What tools do you currently use?

1. How many have run an experimental design before?
2. How many have not?
3. How many use textbooks to create their experimental designs?
4. How many use software?
5. Which software do you use?
6. How many use open-source code to create their designs?
7. How many write their own code to create designs?

Modern Screening Design of Experiments

Get More Information from Fewer Trials

JMP Corporate

Topics

1. Review of *Screening* DOE vs. *Response Surface* DOE
2. *Custom DOE* platform will nearly always give a useable design,
   BUT in certain design situations, you can do better…
3. Discussion of *Definitive Screening Designs* (DSD) vs. classic designs like *Plackett-Burman* (PB) and *Fractional Factorial* (FF)
4. When you have continuous and **categorical factors at just 2 levels**, consider creating *Orthogonal Mixed-Level* (OML) design instead of DSD.
5. When you have continuous and **categorical factors at greater than 2 levels**, consider creating *Orthogonal Arrays* (OA), or *Nearly Orthogonal Arrays* (NOA)
6. When you have many factors (e.g., > 20) consider using *Group Orthogonal SuperSaturated Designs (*GOSSD)
7. All screening designs can be *augmented* into RSM designs.
   This is not always practical if factor ranges are changed.

Don’t Let All the Design Choices Confuse You (Pg. 1)

1. *Custom DOE* will nearly always give you a useable design – it can:
   a. Handle all types of factors
   b. Use categorical or discrete numeric factors at any number of levels
   c. Use inequality constraints and/or disallowed combinations
   d. Perform augmentation (even with constraints to repair broken designs)

2. *Custom DOE* is NOT always an appropriate choice for creating Space-Filling designs for use with complex computer simulation experiments.

3. If you can use a ***Definitive Screening Design***, it is recommended you do so
   a. If only a few factors are active it may collapse to support a response-surface model
   b. Often no more work than classical screening designs that don’t test for curvature
   c. Can’t use with categorical factors with > 2 levels, or with constraints, or HTC factors
   d. *Easy DOE* creates DSDs when appropriate under model choice 2

Don't Let All the Design Choices Confuse You (pg. 2)

Designs on the Path: *DOE > Classical > Screening Design > XXX*

4. When you have continuous and categorical factors at just 2 levels, consider creating *Orthogonal Mixed-Level* (OML) design instead of a *Definitive Screening Design* (DSD)

5. When you have continuous & categorical factors at greater than 2 levels, consider creating *Orthogonal Arrays* (OA), or *Nearly Orthogonal Arrays* (NOA) instead of *Custom DOE (D-optimal)*

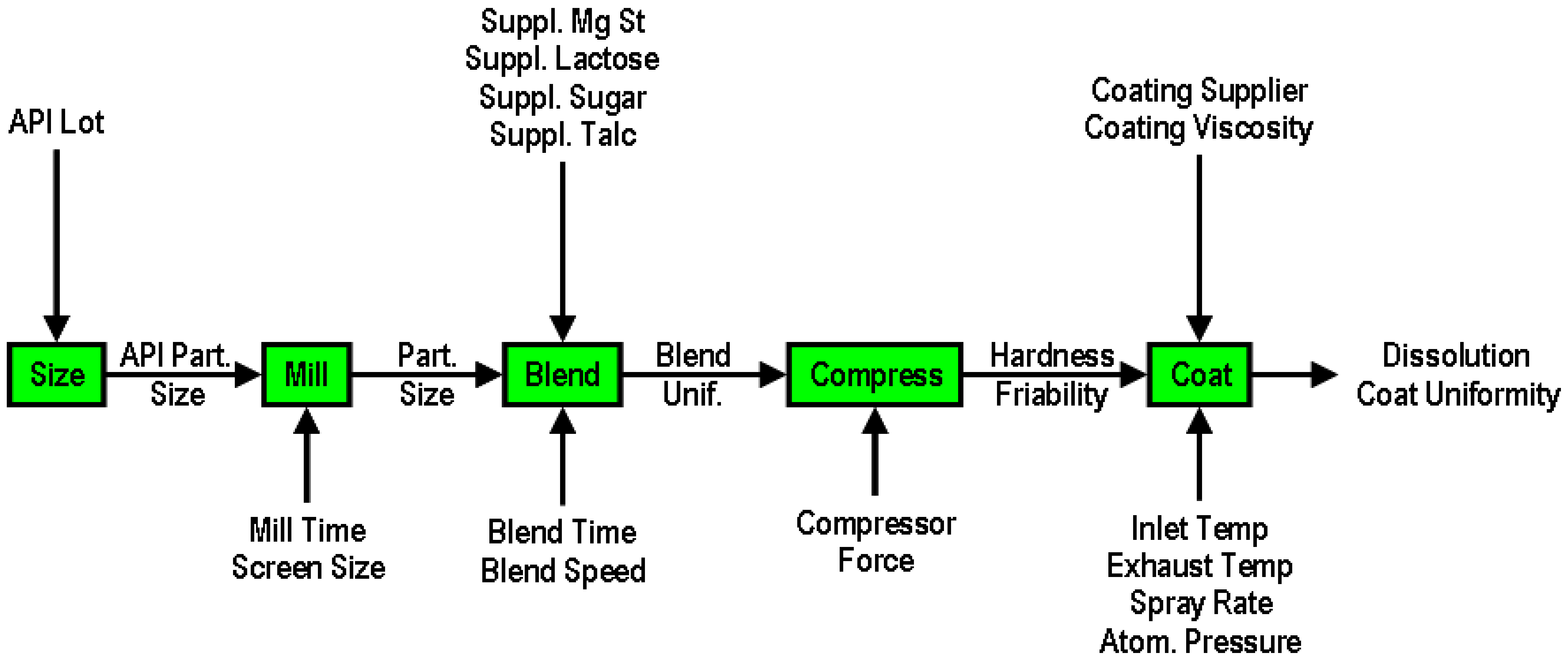
6. Want classic designs like *Plackett-Burman* or *Fractional Factorial*? In many cases, *Custom DOE* and *Easy DOE* create these designs when appropriate.

Designs on the Path: *DOE > Special Purpose > Group Orthogonal Screening*

7. More than 20 factors? Consider *Group Orthogonal SuperSaturated Designs* instead of other screening methods for continuous & 2-level categorical factors.

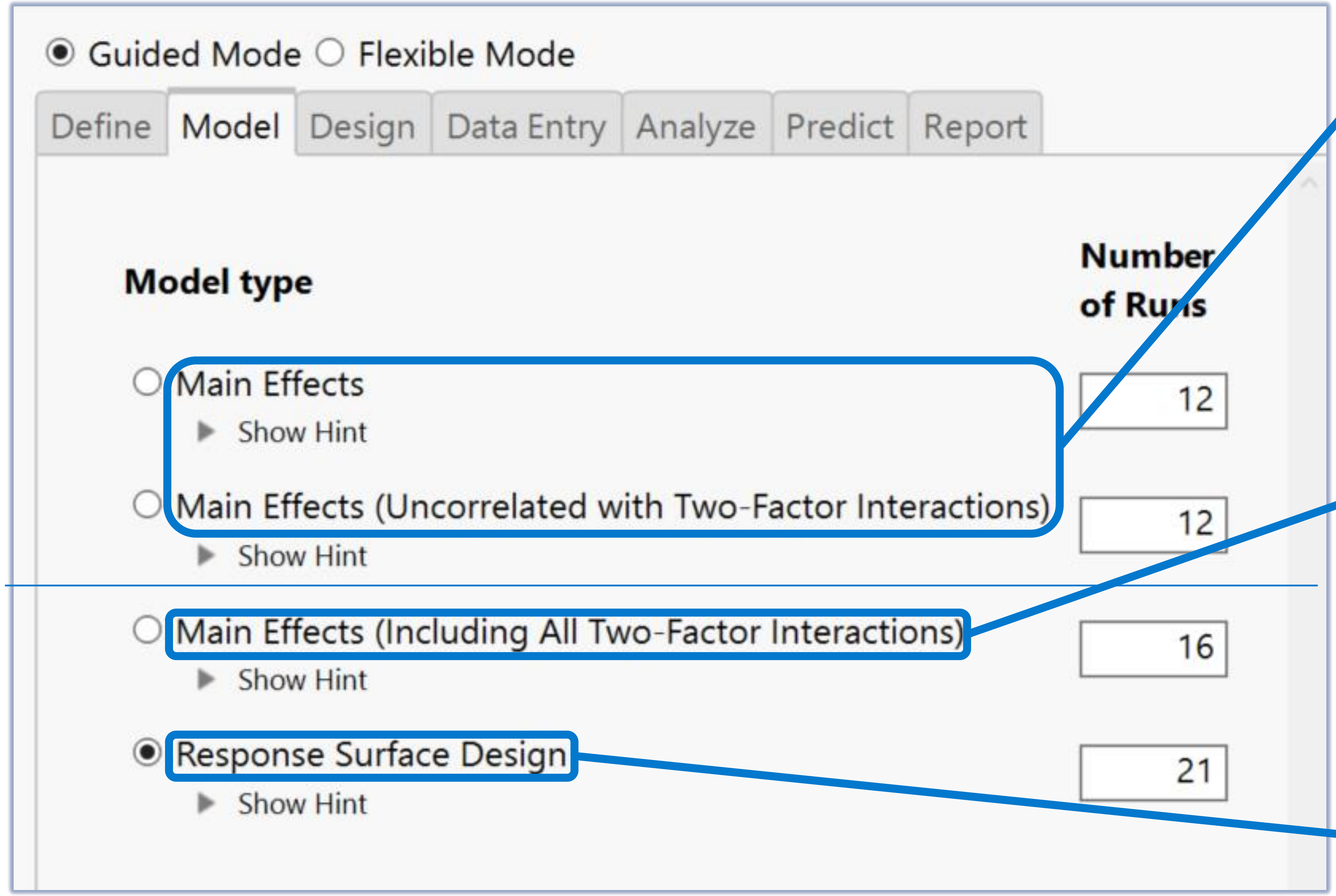
Classic Definition of DOE

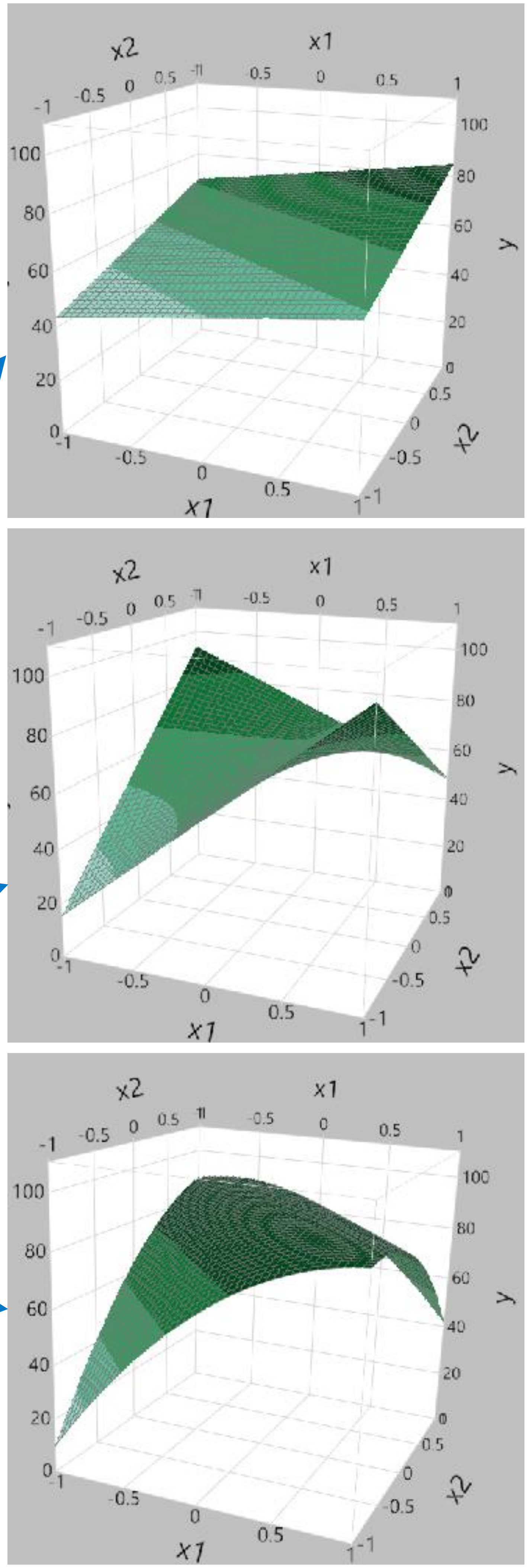
Purposeful control of the inputs (factors) in such a way as to deduce their relationships (if any) with the output (responses).

