HOW TO SAVE RUNS, YET REVEAL BREAKTHROUGH INTERACTIONS, BY DOING ONLY A SEMIFOLDOVER ON MEDIUM-RESOLUTION SCREENING DESIGNS

Mark J. Anderson
Principal
Stat-Ease, Inc.
2021 East Hennepin Ave, Suite 480
Minneapolis, MN 55413
Mark@StatEase.com

Patrick J. Whitcomb
Principal
Stat-Ease, Inc.
2021 East Hennepin Ave, Suite 480
Minneapolis, MN 55413
Pat@StatEase.com

KEY WORDS

Design of experiments, foldover, semifoldover.

SUMMARY

Via case studies, this paper reviews the strategy of foldover on low-resolution (III) two-level fractional factorials and demonstrates how to reduce experimental runs by making use of semifoldover methods to augment medium-resolution (IV) designs.

INTRODUCTION

Two-level factorial design of experiments (DOE) work very well as screening tools. If performed properly, they can reveal the vital few factors that significantly affect your process. To save on costly runs, experimenters often perform only a fraction of all the possible combinations. There are many varieties of fractional two-level designs, such as Taguchi or Plackett-Burman, but we will restrict our discussion to the standard ones that statisticians symbolize as “2^{k-p}”, where k refers to the number of factors and p is the fraction. Regardless of how you do it, cutting out runs reduces the ability of the design to resolve all possible effects, specifically the higher-order interactions. Minimal-run designs, such as seven factors in eight runs (2^{7-4}) – a 1/16th (2^{-4}) fraction, can only estimate main effects. Statisticians label these low-quality designs as “resolution III” to indicate that main effects will be aliased with two-factor interactions. Resolution III designs often produce significant improvements, but it’s like kicking your PC (or slapping the monitor) to make it work: You won’t discover what really caused the failure.

To improve design resolution, many experimenters do a complete “foldover” when results look significant. This requires adding a second block of runs with the all the factors reversed (low levels go to high and vice-versa). For example, you can take the seven-factor in eight-run design and do a second block of eight runs with all levels opposite of the first block. It is well known (Montgomery 2001) that a complete foldover of a resolution III fractional two-level factorial makes it resolution IV - main effects now aliased only with highly unlikely three-factor or higher-order interactions. (To help you grasp the concept of resolution, think of main effects as 1 factor and add this to the number of factors it will be aliased with. In resolution III it’s a 1-to-2 relation, which adds to 3. Resolution IV indicates a 1-to-3 aliasing (1+3=4). A resolution V design aliases main effects only with four-factors (1+4=5).)

Because of their ability to more clearly reveal main effects, resolution IV designs work much better than resolution III for screening purposes. However, two-factor interactions remain aliased with each other (2+2=4). You might be tempted to try a complete foldover of a resolution IV fractional factorial, but be forewarned that this may create a replicate of the existing runs, rather than a unique fraction. (Montgomery, page 347). It’s better to do a foldover only on a single factor (adding a second fraction in which the signs for only one factor are reversed). This de-aliases the two-factor interactions (2 fi’s) involving the foldover factor from other 2fi’s, but at a cost of double the original runs. It’s much more efficient to do a partial foldover, referred to as a “semifoldover” (Mee and Peralta 2000), that involves only half the original block of runs. Rather than getting into all of the statistical details on foldover methods, we will keep things as simple as possible by illustrating the methods via case studies.

EXAMPLE OF FULL FOLDOVER FOR LOW-RESOLUTION (III) DESIGNS

The following case study (Anderson and Whitcomb 2000) provides a good primer on the use of traditional foldover methods. One of the authors became envious of the skating ability of his co-author. The 'wannabe' skater secretly got together with a local manufacturer of in-line skates and borrowed their highest-tech prototype. Bewildered by all the options, he decided to try various combinations:

A. Pad goes inside skate to elevate the heel: Out –, In +
B. Bearing constructed either from old material (–) or new alloy (+)
C. Gloves made specially for in-line skating to protect wrists: On –, Off +
D. Helmet fits with logo to back (–) or front (+) - can't tell which is correct, so try both
E. Wheels made of hard (–) or soft (+) polymer
F. Covers go on wheels to make them look faster: On –, Off +
G. Neon lighting (from generator on skates) for night-time use: Off –, On +

The special gear would have to be returned fairly quickly, so a quick-and-dirty fractional factorial was set up. If anything proved to be statistically significant, the result would be a faster time around the track (with ego-gratification to the skater). Table 1 shows the initial design, a 2^{7-4} resolution III fractional factorial, and the resulting times around the track.

