Screening Designs and Design Evaluation

Bradley Jones
Science of Test
April 2017

Statistical Facts of Life

1. All real processes experience variability
2. Statistical models are approximations

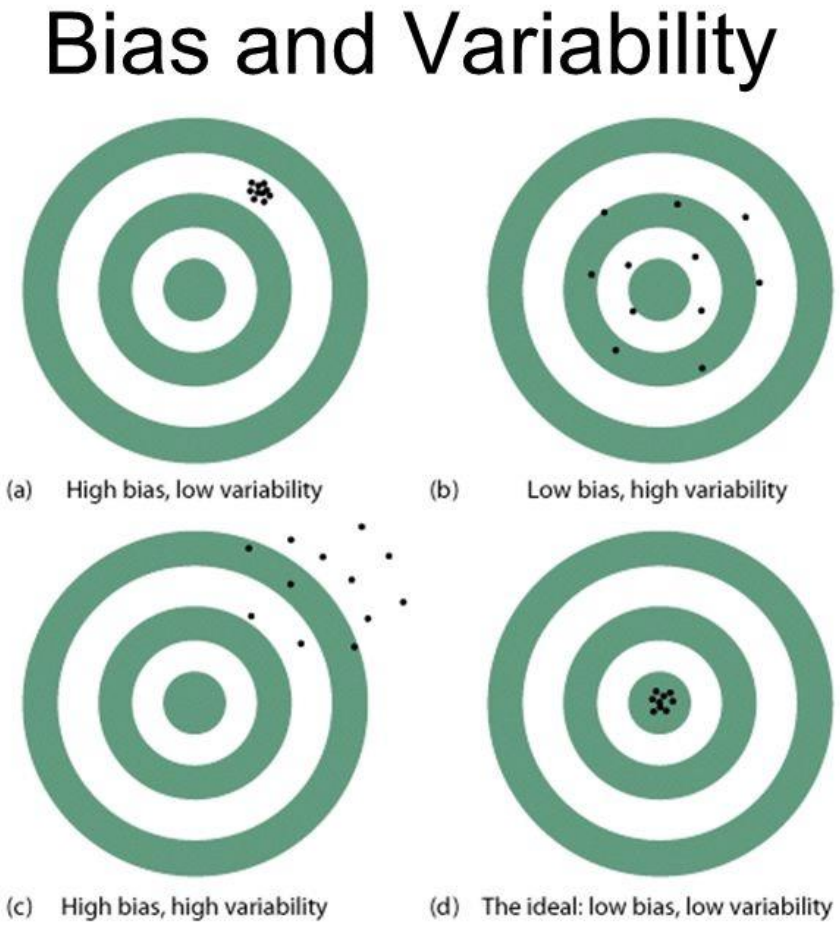
What is the purpose of designing experiments?

Model -> Design

Goals

1. Make the variance of unknowns in a model as small as possible
2. Avoid bias in estimates of unknowns.

Visualizing Bias and Variance



What makes a design good?

1. Low variance of the coefficients.
2. Low variance of predicted responses.
3. Minimal aliasing of terms in the model from likely effects that are not in the model (0.5 or less).
4. Likely effects that are not in the model are not confounded (smaller correlations are better).

The first two deal with variance – the last two with bias.
Reducing variance and bias are fundamental goals.

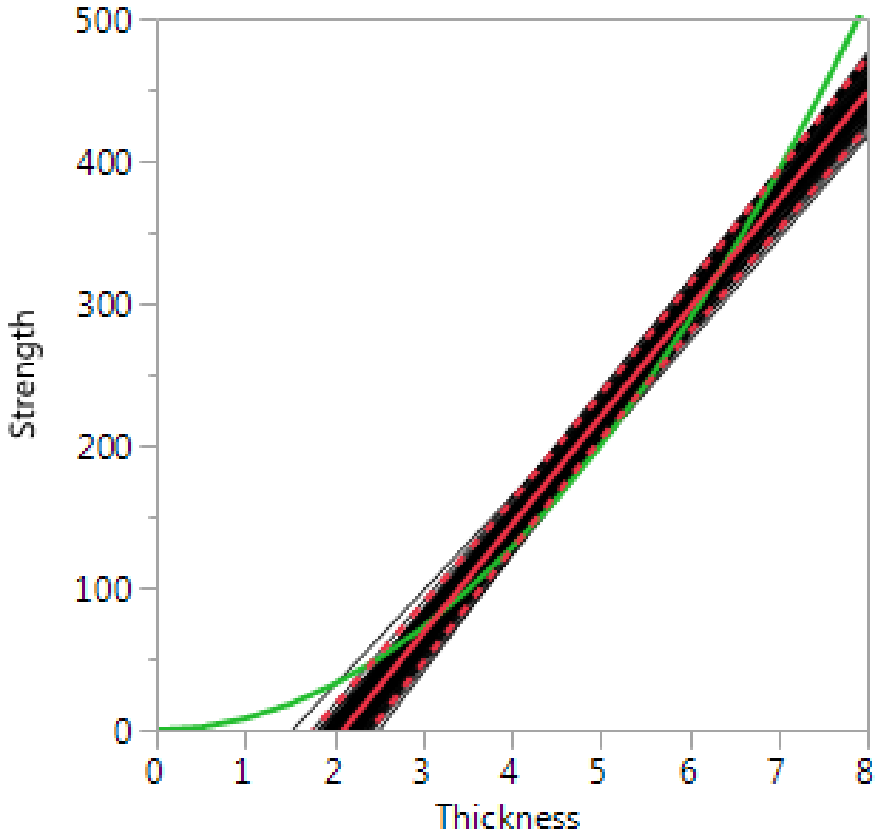
How to lower variance using design

1. Replicate or otherwise add runs to a design
2. For continuous factors, increase the range between high and low settings
3. Use blocking, i.e. run your experiment in homogeneous groups of runs
4. Reduce or eliminate correlation between factors

How to lower bias using design

1. Randomize the run order
2. Avoid confounding

JMP Variance/Bias Demonstration



Summary

1. Variance and bias are fundamental criteria for evaluating designs.
2. Designed experiments aim to reduce variance or both variance and bias compared to trial and error testing.

Screening or finding the “vital few” important factors

Screening Experiments

Goals

1. Define Screening
2. Introduce three types of screening design – pros and cons
3. Compare a textbook screening design to a modern design.

What is a screening experiment?

A major use of fractional factorials is in screening experiments – experiments in which many factors are considered and the objective is to identify those factors (if any) that have large effects.

Montgomery, D.C. (2005) Design and Analysis of Experiments, Wiley, New York, page 283.

More about screening experiments

Screening experiments generally involve many factors (>5)

1. Traditionally, each factor has two levels.
2. The *a priori* model is either main effects only or main effects plus two-factor interactions.
3. The number of experimental runs is kept small to make them inexpensive and fast.
4. Screening experiments are the most commonly run industrial applications.

In factor screening experiments…

We start with little prior knowledge and a large initial set of potential factors influencing the response

Our purpose is to identify the smaller set of active factors.

Primary goal : identify active main effects (MEs)

Secondary goal: identify a few active 2^{nd} order effects

What makes a factor important?

1. Having a large main effect
2. Being involved in a two-factor interaction effect
3. Having a large nonlinear effect

Typical Screening Assumptions

Sparsity of effects – fewer than half of the factors will be active.

Hierarchy of effects – main effects larger and more likely than 2^{nd} order effect

Effect heredity – two-factor interactions are much more likely if both main effects are active

Supporting article in Complexity (2006)

Regularities in Data from Factorial Experiments

XIANG LI,[1] NANDAN SUDARSANAM,[2] AND DANIEL D. FREY[1,2]
Massachusetts Institute of Technology, [1]Department of Mechanical Engineering; and [2]Engineering Systems Division, Cambridge, Massachusetts 02139

DOI 10.1002/cplx

Complexity article method and results

Full-factorial designs from three to seven factors

113 sets of responses re-analyzed to check screening assumptions

Findings:

1) Percentage of Main Effects was 41%
2) Main effects averaged twice as large and three times as many as 2FIs
3) Percentage of 2FI was 11% overall
4) Conditional percentages
    1) Both main effects active 33%
    2) One main effect active 4.5%
    3) No main effect active 0.5%

Number of active 2FIs about 1/3 of the number of active MEs

Percentage of Potential Effects in 113 Experiments That Were Active as Determined by the Lenth Method

	Main Effects	Two-Factor Interactions
No. of effects	410	569
No. of active effects	170	63
Percentage of effects that were active (%)	41	11
Confidence intervals ($\alpha = 0.05$) on the percentage of effects that were active (%)	37–46	9–14

So, if there are 6 factors, for example, you would expect 2 or 3 active main effects and one or two 2FIs

Traditional Screening Designs

Resolution III Fractional Factorial

Resolution IV Fractional Factorial

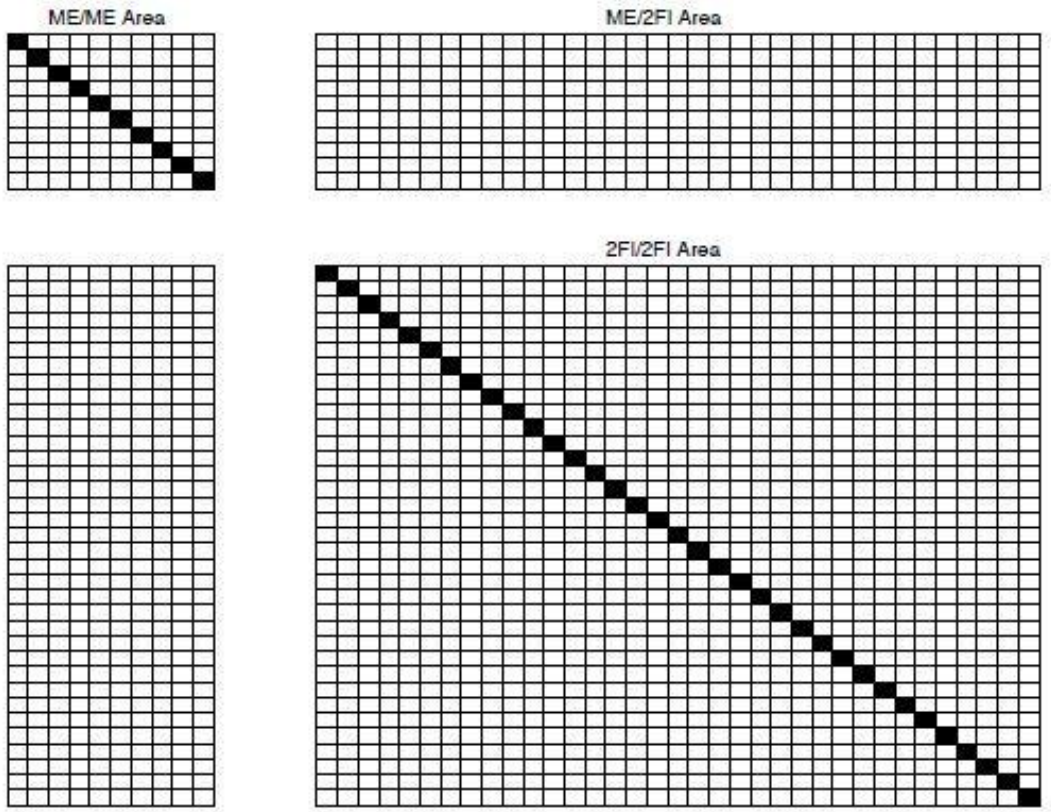
Plackett-Burman

Industrial experience indicates that most designs are not replicated. Budget constraints limit the allowable number of runs.

