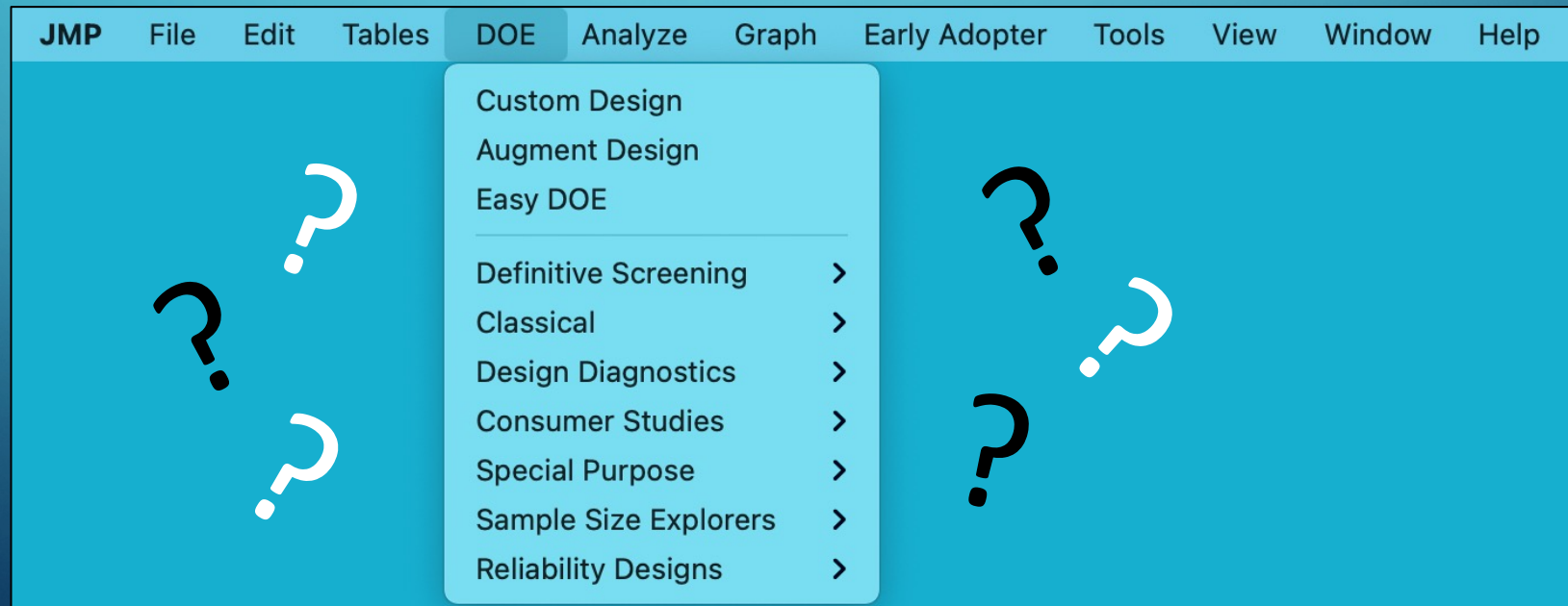


# DESIGN OF EXPERIMENT'S CRUCIAL STEP 0: CHOOSING THE RIGHT DOE OPTION

JMP Discovery Summit Americas 2023

Christine M. Anderson-Cook

candcook@gmail.com



# CHOICES, CHOICES, CHOICES ....

DOE Analyze Graph

- Custom Design
- Augment Design
- Easy DOE
- Definitive Screening >
- Classical >**
  - Two Level Screening >
  - Response Surface Design
  - Full Factorial Design
  - Mixture Design
  - Taguchi Arrays
- Design Diagnostics >
- Consumer Studies >
- Special Purpose >
- Sample Size Explorers >
- Reliability Designs >

Definitive Screening >

- Classical >
- Design Diagnostics >
- Consumer Studies >
- Special Purpose >**
  - Covering Array
  - Space Filling Design
  - Accelerated Life Test Design
  - Nonlinear Design
  - Balanced Incomplete Block Design
  - MSA Design
  - Group Orthogonal Supersaturated >
- Sample Size Explorers >
- Reliability Designs >

When you are an expert ....

When you are getting started ....

# OUTLINE

## 1. What are we trying to accomplish with our experiment?

Understanding what the goal is of the experiment will help to match possible choices to our experiment

**NEEDS SUBJECT MATTER EXPERTISE!**

## 2. What are the common DOE choices in JMP?

Quick walk through of some of JMP's most popular design of experiment choices

## 3. Key Questions to consider before generating a design

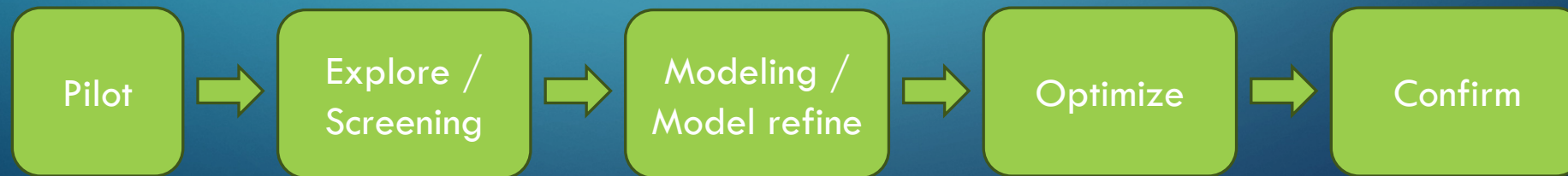
What do we already know about the factors, responses and their relationship?

What are the constraints under which we need to operate?

