Comparison of Industrial Screening Experiment Designs

Lou Johnson – JMC Data Experts Inc

ASQ World Conference – Session M05 – April 30, 2018

ASQ
The Global Voice of Quality

Learning Objectives

- Understand the strengths and weaknesses of definitive screening, fractional factorial, and Plackett-Burman designs for screening input variables to your process.
- Determine the trade-offs between run efficiency, power, and confounding for definitive screening, fractional factorial, and Plackett-Burman screening designs.
- Compare the performance of these three screening designs when executed on the same seven-factor variable space.
- Apply the concepts of run efficiency, power, and confounding to select the screening experiment design applicable to your process scenario.

ASQ

Non-disclosure Statement

"YOU CAN TELL ME ANYTHING, MRS. ROJAK. I'D NEVER VIOLATE THE SACRED FISH STORE-CUSTOMER RELATIONSHIP."

All industrial experiments, results and scenarios are based on the authors' actual experiences. Data units, variable names, etc. have been changed for demonstration purposes only.

The Problem with Screening Designs

Peter Wilson
Post Doctoral Research Fellow at WMG, University of Warwick ••• 3w

Definitive screening designs with a hard to change factor

I have a process I'm trying to screen and then optimise. This process has 10 factors, and one of them is hard to change (full day turn around to change factor). I've been using Minitab, any ideas on how best to approach my problem?

I want a 10 factor Definitive Screening Design

Jean-Luc HEINRICH Warning !!! Definitive screening designs has been tested by different experts and found not so robust as claimed by authors.

Watch out for those Definitive Screening Designs

Josef Betschart From my experience, you will not get a reliable result which will meet accadademic standards if you do an experiment with 10 factors and 40 runs!!!!
I would recomend to gain more process knowledge in order to reduce factors. From my experience I learned that the biggest potential lays hidden in interaction of factors. If you still want to run this DoE and you can't exclude some interactions you must conduct a minimum of 70 runs.

You need 40 runs and Res IV because of interactions

Phil Kay Hi Josef, it sounds like you have had some interesting experience in a situation that is in in some way similar to Peter's. But perhaps this is a case of "excess of conviction relative to knowledge" (one of Tim Geithner's favourite phrases, according to the book Stess Test). I don't

You might not know what your talking about

Harry Rever Hello Peter, I'd suggest using a 16 run Plackett-Burman design. This would allow you to either add a few more test factors or simply use the blank columns to help with the analysis and you would get some information about interactions. (you could use a 12 run PB design but you would not get any information about interactions). Simply block on your hard to change factor. In other words, for the 16 run PB design, simply randomize the 8 "-" level runs first, run those treatments, then randomize the 8 "+" level runs, and run those second. This "selected randomization" is something you have to do with hard to change factors. Depending on your available data, I'd try to run 2 or 3 replicates, if possible. I've done this many times and it works great. Then, refine using a full factorial design to validate results and fine tune results.

Do a Plackett-Burman in just 16 runs

Phil Kay Frank, your question is getting away from the original question. Probably a good one to post in the JMP Community. However, in general you can create a DSD for the easy-to-change factors and use the resulting table as covariate factors in th... **Show more**

I think we are getting away from the original question.

ASQ

JMC Data Experts

Agenda

- Criteria to Compare Screening Designs
- Case Studies – Examples of each Design
  - Resolution IV Factorial; Calcium Smelting Process
  - Plackett-Burman; Post Treatment of Metal Cast Parts
  - Definitive Screening; pDNA Fermentation Process Yield
- Green Bean Seedling Experiment Comparison
- Strengths of Each Design
- Five Tips for any Screening Experiment
- Question and Answer Discussion

Comparing Apples and Oranges



I love bananas!

Oranges have Vitamin C

Fall apples are cheaper

Pears Nutrition Facts

Serving Size: 3.5 ounces (100 grams), raw

	Amt. Per Serving
Calories	58
Calories from Fat	1
Total Fat	0 g

Pears have low calories

Four Criteria for Comparing Designs*

Screening Experiment – smallest possible factorial experiments designed to determine the largest main effects for 6 - 20 factors assuming three-way interactions are negligible and many of the factors are unimportant.