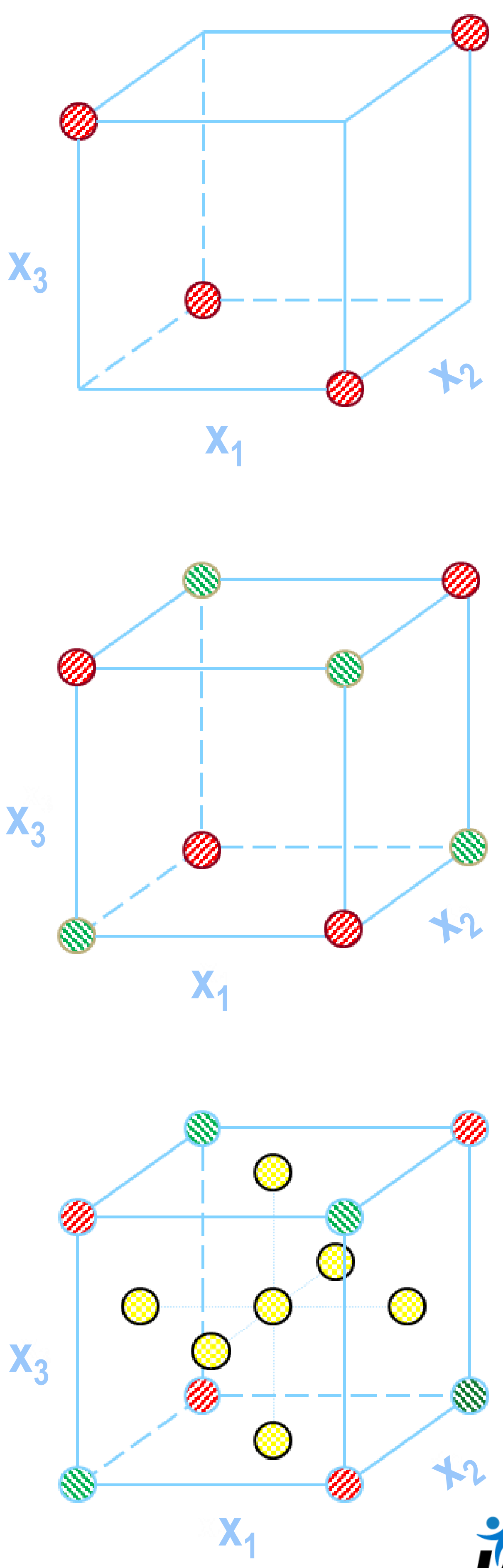


Alternative Definition of DOE

A DOE is a specific collection of trials run to support a *proposed* model.







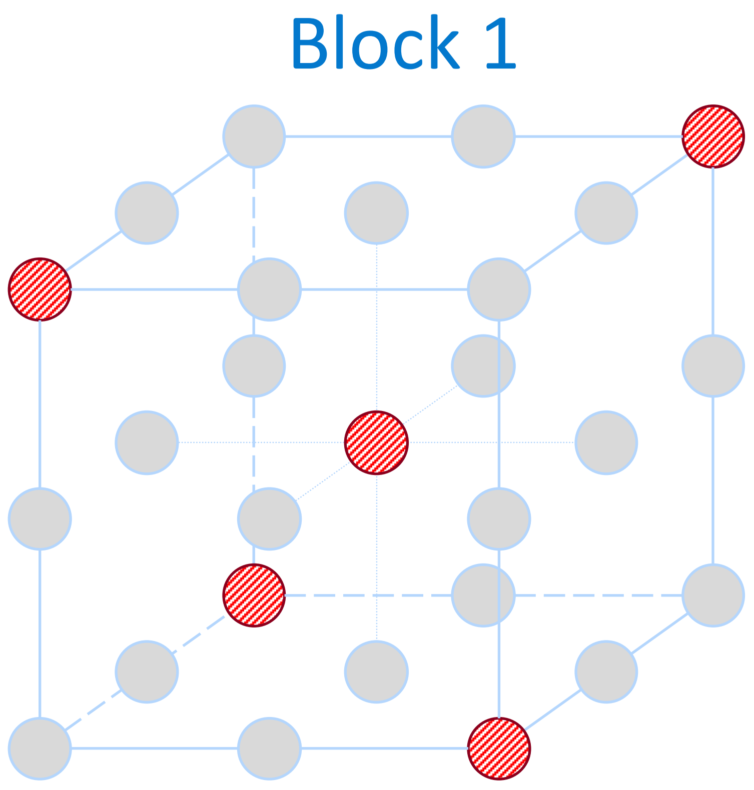
As complexity to be supported increases, so do the number of runs

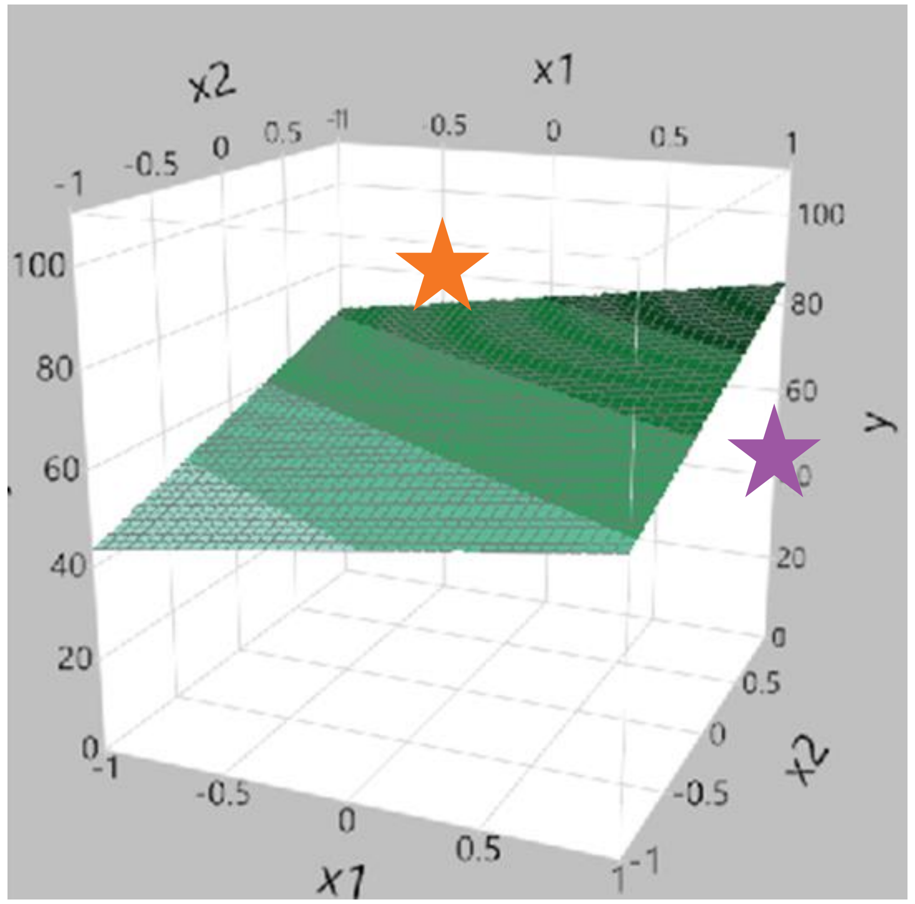
Alternative Definition of DOE

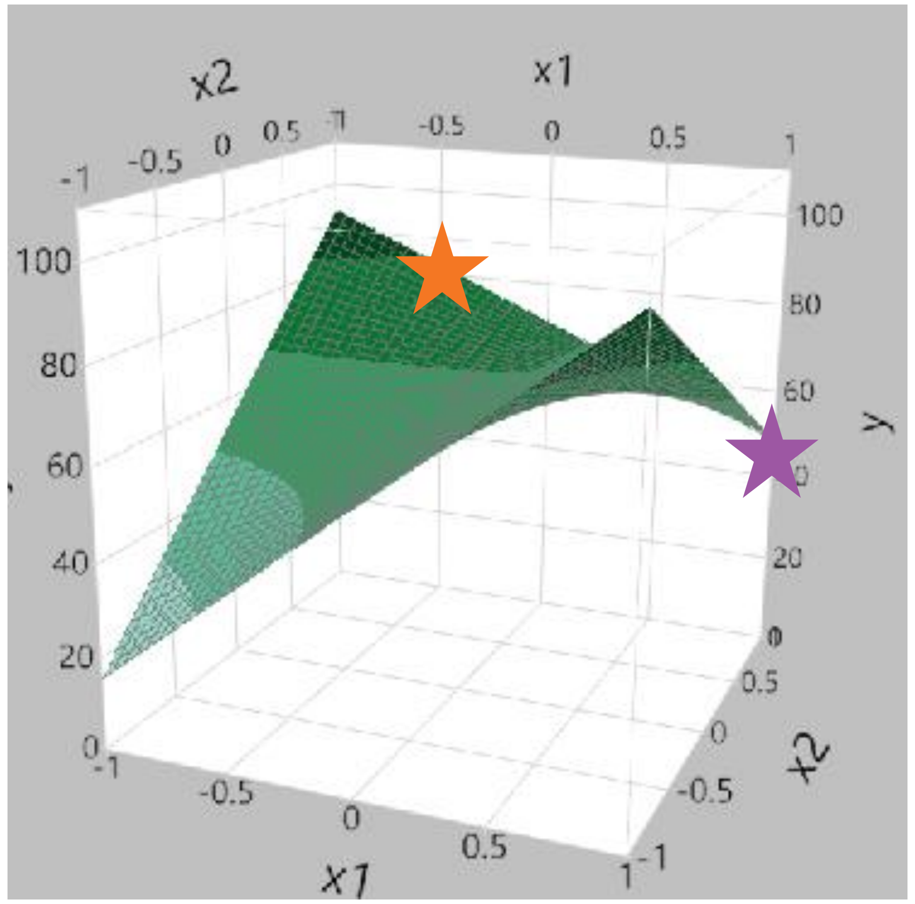
A DOE is a specific collection of trials run to support a *proposed* model.

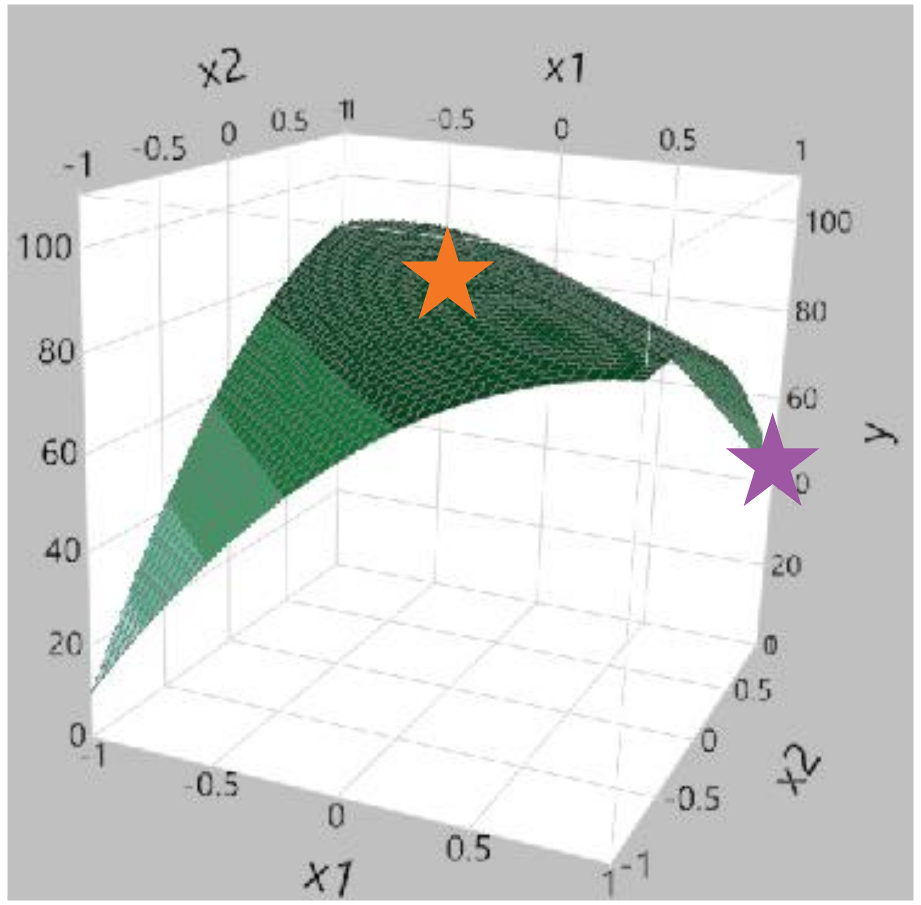
- If proposed model is ***simple*** **-** main effects or *1st **order*** terms (x_1, x_2, x_3, etc.) - the design is called a *__screening__* DOE
  - Goals include **ranking factor importance**, or **"finding a winner"** **quickly, but NOT making predictions**
  - Used with many (> 6?) factors at start of process characterization
- If the proposed model is ***more complex***, with *2nd **order*** terms so that it includes two-way interaction terms (x_1x_2, x_1x_3, x_2x_3, etc.) and in the case of continuous factors, squared terms (x_1^2, x_2^2, x_3^2, etc.), the design is called a *__response-surface__* DOE
  - Goals include **developing a predictive model** of the process, or **characterizing the operating region** of the design space
  - Used with a few (< 6?) factors often after a screening DOE

Classic Response-Surface DOE in a Nutshell



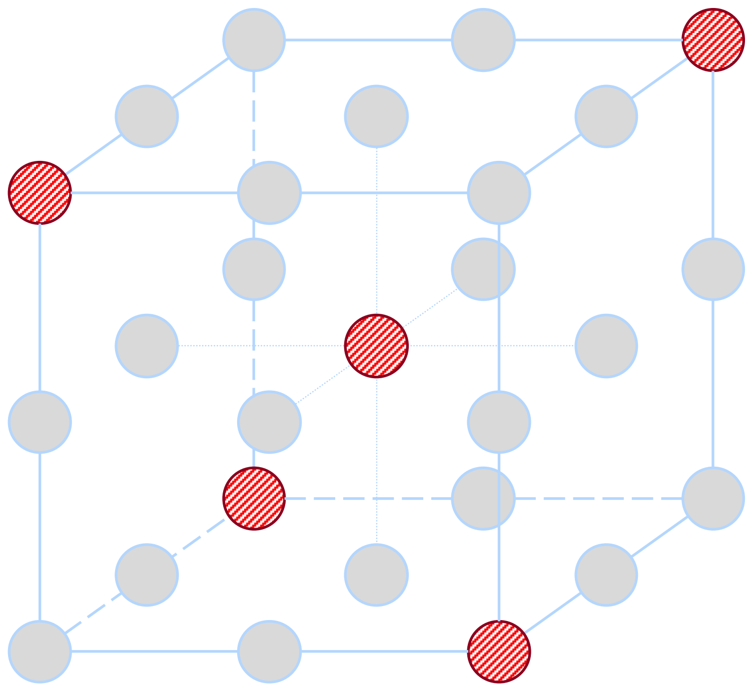






Polynomial Models Used to Calculate Surfaces

Block 1

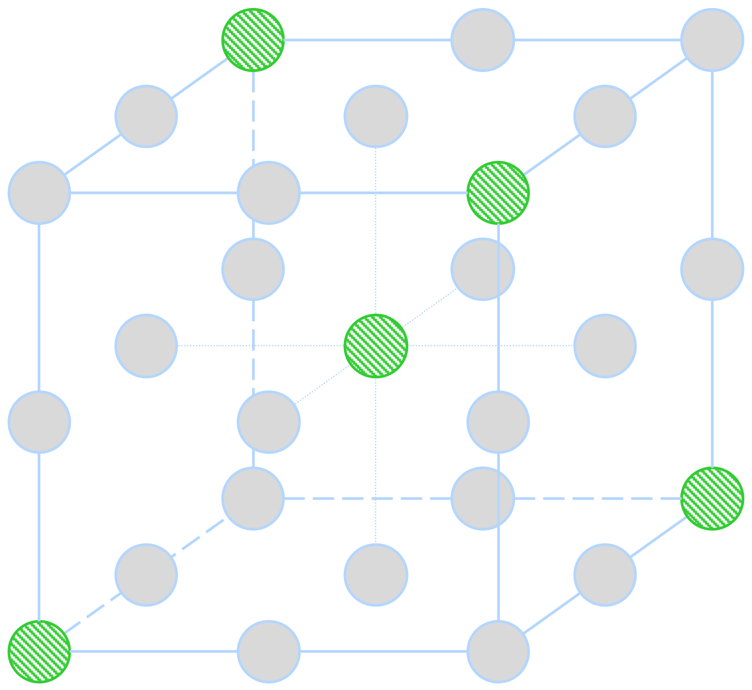


$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$

Run this block 1st to:

(i) estimate the main effects
(ii) use center point to check for curvature.