**NEEDS SUBJECT MATTER EXPERTISE!**

# WHAT ARE WE TRYING TO DO? (EXPERIMENTAL OBJECTIVES)

1. **Pilot Study** – make sure that data quality and input space of interest is suitable
2. **Exploration / Screening** – identify important factors (eliminate those not important), see basic relationships
3. **Modeling** – capture relationship between inputs and response(s) in functional form
4. **Model Refinement** – make sure model has sufficient precision for what is needed
5. **Optimize** – use model to solve problem, optimize system performance
6. **Confirm** – verify results are robust in environment that they will be used



# OUTLINE

## 1. What are we trying to accomplish with our experiment?

Understanding what the goal is of the experiment will help to match possible choices to our experiment

## 2. What are the common DOE choices in JMP?

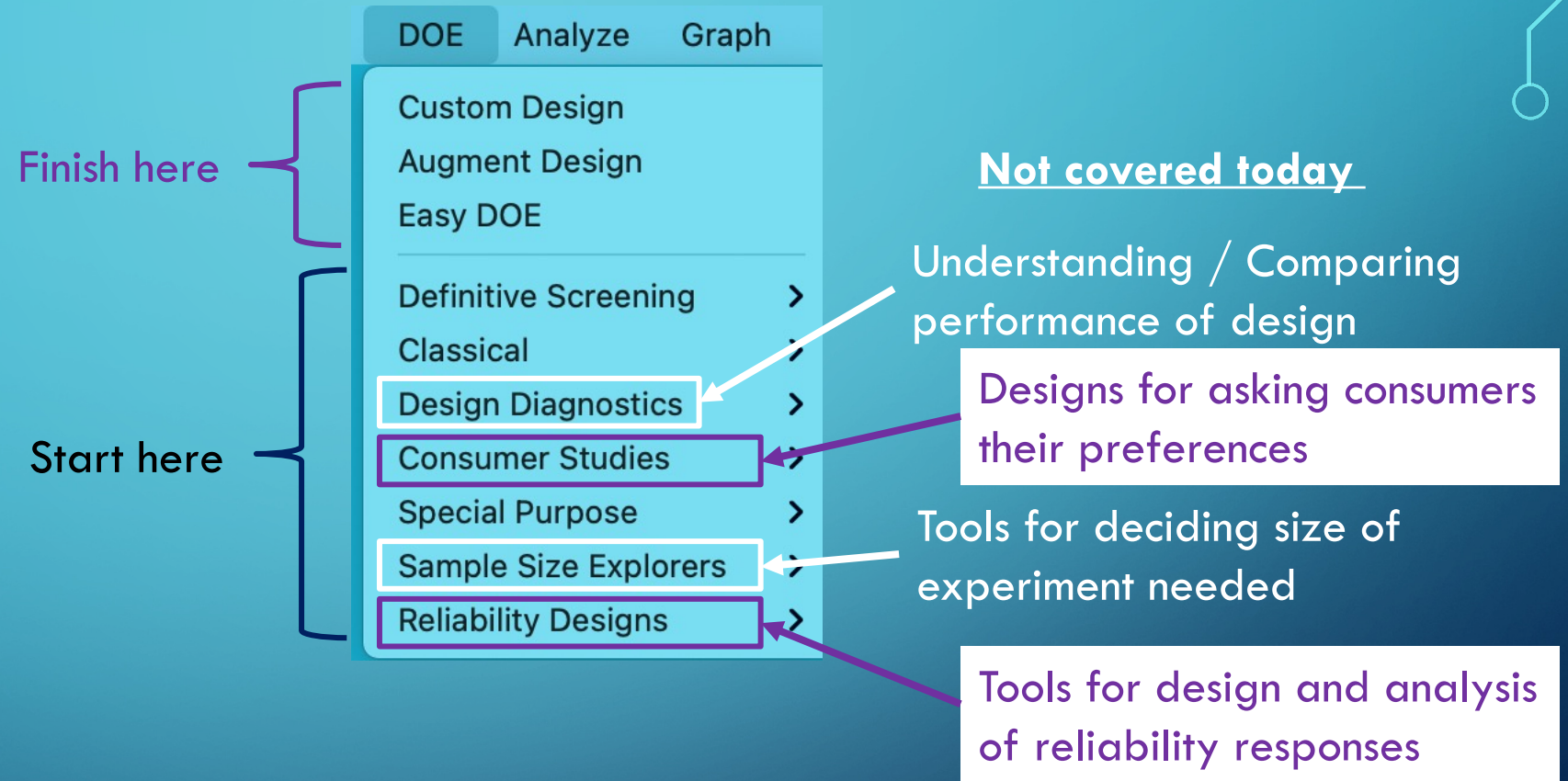
Quick walk through of some of JMP's most popular design of experiment choices

## 3. Key Questions to consider before generating a design

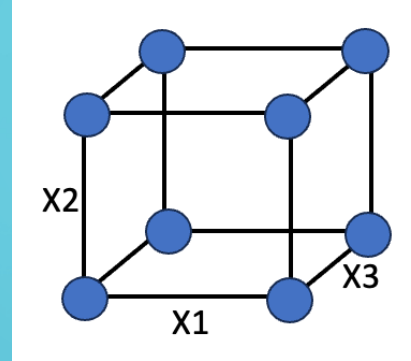
What do we already know about the factors, responses and their relationship?

What are the constraints under which we need to operate?