- **Confounding –** confusion caused by overlapping effects. This is required to reduce the number of runs.
- **Run Efficiency** – the number of experimental runs and their information value
- **Projection** – the value of the information about the important factors after removing unimportant factors
- **Power** – the probability your experiment will detect an effect if that effect actually exists

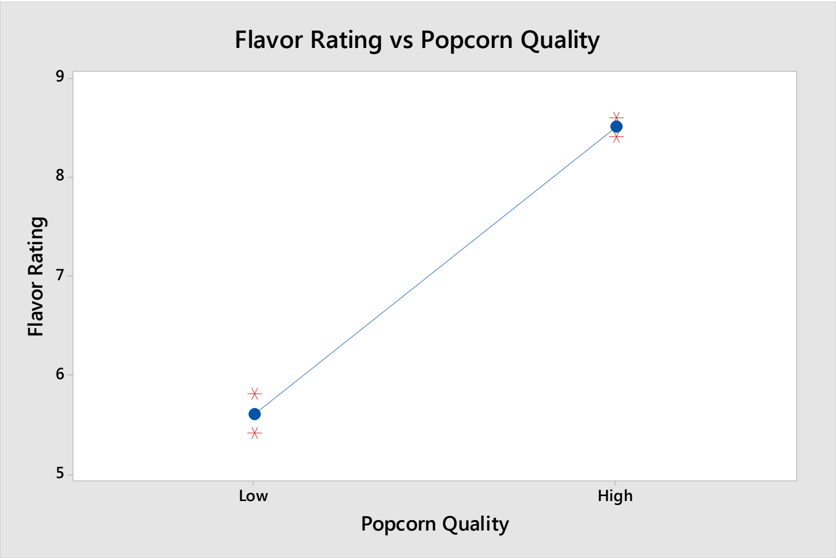
*Jones, B and Nachtsheim, C. (2011)

Defining Projection for a Screening Design

Projection – assuming that some effects are going to be removed from the analysis because they are unimportant, what does the dataset tell us about the factors that remain.

Popcorn Worksheet - Projections ***

↓	C1-T	C2-T	C3
	Quality	Rub	Flavor
1	High	Some	8.6
2	Low	Some	5.4
3	High	None	8.4
4	Low	None	5.8



Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value
Constant		7.050	*	*	*
Quality	2.900	1.450	*	*	*
Rub	-0.10000	-0.05000	*	*	*
Quality*Rub	0.3000	0.1500	*	*	*

Coded Coefficients

Term	Effect	Coef	SE Coef	T-Value	P-Value
Constant		7.050	0.112	63.06	0.000
Quality	2.900	1.450	0.112	12.97	0.006

ASQ

JMC Data Experts

Effect of the Design on Power

Power – an effect may test as insignificant if its size was too small, the noise was so large it was covered up or the data points collected did not do a good job of measuring the effect. Power is a combination of all three of these considerations.

How much information (data points) is collected

How well the data points cover the design space

t = Function of : [Effect] / (Standard Deviation)

$df_{error} = [N – (main\ effects) – 1]$

Degrees of Freedom for error come from data points not used to estimate effects and from replicates

Properties of Resolution IV Screening Experiments

- **Confounding –** Main effects are clear. Two-way interactions are confounded with other two-way interactions

  A

  AB + CE + FG   or   CG – EF   (16 or 32 runs)