Block 2

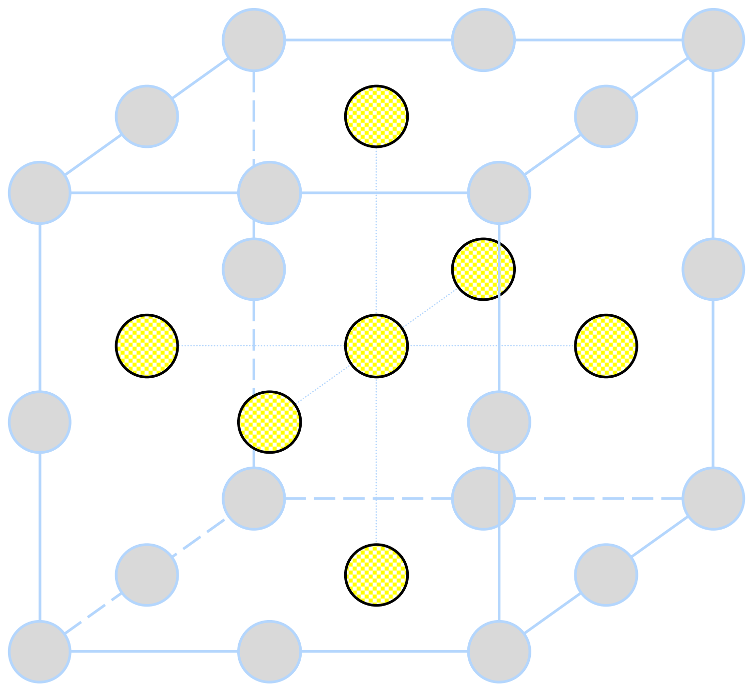


$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$
$$+ a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$$

Run this block 2nd to:

(i) repeat main effects estimate
(ii) check if process has shifted
(iii) add interaction effects to model <u>if needed.</u>

Block 3

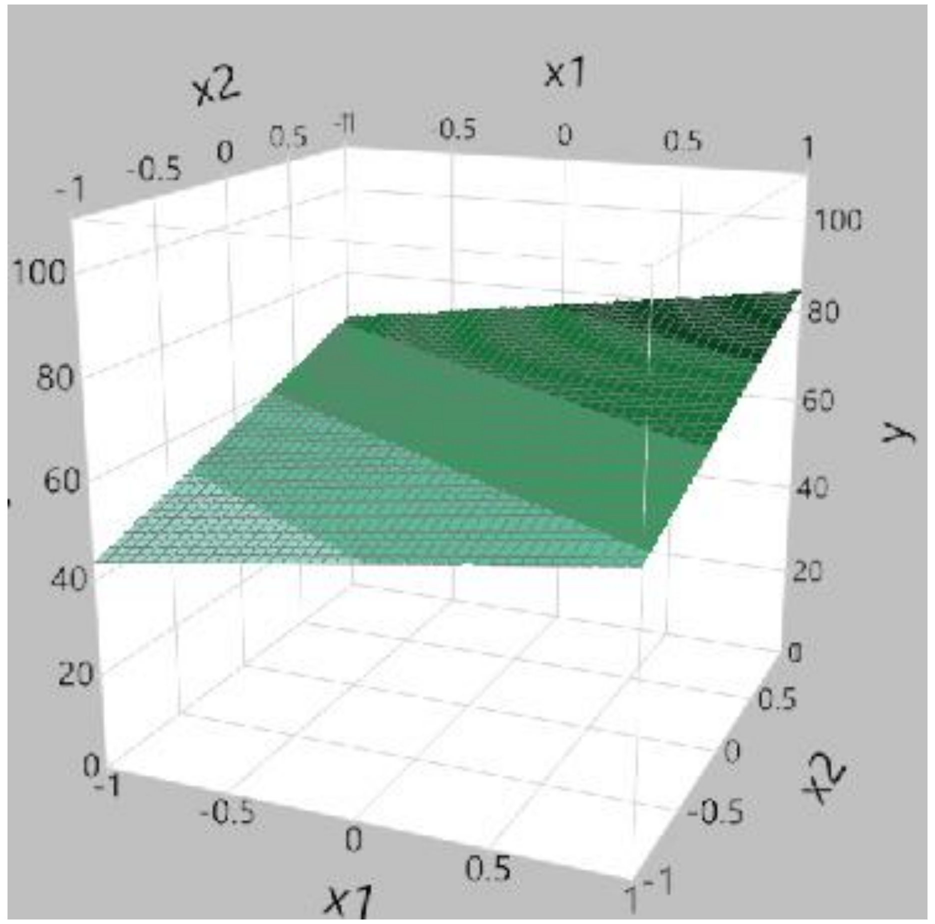


$$y = a_0 + a_1x_1 + a_2x_2 + a_3x_3$$
$$+ a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$$
$$+ a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2$$

Run this block 3rd to:

(i) repeat main and interaction effects estimate
(ii) check if process has shifted
(iii) add curvature effects to model <u>if needed.</u>

Quadratic model is not much bigger than *Interaction* model.
If you have continuous factors, choose full 2nd order, *Quadratic*.

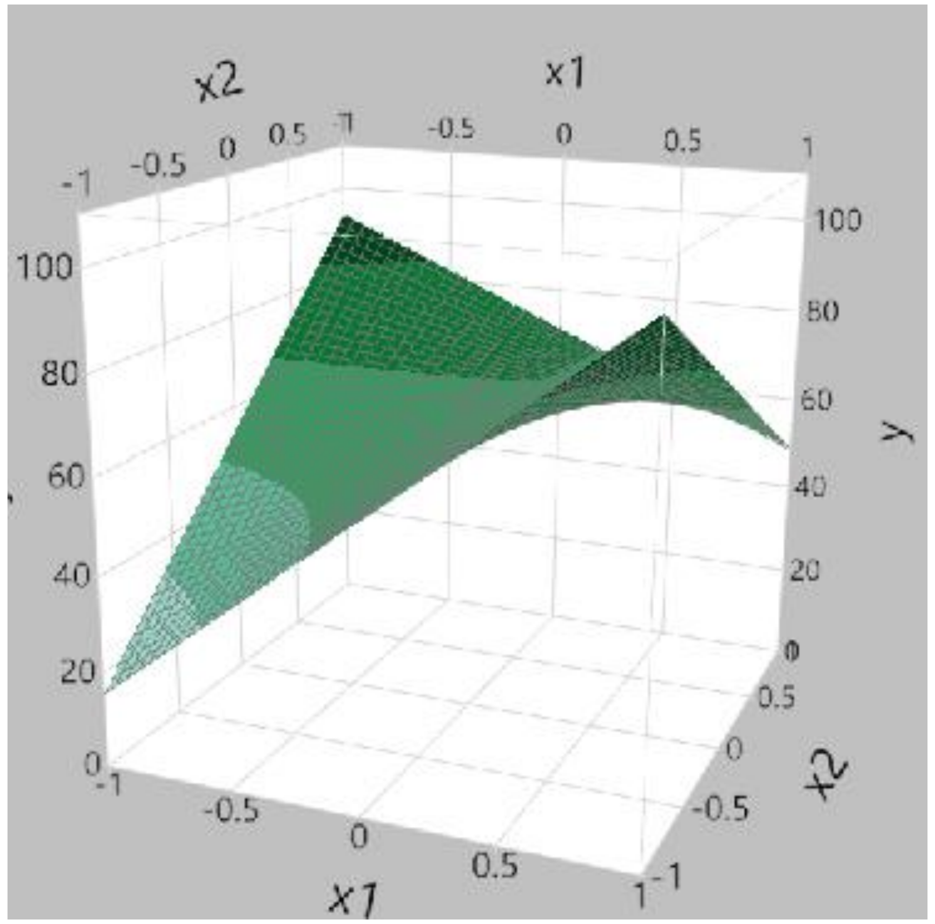


1st Order

$y = a_0 + a_1x_1 + a_2x_2$

For k factors there are
k main effects

3-factor Linear Model has 4 terms (8 corners)
6-factor Linear Model has 7 terms (64 corners)
10-factor Linear Model has 11 terms (1K corners)
20-factor Linear Model has 21 terms (1M corners)



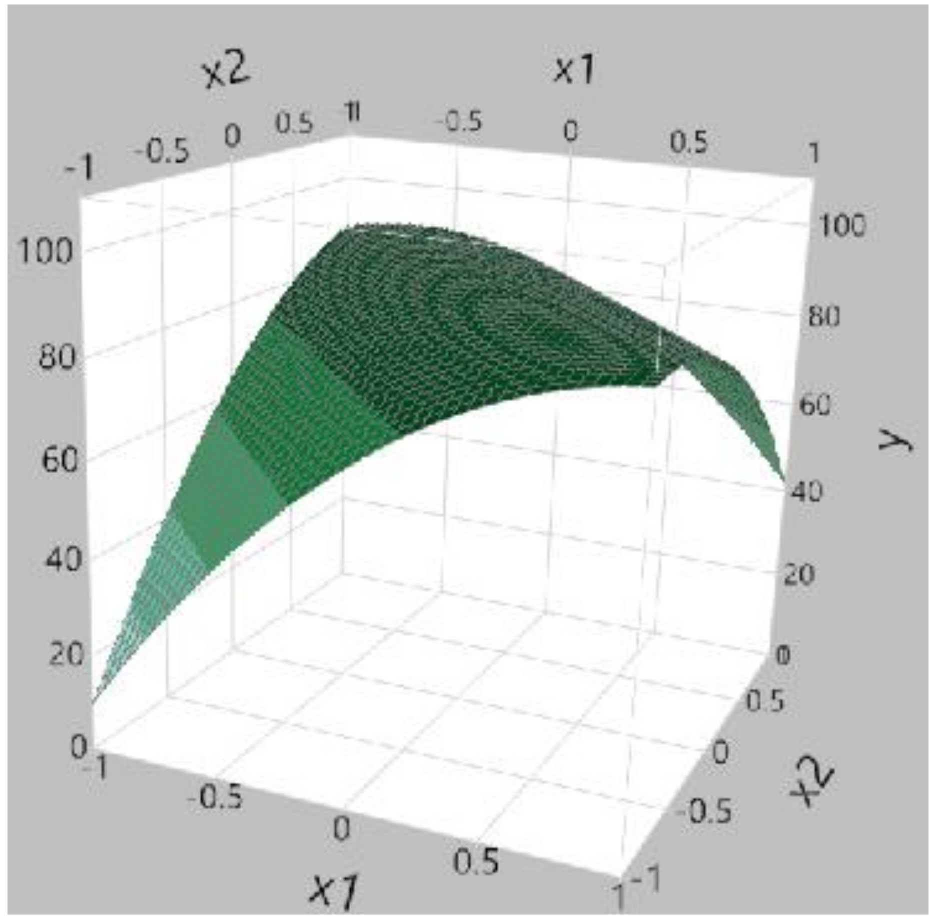
2nd Order

$y = a_0 + a_1x_1 + a_2x_2$

$+ a_{12}x_1x_2$

For k factors there are
k(k-1)/2 interaction effects

3-f Interaction Model has 7 terms (2X ME)
6-f Interaction Model has 22 terms (3X ME)
10-f Interaction Model has 56 terms (5X ME)
20-f Interaction Model has 211 terms (10X ME)



Full 2nd Order

$y = a_0 + a_1x_1 + a_2x_2$

$+ a_{12}x_1x_2$

$+ a_{11}x_1^2 + a_{22}x_2^2$

For k factors there are
k squared effects

3-f Quadratic Model has 10 terms (2.5X ME)
6-f Quadratic Model has 28 terms (4X ME)
10-f Quadratic Model has 66 terms (6X ME)
20-f Quadratic Model has 231 terms (11X ME)

If no squared terms, then optimum can ONLY be a corner!