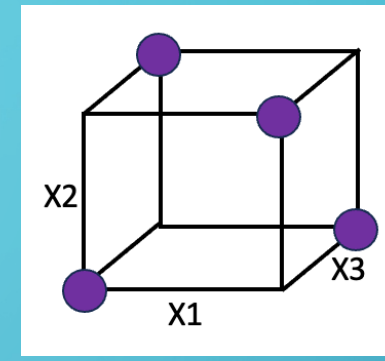
# CHOICES



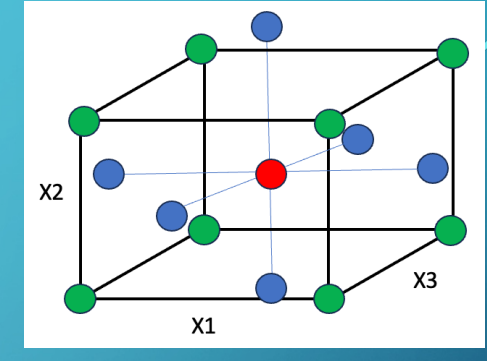
# CLASSICAL DESIGNS – PART 1



Full Factorial



Two-Level Screening



Response Surface

- **Key characteristics:**

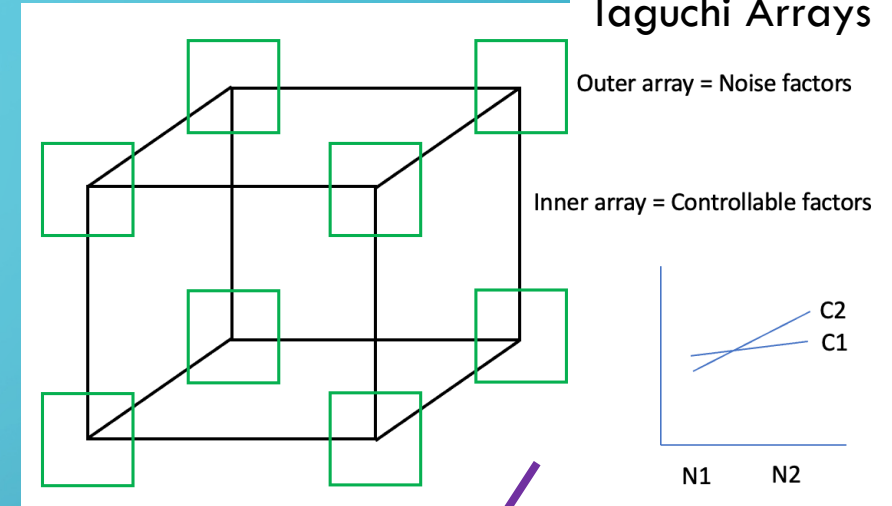
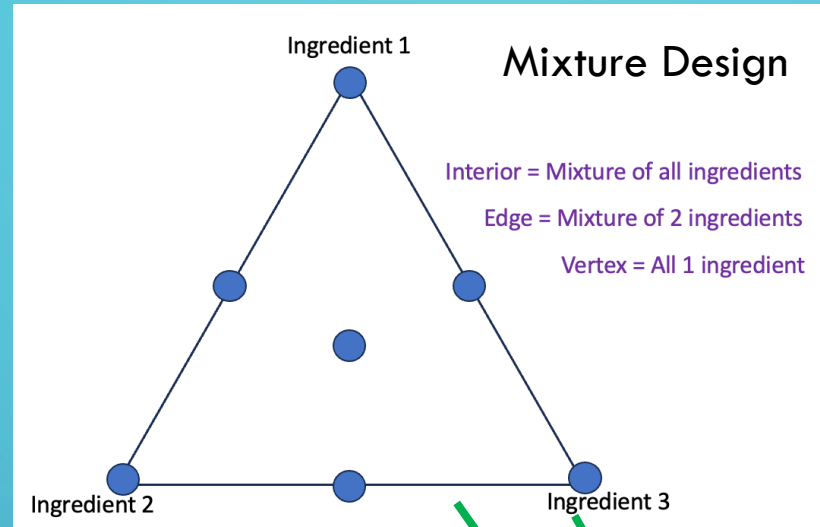
- Designed for standard situations with pre-specified regions and # of runs
- Good overall performance

- **Purpose:**

- Screening & Modeling

1. Pilot Study
2. Exploration / Screening
3. Modeling
4. Model Refinement
5. Optimize
6. Confirm

# CLASSICAL DESIGNS – PART 2



## • Key characteristics:

### • Special situations:

- Mix = when proportions of ingredients matters
- Tag = making process robust to things we can't control in production

## • Purpose:

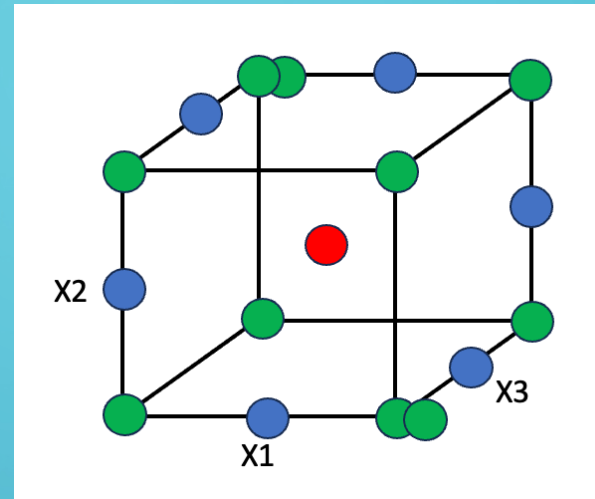
- Mix = Screening & Modeling
- Tag = Optimizing robustness

1. Pilot Study
2. Exploration / Screening
3. Modeling
4. Model Refinement
5. Optimize
6. Confirm



# DEFINITIVE SCREENING DESIGN

- Key characteristics:
  - Allows screening for important factors, but also exploration for curvature
  - Specialized analysis
- Purpose:
  - Screening & Modeling



1. Pilot Study
2. Exploration / Screening
3. Modeling
4. Model Refinement
5. Optimize
6. Confirm

	$\Sigma$	X1	X2	X3
•	1	0	1	1
•	2	0	-1	-1
•	3	1	0	1
•	4	-1	0	-1
•	5	1	-1	0
•	6	-1	1	0
•	7	1	-1	-1
•	8	-1	1	1
•	9	1	1	-1
•	10	-1	-1	1
•	11	1	-1	1
•	12	-1	1	-1
•	13	1	1	-1
•	14	-1	-1	1
•	15	1	1	1
•	16	-1	-1	-1
•	17	0	0	0

# SPECIAL PURPOSE – COVERING ARRAY

← Factors

Runs	1	2	3	4	5	6	7	8	9	10
1	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	0	1	0	0	0	0	1
4	1	0	1	1	0	1	0	1	0	0
5	1	0	0	0	1	1	1	0	0	0
6	0	1	1	0	0	1	0	0	1	0
7	0	0	1	0	1	0	1	1	1	0
8	1	1	0	1	0	0	1	0	1	0
9	0	0	0	1	1	1	0	0	1	1
10	0	0	1	1	0	0	1	0	0	1
11	0	1	0	1	1	0	0	1	0	0
12	1	0	0	0	0	0	0	1	1	1
13	0	1	0	0	0	1	1	1	0	1