- **Run Efficiency** – commonly 8, 16 or 32 runs

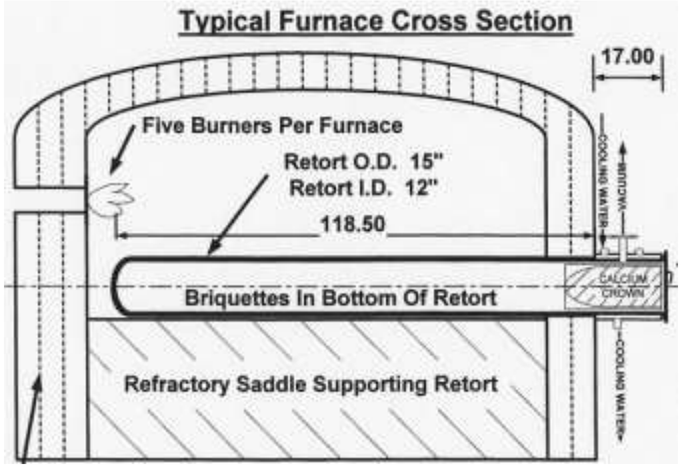
- **Power** – higher than PB or DSD due to a higher number of runs and good coverage of the design space.

- **Projection** –
  - ❖ full factorial in any 3, 4, or 5 factors for 8,16 or 32 runs
  - ❖ Res V in any 5 factors in 16 runs
  - ❖ Res IV in any 6, 7, or 8 factors in 16 runs

Example of a Resolution IV Screening Experiment

Aluminothermic process (High temperature vacuum reduction calcium oxide)

Calcium metal is produced by the aluminothermic method in which aluminum is used as reducing agent and calcium oxide is reduced in vacuum at high temperature. Calcium oxide is obtained from calcium carbonate. Calcium oxide is grounded to small particles and dry blended with the desired amount of finely divided aluminum. To ensure good contact, this mixture is compacted into briquettes or retorts (these are horizontal tubes made up of heat resistant steel). It is heated to 1100-1200°C. The open ends of retort protrude from the furnace and cooled by water jacket to condense the calcium vapor. Retorts are sealed and evacuated to pressure less than 13 Pa. After 24 hours of reaction vacuum is broken by Argon and 99% pure calcium is recovered and calcium aluminate residue is removed.



Experiment Goal – determine process to maximize calcium yield per lb charge.

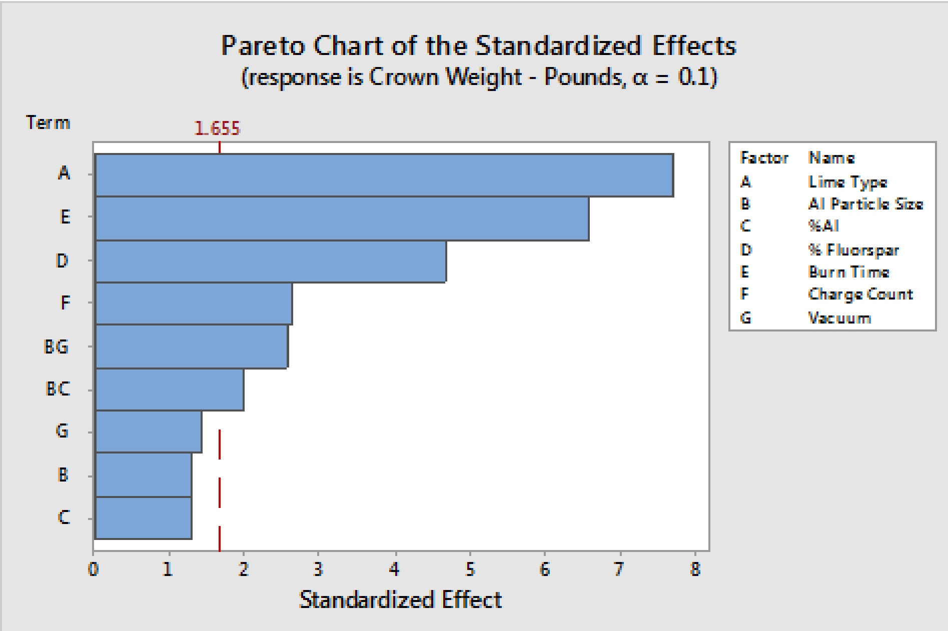
Experimental Factors

1. Lime; 1 - 2%
2. Al Particle Size; Fine - Course
3. Al Content %; 17.7 – 26.2%
4. Flourspar %; 0 – 6%
5. Burn Time; 30 – 60 min
6. Charge; 6 – 7 count
7. Vacuum; 2260 – 2415 mTorr