Table 1: First experiment on in-line skates

Std	*Block*	*A: Pad*	*B: Bearing*	*C: Gloves*	*D: Helmet*	*E: Wheels*	*F: Covers*	*G: Neon*	*Time (sec.)*
1	1	Out	Old	On	Front	Soft	Off	Off	195
2	1	In	Old	On	Back	Hard	Off	On	192
3	1	Out	New	On	Back	Soft	On	On	200
4	1	In	New	On	Front	Hard	On	Off	165
5	1	Out	Old	Off	Front	Hard	On	On	190
6	1	In	Old	Off	Back	Soft	On	Off	195
7	1	Out	New	Off	Back	Hard	Off	Off	166
8	1	In	New	Off	Front	Soft	Off	On	201

The plot of effects shown in Figure 1 reveals several outstanding effects of skate configuration on the resulting times around the track.

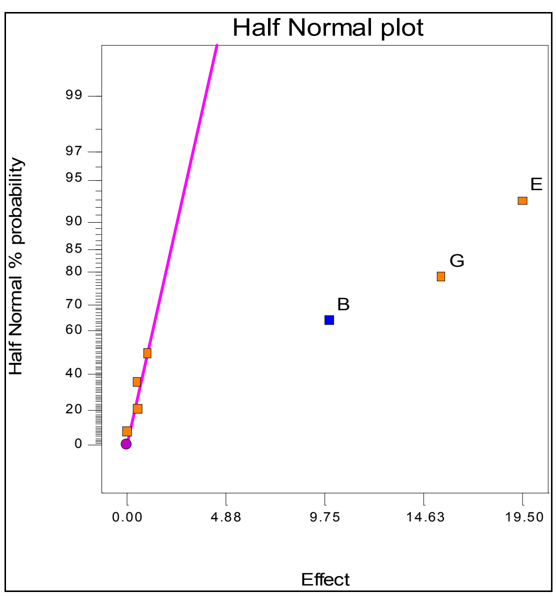

Figure 1: Plot of effects for in-line skating experiment

Analysis of variance (ANOVA) verifies that something significant happened. The cube plot in Figure 2 tells the apparent story.

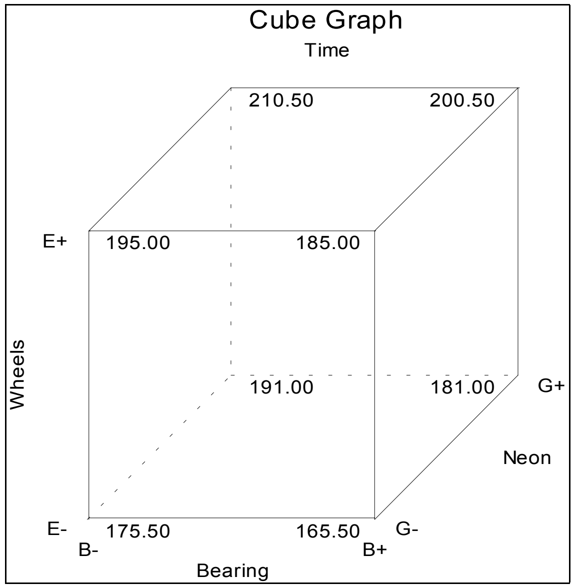

Figure 2: Cube plot of significant effects on skating (time in seconds)

The contrast from worst time (210.5 seconds) to best time (165.5 seconds) represents a very noticeable improvement of 45 seconds. Notice on the cube plot that changing the bearing from the old (B-) to the new material (B+) caused a decrease in time. This makes sense, as does the effect of wheels. You can see that going to the soft wheels (E+) increases the time. These two changes were put into effect immediately. However, factor G, neon lighting, begs belief. This was put in as a factor because the generator needed to power the lights creates some drag. But notice that the times actually decrease when the lights go on (G+). Perhaps there's some sort of psychological impact on the skater, but this is a stretch. More likely, what's listed as G is really an aliased effect.