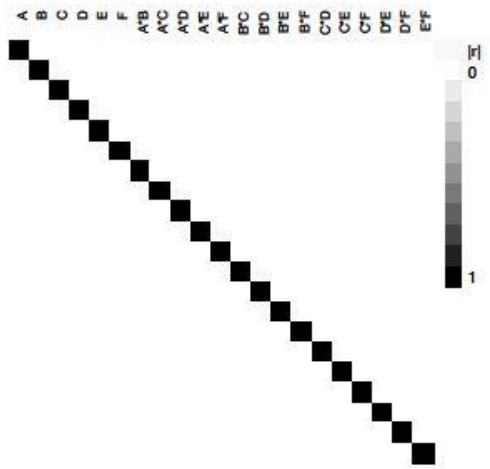
A Graph for Quick Design Comparison

The correlation cell plot…

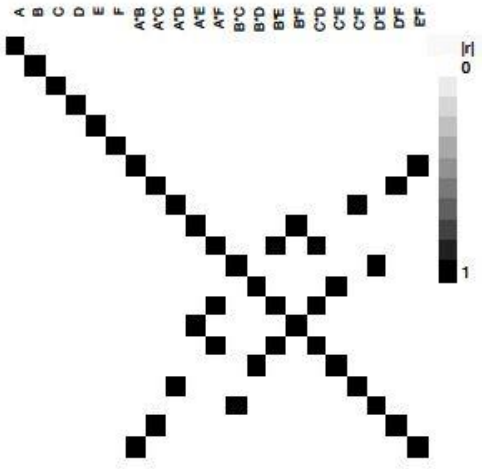
1. Cells off the diagonal are white for orthogonal effects
2. Confounded effects are colored black
3. The greater the correlation, the darker the cell



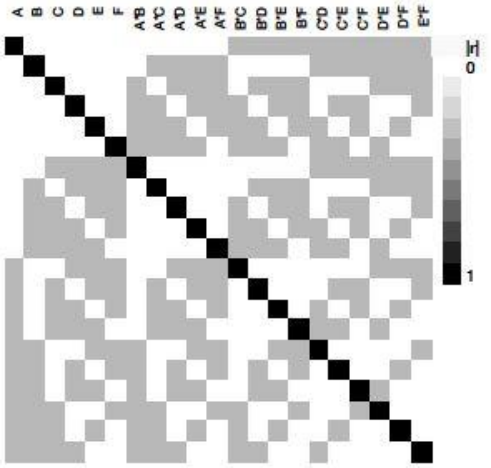
Cell plots of 4 screening designs for 6 factors



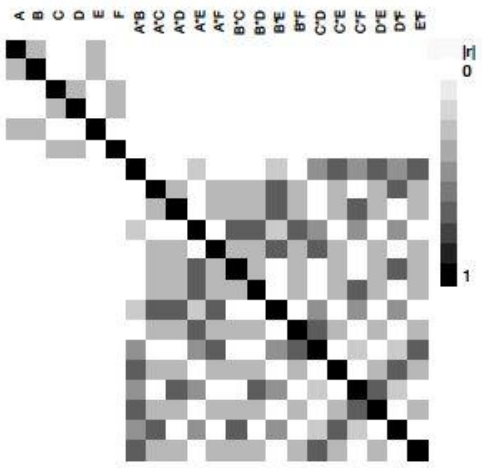
(a) 2^{6-1} resolution VI design: $n = 32$



(b) 2^{6-2} resolution IV design: $n = 16$



(c) Plackett-Burman design: $m = 6$ and $n = 12$



(d) EFD: $m = 6$ and $n = 12$

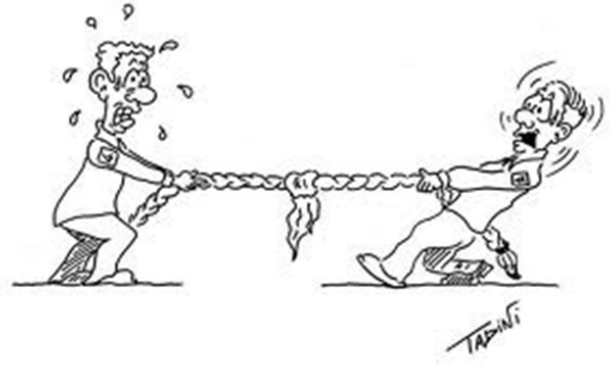
Desirable design features

orthogonality of the MEs,

orthogonality of MEs and 2FIs,

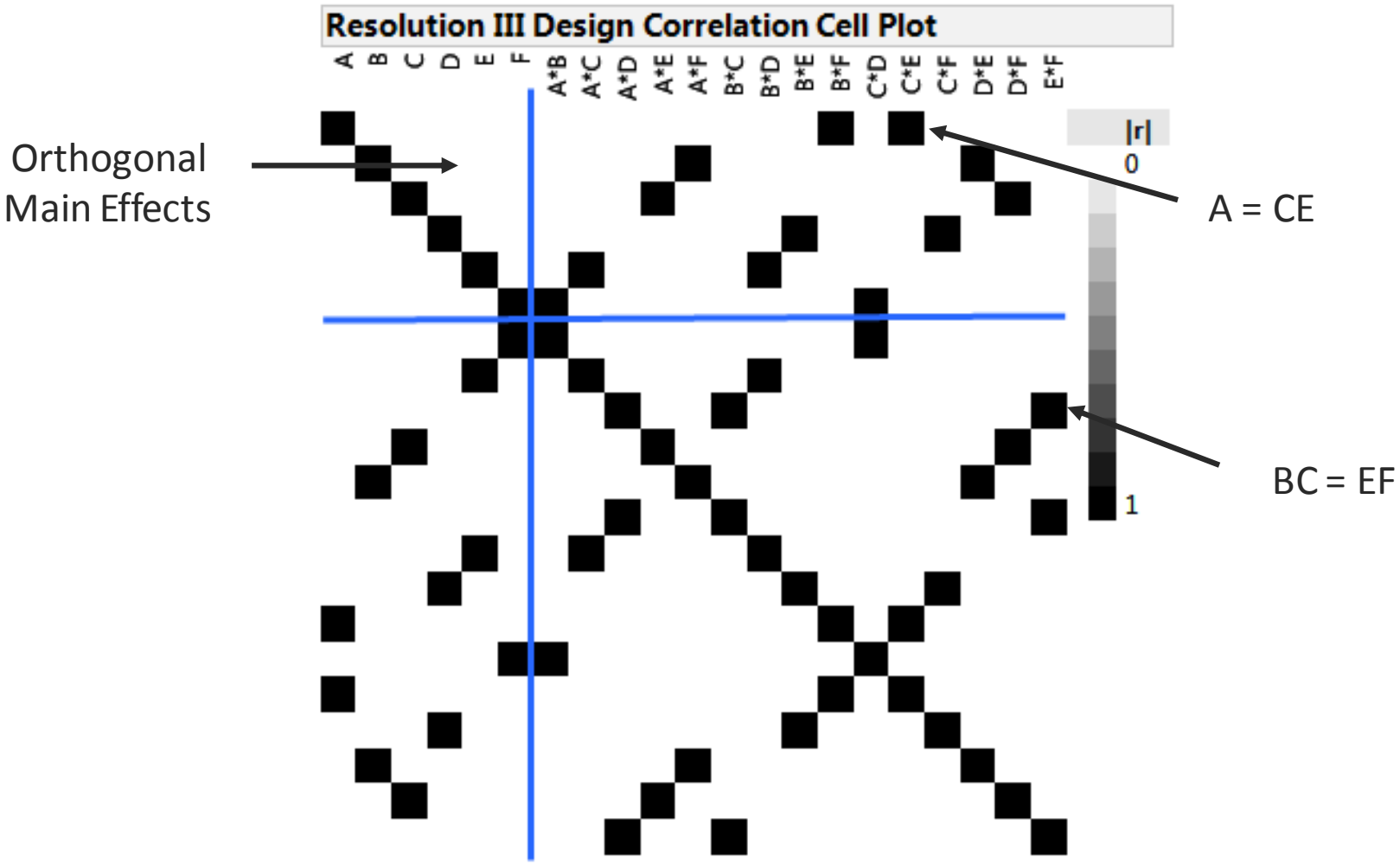
orthogonality of 2FIs with each other,

small run size



Problems with Standard Screening Designs

Resolution III fractional factorial designs confound main effects with two-factor interactions and two-factor interactions with each other.

