caught well before they become too costly.

As another way to analyze this data, a formal main-effects statistical model can also be fitted to observed response rates. The resulting normal plot of the effect estimates is shown in Figure 3. The effects that fall away from the pseudo-standard error (PSE) line are generally significant and confirm earlier conclusions.

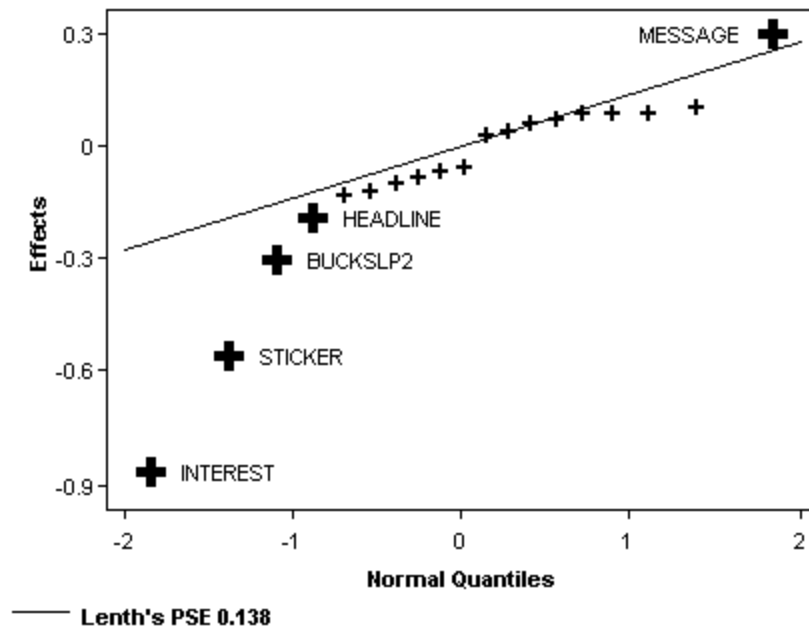


Figure 3. Normal Plot of Effect Estimates

Example 3: Computer Simulation Experiment

As noted in the first section, statistical design of experiments began with applications in agriculture and later in manufacturing—growing and making actual physical stuff. However, physical experiments such as this can be time-consuming and costly. With the advent of fast and affordable computing power, a versatile new tool in science and technology is the computer experiment. This involves developing a mathematical model of the elements of a physical system and then modeling its general characteristics with a much simpler summary model. To account for variation of the true mathematical model around the simpler one, the summary model includes one or more random components. Thus, physical experiments can be performed by computer simulation. Although such computer experiments cannot replace physical experiments, they do provide a good approximation for understanding the general features of your process.

As an example, consider an experimental study conducted by the National Institute of Standards and Technology (NIST) Building and Fire Research Laboratory concerning the reflective coatings for buildings (Zarr 1998). A reflective coating applied to the exterior of a building is designed to reflect solar radiation, thereby reducing the cooling requirements for a conditioned space. Below are some the questions the experimenters were interested in answering:

- How do building parameters such as roof reflectance, building insulation, attic ventilation, and attic size affect the heating and cooling loads of the building in different climates? Which parameter is the most important?
- Does the presence of a reflective roof cause an undesired increase in heating load during the winter heating season?

The input parameters to be controlled were: geographic location, roof reflectance index, ceiling insulation resistance, attic ventilation rate, and attic framing area. Table 4 shows the factors and levels of interest.