The alias structure for the 2^{7-4} resolution III fractional factorial design is shown on Table 2. Notice that G = G +AF +BE +CD. This tells you that what's labeled arbitrarily as G on the effects graphs could be due to the interactions AF and/or BE and/or CD. All three of these interactions correlate completely with factor G. A likely possibility is that G is an alias for the interaction BE, because both parents (B and E) already appear significant.

Table 2: Alias structure for 2^{7-4} design (seven factors in eight runs)

Labeled as	Actually
A	A + BD + CE + FG
B	B + AD + CF + EG
C	C + AE + BF + DG
D	D + AB + CG + EF
E	E + AC + BG + DF
F	F + AG + BC + DE
G	G + AF + BE + CD

Given the confusing results, the skate manufacturer grants an extension on use of their prototype gear. Rolling right along, the experimenter decides to do a complete foldover on the initial design. Table 3 shows the factor levels and results for this followup design. Notice that it's designated as block number 2. The statistical analysis removes any block differences before calculating effects, thus negating any changes due to equipment wear, skater fatigue or environmental factors such as temperature or humidity at the track.

Table 3: Followup design (foldover) on in-line skates

Std	***Block***	***A: Pad***	***B: Bearing***	***C: Gloves***	***D: Helmet***	***E: Wheels***	***F: Covers***	***G: Neon***	***Time (sec.)***
9	2	In	New	Off	Back	Hard	On	On	175
10	2	Out	New	Off	Front	Soft	On	Off	211
11	2	In	Old	Off	Front	Hard	Off	Off	202
12	2	Out	Old	Off	Back	Soft	Off	On	205
13	2	In	New	On	Back	Soft	Off	Off	212
14	2	Out	New	On	Front	Hard	Off	On	175
15	2	In	Old	On	Front	Soft	On	On	204
16	2	Out	Old	On	Back	Hard	On	Off	201

As suspected by the experimenter, the foldover revealed he was spinning his wheels over the effect of neon lighting (G). The plot of effects in Figure 3 shows the interaction AF instead. This seemed odd because neither of the parent effects came out significant, but it couldn't be ruled out on this basis alone.

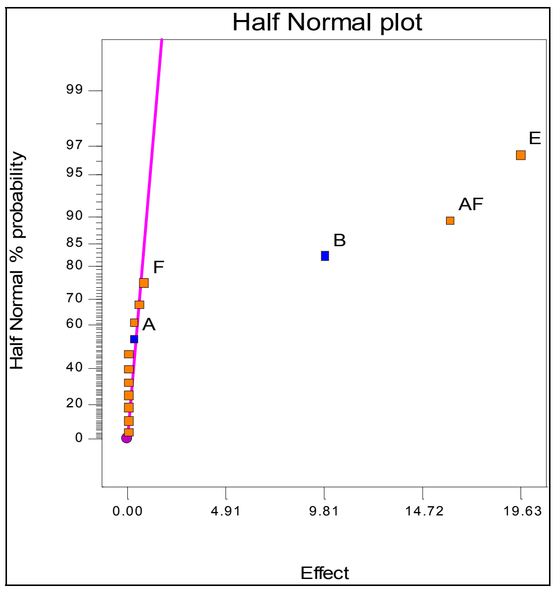

Figure 3: Plot of effects after foldover on skating experiment

The statistical ANOVA (not shown) confirmed the significance of the plot of effects. However, the experimenter could not come up with a plausible reason for an interaction involving the heel pad (A) and the wheel covers (F), so he inspected the alias structure and found that [AF] = AF + BE + CD. (Remember that folding over a resolution III design improves it only to resolution IV, which means that two-factor interactions remain aliased with each other.) The interaction of C (gloves on or off) and D (helmet logo back or front) is just as implausible as AF, so the experimenter rejected this also. The only other possibility, interaction BE, made sense, not only in a physical sense, but because it derives from parent effects already in the predictive model. This is shown in Figure 4, which shows consistently poorer times with the soft wheels (E+). When coupled with the hard wheels (E-), the new bearings made of high-tech alloy (B+) gave a big boost to the skater, who now speeded around the track in record time, at least for him.