Example: All combinations of levels for any 3 factors appear at least once

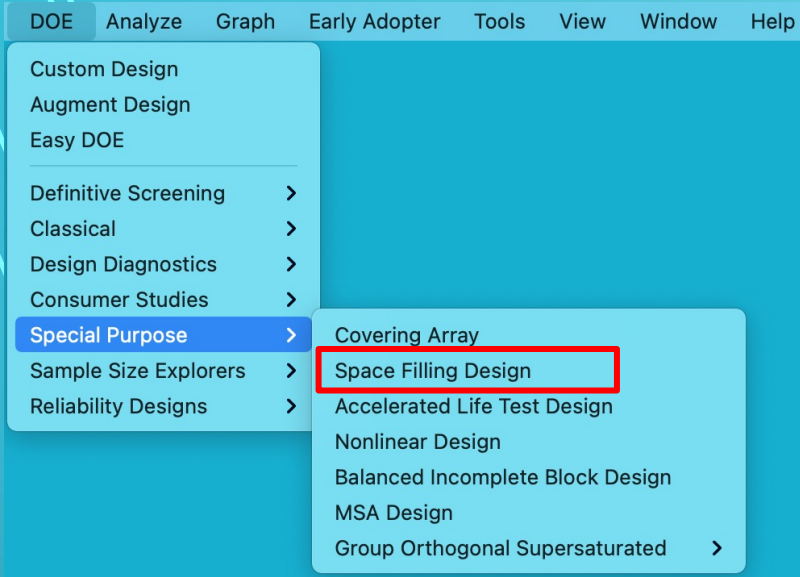
- **Key characteristics:**

- Want to test if combinations of factors cause problems
- Ideal for testing of software across very large number of combinations

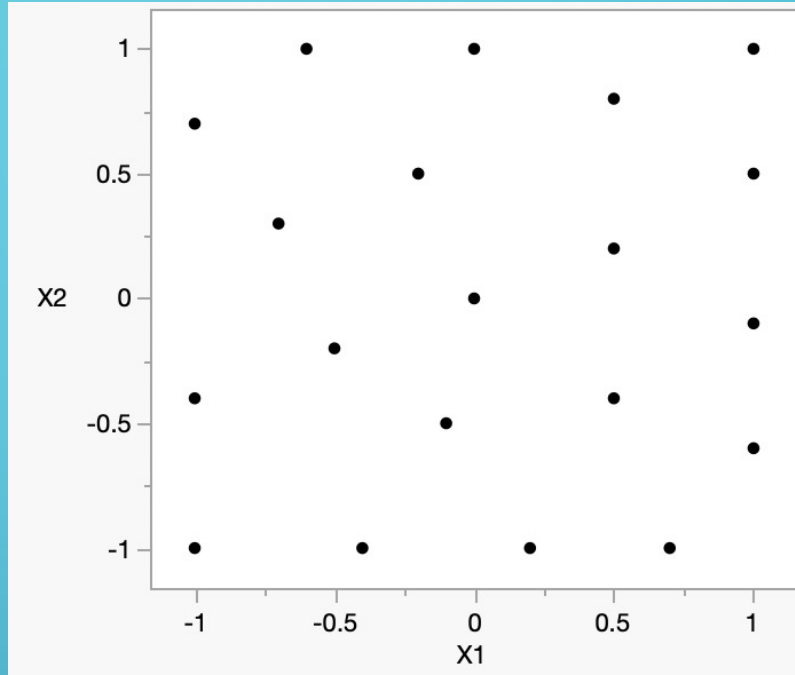
- **Purpose:**

- Exploration

1. Pilot Study
2. Exploration / Screening
3. Modeling
4. Model Refinement
5. Optimize
6. Confirm



# SPECIAL PURPOSE – SPACE FILLING



- Sphere Packing
- Latin Hypercube
- Uniform
- Minimum Potential
- Maximum Entropy
- Gaussian Process IMSE Optimal
- Fast Flexible Filling

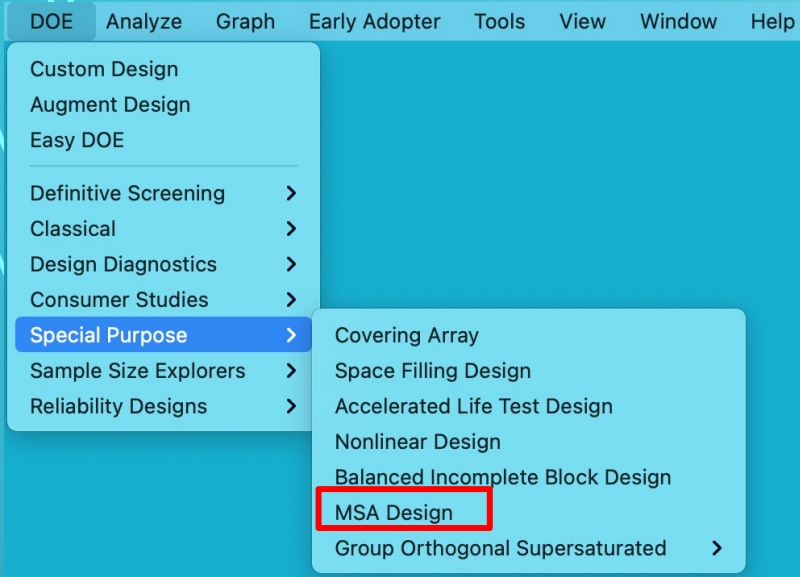
## • Key characteristics:

- Good option if little is known about underlying relationship
- Common in computer experiments