Factors	Levels
LOC (geographic location)	Miami, Phoenix, Birmingham, DC, Portland, Bismarck
REFLECT (roof solar reflectance)	0.10, 0.30, 0.45, 0.60, 0.80
RESIST (ceiling thermal resistance)	0, 11, 19, 30, 38, 49
ATICVENT (attic ventilation rate)	0.5, 2.3, 4.6, 6.9, 9.2
ATICAREA (attic framing area)	333, 417, 501

Table 4. Factors and Levels in Reflective Coating Study

If an experiment on these factors were to be run physically, the engineers would have to build or find houses at the correct location with certain exact specifications, and then modify the physical attributes to match the levels of interest. Obviously, this would be time-consuming and expensive, and in fact probably infeasible. Instead NIST engineers decided to use historical data on weather and energy use to simulate the experiments. They performed a full-factorial experiment in $6 \times 5 \times 6 \times 5 \times 3 = 2700$ runs, which required around 31 hours to simulate on a PC. The following sections explore different design options the engineers might have employed to significantly reduce this computation time.

Design Choice 1

As part of a strategy of sequential experimentation, instead of running a full-factorial experiment, you can run a screening experiment as a first step. This helps to identify the important factors and their corresponding levels of interest, and it weeds out the nonsignificant factors. Usually only two levels are chosen for each factor in a screening design. All the factors in the NIST scenario have more than two levels, but you can adapt this to a screening design in the following way. Besides Location, all factors are quantitative, so you can use the minimum and the maximum of the range as the two levels, low and high, of each factor. For Location you can pick two of the six cities—say, the ones that have the minimum and maximum average annual temperature.

With five two-level factors, a good option is the fractional-factorial 2^{5-1} design, which is a half-fraction in 16 runs. This design has resolution V—that is, all 5 main effects and all 10 two-factor interactions are simultaneously estimable with full efficiency. Assuming there are no quadratic effects, this design works well to identify the main effects and interactions. The 2^{5-1} design for this case is shown in Table 5.

RUN	LOC	REFLECT	RESIST	ATICVENT	ATICAREA
1	Portland	0.1	0	0.5	500
2	Phoenix	0.1	0	0.5	300
3	Portland	0.8	0	0.5	300
4	Phoenix	0.8	0	0.5	500
5	Portland	0.1	49	0.5	300
6	Phoenix	0.1	49	0.5	500
7	Portland	0.8	49	0.5	500
8	Phoenix	0.8	49	0.5	300
9	Portland	0.1	0	9.2	300
10	Phoenix	0.1	0	9.2	500
11	Portland	0.8	0	9.2	500
12	Phoenix	0.8	0	9.2	300
13	Portland	0.1	49	9.2	500
14	Phoenix	0.1	49	9.2	300
15	Portland	0.8	49	9.2	300
16	Phoenix	0.8	49	9.2	500

Table 5. Fractional-Factorial Design

Design Choice 2

On the other hand, if you suspect there might be more complicated effects, then reducing the number of levels studied for each factor might not be appropriate. Instead, suppose you want a design that preserves the original factor levels, over the original experimental design space. You could explore the option of using an orthogonal array. Ideally you would like to use a design that is orthogonal and balanced, with all levels of all factors occurring equally often and all parameter estimates uncorrelated.

The SAS/STAT® market research experimental design macros (Kuhfeld 2009) are powerful tools for constructing orthogonal and near-orthogonal experimental designs. The %MktEx macro can construct over 117,000 orthogonal arrays from its catalog. If an orthogonal array does not exist for your

application, the %MktEx macro can find an efficient experimental design by using a coordinate exchange or candidate set search algorithm.

The following statement specifies the five factors and their levels, corresponding to those in the building simulation study, to construct a 180-run orthogonal array which can estimate all the main effects and the two-factor interactions:

```
%mktex(6 5 6 5 3, interact=@2, n=180)
```

The %Mktex macro uses the best method available to find the best combination of a subset of 180 runs among a full-factorial design $6^2 5^2 3^1$ of 2,700 runs. For this particular $6^2 5^2 3^1$ factor structure, no standard design exists, so construction by design optimization is warranted. The %Mktex macro uses the SAS/QC OPTEX procedure the underneath and also uses the coordinate-exchange method to find the appropriate design. Table 6 shows a partial listing of the design (the factor names and levels were uncoded using an additional step that is not shown):