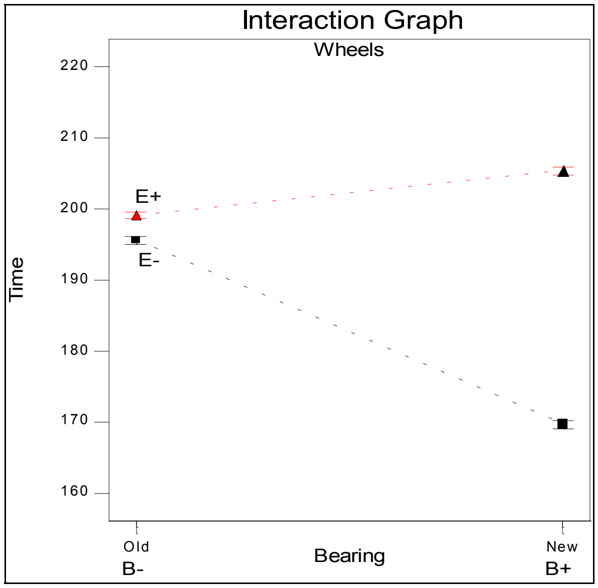

Figure 4: Interaction of B (bearing) and E (hard (-) vs soft (+) wheel)

Confirmation runs provided support for the interaction of bearings and wheels. The two-part DOE proved to be a big success except for one thing: How to stop without crashing. If only someone would come up with reliable brakes for in-line skates!

CASE STUDY ON SEMIFOLDOVER METHODS FOR MEDIUM-RESOLUTION (IV) DESIGNS

As shown in the previous example, foldover methods work great for improving resolution III two-level factorial designs, but what would be a good strategy for resolution IV designs? We suggest that you consider a semifoldover in this case. To illustrate this more efficient design-augmentation method, we will apply it to experimentation on a simulated spin coater (see Figure 5) that applies a photo resist to a silicon wafer. The key response is coating thickness, which must be closely controlled. The following six factors will be varied, each at two levels: A – spin speed, B- acceleration, C – volume, D – spin time, E – resist (there are two methods of preparing the batches) and F – exhaust cover (on or off).

Figure 5. Typical spin coater with control panel for adjusting variables

Let's begin by doing a screening experiment – a 1/4th fractional factorial (2^{6-2}) shown in Table 4. We generated the data on the thickness response with a simulator that included a component for experimental error, thus providing some realism to the exercise. However, since the case study is intended for illustrative purposes only, you should focus on the methods employed, rather than the specific results, and extrapolate them to your own process.

Table 4. First spin-coater experiment (2^{6-2} fractional factorial design)

Standard Order	*Run #*	*Block*	*A Speed RPM*	*B Accel.*	*C Vol. cc*	*D Time sec*	*E Resist Batch*	*F Exhaust Cover*	*Response Thickness Mil*
1	7	1	6650	5	3	6	1	Off	4524
2	8	1	7350	5	3	6	2	Off	4657
3	1	1	6650	20	3	6	2	On	4293
4	14	1	7350	20	3	6	1	On	4516
5	13	1	6650	5	5	6	2	On	4508
6	10	1	7350	5	5	6	1	On	4432
7	11	1	6650	20	5	6	1	Off	4197
8	12	1	7350	20	5	6	2	Off	4517
9	16	1	6650	5	3	14	1	On	4521
10	15	1	7350	5	3	14	2	On	4610
11	3	1	6650	20	3	14	2	Off	4297
12	6	1	7350	20	3	14	1	Off	4560
13	4	1	6650	5	5	14	2	Off	4487
14	2	1	7350	5	5	14	1	Off	4487
15	9	1	6650	20	5	14	1	On	4197
16	5	1	7350	20	5	14	2	On	4509

By inspecting the alias structure for this medium resolution (IV) design (Table 5) you can see that main effects get confounded only with interactions of three or more factors, which can safely be ignored. That's good. However, all two-factor interactions are aliased with at least one other two-factor interaction (2fi). That's bad if any of these interactions significantly affect the coating thickness.