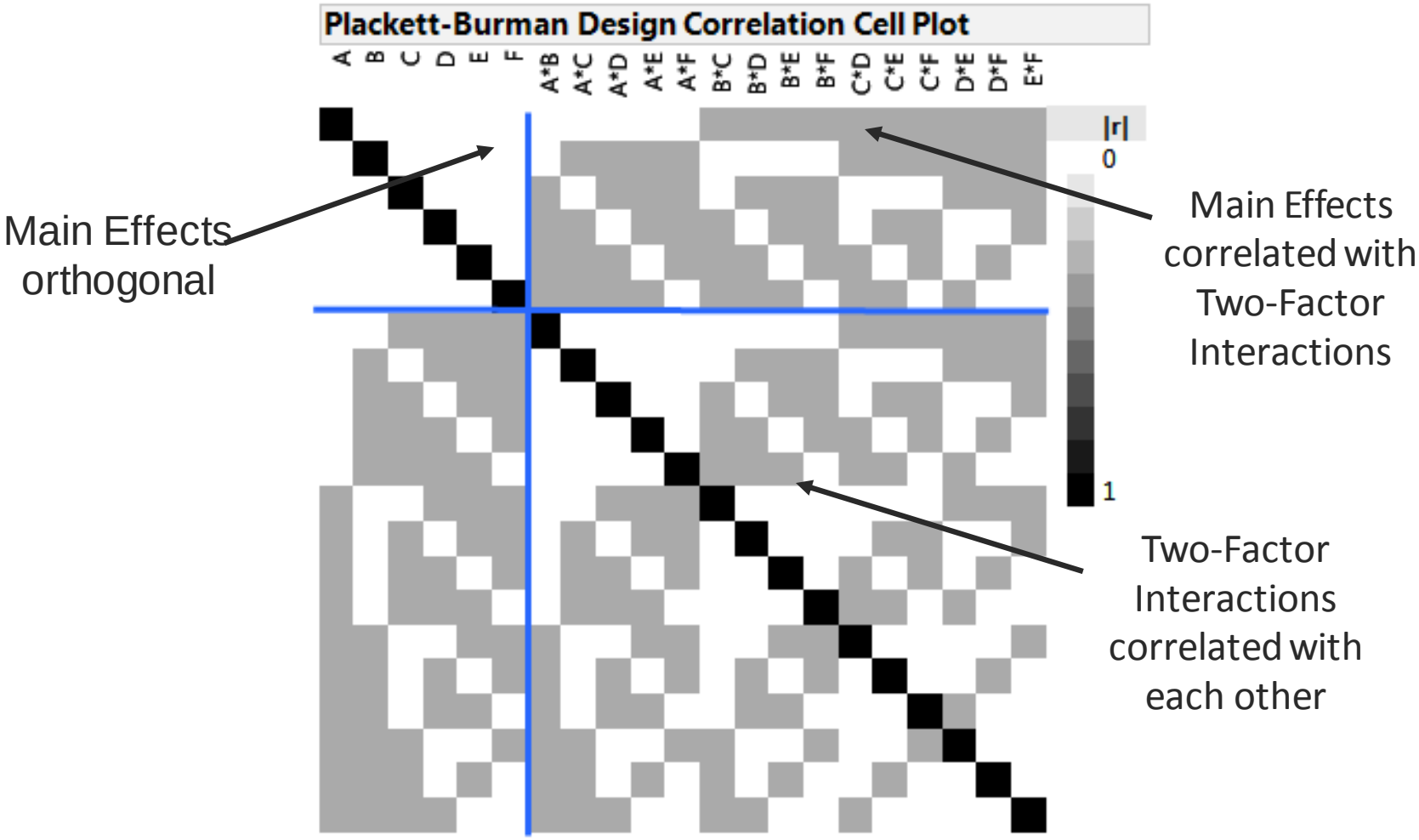


Possible Consequences

1. An active two-factor interaction may cause its confounded main effect to look insignificant when it is actually active.
2. Conversely, an active two-factor interaction can cause its confounded main effect to look significant when it is not.
3. Active two-factor interactions are generally not uniquely identifiable.

Problems with Standard Screening Designs

Nonregular orthogonal designs such as Plackett-Burman designs have “complex aliasing” of the main effects by two-factor interactions. Two-factor interactions are also correlated.

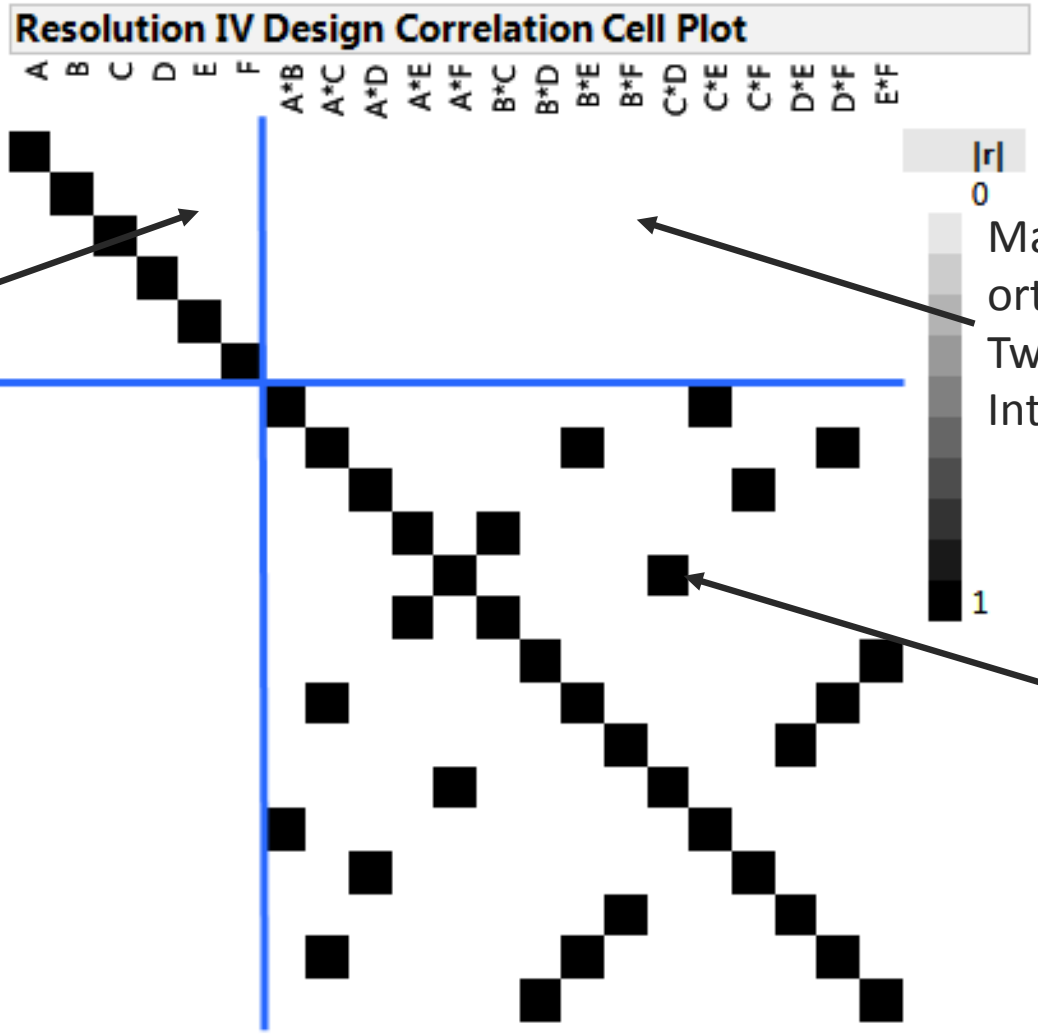


Possible Consequences

1. An active two-factor interaction may cause its correlated main effect to look insignificant when it is actually active.
2. Conversely, an active two-factor interaction can cause its correlated main effect to look significant when it is not.

Problems with Standard Screening Designs

Resolution IV fractional factorial designs confound two-factor interactions with each other.



Orthogonal Main Effects

Main Effects orthogonal to Two-Factor Interactions

Two-Factor Interactions are confounded with each other.

Possible Consequence

An active two-factor interaction may not be uniquely identifiable, so further runs may be necessary to resolve the active effects.

Further Problems with Standard Screening Designs

1. Adding center points to detect curvature works but does not identify which factors are responsible.
2. Blocking is restrictive – the number of blocks may only be a power of two.

Screening Conundrum – too many potential effects

Number of Factors	Number of Two-Factor Interactions
6	15
7	21
8	28
9	36
10	45

Two-factor interactions matter but for standard screening scenarios there are so many of them that estimating them all is too expensive.

Screening Design Example

Scenario

1. There are 6 factors.
2. The model consists of all the main effects.
3. Budget for 12 experimental runs.

Engineering Data with Responses

Run	Methanol	Ethanol	Propanol	Butanol	pH	Time	Yield (mg)
1	0	0	0	10	6	1	10.94
2	0	10	0	0	9	1	15.79
3	0	10	0	10	9	2	25.96
4	10	10	10	0	6	1	35.92
5	0	0	10	0	6	2	22.92
6	0	10	10	10	6	1	23.54
7	10	10	0	0	6	2	47.44
8	10	0	0	0	9	1	19.80
9	10	0	10	10	9	1	29.48
10	0	0	10	0	9	2	17.13
11	10	10	10	10	9	2	43.75
12	10	0	0	10	6	2	40.86

Bie, Xiaomei, et. al. “Screening the main factors affecting extraction of the antimicrobial substance from Bacillus sp. fmbJ using the Plackett–Burman method “ World Journal of Microbiology & Biotechnology (2005) 21: 925–928 Springer 2005

Main – Effects Models

All Main Effects

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	27.79	0.71	39.0	<.0001*
Methanol(0,10)	8.41	0.71	11.8	<.0001*
Ethanol(0,10)	4.27	0.71	6.0	0.0019*
Propanol(0,10)	1.00	0.71	1.4	0.2214
Butanol(0,10)	1.29	0.71	1.8	0.1292
pH(6,9)	-2.48	0.71	-3.5	0.0178*
Time(1,2)	5.22	0.71	7.3	0.0007*

Model in Published Paper

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	27.79	0.86	32.2	<.0001*
Methanol(0,10)	8.41	0.86	9.8	<.0001*
Ethanol(0,10)	4.27	0.86	5.0	0.0017*
pH(6,9)	-2.48	0.86	-2.9	0.0240*
Time(1,2)	5.22	0.86	6.0	0.0005*

Researchers removed two main effects that were not statistically significant.

JMP Screening Design Demonstration

All Main Effects Model

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	27.79	0.71	39.0	<.0001*
Methanol(0,10)	8.41	0.71	11.8	<.0001*
Ethanol(0,10)	4.27	0.71	6.0	0.0019*
Propanol(0,10)	1.00	0.71	1.4	0.2214
Butanol(0,10)	1.29	0.71	1.8	0.1292
pH(6,9)	-2.48	0.71	-3.5	0.0178*
Time(1,2)	5.22	0.71	7.3	0.0007*

Alias Matrix

Effect	Methanol*Time	Propanol*Time
Intercept	0	0
Methanol	0	-0.33
Ethanol	0.333	-0.33
Propanol	-0.33	0
Butanol	0.333	-0.33
pH	-0.33	0.333
Time	0	0

Definitive Screening Design Tutorial

1. What is a definitive screening design (DSD)?
2. Conference matrix based DSDs
3. Analyzing DSDs
4. Adding two-level categorical factors
5. Blocking schemes for DSDs

A six-factor screening example

Experimenter has six factors to investigate and is concerned about two-factor interactions

Can only afford 12 runs

Standard advice: Plackett-Burman 12-run design.