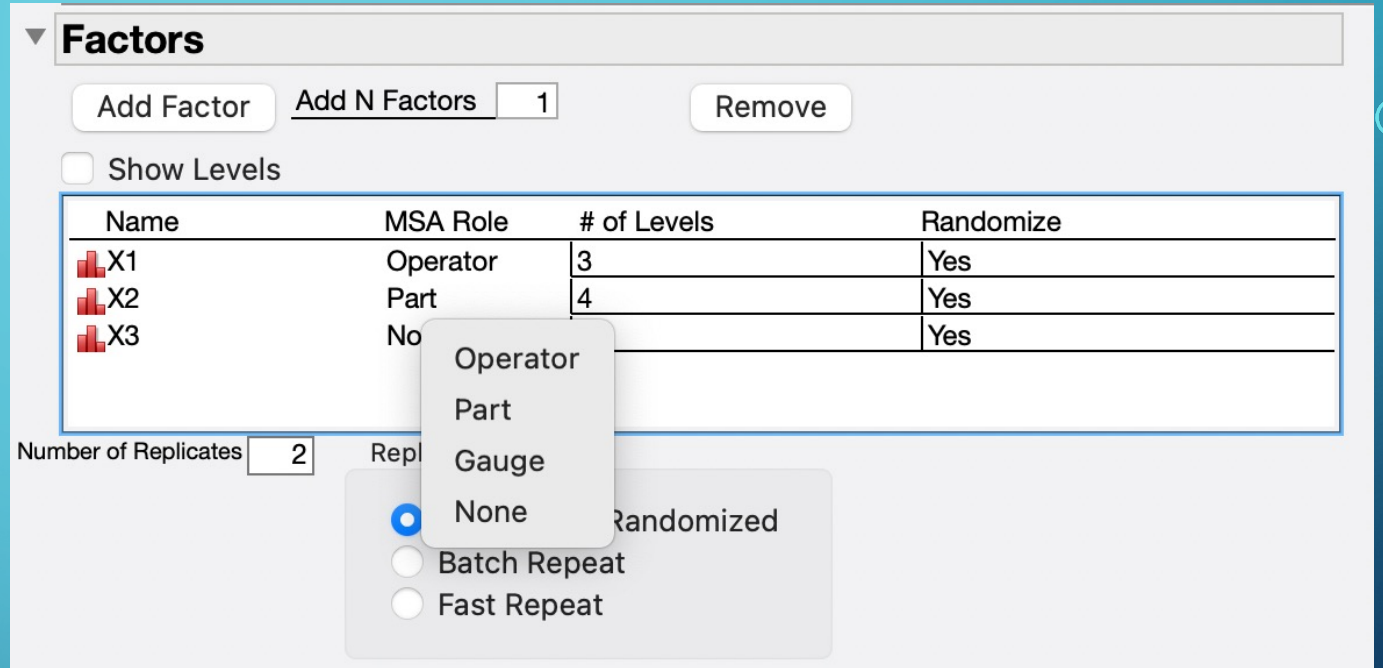
## • Purpose:

- Exploration & Modeling

1. Pilot Study
2. Exploration / Screening
3. Modeling
4. Model Refinement
5. Optimize
6. Confirm



# SPECIAL PURPOSE – MSA DESIGN (MEASUREMENT SYSTEM ANALYSIS)



- **Key characteristics:**

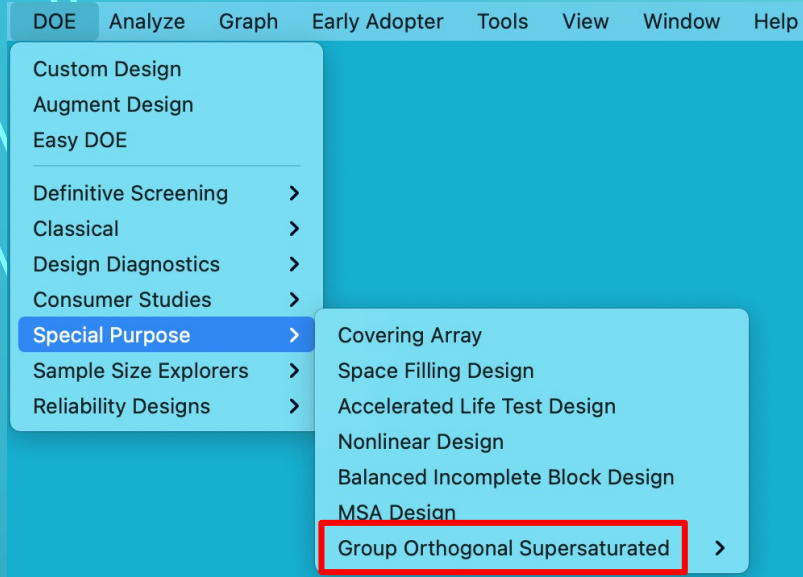
- Design to characterize sources of variability
- Assess precision of measurement system

- **Purpose:**

- Pilot study – data adequacy

1. **Pilot Study**
2. Exploration / Screening
3. Modeling
4. Model Refinement
5. Optimize
6. Confirm

# SPECIAL PURPOSE – GROUP ORTHOGONAL SUPERSATURATED



### Group Orthogonal Supersaturated Design

Number of Runs   
Must be a multiple of 2 (or preferably 4).

Number of Factors   
Add 1 for the Intercept.

Structure

Number of Groups	Group Size	Number of Parameters
2	4	8

Note: Select an option above.

Factors

	Role	Lower	Upper
Intercept	Continuous	1	1
X1	Continuous	-1	1
X2	Continuous	-1	1
X3	Continuous	-1	1
X4	Continuous	-1	1
X5	Continuous	-1	1
X6	Continuous	-1	1
X7	Continuous	-1	1

Group Structure

Group 1

X1  
X2  
X3

Group 2

X4  
X5  
X6  
X7

- Key characteristics:

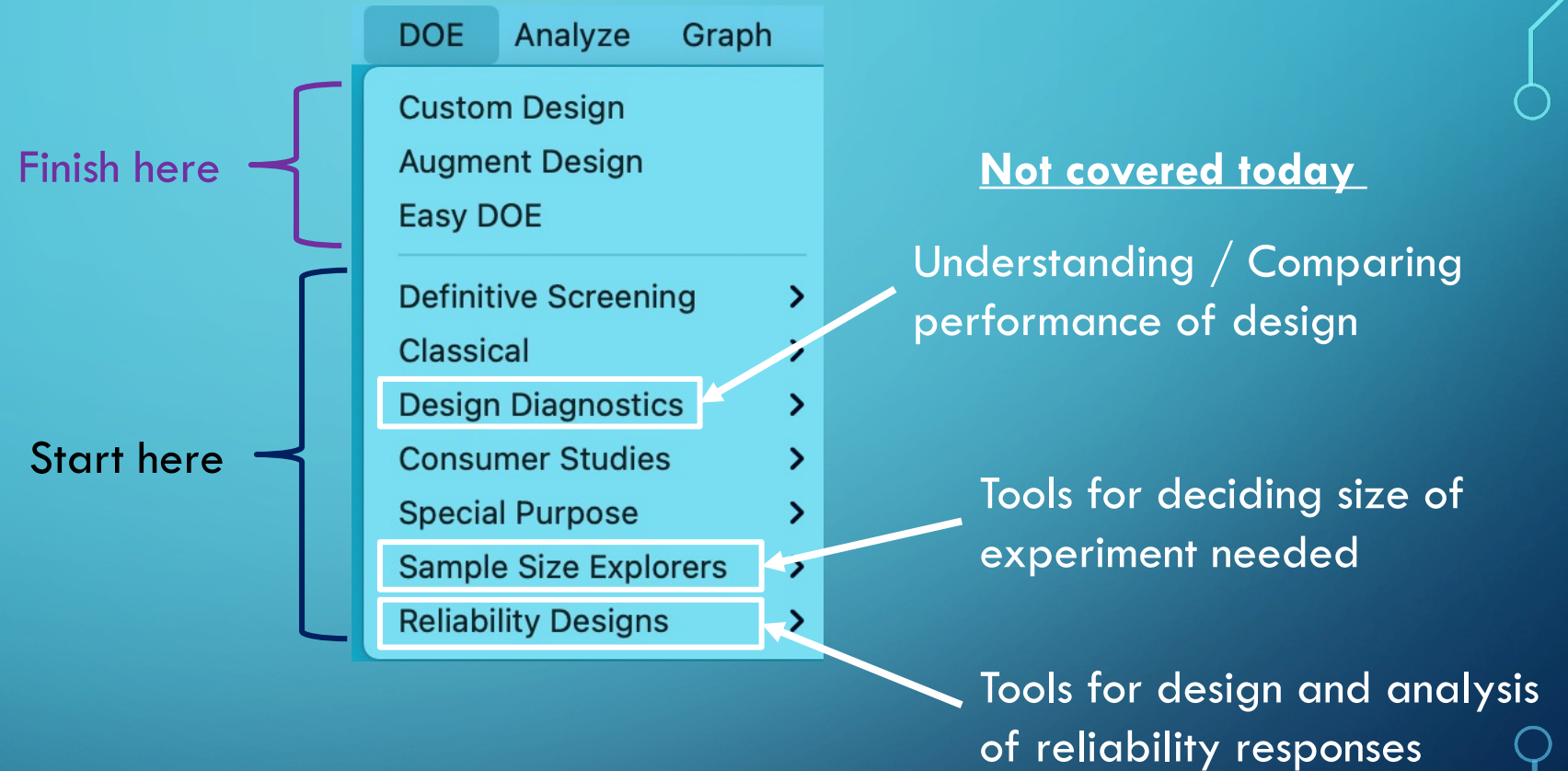
- When # factors > # runs
- Complementary analysis to identify important factors
- Important to not have too many active factors

- Purpose:

- Exploration / Screening

1. Pilot Study
2. Exploration / Screening
3. Modeling
4. Model Refinement
5. Optimize
6. Confirm

# CHOICES



# CUSTOM DESIGN TO THE RESCUE!!

DOE - Custom Design

Custom Design

Responses

Add Response Remove Number of Responses...

Response Name	Goal	Lower Limit	Upper Limit	Importance	Lower Detection Limit	Upper Detection Limit	Units
Y	Maximize	.	.	.	.	.	

Factors

Add Factor Remove Add N Factors 1

Factor Name	Level	Changes	Values	Units
-------------	-------	---------	--------	-------

Advanced Runs

Load a set of candidate runs for covariates from the current data table.

Specify Factors

Add a factor by clicking the Add Factor button. Double click on a factor name or level to edit it.

Continue

# THINK SEQUENTIALLY – AUGMENT!

	X1	X2	X3	Y
1	1	-1	-1	7
2	1	-1	1	15
3	0	0	-1	2
4	-1	1	1	5
5	-1	-1	1	9
6	0	0	1	10
7	1	0	0	9
8	-1	-1	-1	1
9	-1	1	-1	-3
10	0	-1	0	8
11	1	1	1	11
12	0	1	0	4
13	1	1	-1	3
14	0	0	0	6
15	-1	0	0	3



**Augment Design**  
Adds runs to an existing design in such a way that the resulting design is optimal.

Select Columns: 4 Columns  
X1  
X2  
X3  
Y

Cast Selected Columns into Roles

Y, Response: Y  
*optional numeric*

X, Factor: X1  
X2  
X3

Action: OK, Cancel, Remove, Recall, Help



**Augment Design**

**Factors**

Name	Role	Changes	Values	Units
X1	Continuous	Easy	-1 1	
X2	Continuous	Easy	-1 1	
X3	Continuous	Easy	-1 1	

Group new runs into separate block

**Define Factor Constraints**

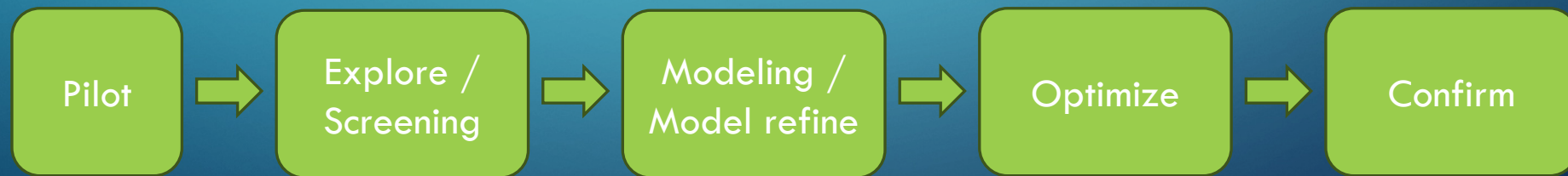
Augmentation Choices

Replicate Add Centerpoints Fold Over Add Axial Space Filling **Augment**



# WHAT ARE WE TRYING TO DO? (EXPERIMENTAL OBJECTIVES)

1. **Pilot Study** – make sure that data quality and input space of interest is suitable
2. **Exploration / Screening** – identify important factors (eliminate those not important), see basic relationships
3. **Modeling** – capture relationship between inputs and response(s) in functional form
4. **Model Refinement** – make sure model has sufficient precision for what is needed
5. **Optimize** – use model to solve problem, optimize system performance
6. **Confirm** – verify results are robust in environment that they will be used