Table 5. Aliases for 2^{6-2} design ([Estimated Terms] = Aliased Terms)

[Intercept] = Intercept + ABCE + ADEF + BCDF
[A] = A + BCE + DEF + ABCDF
[B] = B + ACE + CDF + ABDEF
[C] = C + ABE + BDF + ACDEF
[D] = D + AEF + BCF + ABCDE
[E] = E + ABC + ADF + BCDEF
[F] = F + ADE + BCD + ABCEF
[AB] = AB + CE + ACDF + BDEF
[AC] = AC + BE + ABDF + CDEF
[AD] = AD + EF + ABCF + BCDE
[AE] = AE + BC + DF + ABCDEF
[AF] = AF + DE + ABCD + BCEF
[BD] = BD + CF + ABEF + ACDE
[BF] = BF + CD + ABDE + ACEF
[ABD] = ABD + ACF + BEF + CDE
[ABF] = ABF + ACD + BDE + CEF

The half-normal plot of effects (Figure 6) reveals significantly large main effects due to factors A, B, C and E. It's safe to assume that these are true effects. However, the AB interaction, which also stands out on the effect plot, is aliased with CE. Therefore we are left wondering whether it is AB, or CE or both that are significant.

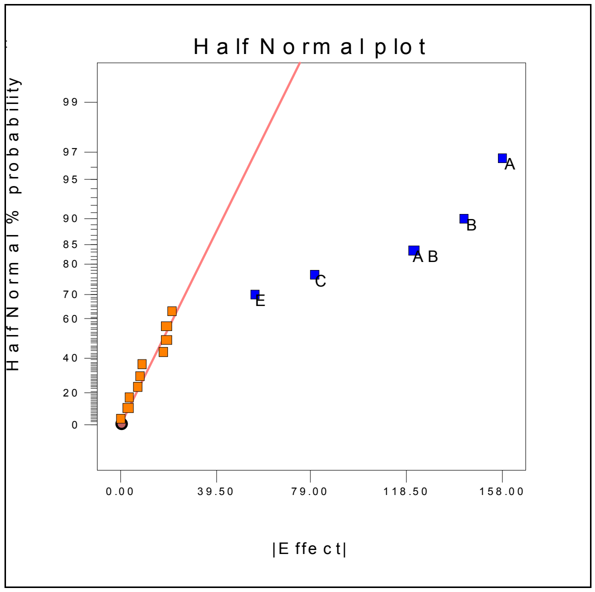

Figure 6. Half-normal plot for first spin-coater experiment

Plots of AB and CE can be seen in Figures 7 and 8. Which one is correct? We can only guess at this point, because there's no way to distinguish between them using the data collected so far.