Real-World Design Issues

Reasons why classical/textbook designs & modern screening designs likely will not work…

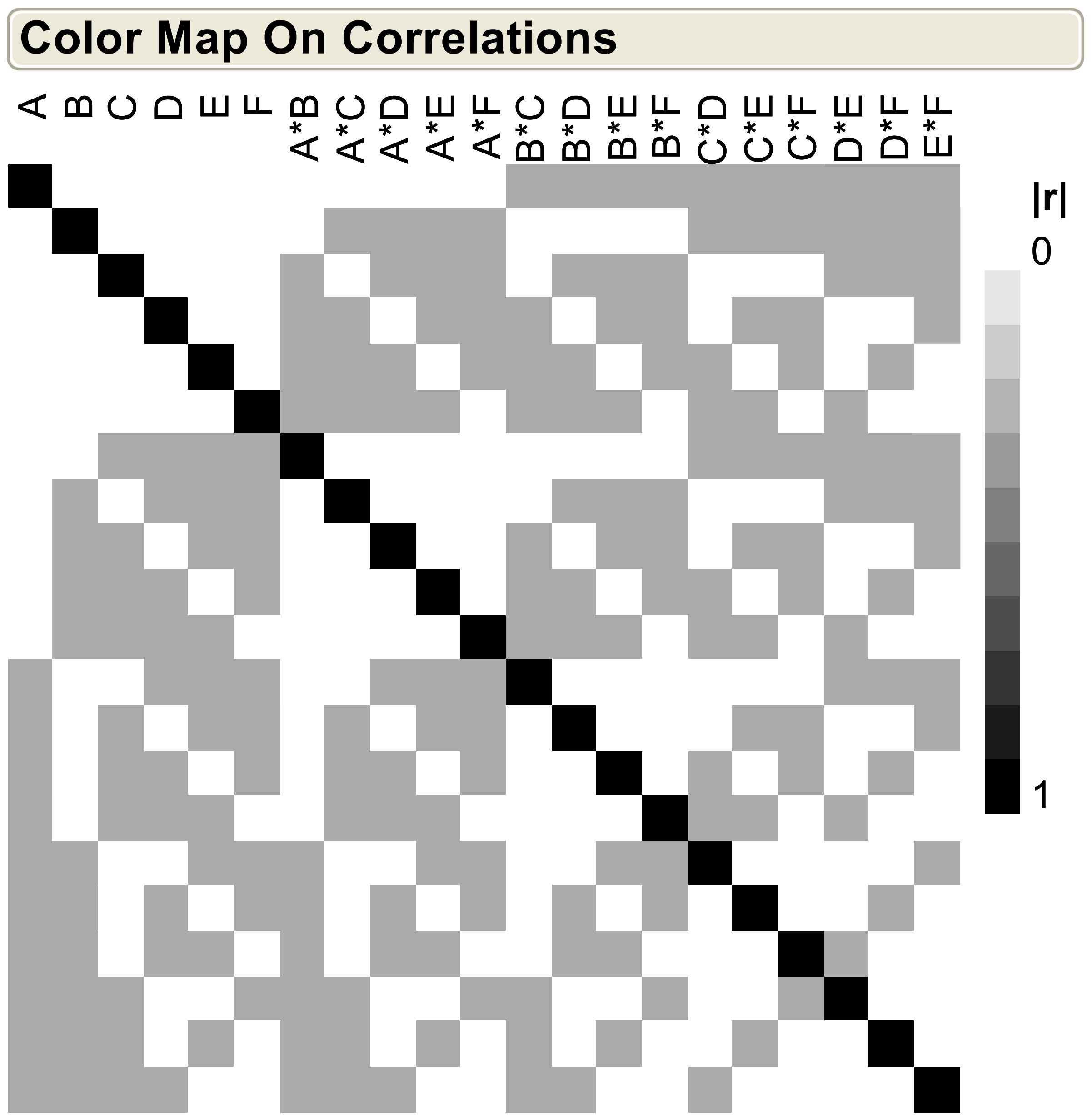
Want to Make Designs Fit the Problem – NOT Make Problems Fit the Designs!

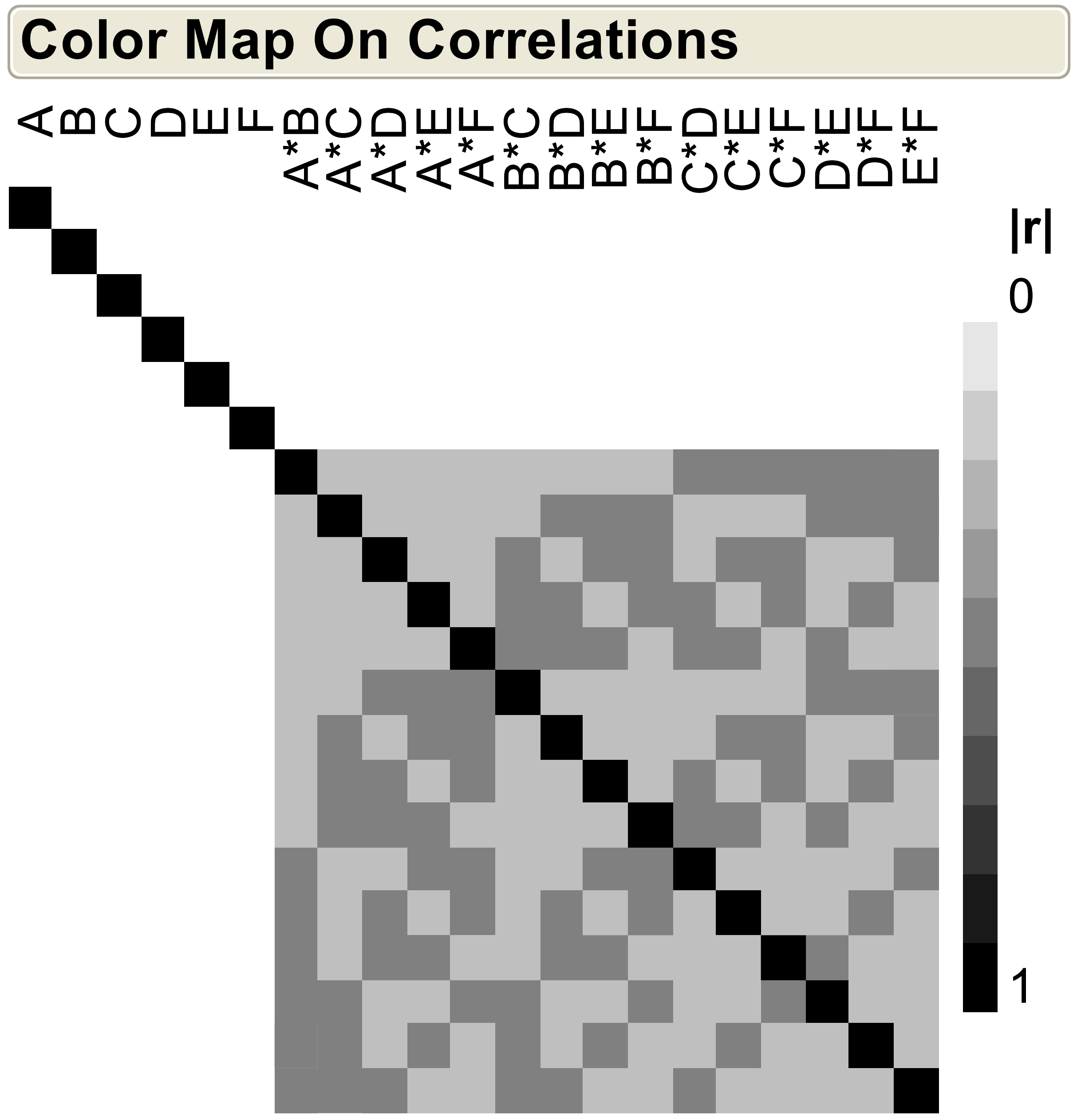
- Work with these different kinds of control variables or factors:
  - Continuous/quantitative? (Finely adjustable like *temperature*, *speed*, *force*)
  - Categorical/qualitative? (Comes in types, like material = *rubber*, *polycarbonate*, *steel* with mixed # of levels; 3 chemical agents, 4 decontaminants, 8 coupon materials…)
  - Mixture/formulation? (Blend different amounts of *ingredients,* and the process performance is dependent on the ***proportions*** of components more than on the amounts)
  - Blocking? e.g., "lots" of the same raw materials, multiple "same" machines, samples that get processed in "groups" – like "eight in a tray," run tests over multiple days – i.e., variables for which we believe there ***shouldn't be a causal effect*** – but we need to know if they do or are correlated with a lurking variable!
- Work with combinations of these four kinds of variables?
- Certain combinations cannot be run? (too costly, unsafe, breaks the process)
- Certain factors are hard-to-change (temperature takes a day to stabilize)
- Would like to add onto existing trials? (really expensive/time consuming to run, or by adding constraints can repair a broken design)

Topics

1. Review of *Screening* DOE vs. *Response Surface* DOE
2. *Custom DOE* platform will nearly always give a useable design,
   BUT in certain design situations, you can do better…
3. Discussion of *Definitive Screening Designs* (DSD) vs. classic designs like *Plackett-Burman* (PB) and *Fractional Factorial* (FF)
4. When you have continuous and **categorical factors at just 2 levels**, consider creating *Orthogonal Mixed-Level* (OML) design instead of DSD.
5. When you have continuous and **categorical factors at greater than 2 levels**, consider creating *Orthogonal Arrays* (OA), or *Nearly Orthogonal Arrays* (NOA)
6. When you have many factors (e.g., > 20) consider using *Group Orthogonal SuperSaturated Designs (*GOSSD)
7. All screening designs can be ***augmented*** into RSM designs.
   This is not always practical if factor ranges are changed.

Color Maps for 6-factor, Plackett-Burman (Left) and Definitive Screening Design (Right)