But the PB design has aliasing between main effects and 2FIs

It is possible to create an alias-optimal design where main effects are completely independent of 2FIs

Alias optimal design, keeping D-efficiency > 91%

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	1	1	-1	1	1	1
2	-1	-1	1	-1	-1	-1
3	1	1	1	-1	1	-1
4	-1	-1	-1	1	-1	1
5	1	1	-1	1	-1	-1
6	-1	-1	1	-1	1	1
7	1	-1	-1	1	1	-1
8	-1	1	1	-1	-1	1
9	1	-1	1	1	-1	1
10	-1	1	-1	-1	1	-1
11	1	-1	-1	-1	-1	1
12	-1	1	1	1	1	-1

D-efficiency: 92% vs. orthogonal design

Cost: 10% longer confidence intervals for MEs

Alias matrix:
Can we get rid of the nonzero guys in row 1?

Effect	A*B	A*C	A*D	A*E	A*F	B*C	B*D	B*E	B*F	C*D	C*E	C*F	D*E	D*F	E*F
Intercept	0	0.333	0.333	0	0	0	0	-0.33	0.333	0.333	0	0	0	0	-0.33
A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
D	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
F	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Answer: Yes, lower the D-efficiency constraint to 85%

New Design 1

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	1	1	-1	1	1	1
2	-1	-1	1	-1	-1	-1
3	1	1	1	-1	1	-1
4	-1	-1	-1	1	-1	1
5	1	1	-1	1	-1	-1
6	-1	-1	1	-1	1	1
7	1	-1	-1	1	1	-1
8	-1	1	1	-1	-1	1
9	1	-1	1	1	-1	1
10	-1	1	-1	-1	1	-1
11	1	-1	-1	-1	-1	1
12	-1	1	1	1	1	-1

D-efficiency: 92%

(Not Orthogonal)

"Size" of Alias Matrix = 6/9

New Design 2

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0

D-efficiency: 85%

(Orthogonal)

"Size" of Alias Matrix = 0

Some Interesting Facts About Design 2

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	0	1	-1	-1	-1	-1
2	0	-1	1	1	1	1
3	1	0	-1	1	1	-1
4	-1	0	1	-1	-1	1
5	-1	-1	0	1	-1	-1
6	1	1	0	-1	1	1
7	-1	1	1	0	1	-1
8	1	-1	-1	0	-1	1
9	1	-1	1	-1	0	-1
10	-1	1	-1	1	0	1
11	1	1	1	1	-1	0
12	-1	-1	-1	-1	1	0

Each pair of rows mirror each other (foldover pairs).

Exactly one center point in each run.

D-efficiency is 85%.

Size of A is zero, so the alias matrix is all zeros!

Hey, pretty cool structure!

OK, Nice Result for Six Factors
Does This Work for Any Number of Factors?

Foldover structure guarantees A = 0 (I will show this later)

So all main effects will be unbiased by active interactions

Center points in each row guarantees your ability to estimate quadratic effects for all factors

If we impose this structure, we're guaranteed a nice design if we pick the plus-or-minus one entries carefully

May not be orthogonal, however

Design structure for definitive screening designs

Table 1: General design structure for m factors

Foldover Pair	Run (i)	Factor Levels				
		$x_{i,1}$	$x_{i,2}$	$x_{i,3}$	$\cdots$	$x_{i,m}$
1	1	0	± 1	± 1	$\cdots$	± 1
	2	0	∓ 1	∓ 1	$\cdots$	∓ 1
2	3	± 1	0	± 1	$\cdots$	± 1
	4	∓ 1	0	∓ 1	$\cdots$	∓ 1
3	5	± 1	± 1	0	$\cdots$	± 1
	6	∓ 1	∓ 1	0	$\cdots$	∓ 1
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
m	$2m-1$	± 1	± 1	± 1	$\cdots$	0
	$2m$	∓ 1	∓ 1	∓ 1	$\cdots$	0
Centerpoint	$2m+1$	0	0	0	$\cdots$	0

A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects

BRADLEY JONES

SAS Institute, Cary, NC 27513

CHRISTOPHER J. NACHTSHEIM

Carlson School of Management, University of Minnesota, Minneapolis, MN 55455

Screening designs are attractive for assessing the relative impact of a large number of factors on a response of interest. Experimenters often prefer quantitative factors with three levels over two-level factors because having three levels allows for some assessment of curvature in the factor–response relationship. Yet, the most familiar screening designs limit each factor to only two levels. We propose a new class of designs that have three levels, provide estimates of main effects that are unbiased by any second-order effect, require only one more than twice as many runs as there are factors, and avoid confounding of any pair of second-order effects. Moreover, for designs having six factors or more, our designs allow for the efficient estimation of the full quadratic model in any three factors. In this respect, our designs may render follow-up experiments unnecessary in many situations, thereby increasing the efficiency of the entire experimentation process. We also provide an algorithm for design construction.

Key Words: Alias; Confounding; Coordinate Exchange Algorithm; D-Efficiency; Response Surface Designs; Robust Designs; Screening Designs.

JOURNAL OF QUALITY TECHNOLOGY , VOL. 43, NO. 1, QICID: 33051, January 2011, pp. 1-15

Design Properties

1. The number of required runs is only one more than twice the number of factors.

Design Properties

1. The number of required runs is only one more than twice the number of factors.
2. Unlike resolution III designs, main effects are completely independent of two-factor interactions.

Design Properties

1. The number of required runs is only one more than twice the number of factors.
2. Unlike resolution III designs, main effects are completely independent of two-factor interactions.
3. Unlike resolution IV designs, two-factor interactions are not completely confounded with other two-factor interactions, although they may be correlated

Design Properties

1. The number of required runs is only one more than twice the number of factors.
2. Unlike resolution III designs, main effects are completely independent of two-factor interactions.
3. Unlike resolution IV designs, two-factor interactions are not completely confounded with other two-factor interactions, although they may be correlated
4. Unlike resolution III, IV and V designs with added center points, all quadratic effects are estimable in models comprised of any number of linear and quadratic main effects terms.

Design Properties (continued)

5. Quadratic effects are orthogonal to main effects and not completely confounded (though correlated) with interaction effects.

Design Properties (continued)

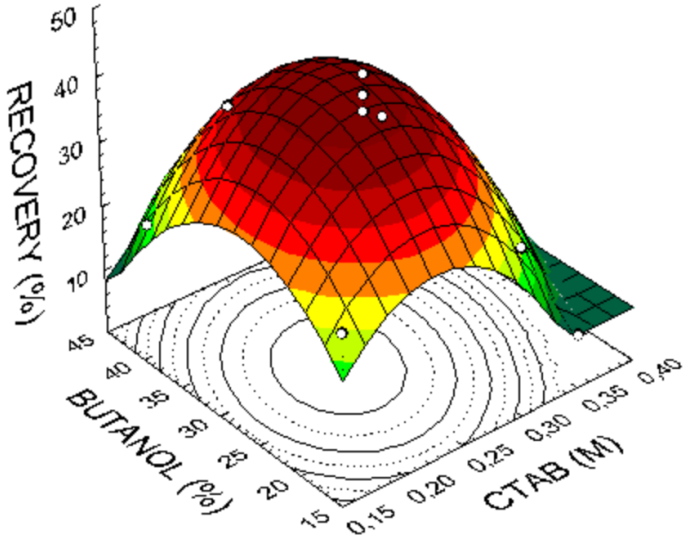
5. Quadratic effects are orthogonal to main effects and not completely confounded (though correlated) with interaction effects.
6. With six or more factors, the designs are capable of estimating all possible full quadratic models involving three or fewer factors with very high levels of statistical efficiency.

Design Properties (continued)

5. Quadratic effects are orthogonal to main effects and not completely confounded (though correlated) with interaction effects.
6. With six or more factors, the designs are capable of estimating all possible full quadratic models involving three or fewer factors with very high levels of statistical efficiency.
7. Designs for 4, 6, 8, and 10 factors are orthogonal.

Screening at Three Levels has Distinct Advantages

1. The world is not linear!
2. We can include current settings in experiments where we are assessing the impact of increases and decreases to the current “best” settings.
3. We may be able to screen and optimize in one fell swoop.



Summary: Why are DSDs “Definitive?”

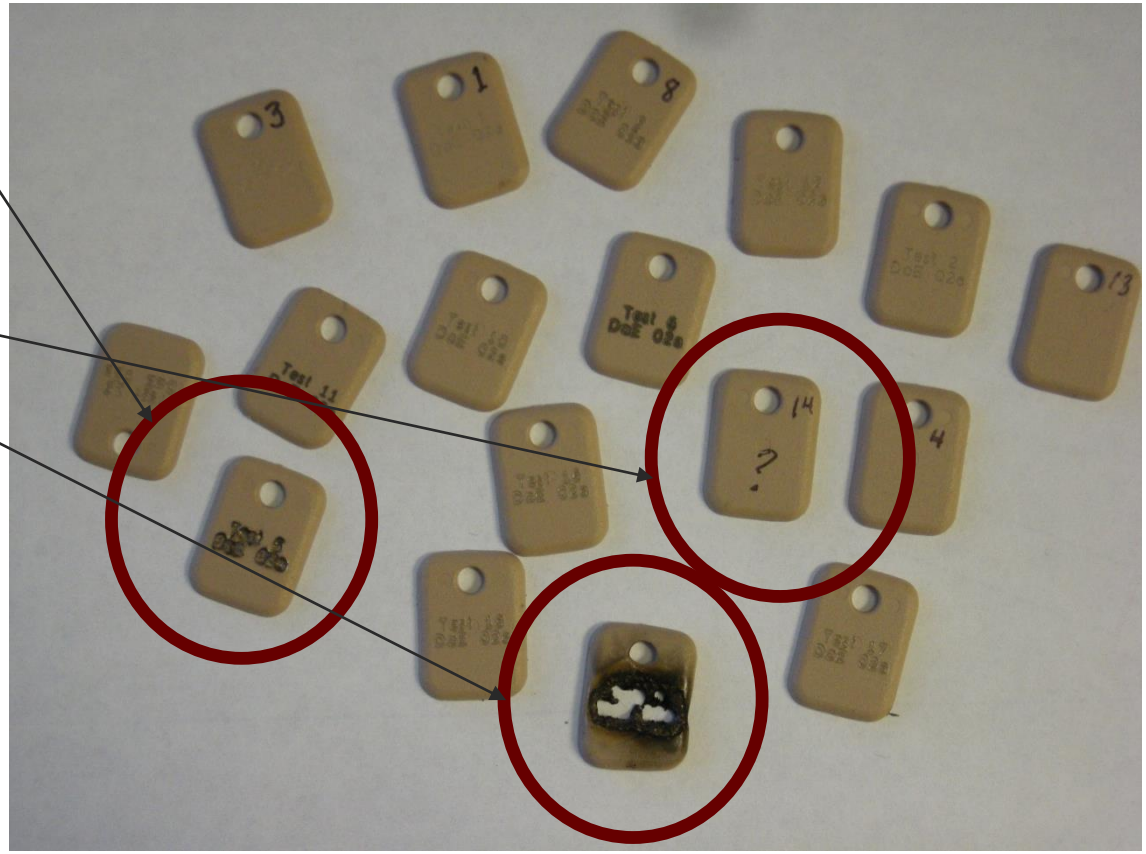
DSDs are “definitive” in the sense that:

- Main effects are free of two-factor interactions
- Quadratic effects are all estimable
- Two-factor interactions not tied to two-factor interactions, hence, no need for follow-up work to resolve confounding issue.
- You may be able to do screening and optimization in one step

A laser etching experiment

Initial experience:
From easy to read,

To not-so-easy to read



Factors and ranges….

Factors	Factor ranges	
	Low level	High level
Mark Speed	8	15
Frequency	1	5
Percent Power	15	55
Repetitions	1	5
Humidity	5%	15%
Plastic	1	3

Standard approach 1: Resolution IV Fractional Factorial Design

2^{6-2} resolution IV fractional factorial design

Number of runs required: 16

Resulting process settings:

Speed = 15 (High)

Frequency = 5 (Low)

Power = 55% (High)

Best process score: 6.31

Standard approach 2:
Central Composite response surface Design (CCD)

CCD is a resolution V fractional factorial (32 runs) with $2m = 12$ runs added to allow estimation of all curvatures plus an overall center-value run (so $n = 45$).

12 added runs:

±1	**0**	**0**	**0**	**0**	**0**
0	**±1**	**0**	**0**	**0**	**0**
0	**0**	**±1**	**0**	**0**	**0**
0	**0**	**0**	**±1**	**0**	**0**
0	**0**	**0**	**0**	**±1**	**0**
0	**0**	**0**	**0**	**0**	**±1**

Results from 45-run CCD

Number of runs required: 45

Resulting process settings:

- Speed = 15 (High)
- Repetitions = 1 (Low)
- Frequency = 1 (Low)
- Humidity = 5 (Low)
- Power = 39.78% (Optimum)
- Plastic = 1.83 (Optimum)

Best process score: 1.80

New Approach: Definitive Screening Design

DSD with all continuous factors

Number of runs required: 13

Resulting process settings:

- Speed = 15 (High)
- Frequency = 1 (Low)
- Power = 38.31% (Optimum)

Best process score: 1.74

Comparison of designs

Design	Runs	Number of active factors found	Best process setting result (smaller is better)
Resolution IV FF	16	3	6.31
Central Composite RSD	45	6	1.8
DSD	13	3	1.74

Verification run at DSD optimum settings

Why did the DSD do so well?

There was a large quadratic effect for Power

The two-level screening design could not detect this

The CCD response surface design found it.

The DSD found it.

The DSD also performed well with only 13 runs because of sparsity---only five active effects.

Overview – Definitive Screening Designs

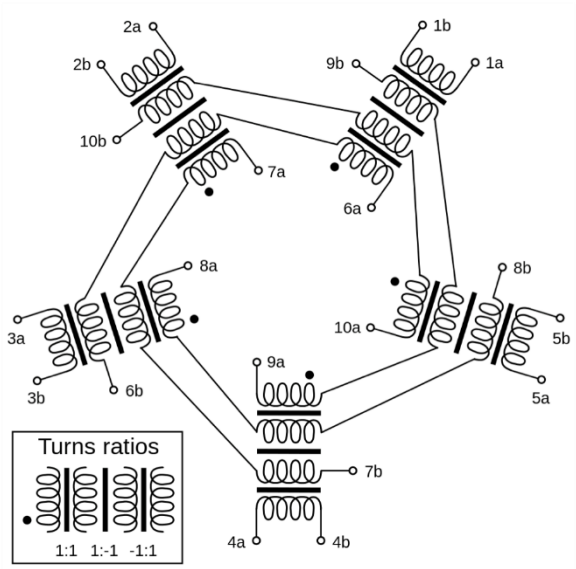
1. What is a definitive screening design (DSD)?
2. Conference matrix based DSDs
3. Analyzing DSDs
4. Adding two-level categorical factors
5. Blocking schemes for DSDs

“Conference Matrix” Approach

An *mxm* square matrix C with 0 diagonal and +1 or -1 off diagonal elements such that:

$$\mathbf{C}^{\mathrm{T}}\mathbf{C} = (m - 1)\mathbf{I}_{m' m}$$

Came from field of telephony.



Conference Matrix of Order 6

$$C = \begin{pmatrix} 0 & +1 & +1 & +1 & +1 & +1 \\ +1 & 0 & +1 & -1 & -1 & +1 \\ +1 & +1 & 0 & +1 & -1 & -1 \\ +1 & -1 & +1 & 0 & +1 & -1 \\ +1 & -1 & -1 & +1 & 0 & +1 \\ +1 & +1 & -1 & -1 & +1 & 0 \end{pmatrix}$$

Here the amazing result (for many even *m*):

Form the augmented matrix:

$$D = \begin{bmatrix} & & & & +C \\ & & & & \\ & & & & -C \\ & & & & 0 \end{bmatrix}$$

…and you get an orthogonal (D-optimal for main effects) definitive screening design! No need to compute**!**

Conference matrix-based DSDs do not exist for odd values of *m*

Feasible design sizes (*n*) are:

Like Plackett-Burman, the designs are available in steps of four, with the exception of *m = 22*.

m	n
6	13
8	17
10	21
12	25
14	29
16	33
18	37
20	41
NA	NA
24	49
26	53
28	57
30	61

What to do if m is odd

DSDs exist for *m* odd, but they are not orthogonal for main effects

For *m* odd:

1. Add one fake factor so that $m' = m + 1$ is even
2. Construct the DSD for $m + 1$ columns
3. Now drop the fake factor column
4. Result is an orthogonal m-factor DSD with $n = 2(m + 1) + 1$

You get an orthogonal design for the price of 2 extra runs

Caveat: On rare occasion you may need to add 4 (or more) runs when m is odd

It may be that there is no conference matrix solution for m + 1 factors

Need to go to the next larger even number ($m + 3$)

Now obtain the orthogonal DSD for $m + 3$ factors

Drop the 3 fake factor columns

Only time this happens for $m < 30$ is $m = 21$. (There is no 22-by-22 conference matrix)

In general, if we add k fake columns to get to the next smallest conference matrix design, the cost is $2k$ runs

JMP uses this to always yield orthogonal DSDs

Example $m = 12,\ n = 2m + 1 = 25$ (Demo)

Example: $m = 7,\ n = 2(m + 1) + 1 = 17$ (Demo)

Example: $m = 21,\ n = 2(m + 3) + 1 = 49$ (Demo)

Power can be increased by adding fake factors

Example

Experimenter has $m = 10$ factors

DSD will require $2m + 1 = 21$ runs

Suppose a power analysis indicated that $n = 32$ runs are required to obtain required power

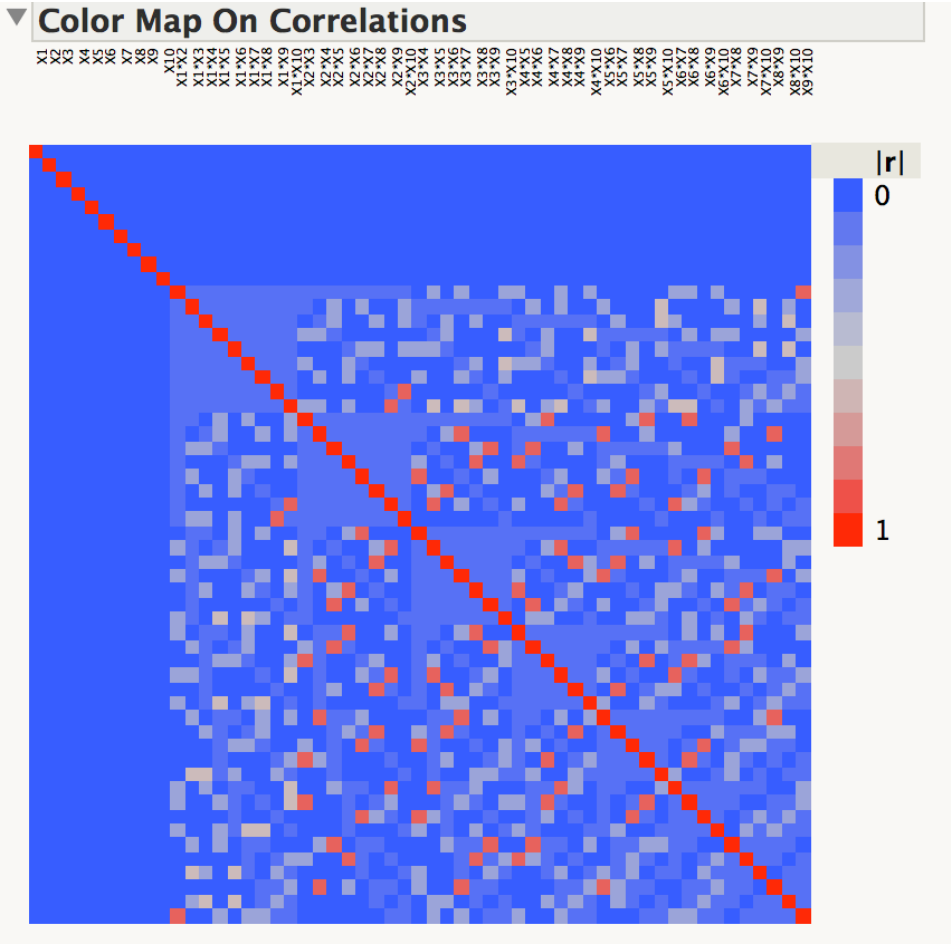
Therefore 11 additional runs are required

Using $k = 6$ fake factors will result in a $2(m+k) + 1 = 33$ run DSD

Create the design based on *16* factors and drop six columns

JMP Demo

33-run 10-factor DSD



Conference Matrix DSDs Summary

Subjects we addressed:

Constructing DSDs using conference matrices

Benefits of “fake” factors

We show cooler benefits in analysis section!