OBS	Loc	Reflect	Resist	Aticvent	Aticarea
1	Bi rmi ngham	0. 10	0	4. 6	333
2	Bi rmi ngham	0. 10	0	9. 2	417
3	Bi rmi ngham	0. 10	19	0. 5	333
4	Bi rmi ngham	0. 10	30	2. 3	501
5	Bi rmi ngham	0. 10	38	6. 9	333
6	Bi rmi ngham	0. 10	49	2. 3	333
7	Bi rmi ngham	0. 30	0	0. 5	501
8	Bi rmi ngham	0. 30	11	2. 3	333
9	Bi rmi ngham	0. 30	19	9. 2	501
10	Bi rmi ngham	0. 30	30	9. 2	417
.
.
.
179	Portl and	0. 80	38	0. 5	501
180	Portl and	0. 80	49	9. 2	333

Table 6. Orthogonal Array

Table 7 shows the output from the %Mkteval evaluation of this design. The table displays the canonical correlations between pairs of coded factors. If a design is orthogonal for main effects, the matrix shows all zeroes off the diagonal. This design is not orthogonal because the matrix shows minor presence of correlation between the main effects. Note that this amount of correlation is unlikely to severely impact the quality of the information obtained from this design. If a 180-run design that enables you to estimate all main effects and interactions between the original factors as classification effects is what you want, this is a perfectly adequate alternative.

	LOC	REFLECT	RESI ST	ATI CVENT	ATI CAREA
LOC	1	0.10	0.08	0.06	0.10
REFLECT	0.10	1	0.11	0.13	0.08
RESI ST	0.08	0.11	1	0.07	0.06
ATI CVENT	0.06	0.13	0.07	1	0.11
ATI CAREA	0.10	0.08	0.06	0.11	1

Table 7. Design Evaluation

Design Choice 3

In Design Choice 2, the %Mktex macro took several minutes to search an appropriate design. Also, it should be noted that %Mktex treats all the variables as classification variables. However, if you want to consider only Location as a classification variable because it has qualitative levels, you can search directly for a nonstandard design with the OPTEX procedure. With this method, you specify a candidate set of design points and a linear model, and then use an optimal design algorithm to choose a particular number of points so that the terms in the model can be estimated as efficiently as possible.

The following SAS statements use the SAS/STAT PLAN procedure to create the full set of 2,700 runs as a candidate set and then use PROC OPTEX to select just 60 of them optimally—a 1/45 fraction. In Design Choice 2 the main effects and interactions are specified, but in this example the quadratic effects are also included to detect the presence of curvature.

```

/*
/ Create complete 6x5x6x5x3 candidate set.
/-----*/
proc plan;
  factors loc=6 reflect=5 resist=6 aticvent=5 aticarea=3 / noprint;
  output out=Full;
run;

/*
/ Select 60 run design D-optimally.
/-----*/
proc optex data=Full coding=orthcan;
  class loc;
  model loc|reflect|resist|aticvent|aticarea@2
        reflect*reflect resist*resist
        aticvent*aticvent aticarea*aticarea;
  generate n=60 method=m_fedorov niter=100 keep=10;
  output out=Design3;
quit;

```

The output obtained from PROC OPTEX and shown in Table 8 illustrates the efficiencies of the design saved in the data set `Design3`. These efficiency numbers represent the goodness of a design relative to the 2,700-run full-factorial design which is both orthogonal and balanced. The best D-efficiency design was found in the first try, and the best A-efficiency design was found in the eighth try. The efficiencies shown in this table are relative to the best hypothetical design. They are not on a scale of 0 to 100 because only one of the five factors is specified as a classification variable.