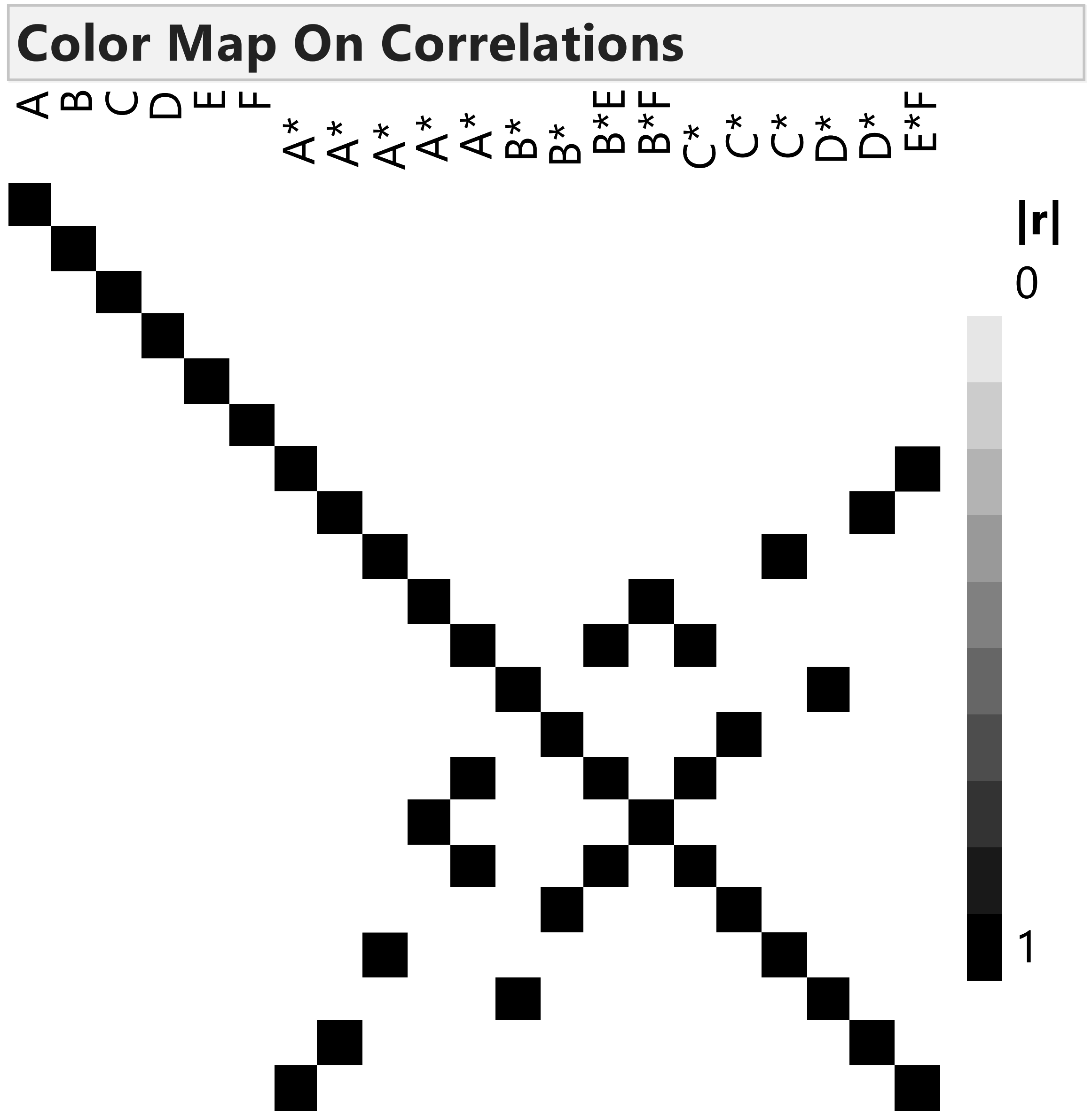


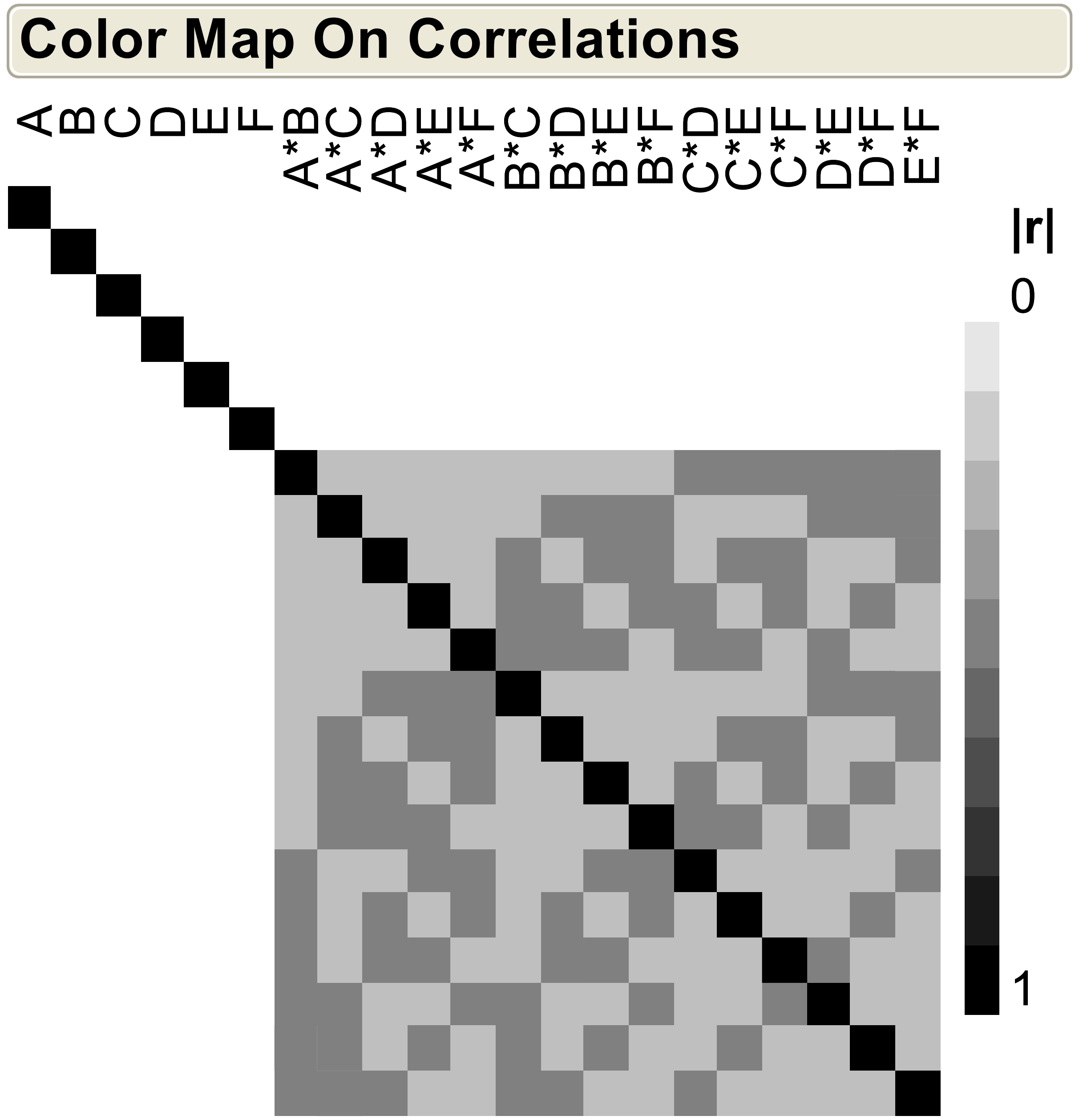


Including center point with Plackett-Burman, these two designs are both 13 trials.

Same size BUT Definitive Screening can test for curvature in each factor

Color Maps for 6-factor, Fractional Factorial (Left) and Definitive Screening Design (Right)





Including center point with FF increases size to 17 trials. 13-trial Definitive Screening Design is **4 fewer tests AND can test for curvature in each factor**
Or, add 4 extra rows to DSD to improve robustness of Fitting Models

Do we give up nothing using DSDs?

- Relative to same size classic 2-level screening designs
  - Confidence intervals increase – typically ≤10%
  - Standard error increases – typically ≤ 10%
  - Power is reduced for main effects – typically ≤ 10% (comparing just ME)
  - Power for squared terms is "low"
    - Still better than power for single center point test for curvature
    - Power is same as larger Central Composite Design supporting full quadratic model
    - Power increases as fewer curvature terms are evaluated – drop least important terms (Factor Sparsity is our friend!)

ANY OTHER WEAKNESSES?

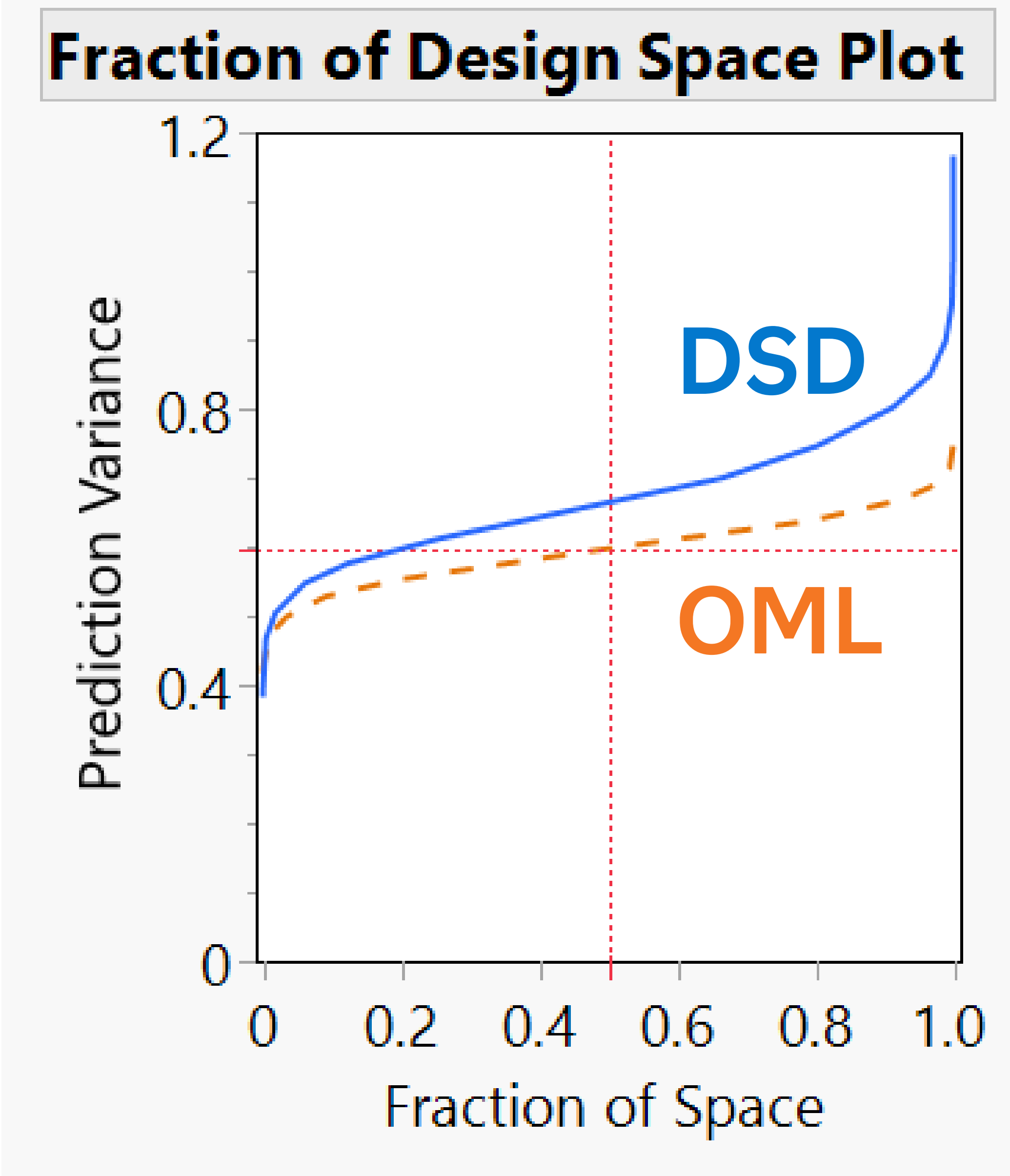
- Factor range for screening may not include optimum
  - So, follow on design will be over different ranges – really can't augment
  - This is more likely with early product development than with designs testing mature systems

Comparing Performance for 18-run Screening Designs

New **OML** + 2cp vs. Older **DSD*** for 4 Continuous & 4 Categorical Factors

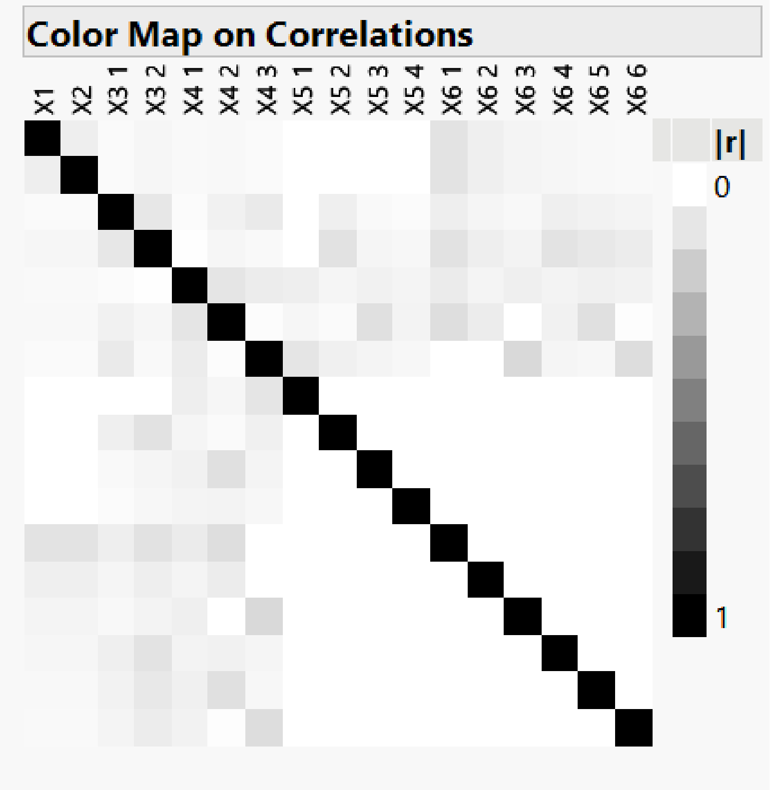
OML has **Better Power** overall – especially for **quadratic** terms,
and **Better Prediction Variance** across Design Space (lower & flatter)

	Term	DSD Power	OML Power	
	Intercept	0.237	0.211	
4 cont.	Temperature	0.828	0.789	4% Lower
	RPM	0.821	0.789	
	Pressure	0.818	0.789	
	Time	0.821	0.789	
4 cat.	Stirring Position	0.895	0.911	2% Higher
	Catalyst	0.891	0.911	
	Concentration	0.891	0.911	
	Feed Rate	0.895	0.911	
4 quad.	Temperature*Temperature	0.245	0.368	**50% Higher**
	RPM*RPM	0.245	0.368	
	Pressure*Pressure	0.245	0.368	
	Time*Time	0.245	0.368	

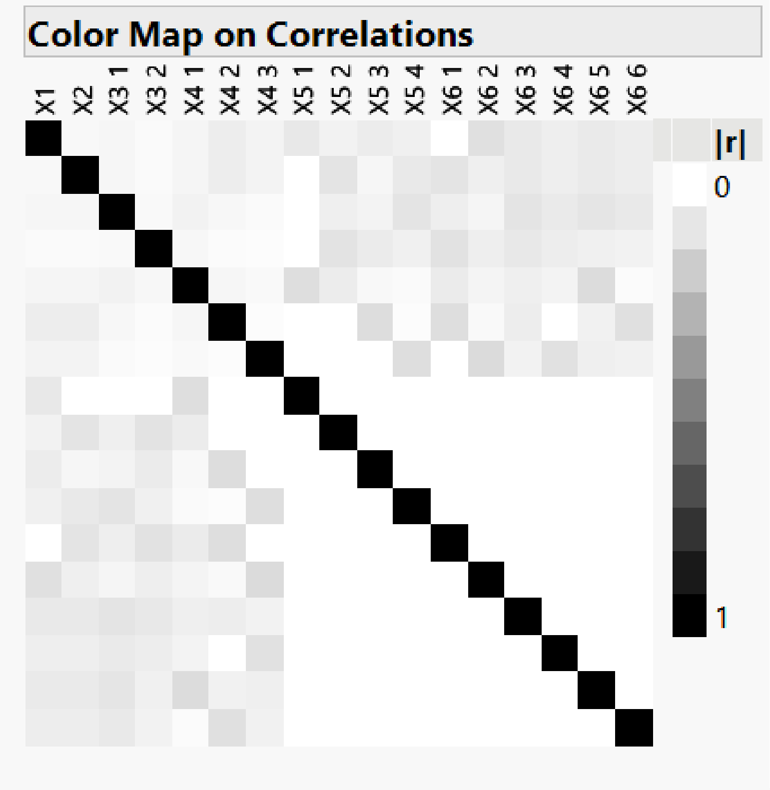


*Definitive Screening Design is better choice if all factors are continuous, because you check for curvature for ALL factors.
OML designs require multiples of 2, 4, or 8 factors to achieve balance, & don't always add cps. for all continuous factors.