Recommendation: Use conference matrix construction unless you have an odd number of factors and runs are expensive

Tutorial Outline

1. What is a definitive screening design (DSD)?
2. Conference matrix based DSDs
3. Analyzing DSDs
4. Adding two-level categorical factors
5. Blocking schemes for DSDs

Simplest idea – Fit the main (linear) effects model

Advantages:

MEs unbiased – you can believe the coefficient estimates

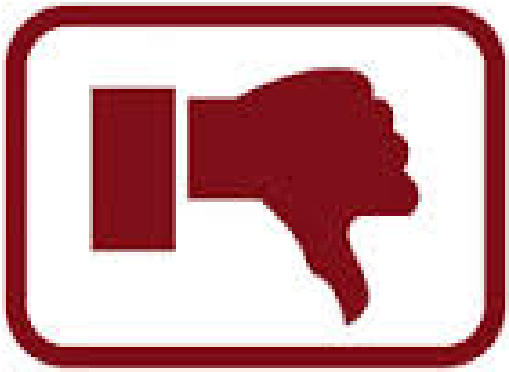
Fit the Main Effects Model

Disadvantages:

Estimate of σ inflated with strong 2FIs or quadratic effects

May make active MEs appear not statistically significant.

You cannot believe the coefficient standard errors.



Analysis Idea #2 – Use Stepwise Regression

Advantages:

Standard analytical tool

Easy to do in JMP

Procedure

Specify the response surface model

Use forward variable selection based on the AICc criterion

The full second order (RSM) model

The response surface model (RSM) is the model consisting of:

1. The intercept term.
2. All main linear effects (for *m* factors, there are m of these)
3. All main quadratic (curvature) effects (*m* of these)
4. All two-factor interactions [there are *m(m-1)/2* of these]

Number of terms in the full RSM:

1 + 2m + m(m-1)/2 = (m+1)(m+2)/2

Example: Six Factor RSM (m = 6)

1. Constant term
2. $m = 6$ main linear effects: $X_1, X_2, X_3, X_4, X_5, X_6$
3. $m = 6$ main quadratic effects: $X_1^2, X_2^2, X_3^2, X_4^2, X_5^2, X_6^2$
4. $m = m(m-1)/2 = 15$ two-factor interactions:

X_1X_2	X_1X_3	X_1X_4	X_1X_5	X_1X_6
	X_2X_3	X_2X_4	X_2X_5	X_2X_6
		X_3X_4	X_3X_5	X_3X_6
			X_4X_5	X_4X_6
				X_5X_6

Total is 1 + 6 + 6 + 15 = 28 model terms

Problem!

The number of model terms (28) is greater than the number of runs in the DSD (13).

Consequences of n = 13:

- We cannot fit the full RSM
- We can estimate at most 13 model terms

One solution: Use Forward Stepwise Selection

Sparsity assumption: not all effects are active

We hope that the number of active effects is substantially fewer than the number of runs.

OK, which effects are active, which are not?

Use a forward stepwise procedure to find out

I prefer minimizing the AICc criterion to decide when to stop. Simulation studies show that it does a better job of finding the active effects when analyzing data with small numbers of runs.

DSD Summary (so far)

Definitive Screening Designs:

1. Easy to generate in JMP
2. Permits screening designs having three levels
3. Main linear effects are independent of (unbiased by) two-factor interactions, should they exist
4. Can fit response surface-type models using stepwise regression (Minimum AICc stopping rule).
5. Possible to screen and optimize in one step

Limitations of DSDs

No design can do everything in one shot. DSDs are no exception.

Limitations include:

1. Stepwise breaks down if there are more than about n/2 active terms in the model

   For example, for six factors, m = 13, if there are more than about 6 active terms, stepwise has difficulty finding the correct model.

   Generally only good if there are just a few two-factor interactions and/or quadratic effects

2. Power is low for finding moderate quadratic effects. The quadratic effect must be large (3 sigma) to have high (>0.9) power.

Addressing the limitations

1. If many terms appear to be active:

   Augment the DSD to identify interactions and quadratic terms.
2. Run a DSD with more than the minimum run size (next)

Published online

Technometrics

Taylor & Francis
Taylor & Francis Group

ISSN: 0040-1706 (Print) 1537-2723 (Online) Journal homepage: http://www.tandfonline.com/loi/utch20

Effective Design-Based Model Selection for Definitive Screening Designs

Bradley Jones & Christopher J. Nachtsheim

To cite this article: Bradley Jones & Christopher J. Nachtsheim (2016): Effective Design-Based Model Selection for Definitive Screening Designs, Technometrics, DOI: 10.1080/00401706.2016.1234979

To link to this article: http://dx.doi.org/10.1080/00401706.2016.1234979

Analysis Idea #3 – New Method

Since main effects and 2nd order effects are orthogonal to each other you can split the response column (Y) into two new columns.

One column for identifying main effects – call it YME

One column for identifying 2nd order effects – call it Y2nd

The two columns are orthogonal to each other and their sum is Y.

Computing the New Responses

1. Fit the main effects model (No Intercept) and save the predicted **values (YME). These are the responses for the main effects model.**
2. **Save the** residuals from the fit above – these residuals are the responses for the 2^{nd} order effects ($Y2^{nd}$).

Digression: Benefits of "Fake" Factors

Adding Fake Factors (factors you don't use) provides a way to estimate variance without repeating center runs!

Why?

Fake factors are orthogonal to the real factors

Fake factors are orthogonal to all the 2nd order effects

Assuming the 3rd and higher order effects are negligible, we can use the fake factor degrees of freedom to create an unbiased estimate of the error variance!

- Note: Use both the real and fake factors when fitting the main effects model in step 1 of the previous slide.

Example: Six real factors and two fake factors

A	B	C	D	E	F	Fake 1	Fake 2	Y	Y2nd	YME
0	1	1	1	1	1	1	1	94.51	101.04	-6.53
0	-1	-1	-1	-1	-1	-1	-1	107.57	101.04	6.53
1	0	1	1	-1	1	-1	-1	94.36	101.175	-6.815
-1	0	-1	-1	1	-1	1	1	107.99	101.175	6.815
1	-1	0	1	1	-1	1	-1	91.80	90.525	1.275
-1	1	0	-1	-1	1	-1	1	89.25	90.525	-1.275
1	-1	-1	0	1	1	-1	1	93.70	94.485	-0.785
-1	1	1	0	-1	-1	1	-1	95.27	94.485	0.785
1	1	-1	-1	0	1	1	-1	89.55	88.71	0.84
-1	-1	1	1	0	-1	-1	1	87.87	88.71	-0.84
1	-1	1	-1	-1	0	1	1	94.58	95.235	-0.655
-1	1	-1	1	1	0	-1	-1	95.89	95.235	0.655
1	1	-1	1	-1	-1	0	1	93.23	89.58	3.65
-1	-1	1	-1	1	1	0	-1	85.93	89.58	-3.65
1	1	1	-1	1	-1	-1	0	98.11	95.815	2.295
-1	-1	-1	1	-1	1	1	0	93.52	95.815	-2.295
0	0	0	0	0	0	0	0	99.75	99.75	0

Adds 4 runs – 2 error df

Example Column Correlations

Correlations

	Y	Y2nd	YME
Y	1.0000	0.7828	0.6223
Y2nd	0.7828	1.0000	0.0000
YME	0.6223	0.0000	1.0000

Note that the two new responses are orthogonal to each other.

YME
-6.53
6.53
-6.815
6.815
1.275
-1.275
-0.785
0.785
0.84
-0.84
-0.655
0.655
3.65
-3.65
2.295
-2.295
0

Examining the Main Effects Response (YME)

Note responses for each foldover pair sum to zero.

The response for the center run is zero.

There are 17 rows but only 8 independent values

(degrees of freedom – df)

There are 6 real factors but 8 df, so there are

$8 - 6 = 2$ df for estimating σ^2

Y2nd
101.04
101.04
101.175
101.175
90.525
90.525
94.485
94.485
88.71
88.71
95.235
95.235
89.58
89.58
95.815
95.815
99.75

Examining the 2nd Order Response (Y2nd)

Responses for each foldover pair are the same.

There are 17 rows but only 9 independent values

(degrees of freedom – df)

After estimating the Intercept, there are 8 df left for estimating 2nd order effects.

Analysis – Identify Active Main Effects

1. Recall that the residuals from fitting the Main Effects data to the real factors have 2 degrees of freedom.
2. To estimate σ^2 , sum the squared residuals from this fit and divide the result by 2.
3. Using this estimate, do t-tests of each coefficient
4. If the resulting p-value for an effect is small (<0.05 say), conclude that effect is active.

Digression: Model Heredity Assumption

The heredity assumption stipulates that 2nd order effects only occur when the associated main effects are active.

Example 1: If main effects A and B are in the model you can consider the two-factor interaction AB

Example 2: B must be in the model before considering the quadratic effect B^2

While there is no physical law requiring that models exhibit heredity, there is empirical evidence that such models are much more probable in real systems.

Advantage of the Heredity Assumption

The set of possible models using the heredity assumption may be much smaller than allowing any 2^{nd} order effect to appear in the model

Example: Suppose your main effects analysis yields 3 active main effects (C, D, F say). Then the allowable 2^{nd} order terms are CD, CF, DF, C^2, D^2, F^2

We have 8 degrees of freedom and only 6 effects, so it is possible to identify all 6 if they are active.

If we allow consideration all 2^{nd} order effects, there are 15 two-factor interactions and 6 quadratic terms – or 21 terms in all.