# EASY DOE – HELP WITH DESIGN AND ANALYSIS

Guided Mode  Flexible Mode

Define Model Design Data Entry Analyze Predict Report

- **Key characteristics:**

- Assumes you want to do screening or modeling
- Will guide you through all steps of building, running and analyzing the experiment

- **Purpose:**

- Exploration & Modeling

Guided Mode Flexible Mode

Define Model Design Data Entry Analyze Predict Report

▶ Responses

▼ Factors

Role

The factor  can  
▶ S

lies  
How  
▶ S

Model type

Number of Runs

Guided Mode Flexible Mode

Define Model Design Data Entry Analyze Predict Report

Response(s)		Design			Number of Runs	
1/0 Cols	Y	3/0 Cols	X1	X2		X3
12/0 Rows	1	1	-1	1	1	8
	2	2	-1	1	-1	12
	3	3	-1	-1	1	12
	4	4	-1	-1	-1	16
	5	5	-1	1	-1	12
	6	6	-1	1	1	12
	7	7	1	1	1	12
	8	8	1	-1	1	12
	9	9	1	1	-1	12
	10	10	1	-1	-1	12
	11	11	1	-1	-1	12
	12	12	1	-1	1	16

Export Data Load Response

# OUTLINE

## 1. What are we trying to accomplish with our experiment?

Understanding what the goal is of the experiment will help to match possible choices to our experiment

## 2. What are the common DOE choices in JMP?

Quick walk through of some of JMP's most popular design of experiment choices

## 3. Key Questions to consider before generating a design

What do we already know about the factors, responses and their relationship?

What are the constraints under which we need to operate?

# WHAT DO WE ALREADY KNOW?

## Factors:

- Which ones?
  - Too many → experiment will need to be large to understand them all
  - Too few → possibility of missing something important
- What type are they?

Continuous  
Discrete Numeric  
Categorical  
Blocking  
Covariate  
Mixture  
Constant  
Uncontrolled

- What ranges / values for each?
  - Too big → difficult to understand what is happening, miss a subtle feature
  - Too small → miss target location, effect of factors look very small
  - Wrong location → miss target location

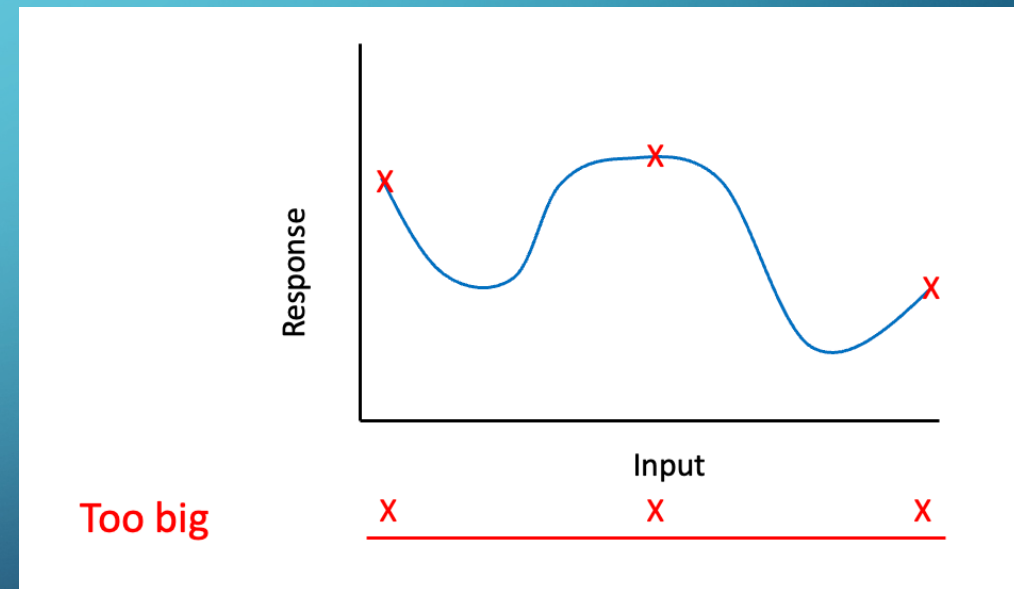
# WHAT DO WE ALREADY KNOW?

## Factors:

- Which ones?
  - Too many → experiment will need to be large to understand them all
  - Too few → possibility of missing something important
- What type are they?

Continuous  
Discrete Numeric  
Categorical  
Blocking  
Covariate  
Mixture  
Constant  
Uncontrolled

- What ranges / values for each?
  - Too big → difficult to understand what is happening, miss a subtle feature
  - Too small → miss target location, effect of factors look very small
  - Wrong location → miss target location



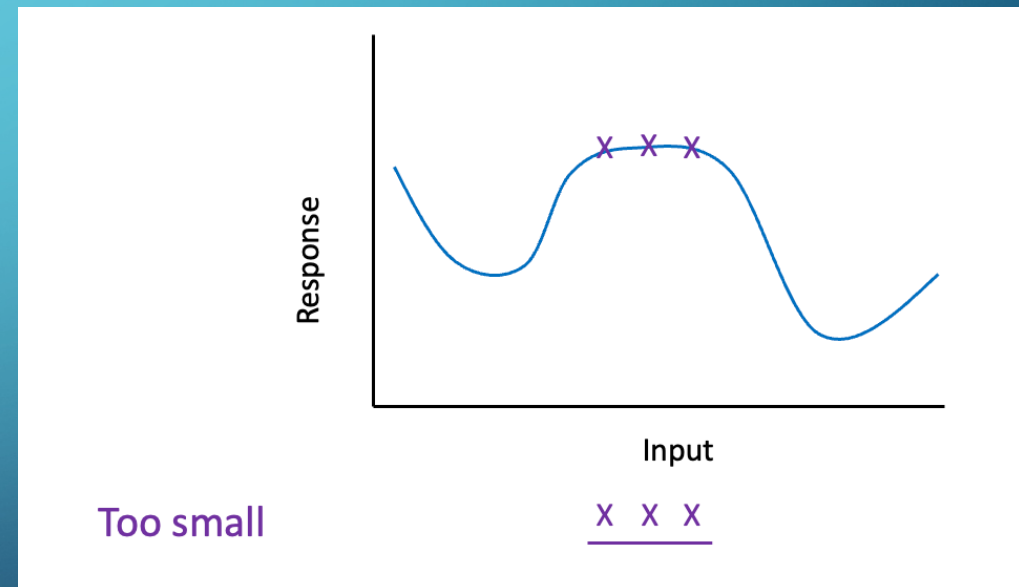
# WHAT DO WE ALREADY KNOW?