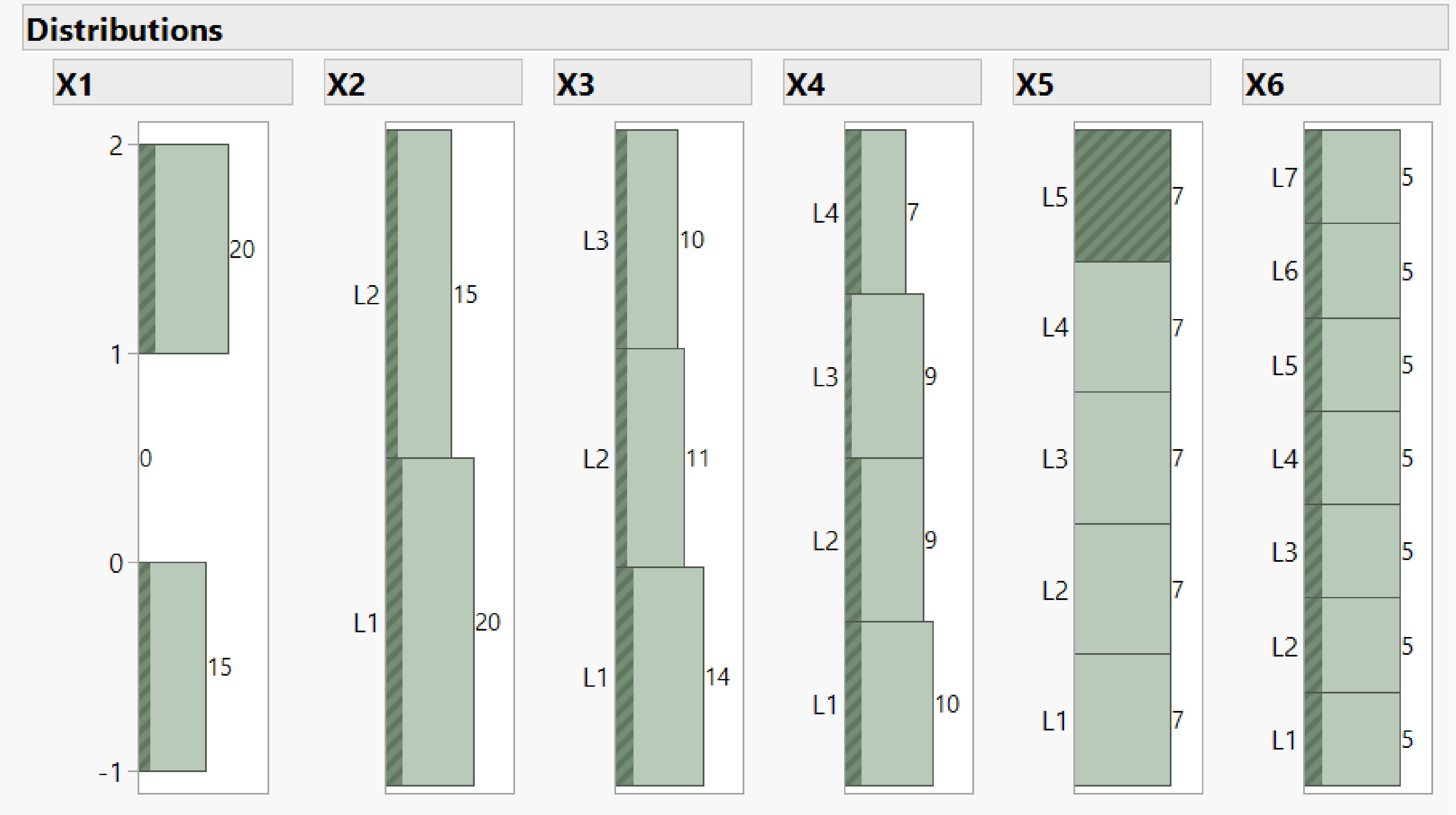
Custom DOE (D-optimal) vs. NOA for multi-level categorical factors - NOA has better balance of number of levels

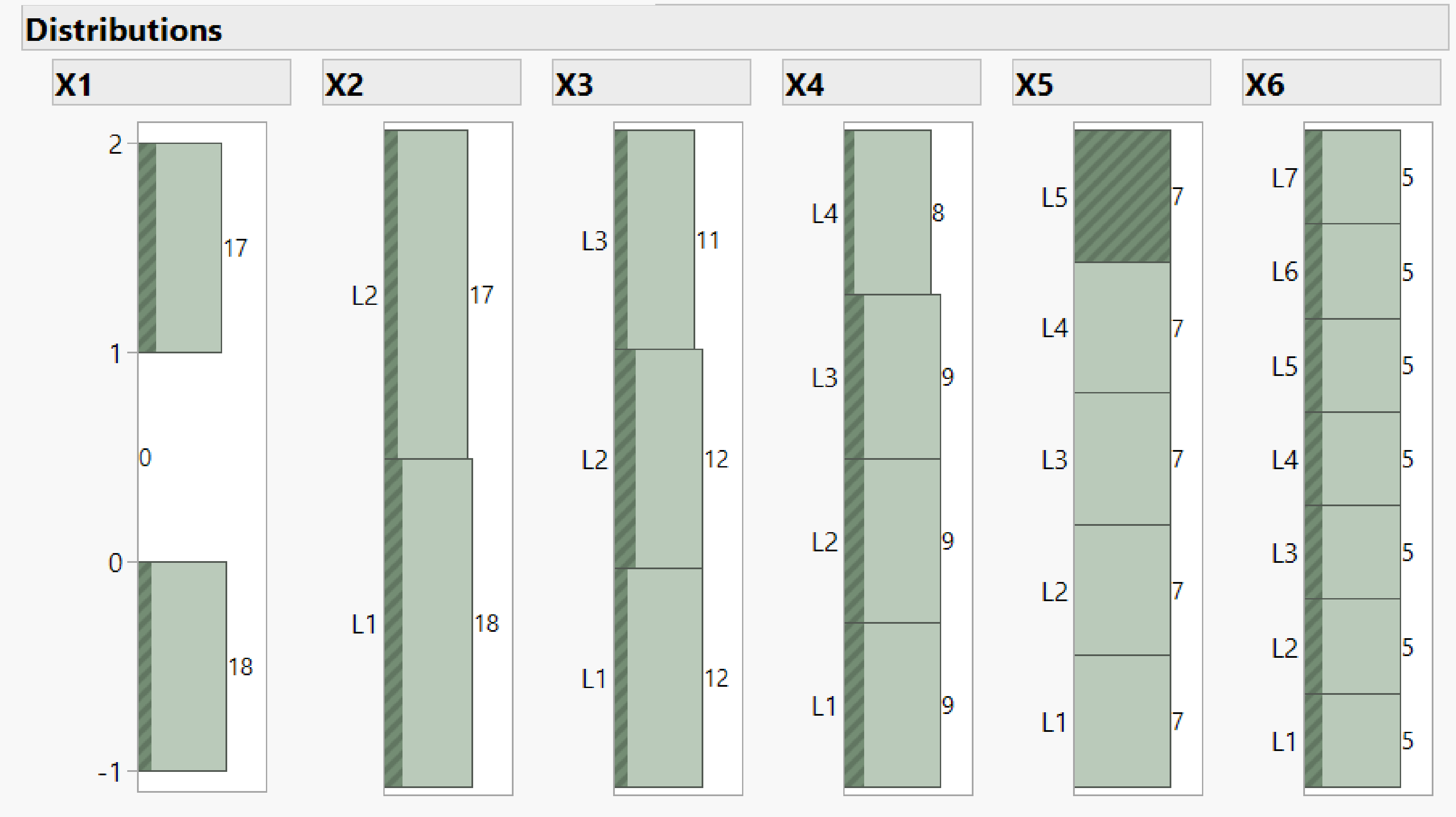


Correlations among Main Effects are all less than 0.15

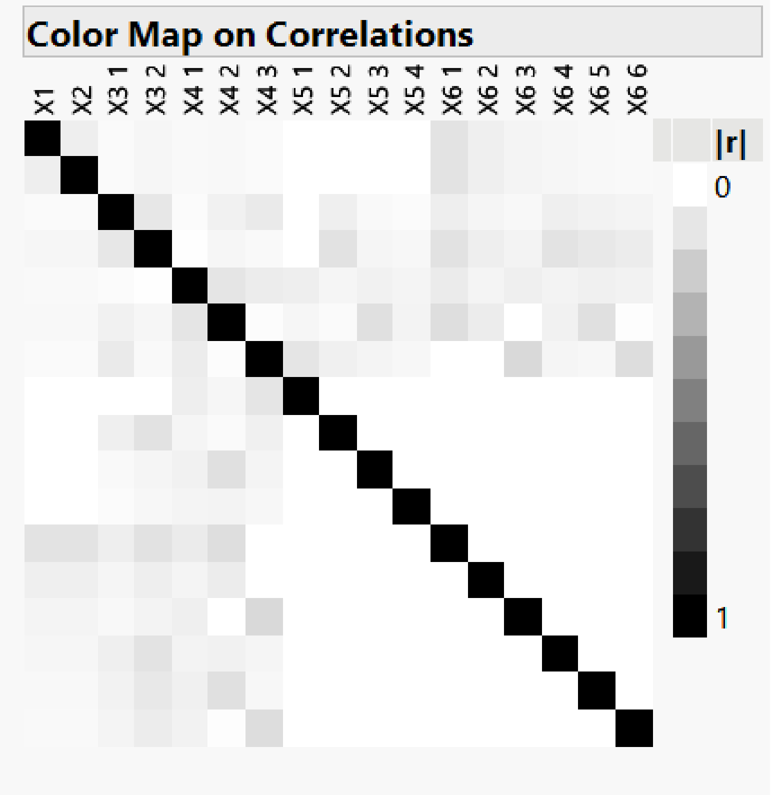


Correlations among Main Effects are all less than 0.15

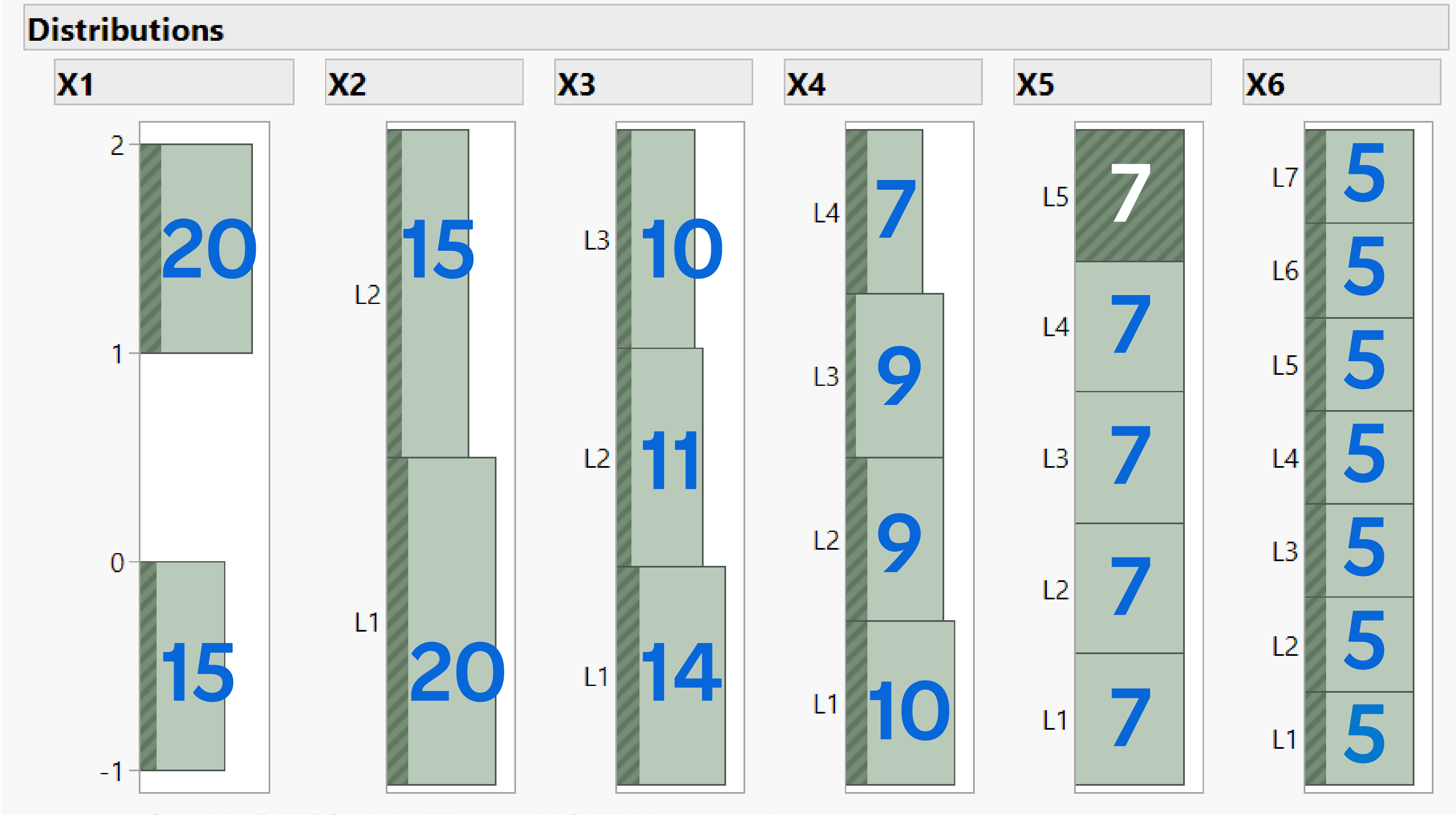


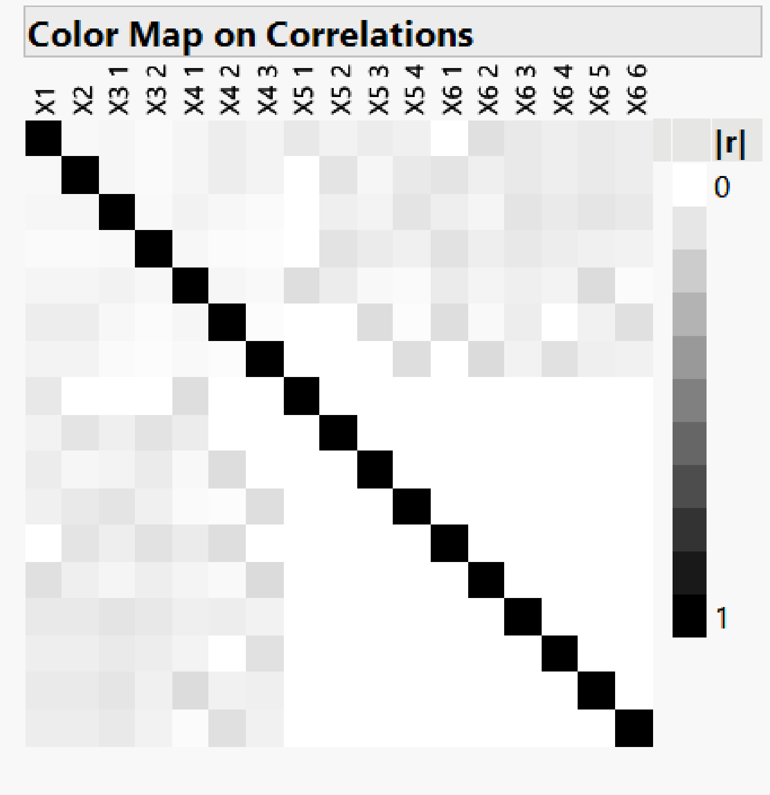


Custom DOE (D-optimal) vs. NOA for multi-level categorical factors - NOA has better balance of number of levels