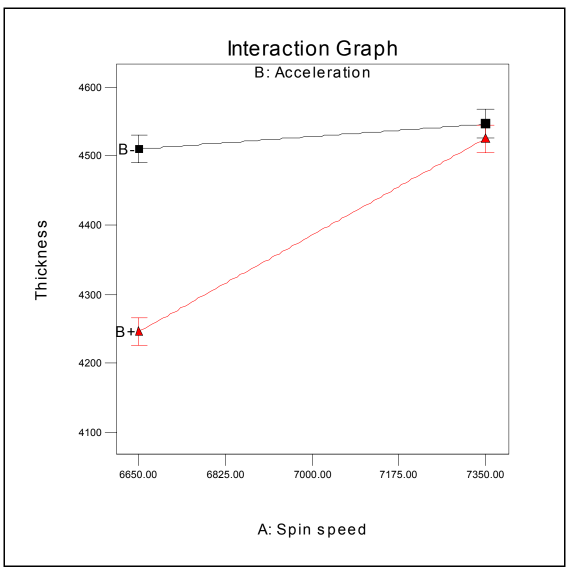

Figure 7: AB interaction plot

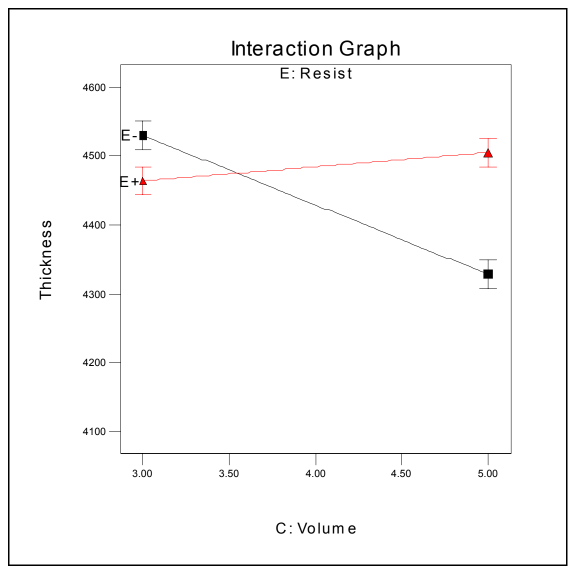

Figure 8: CE interaction plot

To achieve resolution of the aliased interactions (AB versus CE), let's try a single-factor 16-run foldover on A. See Table 6 for the results.

Table 6. Second spin-coater experiment (foldover on factor A only of 2^{6-2} fractional factorial design)

Standard Order	*Run #*	*Block*	*A Speed RPM*	*B Accel.*	*C Vol. cc*	*D Time sec*	*E Resist Batch*	*F Exhaust Cover*	*Response Thickness Mil*
17	20	2	7350	5	3	6	1	Off	4614
18	23	2	6650	5	3	6	2	Off	4447
19	21	2	7350	20	3	6	2	On	4475
20	17	2	6650	20	3	6	1	On	4282
21	30	2	7350	5	5	6	2	On	4612
22	28	2	6650	5	5	6	1	On	4327
23	26	2	7350	20	5	6	1	Off	4327
24	22	2	6650	20	5	6	2	Off	4427
25	29	2	7350	5	3	14	1	On	4657
26	24	2	6650	5	3	14	2	On	4528
27	25	2	7350	20	3	14	2	Off	4488
28	27	2	6650	20	3	14	1	Off	4312
29	31	2	7350	5	5	14	2	Off	4620
30	19	2	6650	5	5	14	1	Off	4336
31	32	2	7350	20	5	14	1	On	4342
32	18	2	6650	20	5	14	2	On	4306

Compare Table 4 to Table 6: Notice how only the levels of A change from block 1 to block 2 – they've gone opposite. The other factor levels remain the same in both blocks.

As you can see in Table 7, all 2fi's involving factor A now become clear of other 2fi's. Most importantly, AB is no longer aliased CE. These two interactions are now aliased only with unlikely four-factor interactions. To save space, we don't show this on the table, but the completed design (with foldover) also gives estimates of 10 three-factor interactions (3fi's) and 2 four-factor interactions (4fi's). These higher-order interactions won't likely be active, so it's really a waste to invest in the 12 runs needed to estimate them.

Table 7. Aliases of main effects and 2fi's for 2^{6-2} design after foldover on factor A

[Intercept] = Intercept + BCDF
[Block 1] = Block 1 + ABCE + ADEF
[Block 2] = Block 2 - ABCE - ADEF
[A] = A + ABCDF
[B] = B + CDF
[C] = C + BDF
[D] = D + BCF
[E] = E + BCDEF
[F] = F + BCD
[AB] = AB + ACDF
[AC] = AC + ABDF
[AD] = AD + ABCF
[AE] = AE + ABCDEF
[AF] = AF + ABCD
[BC] = BC + DF
[BD] = BD + CF
[BE] = BE + CDEF
[BF] = BF + CD
[CE] = CE + BDEF
[DE] = DE + BCEF
[EF] = EF + BCDE

Figure 9 shows the half-normal plot for the second design. It reveals CE as the real interaction, not AB.