## Factors:

- Which ones?
  - Too many → experiment will need to be large to understand them all
  - Too few → possibility of missing something important
- What type are they?

Continuous  
Discrete Numeric  
Categorical  
Blocking  
Covariate  
Mixture  
Constant  
Uncontrolled

- What ranges / values for each?
  - Too big → difficult to understand what is happening, miss a subtle feature
  - Too small → miss target location, effect of factors look very small
  - Wrong location → miss target location



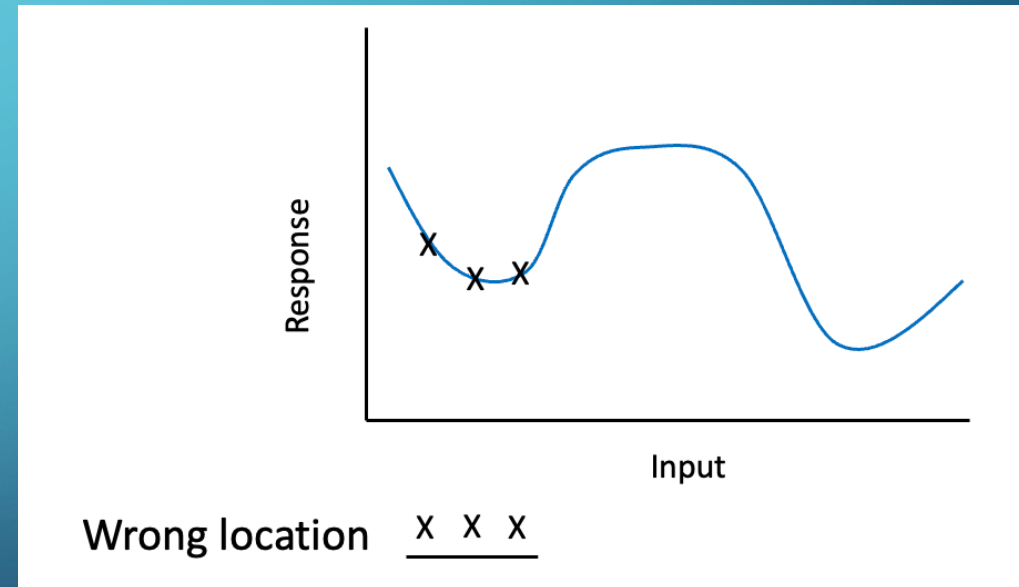
# WHAT DO WE ALREADY KNOW?

## Factors:

- Which ones?
  - Too many → experiment will need to be large to understand them all
  - Too few → possibility of missing something important
- What type are they?

Continuous  
Discrete Numeric  
Categorical  
Blocking  
Covariate  
Mixture  
Constant  
Uncontrolled

- What ranges / values for each?
  - Too big → difficult to understand what is happening, miss a subtle feature
  - Too small → miss target location, effect of factors look very small
  - Wrong location → miss target location



# WHAT DO WE ALREADY KNOW?

Relationship between inputs and responses:

- Which responses are we interested in? [COMMON MISTAKE: forgetting an important response]
- What do we know about the relationships?
  - Continuous
  - Smooth
  - Complexity
    - First-order (main effects) – common for screening
    - Interactions
    - Curvature

[COMMON MISTAKES: - assuming you know too much  
- not designing for most complicated relationship]

- First key decision: model-based (confident in smooth continuous, not too big a region)  
or space-filling (not sure what to expect, large region, protects against surprises)

Anderson-Cook and Lu, 2021



# CONSTRAINTS?

- Input regions / combinations where responses not possible?  
not of interest?
- Budget?
  - Time
  - Cost

} Better to think in terms of ranges for # of runs, and then compare several designs

## Define Factor Constraints

- None
- Specify Linear Constraints
- Use Disallowed Combinations Filter
- Use Disallowed Combinations Script

The screenshot shows the JMP Design Explorer interface. It is divided into several sections:

- Factors:** A table with columns for Name, Role, Values, and Units. It lists three factors: X1, X2, and X3, all with a Continuous role and values ranging from -1 to 1.
- Model:** A table with columns for Name and Estimability. It lists terms from Intercept to X1\*X3, all with a Necessary estimability.
- Alias Terms:** A section for defining alias terms.
- Design Explorer Options:** A section for selecting options for a single design or all combinations. It includes a Criterion dropdown (set to D-Optimality), a Runs input (16), Center Points (0), Replicates (0), and Random Starts (5). There are also checkboxes for D-Optimality, A-Optimality, I-Optimality, and Alias Optimality, and a 'Generate All Designs' button.

JMP White Paper: Benefits of considering several different design sizes  
Anderson-Cook, 2022

# HELPFUL RESOURCES

Types of Designed Experiments:

[https://www.jmp.com/en\\_in/statistics-knowledge-portal/what-is-design-of-experiments/types-of-design-of-experiments.html](https://www.jmp.com/en_in/statistics-knowledge-portal/what-is-design-of-experiments/types-of-design-of-experiments.html)

Model-Based versus Space-Filling:

Anderson-Cook, C.M. Lu, L. (2021) “The First Fork in the Road” Quality Progress 54(11) 48-51.

Considering and comparing several design sizes and types:

JMP White Paper: Anderson-Cook, 2022

[https://www.jmp.com/en\\_us/whitepapers/jmp/choosing-the-right-design.html](https://www.jmp.com/en_us/whitepapers/jmp/choosing-the-right-design.html)

The Why and How of Asking Good Questions:

JMP White Paper: Anderson-Cook, 2023

[www.jmp.com/asking-good-questions](http://www.jmp.com/asking-good-questions)