Design Number	D-Efficiency	A-Efficiency	G-Efficiency	Average Prediction Standard Error
1	147.0331	103.7270	85.9147	0.8017
2	146.8678	102.8669	81.3364	0.8050
3	146.8678	102.8669	81.3364	0.8050
4	146.8678	102.8669	81.3364	0.8050
5	146.8678	102.8669	81.3364	0.8050
6	146.8678	102.8669	81.3364	0.8050
7	146.8180	103.0231	84.1224	0.8044
8	146.7163	107.8359	88.2064	0.7863
9	146.6867	104.3503	84.0153	0.7993
10	146.6535	101.8050	87.6281	0.8092

Table 8. Design Efficiencies

Table 9 shows the frequency distributions of the factor levels for the final design that the OPTEx procedure found. Note that only the levels of the classification factor LOC are approximately uniformly sampled. For the other factors, modeled with quadratic effects, only extreme and central levels are included, the extreme levels with more weight. This is in accord with the ideal optimum design for a quadratic model over a rectangular candidate space, which only requires three levels per factor.

Loc	Freq	%	Reflect	Freq	%	Resist	Freq	%	Aticvent	Freq	%	Aticarea	Freq	%
1	8	13.33	1	27	45.00	1	26	43.33	1	26	43.33	1	27	45.00
2	8	13.33	2	0	0.00	2	0	0.00	2	0	0.00	2	6	10.00
3	11	18.33	3	6	10.00	3	4	6.67	3	7	11.67	3	27	45.00
4	9	15.00	4	0	0.00	4	3	5.00	4	0	0.00			
5	12	20.00	5	27	45.00	5	0	0.00	5	27	45.00			
6	12	20.00				6	27	45.00						

Table 9. Frequency Distribution of Factor Levels

Design Choice 4

This design choice offers another way of treating the Location factor. If you examine the factor variables, it seems that the Location variable was considered in order to account for the range of outdoor temperature, which is likely to contribute in obvious and relatively uninteresting ways to the variability observed in the response data. For making inferences over the whole range of environmental conditions, this makes Location a *block effect*, rather than being an effect of direct interest. To adjust for the variability between locations, Location should be treated as a blocking variable. By doing so, you are assuming that Location probably has an effect on the response but that effect is only a nuisance in estimating other effects of direct interest. Furthermore, if you are interested in studying the response surface, the design should be able to estimate the quadratic effects in addition to the main effects and two-factor interactions.

The following SAS statements again use PROC OPTEx on the full candidate set of 450 runs. (Location is excluded because it is the blocking variable.) This time the statements ask that the 180-run design be optimally blocked into 30 blocks of 6 runs each.

```

/*
/ Create complete 5x6x5x3 candidate set.
/-----*/
proc plan;
  factors reflect=5 resist=6 aticvent=5 aticarea=3 / noprint;
  output out=Full;
run;
/*
/ Create a response surface blocked design in 6 blocks with 30 runs
/-----*/
proc optex data=Full coding=orthcan;
  model reflect|resist|aticvent|aticarea@2
        reflect*reflect resist*resist
        aticvent*aticvent aticarea*aticarea;
  blocks structure=(6)30;
  output out=Design4 blockname=loc;
quit;

```

Table 10 shows the partial listing of the response surface design (the variable levels were uncoded using an additional data step not shown here):

	Loc	Reflect	Resist	Aticvent	Aticarea
1	Miami	0.10	19	9.2	501
2	Miami	0.45	30	4.6	417
3	Miami	0.45	0	9.2	417
.
.
30	Miami	0.10	49	0.5	333
31	Phoenix	0.45	49	9.2	501
32	Phoenix	0.45	49	9.2	417
33	Phoenix	0.80	0	9.2	333
.
.
60	Phoenix	0.10	19	4.6	417
61	Birmingham	0.80	49	9.2	333
62	Birmingham	0.10	49	9.2	501
63	Birmingham	0.10	0	9.2	333
.
.
90	Birmingham	0.80	30	4.6	333
91	DC	0.80	49	0.5	333
92	DC	0.10	49	9.2	501
93	DC	0.80	0	9.2	333
.
.
120	DC	0.80	49	9.2	333
121	Portland	0.80	19	9.2	333
122	Portland	0.80	49	0.5	417
123	Portland	0.10	0	9.2	501
.
.
150	Portland	0.45	30	9.2	501
151	Bismarck	0.10	0	0.5	417
152	Bismarck	0.80	0	0.5	333
153	Bismarck	0.80	0	9.2	501
.
.
180	Bismarck	0.10	19	9.2	333