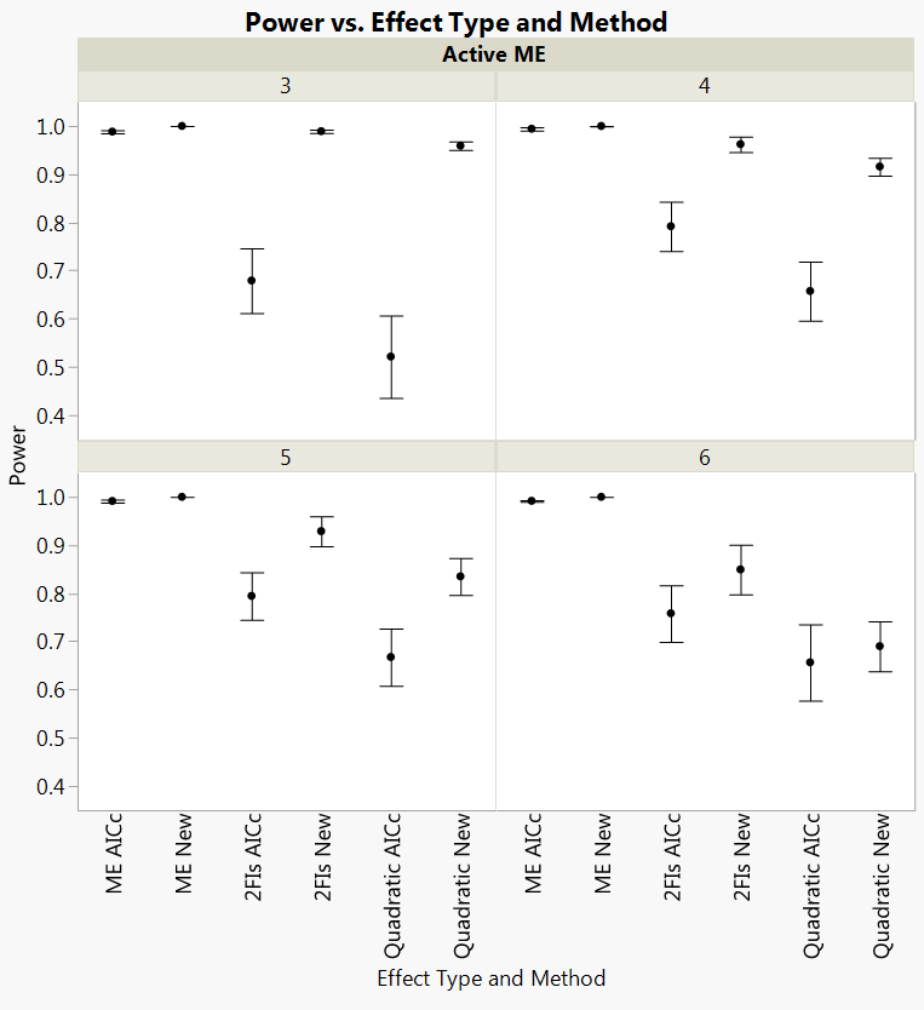
There are 2^{21} or more 2 million possible models – a much harder model selection problem.

Analysis – Identifying 2nd Order Effects

Form all the 2nd order terms involving the active main effects

Do all subsets regression up to the point where the MSE of the best 2nd order model for a given number of terms is not significantly larger than your estimate of σ^2

Simulation Comparisons New Method vs. Stepwise



Comparison for DSD with 6 factors and 17 runs (i.e. 2 fake factors)

Power for detecting 2FIs and Quadratic effects is much higher for the new method especially when fewer MEs are active

Analyzing DSDs Conclusion

Stepwise with AICc works adequately if there are few active 2nd order effects (one or two 2FIs and/or one quadratic effect)

The new method is now the default method in JMP 13 because it performs better than any existing alternative analytical procedure I know.

Tutorial Outline

1. What is a definitive screening design (DSD)?
2. Conference matrix based DSDs
3. Analyzing DSDs
4. Adding two-level categorical factors
5. Blocking schemes for DSDs

DSDs with two-level categorical factors

1. Two-level factors: why is this a problem?
2. Conference matrix construction method
3. A real example

Many design problems involve categorical factors

Examples:

Two operators

Two production lines

Drug and placebo

Two catalysts

Two machines

Two plastic materials

Etc., etc.

Building DSDs with added two-level categorical factors

1. Start by repeating the steps for creating a design for continuous factors only except add two rows of zeros instead of one.

2. Change the zeros in the columns for the categorical factors to +1 or– two cases

    a) If the zeros are the added rows of zeros make all the categorical factors –1 for the 1st row and +1 for the 2nd row.

    b) If the zeros are from the conference matrix and its fold over, make the factor –1 for the 1st row and +1 for the 2nd row.

Example – 4 continuous and 2 categorical factors

Conference Matrix with fold over appended plus two zero rows.

A	B	C	D	E	F
0	+1	+1	+1	+1	+1
+1	0	+1	–1	–1	+1
+1	+1	0	+1	–1	–1
+1	–1	+1	0	+1	–1
+1	–1	–1	+1	0	+1
+1	+1	–1	–1	+1	0
0	–1	–1	–1	–1	–1
–1	0	–1	+1	+1	–1
–1	–1	0	–1	+1	+1
–1	+1	–1	0	–1	+1
–1	+1	+1	–1	0	–1
–1	–1	+1	+1	–1	0
0	0	0	0	0	0
0	0	0	0	0	0

Example *continued…*

Changed categorical factors (E & F) with zero entries.

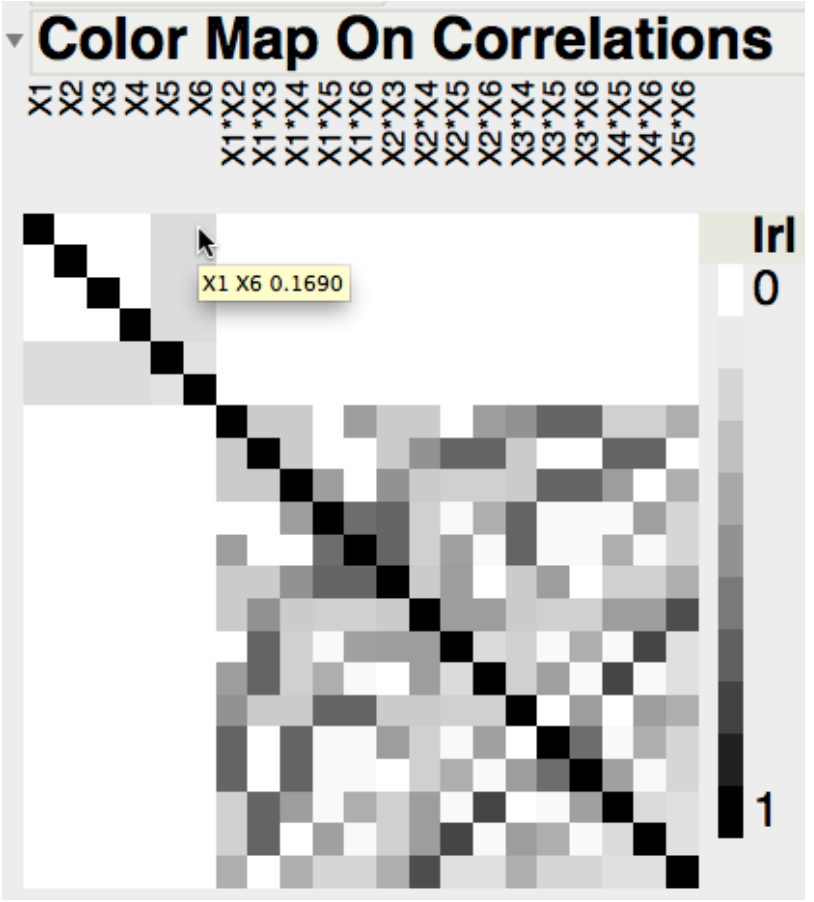
A	B	C	D	E	F
0	+1	+1	+1	+1	+1
+1	0	+1	–1	–1	+1
+1	+1	0	+1	–1	–1
+1	–1	+1	0	+1	–1
+1	–1	–1	+1	–1	+1
+1	+1	–1	–1	+1	–1
0	–1	–1	–1	–1	–1
–1	0	–1	+1	+1	–1
–1	–1	0	–1	+1	+1
–1	+1	–1	0	–1	+1
–1	+1	+1	–1	+1	–1
–1	–1	+1	+1	–1	+1
0	0	0	0	–1	–1
0	0	0	0	+1	+1

Statistical Properties – added two-level categorical factors

1. Main effects of categorical factors have small correlations with other factors.
2. The number of required runs is only two more than twice the number of rows in the conference matrix chosen.
3. Main effects are independent of two-factor interactions.
4. Two-factor interactions are not completely confounded with other two-factor interactions, although they may be correlated
5. All quadratic effects are estimable in models comprised of any number of linear and quadratic main effects terms.
6. Quadratic effects are orthogonal to main effects and not completely confounded (though correlated) with interaction effects.

The “OK” news is that the design is not completely orthogonal for main effects—here m = 4 and c = 2

Two-level factors are correlated with each other and with the continuous factors



Example: Extracting food solids from peanuts in solution, n = 18

	Factor	Low	High
1	Water pH level	6.95	8.0
2	Water temp	20C	60C
3	Extraction time	15	40
4	Water-Peanuts Ratio	5	9
5	Agitation speed	5,000	10,000
6	Hydrolyzed?	N	Y
7	Presoaking?	N	Y

Classical Approach: 2^{7-3} Resolution IV FF

Problem:

Seven two-level factors in 16 runs.

Main effects are critical

Partial information about two-factor interactions is desirable.

Standard approach: Resolution IV fractional factorial design

JMP Results: Ambiguity!