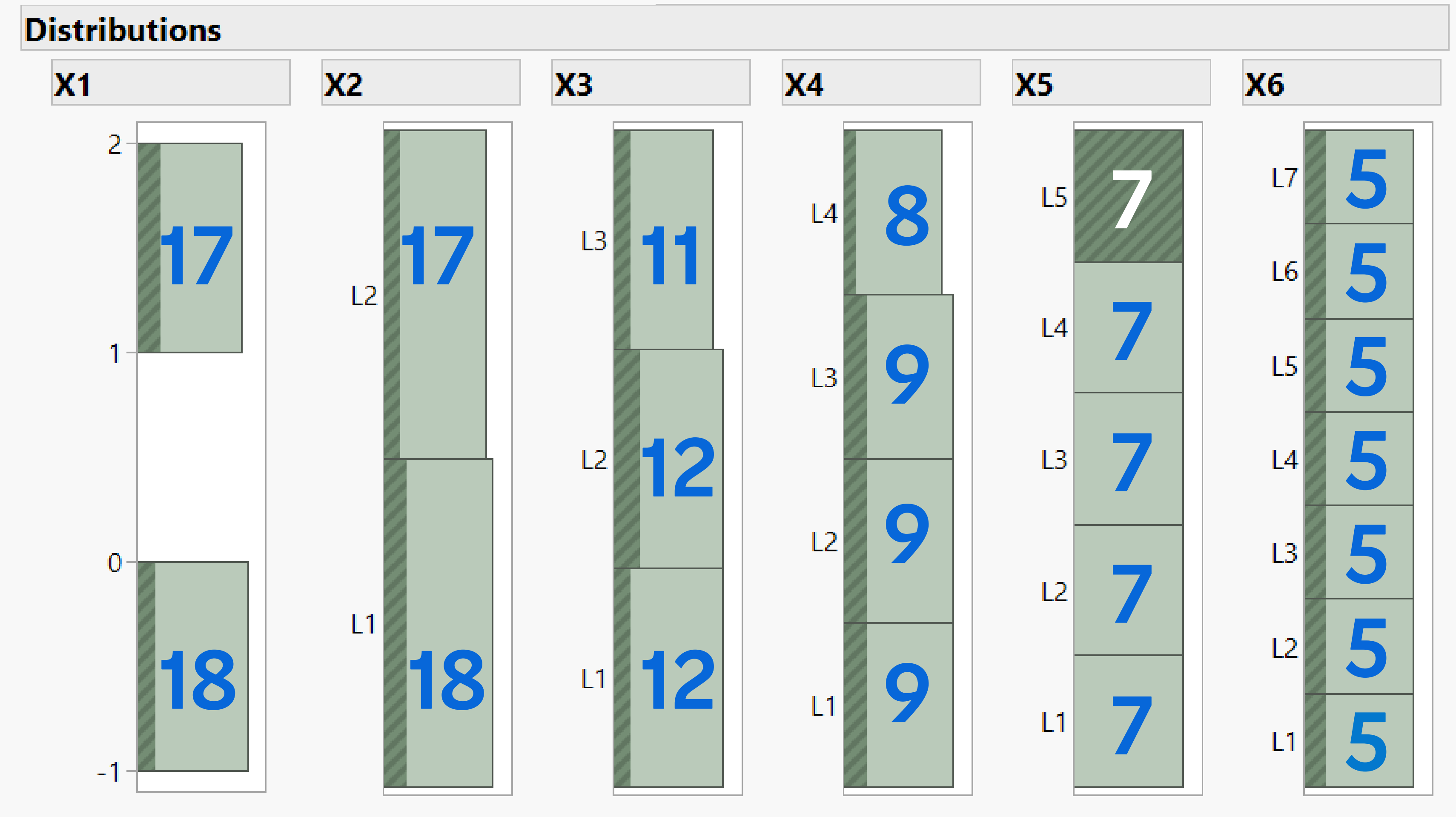


Correlations among Main Effects are all less than 0.15

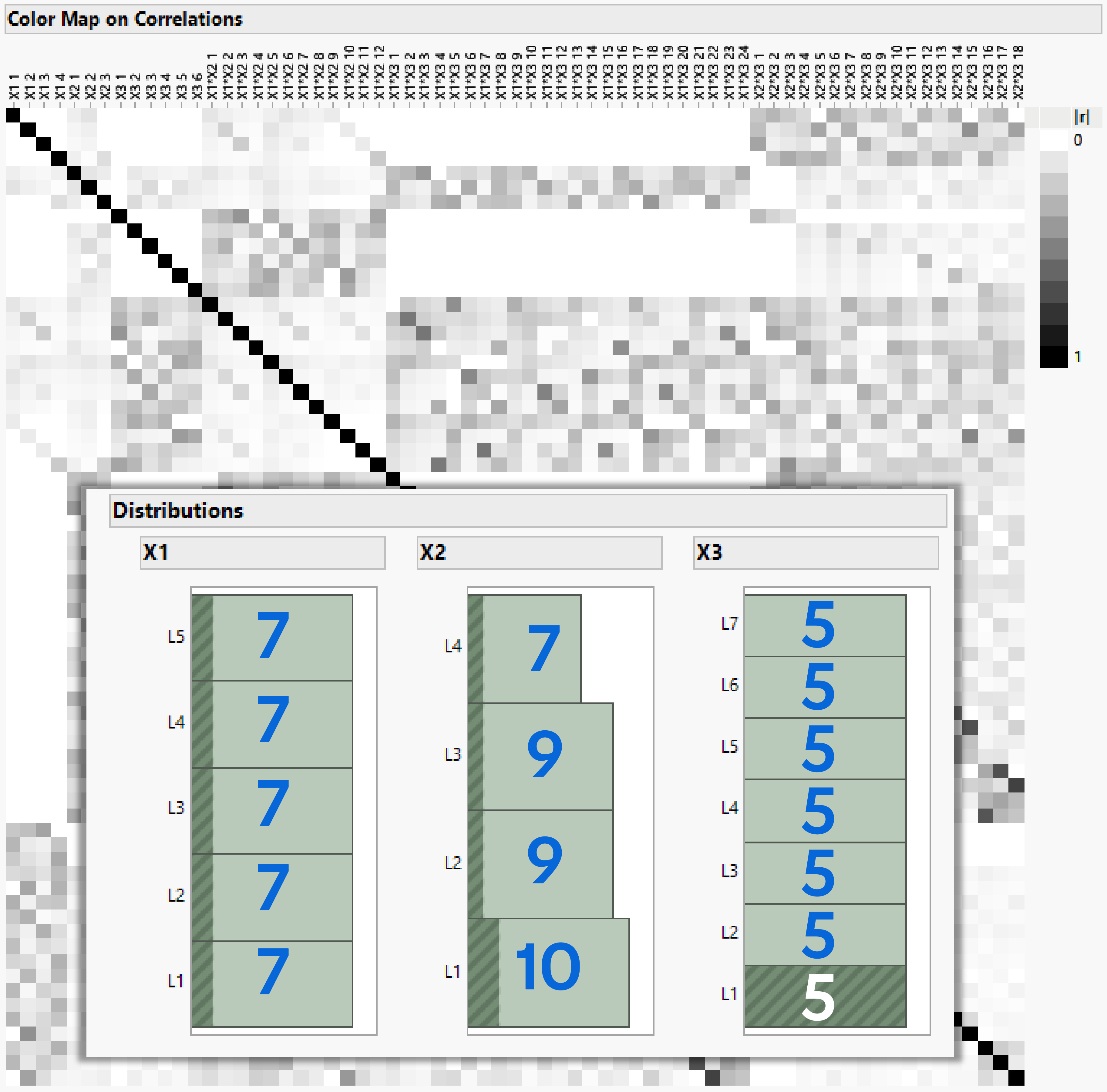


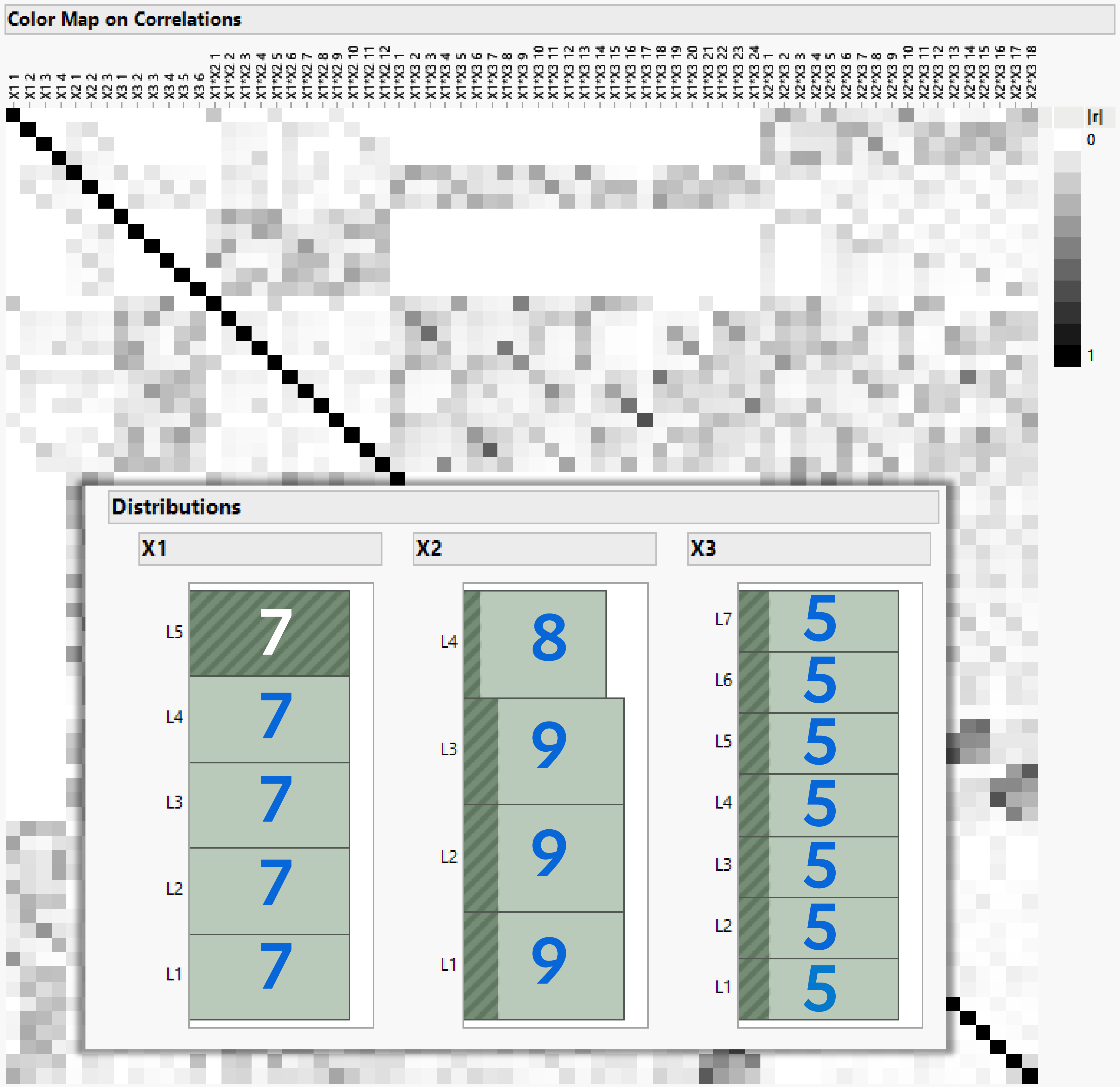


Correlations among Main Effects are all less than 0.15



Custom DOE vs. NOA for multi-level categorical factors
NOA has better balance of number of levels