Table 10. Partial Listing of Blocked Optimal Response Surface Design

Summary

In *Experimentation Matters*, Stefan Thomke explores the many new kinds of experiments that are possible nowadays, and he gives principles that enterprises should heed in order to benefit from these new technologies. Citing new possibilities for computer modeling and simulation (as in Example 3) and for rapid prototyping by combinatorial methods (similar to the other two examples), Thomke highlights a key guideline: experiment early and often, without fear of failure. “Effective experiments are supposed to reveal what does not work *early*,” writes Thomke; the ideal is to “fail often to succeed sooner.” With this thought in mind, and with the examples presented in this paper for inspiration, we encourage you to look for novel experimental opportunities in your own enterprise. As you do so, SAS software is your critical tool for finding fruitful failures, the kind of short-term failures that lead to long-term success.

References

Box, G. E. P., Hunter, W. G., and Hunter, J. S. 2005. *Statistics for Experimenters*, Second Edition. New York: John Wiley & Sons.

Fisher, R. A. 1971. *Design of Experiments*, Ninth Edition. New York: MacmillanBell, G. H., Ledolter, J., and Swersey A. J. 2006. “Experimental Design on the Front Lines of Marketing: Testing New Ideas to Increase Direct Mail Sales.” *International Journal of Marketing Research*, 23, 309–319.

Green Revolution. 2009. http://en.wikipedia.org/wiki/Green_Revolution

Kuhfeld, W. F. 2009. “Marketing Research Methods in SAS,” MR-2009, Technical Report, SAS Institute Inc., http://support.sas.com/resources/papers/tnote/tnote_marketresearch.html.

Kuhfeld, W. F., Tobias, R. D., and Garratt, M. 1994. “Efficient Experimental Design with Marketing Research Applications.” *Journal of Marketing Research*, 31(4), 545–557. Netcraft 2009. http://news.netcraft.com/archives/2009/02/18/february_2009_web_server_survey.html

Peters, T. and Waterman, R. 1982. *In Search of Excellence*. New York: Harper and Row.

Plackett, R. L. and Burman, J. P. 1946. “The Design of Optimum Multifactorial Experiments.” *Biometrika*, 33(4), 305–325.

Ramirez, J. G. and Weisz, J. T. 2009. “Designing Multi-Step Fractional Factorial Split Plots: A Combined JMP® and SAS® User Application.” *Proceedings of the SAS Global Forum 2009 Conference*. Cary, NC: SAS Institute, Inc.

SAS Institute Inc. 2008. *SAS/QC® 9.2 User’s Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2008. *SAS/STAT® 9.2 User’s Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2008. *Getting Started with the SAS® 9.2 ADX Interface for Design of Experiments*. Cary, NC: SAS Institute Inc.

Thomke, S. 2006. *Experimentation Matters*. Boston, MA: Harvard Business School Press.

Zarr, R. R. 1998. *Analytical Study of Residential Buildings with Reflective Roofs*. NISTIR 6228, Building and Fire Research Laboratory, National Institute of Standards and Technology, Gaithersburg, MD

Contact Information

Your comments and questions are valued and encouraged. Contact the author at:

Sharad Prabhu
SAS Institute Inc.
SAS Campus Drive
Cary, NC 27513
USA
Office: (919) 531-6663
Fax: (919) 677-4444
Email: Sharad.Prabhu@sas.com

This paper is presented at SAS Global Forum 2009 SAS Presents. The correct bibliographic citation for this paper is as follows:

Prabhu, S., Tobias, R., and Balan, T. 2009. "Experimenting Outside the Box: Using SAS/QC® for Modern Applications of Experimental Design." *Proceedings of the SAS Global Forum 2009 Conference*. Cary, NC: SAS Institute, Inc.

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.