Contrasts

Term	Contrast		Lenth t-Ratio	Individual p-Value	Simultaneous p-Value	Aliases
Ratio	-1.95810		-11.36	<.0001*	0.0004*	Pre-soak?*pH*Water Temp
Pre-soak?	-0.76222		-4.42	0.0041*	0.0396*	
Agitation Speed	0.68295		3.96	0.0064*	0.0598	
Hydrolyze?	-0.41454		-2.41	0.0319*	0.3035	
pH	0.11962		0.69	0.4718	1.0000	Ratio*Pre-soak?*Water Temp
Water Temp	0.00295		0.02	0.9888	1.0000	Ratio*Pre-soak?*pH
Extraction Time	0.00113		0.01	0.9959	1.0000	
Ratio*Ratio	0.01944	*	0.11	0.9192	1.0000	
Ratio*Pre-soak?	0.33411		1.94	0.0697	0.5543	pH*Water Temp
Ratio*Agitation Speed	0.11019		0.64	0.5591	1.0000	
Pre-soak?*Agitation Speed	-0.17147		-0.99	0.3061	0.9963	
Ratio*Hydrolyze?	0.29993	*	1.74	0.0922	0.6848	Agitation Speed*pH
Pre-soak?*Hydrolyze?	-0.06187	*	-0.36	0.7398	1.0000	Agitation Speed*Water Temp
Agitation Speed*Hydrolyze?	-0.29757	*	-1.73	0.0942	0.6955	Ratio*pH, Pre-soak?*Water Temp
Pre-soak?*pH	0.01945		0.11	0.9192	1.0000	Ratio*Water Temp
Pre-soak?*Extraction Time	0.14833	*	0.86	0.3662	0.9997	
Ratio*Pre-soak?*Hydrolyze?	-0.10430	*	-0.61	0.5814	1.0000	Pre-soak?*Agitation Speed*pH, Ratio*Agitation Speed*Water Temp

Ratio x Presoak = pH x Temp

Ratio x Hydrollze = AgitationSpeed x pH

AgitSpeed x Hydrolize = Ratio x pH = PreSoak x Temp

Now redesign experiment as DSD with five continuous and two categorical factors, n = 18

		Factor	Low	High
5 continuous factors	1	Water pH level	6.95	8.0
	2	Water temp	20C	60C
	3	Extraction time	15	40
	4	Water-Peanuts Ratio	5	9
	5	Agitation speed	5,000	10,000
2 categorical factors	6	Hydrolyzed?	N	Y
	7	Presoaking?	N	Y

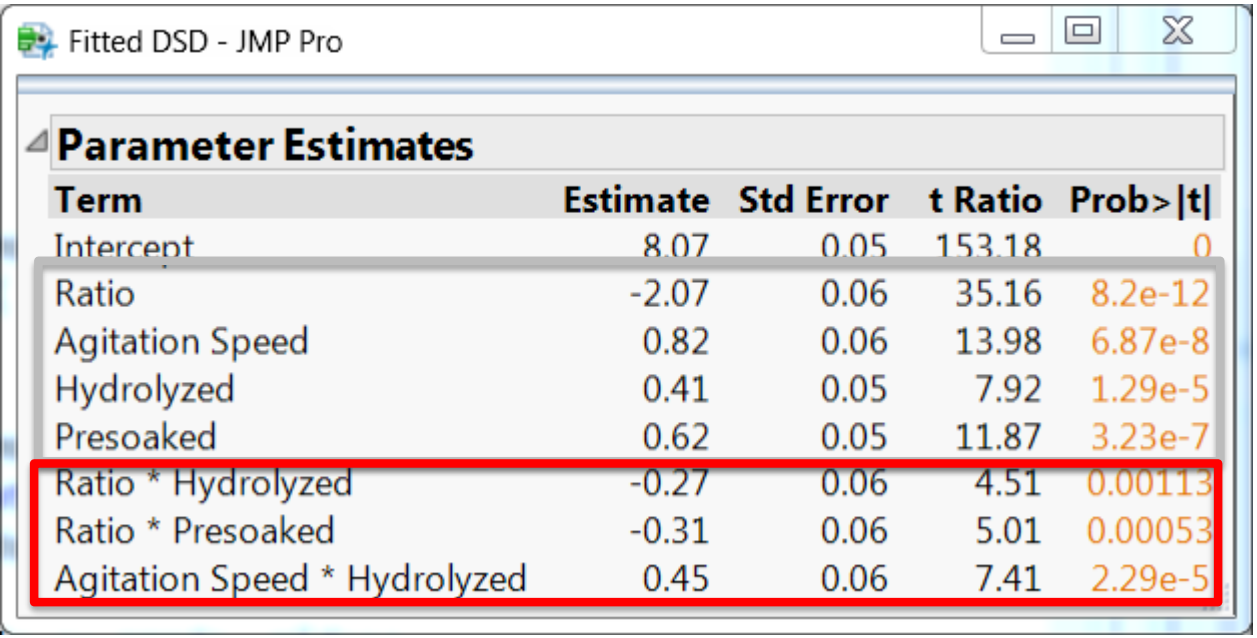
Can a Mixed-Level DSD do Better?

Recall our conclusion: There are three active interaction effects, but cannot identify them!

Using the same number of runs (18) the definitive screening design will not only identify the same main effects, but it will:

- Investigate all curvatures (concluding there are none)
- Investigate all interactions and definitively identify the three active interactions

The new method finds the right model.



The four main effects ✓

The three 2fis ✓

Tutorial Outline

1. What is a definitive screening design (DSD)?
2. Analyzing DSDs
3. Conference matrix based DSDs
4. Adding two-level categorical factors
5. Blocking schemes for DSDs

Blocking Schemes for DSDs

1. Simple solution
2. Implementation in JMP
3. Med-tech experiment

CrossMark click for updates

Blocking Schemes for Definitive Screening Designs

DOI: 10.1080/00401706.2015.1013777

Bradley Jones[a] & Christopher J. Nachtsheim[b]

Publishing models and article dates explained

Received: 16 Jul 2013
Accepted: 25 Jan 2015
Accepted author version posted online: 11 Mar 2015

Building DSDs with orthogonal blocks

1. Start by repeating the steps for creating a design for continuous factors only except add as many rows of zeros as there are blocks. Put the design into standard order.
2. Assign the first fold over pair to the 1st block, the 2nd fold over pair to the 2nd block and so on until you get to the last block, then start over again assigning the next fold over pair to the 1st block.
3. Assign each center run to a separate block.

Example – six factors and two blocks

DSD in Standard Order.

A	B	C	D	E	F
0	+1	+1	+1	+1	+1
0	–1	–1	–1	–1	–1
+1	0	–1	+1	+1	–1
–1	0	+1	–1	–1	+1
+1	–1	0	–1	+1	+1
–1	+1	0	+1	–1	–1
+1	+1	–1	0	–1	+1
–1	–1	+1	0	+1	–1
+1	+1	+1	–1	0	–1
–1	–1	–1	+1	0	+1
+1	–1	+1	+1	–1	0
–1	+1	–1	–1	+1	0
0	0	0	0	0	0
0	0	0	0	0	0

Blocking Example 1 – six factors and two blocks

Blocked DSD in Standard Order.

Block	A	B	C	D	E	F
1	0	+1	+1	+1	+1	+1
1	0	–1	–1	–1	–1	–1
2	+1	0	–1	+1	+1	–1
2	–1	0	+1	–1	–1	+1
1	+1	–1	0	–1	+1	+1
1	–1	+1	0	+1	–1	–1
2	+1	+1	–1	0	–1	+1
2	–1	–1	+1	0	+1	–1
1	+1	+1	+1	–1	0	–1
1	–1	–1	–1	+1	0	+1
2	+1	–1	+1	+1	–1	0
2	–1	+1	–1	–1	+1	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0

Blocking Example 2 – six factors and three blocks

Block	A	B	C	D	E	F
1	0	+1	+1	+1	+1	+1
1	0	–1	–1	–1	–1	–1
2	+1	0	–1	+1	+1	–1
2	–1	0	+1	–1	–1	+1
3	+1	–1	0	–1	+1	+1
3	–1	+1	0	+1	–1	–1
1	+1	+1	–1	0	–1	+1
1	–1	–1	+1	0	+1	–1
2	+1	+1	+1	–1	0	–1
2	–1	–1	–1	+1	0	+1
3	+1	–1	+1	+1	–1	0
3	–1	+1	–1	–1	+1	0
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	0	0	0	0	0	0

Orthogonal blocking schemes are flexible!!!

1. Number of blocks can range from 2 to m
2. Example: 8 factor DSD
   - Minimum: 2 blocks
   - Maximum: 8 blocks
3. Recommended: add a zero row to every block to ensure estimates of quadratic effects

m is the number of factors

m	n	B	Blocksizes----------						
4	10	2	5	5					
4	11	3	5	3	3				
4	12	4	3	3	3	3			
6 or (5)	14 (13)	2	7 (6)	7					
6 or (5)	15 (14)	3	5 (4)	5	5				
6 or (5)	16 (15)	4	5 (4)	5	3	3			
6 or (5)	17 (16)	5	5 (4)	3	3	3	3		
6 or (5)	18 (17)	6	3 (2)	3	3	3	3	3	
8 or (7)	18 (17)	2	9 (8)	9					
8 or (7)	19 (18)	3	7 (6)	7	5				
8 or (7)	20 (19)	4	5 (4)	5	5	5			
8 or (7)	21 (20)	5	5 (4)	5	5	3	3		
8 or (7)	22 (21)	6	5 (4)	5	3	3	3	3	
8 or (7)	23 (22)	7	5 (4)	3	3	3	3	3	3
8 or (7)	24 (23)	8	3 (2)	3	3	3	3	3	3	3

Blocking Summary

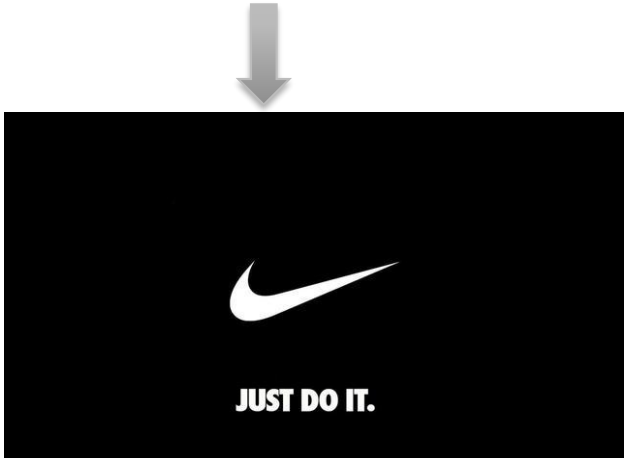
1. Developed blocking schemes for definitive screening designs with or without added categorical factors.
2. Block effects are orthogonal to main (linear) effects.
3. Schemes are extremely flexible, numbers of blocks can range from two up to the number of factors.
4. Illustrated use in the laser etch experiment and Med Tech experiment, the latter having random blocks

Tutorial Summary

Subjects addressed:

1. Screening Design Overview
2. Introduction to DSDs
3. Analyzing DSDs
4. DSDs using conference matrices
5. Adding Categorical Factors to DSDs
6. Blocking DSDs

Recommendation



DSDs provide features that make them best in class for screening.