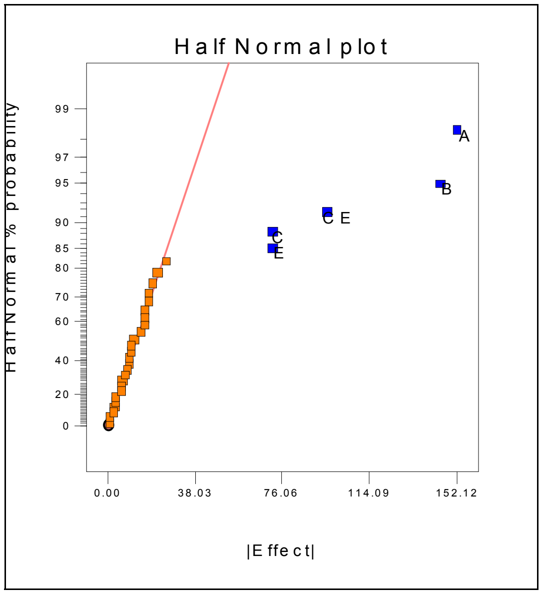

Figure 9: Half-normal plot of effects after folding on single-factor A

Clearly it is the CE interaction that is significant, not AB as shown earlier from the first block of experiments.

Comparing aliasing before and after foldover (Tables 5 and 7), note that before foldover, the fifteen two-factor interactions occur in 6 strings of two, plus 1 string of three aliased together, for a total of 7 estimable 2fi's. After foldover, 9 of the 2fi's are free of other 2fi's, while six 2fi's remain aliased in 3 strings of two each, which adds up to 12 two-factor interactions that can be estimated. Therefore, folding frees up only 5 additional two-factor interactions at a cost of double the runs. This is not a very efficient use of resources. In this case, a semifold (adding only eight new runs rather than sixteen), provides sufficient degrees of freedom to estimate the five additional 2fi's. This will generally be the case for any resolution IV fractional factorial - a semifold supports estimation of as many two-factor interactions as a full foldover.

To ensure a beneficial semifold from a resolution IV design, we recommend that you follow these two steps:

1. Lay out a single-factor foldover from the original design. (Suggestion: choose a factor that's involved in the largest significant two-factor interaction that's aliased with other 2fi(s).)
2. Perform only half of the foldover runs by selecting those where the chosen factor is either at its low level or high level, whichever you believe will generate the most desirable response(s).

Table 8 shows the result of this semifolding process on the spin coater example.

Table 8: Semifold on spin-coater design with factor A set at its low level

Standard Order	*Run #*	*Block*	*A Speed RPM*	*B Accel.*	*C Vol. cc*	*D Time sec*	*E Resist Batch*	*F Exhaust Cover*	*Response Thickness Mil*
17	21	2	6650	5	3	6	2	Off	4447
18	17	2	6650	20	3	6	1	On	4282
19	24	2	6650	5	5	6	1	On	4327
20	20	2	6650	20	5	6	2	Off	4427
21	22	2	6650	5	3	14	2	On	4528
22	23	2	6650	20	3	14	1	Off	4312
23	19	2	6650	5	5	14	1	Off	4336
24	18	2	6650	20	5	14	2	On	4306

We simply selected the runs from Table 6 (the single-factor foldover on factor A) where factor A is at its low level, while deleting from the foldover block all of the runs at the high level of factor A. (The first design indicated that the low level of factor A would give better results than the high level.)

The aliases for the semifolded design can be seen in Table 9. Note that twelve two-factor interactions can be estimated, the same as the folded design.