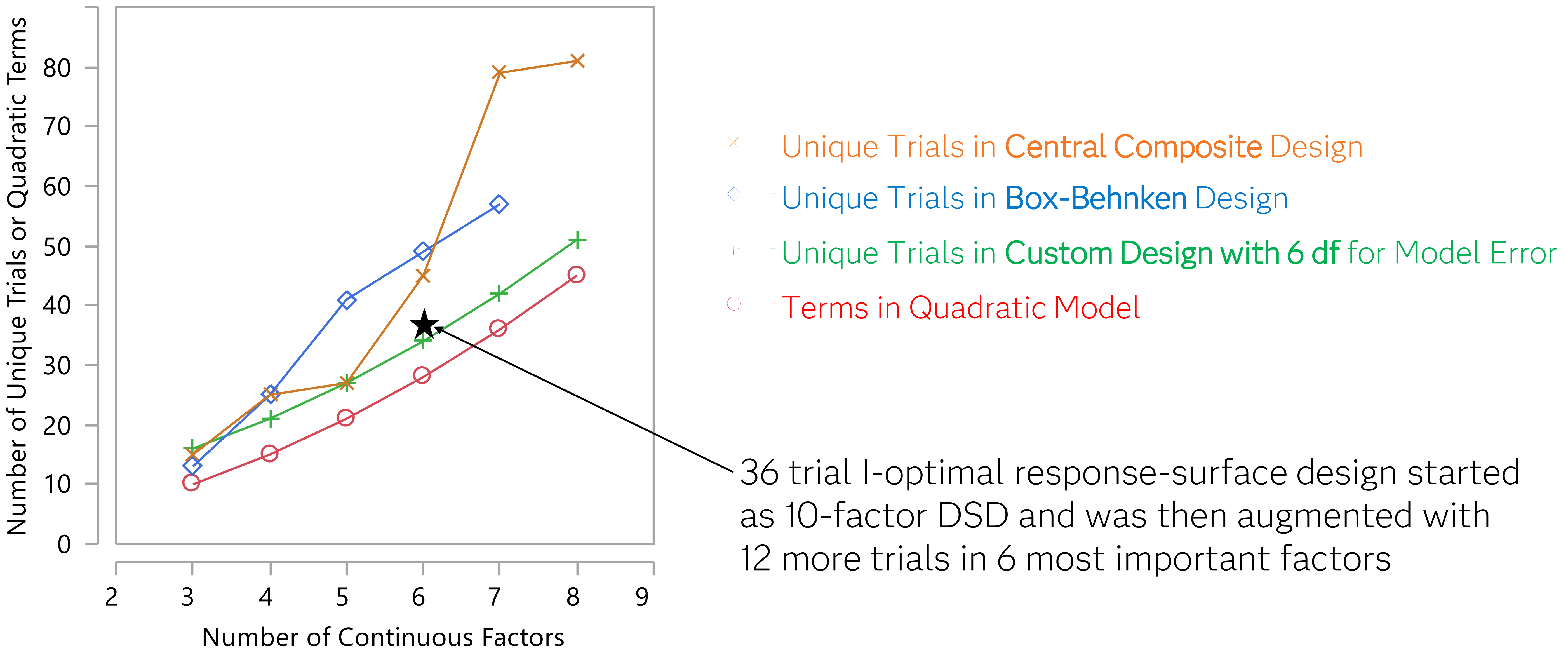




Group Orthogonal SuperSaturated Design (GOSSD)

- Walk through a detailed example for a 28-factor design in 24 trials.
    - Compare GOSSD to older SSD methods using “half” designs
    - Show how to make the *Group Orthogonal Supersaturated Design*
    - Uses a simulator to produce data for 6 factors with effect sizes relative to the noise of 1/2X, 1X, 2X & 3X. Think: very hard, moderately hard, easy, & very easy to detect.
    - Analyze data using *Fit Group Orthogonal Supersaturated*
    - Fit GOSSD will give a message when correlation may be masking effects.
      “Group 5 has one or more effects that may be active. The design should be augmented to remove this ambiguity.”
    - Show the impact of augmenting the GOSSD with 8 trials and reanalyzing.
- Recommended strategy is to place your “most-likely-to-be-significant” factors into separate groups, so they will be orthogonal.

Number of Unique Trials for 3 Response-Surface Designs and Number of Quadratic Model Terms vs. Number of Continuous Factors



If generally running 3, 4 or 5-factor fractional-factorial designs…

1. How many interactions are you not investigating?
2. How many more trials needed to fit curvature?
3. Consider two stages: Definitive Screening + Augmentation

All analyses rank factors A, B & C as top 3

Factor F appears to be most likely fourth factor

- Linear terms only – fourth factor is F
- Linear + Squared terms – fourth factor is D
- Stepwise with min AICc stopping rule – fourth factor is F
- Stepwise with max K-Fold R-Square stopping rule – fourth factor is F
- Stepwise with max Validation R-Square as stopping rule – fourth factor is F
- All possible models – fourth factor is G
- Fit DSD – fourth factor is F

- When D & F are in same 5-factor (with A, B, & C) stepwise model, D drops out
- When G & F are in same 5-factor (with A, B, & C) stepwise model, G drops out
- When D & G are in same 5-factor (with A, B, & C) stepwise model, both drop out

- There is an important difference between saying, *"Factor F has no effect."* and, *"Given the amount of data taken an effect for factor F was not detected."*

- Augmenting design to support 6-factor quadratic model in A, B, C, D, F & G will
  - help resolve the relative contributions of D, F & G
  - increase the power for all terms – but especially - the squared terms

If more than a few factors are significant, then augment design to support 2nd order model

	A	B	C	D	F	G	Block	Yield @ Time t
14	0	0	0	0	0	0	1	7.49
15	1	1	-1	1	-1	1	1	0.98
16	1	1	1	-1	-1	0	1	0.86
17	-1	1	-1	-1	1	1	1	1.25
18	1	-1	1	1	-1	-1	1	1.03
19	1	1	0	-1	1	-1	1	1.07
20	0	0	0	0	0	0	1	7.33
21	1	-1	-1	0	1	-1	1	2.61
22	-1	-1	0	1	-1	1	1	11.39
23	-1	0	1	-1	1	1	1	12.96
24	1	1	-1	1	1	1	1	1.18
25	1	0	1	1	-1	1	2	•
26	1	-1	0	1	1	0	2	•
27	1	-1	-1	1	0	1	2	•
28	1	-1	0	-1	0	-1	2	•
29	1	0	-1	-1	1	0	2	•
30	1	1	0	-1	0	1	2	•
31	1	0	1	0	1	-1	2	•
32	-1	-1	0	0	1	1	2	•
33	0	0	1	1	-1	-1	2	•
34	-1	-1	1	0	0	0	2	•
35	0	1	1	0	1	0	2	•
36	0	1	-1	1	1	-1	2	•

NOTE: First 13 rows of original design are not shown.

These 12 trials added onto original 24 trials to support full quadratic model in 6 most important factors plus a block effect between original and augmented trials

Power for squared terms in 2nd order model is increased to near that of 6-factor RSM designs

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.273
Block	1	0.983
A	1	0.965
B	-1	0.966
C	1	0.976
D	-1	0.969
F	1	0.975
G	-1	0.961
A*B	1	0.887
A*C	-1	0.881
A*D	1	0.825
A*F	-1	0.915
A*G	1	0.732
B*C	-1	0.728
B*D	1	0.853
B*F	-1	0.859
B*G	1	0.724
C*D	-1	0.872
C*F	1	0.838
C*G	-1	0.778
D*F	1	0.847
D*G	-1	0.838
F*G	1	0.86
A*A	1	0.299
B*B	-1	0.361
C*C	1	0.362
D*D	-1	0.309
F*F	1	0.384
G*G	-1	0.347

Power Analysis

Significance Level 0.05

Anticipated RMSE 1

Parameter	Anticipated Coefficients	Power
Intercept	1	0.364
A	1	0.998
B	-1	0.998
C	1	0.998
D	-1	0.998
F	1	0.998
G	-1	0.998
A*A	1	0.527
B*B	-1	0.599
C*C	1	0.582
D*D	-1	0.541
F*F	1	0.573
G*G	-1	0.568

70	A	B	C	D	F	G	Block	Yield @ Time t
14	0	0	0	0	0	0	1	7.49
15	1	1	-1	1	-1	1	1	0.98
16	1	1	1	-1	-1	0	1	0.86
17	-1	1	-1	-1	1	1	1	1.25
18	1	-1	1	1	-1	-1	1	1.03
19	1	1	0	-1	1	-1	1	1.07
20	0	0	0	0	0	0	1	7.33
21	1	-1	-1	0	1	-1	1	2.61
22	-1	-1	0	1	-1	1	1	11.39
23	-1	0	1	-1	1	1	1	12.96
24	1	1	-1	1	1	1	1	1.18
25	1	0	1	1	-1	1	2	•
26	1	-1	0	1	1	0	2	•
27	1	-1	-1	1	0	1	2	•
28	1	-1	0	-1	0	-1	2	•
29	1	0	-1	-1	1	0	2	•
30	1	1	0	-1	0	1	2	•
31	1	0	1	0	1	-1	2	•
32	-1	-1	0	0	1	1	2	•
33	0	0	1	1	-1	-1	2	•
34	-1	-1	1	0	0	0	2	•
35	0	1	1	0	1	0	2	•
36	0	1	-1	1	1	-1	2	•

Compare FDS Plots & Diagnostics for designs augmented to support full quadratic model in 6 factors

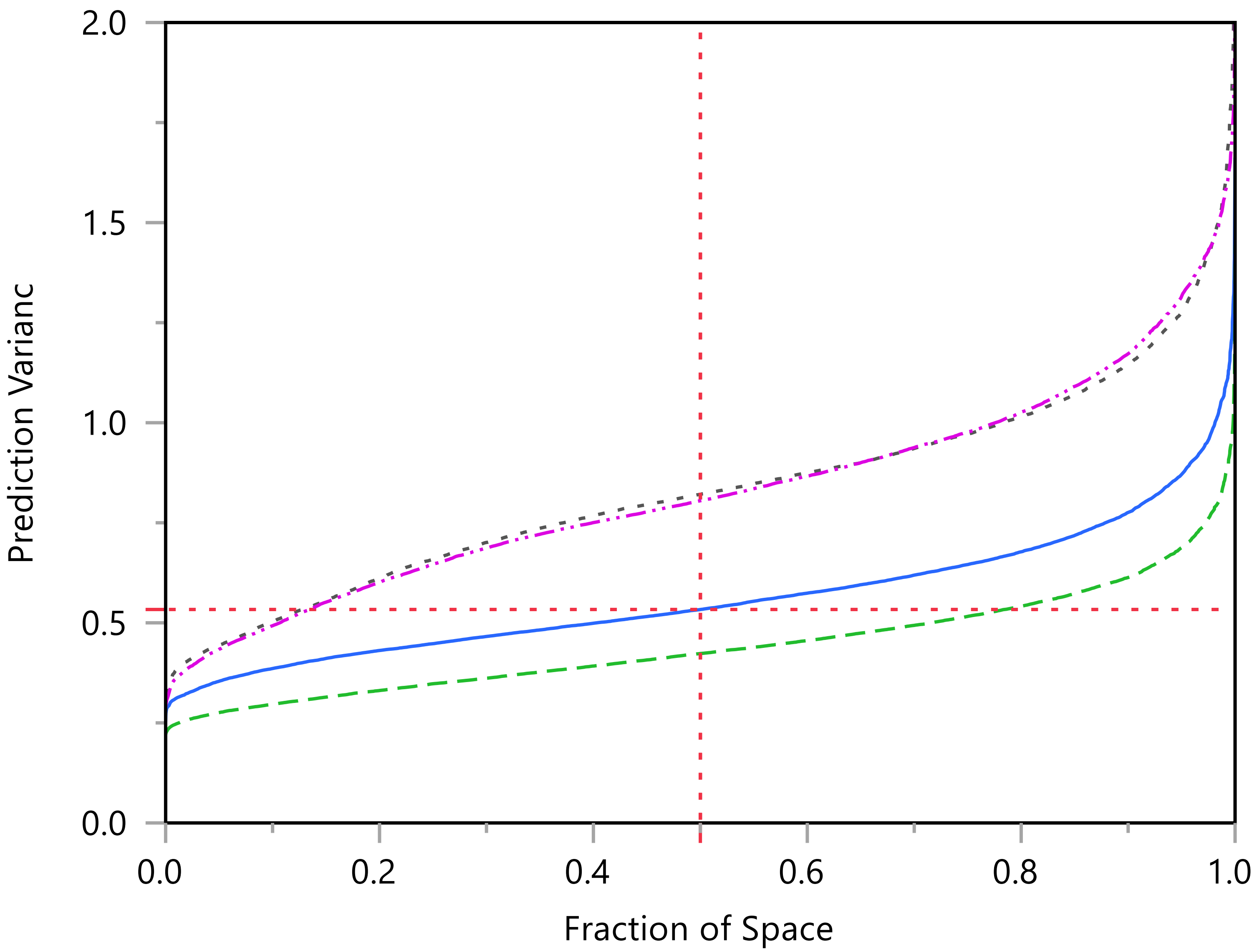
Top: 10-factor Fractional Factorial + C.P.
33 + 9 = 42 total trials

Upper Middle: 10-factor Plackett-Burman + C.P.
25 + 11 = 36 total trials

Lower Middle: 10-factor Definitive Screening Design
21 + 15 = 36 total trials

Bottom: 6-factor I-optimal DOE for full Quadratic model
34 total trials

Fraction of Design Space Plot



Design Diagnostics — aug. FF

I Optimal Design	
D Efficiency	40.729
G Efficiency	56.09719
A Efficiency	12.41717
Average Variance of Prediction	0.82307
Design Creation Time (seconds)	0.05

Design Diagnostics — aug. PB

I Optimal Design	
D Efficiency	38.46605
G Efficiency	54.33992
A Efficiency	14.61968
Average Variance of Prediction	0.833744
Design Creation Time (seconds)	0.05

Design Diagnostics — aug. DSD

I Optimal Design	
D Efficiency	42.15506
G Efficiency	69.61262
A Efficiency	22.27027
Average Variance of Prediction	0.563765
Design Creation Time (seconds)	0.066667

Design Diagnostics — I-opt

I Optimal Design	
D Efficiency	42.94028
G Efficiency	75.52931
A Efficiency	27.20305
Average Variance of Prediction	0.44424
Design Creation Time (seconds)	0.066667

JMP STATISTICAL DISCOVERY