Table 9: Aliases for semifolded design

[Intercept] = Intercept + BCE + DEF + BCDF
[Block 1] = Block 1 - BCE - DEF + ABCE + ADEF
[Block 2] = Block 2 + BCE + DEF - ABCE - ADEF
[A] = A + BCE + DEF + ABCDF
[B] = B + ACE + CDF + ABDEF
[C] = C + ABE + BDF + ACDEF
[D] = D + AEF + BCF + ABCDE
[E] = E + ABC + ADF + BCDEF
[F] = F + ADE + BCD + ABCEF
[AB] = AB + ACE + ACDF + ABDEF
[AC] = AC + ABE + ABDF + ACDEF
[AD] = AD + AEF + ABCF + ABCDE
[AE] = AE + ABC + ADF + ABCDEF
[AF] = AF + ADE + ABCD + ABCEF
[BC] = BC + DF - ABC - ADF
[BD] = BD + CF + ABEF + ACDE
[BE] = BE - ABE + CDEF - ACDEF
[BF] = BF + CD + ABDE + ACEF
[CE] = CE - ACE + BDEF - ABDEF
[DE] = DE - ADE + BCEF - ABCEF
[EF] = EF - AEF + BCDE – ABCDE

Figure 10 shows the half-normal plot for the effects from the semifolded design. As with the folded design, the CE interaction is isolated as the significant two-factor interaction, so this approach proved to be successful.

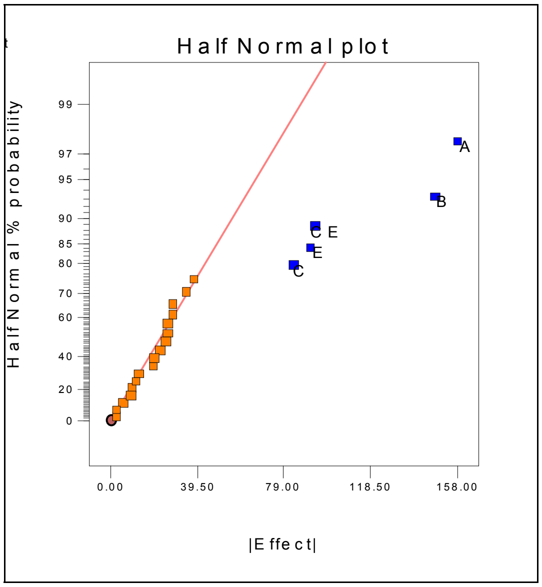

Figure 10: Half-normal plot of effects for semifolded design

Cutting down the design saves time and materials but it comes at a cost: Because the semifold adds only eight runs versus the sixteen for a full foldover, the power is reduced somewhat, so very small effects might be overlooked. This did not happen in the spin coater case, but if experiments can be done quickly at little expense, and you think you need more power, do a standard foldover. Otherwise, you will be better off doing the semifold.

CONCLUSION

Doing a complete foldover is standard operating procedure for experimenters who try to screen factors via resolution III two-level fractional factorial designs. This is guaranteed to improve the resulting design to resolution IV. We suggest that you go directly to a resolution IV designs for screening purposes. What you do as a result of running this higher-resolution design depends on which, if any, effects come out significant. Here's a general strategy for resolution IV screening designs:

Scenario 1 - Nothing significant: Look for other factors that affect your response(s).

Scenario 2 - Only main effects significant: Change these factors to their best levels.

Scenario 3 -Two-factor interaction(s) significant: De-alias by performing a semifold.

By following this strategy you will increase your odds of uncovering breakthrough main effects and interactions at a relatively minimal cost in experimental runs.

REFERENCES

Anderson, M. J. and P. J. Whitcomb, 2000. *DOE Simplified, Practical Tools for Experimentation*. Portland, Oregon: Productivity, Inc.

Mee, R. W. and M. Peralta, 2000. Semifolding 2^{k-p} Designs, *Technometrics*, Vol. 42, No.2 (May): 122 – 143.

Montgomery, D. C., 2001. *Design and Analysis of Experiments*, 5th edition. New York: John Wiley and Sons, Inc.

Sitter, R. R., J Chen and M. Feder,1997. Fractional Resolution and Minimum Aberration in Blocked 2^{n-k} Designs, *Technometrics*, 1997, Vol. 39, No. 4 (November): 382-390.

Sun, D. X., C. F. J. Wu and Y Chen, 1997. Optimal Blocking Schemes for 2^n and 2^{n-p} Designs, *Technometrics*, Vol. 39, No. 3 (August): 298-307.