Experiment Plan

- Resolution IV – $2^{(7-3)}$ Fractional Factorial, 16 runs

Results of ResIV FF Calcium Smelting Example

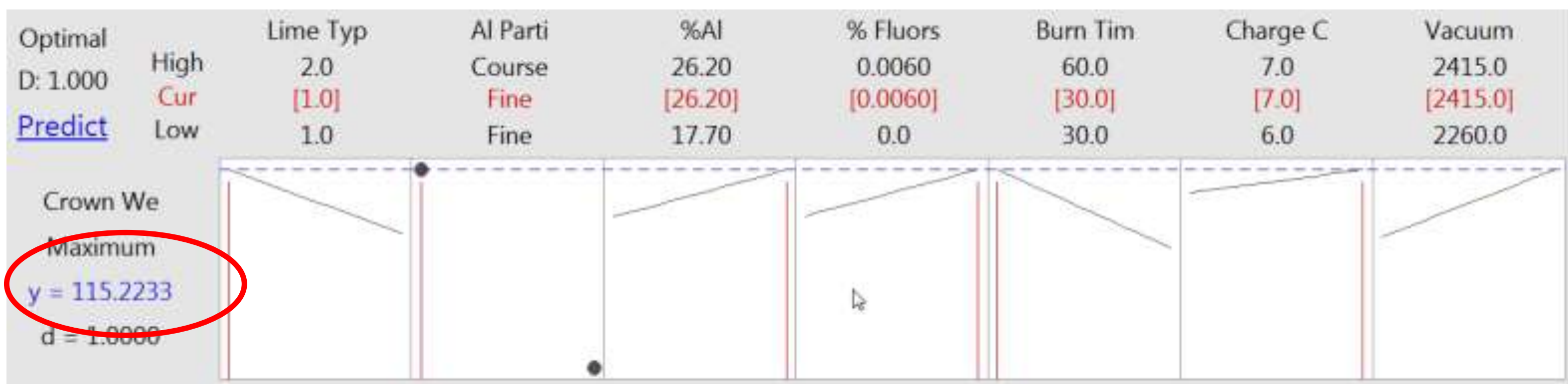


- Resolution IV – 2^{7-3} Fractional Factorial, 16 runs

Four main effects are significant, Lime Type, Burn Time, % Flourspar, Charge Count. Plus two interactions.

Using process below, Yield = 115.2 lb compared to 106.7lb historical AVG

8% Yield Improvement



Properties of Plackett-Burman Experiments

- **Confounding –** Main effects *partially* confounded with two-way interactions. Two-way interactions can not be estimated.

  A - 0.33 BC + 0.33 BD + 0.33 BE - 0.33 BF - 0.33 BG …
  Two-way interactions partially confounded w/ each other.

- **Run Efficiency** – 8 + 4x (x = 1,2,3..), commonly 8, 12, 20.
  at least N runs for (N – 1) factors

- **Power** – less than Res IV FF due to low df for error estimate and bias of main effect estimates as shown above.

- **Projection** –
  - ❖ full factorial in any 2 factors regardless of N
  - ❖ Partial confounding of main & two-ways for 3,4 or 5 factors (each with 1/3 two-way interaction)

Placket-Burman Example

C1	C2	C3	C4	C5-T	C6-T	C7-T	C8-T	C9-T	C10-T	C11-T	C12	C13
StdOrder	RunOrder	PtType	Blocks	Initial	Bead Size	Pressure Trt	Heat Trt	Cooling	Polish	Final Trt	Life	
2	1	1	-1	Treated	Small	HIP	Age	Slow	Mechanical	Peen	5.818	
9	2	1	1	As is	Large	HIP	Anneal	Rapid	Mechanical	Peen	5.863	
14	3	1	1	Treated	Large	None	Age	Rapid	Mechanical	None	6.058	
4	4	1	1	Treated	Large	HIP	Anneal	Slow	Chemical	Peen	5.899	
12	5	1	1	As is	Large	HIP	Age	Slow	Chemical	None	4.625	
16	6	[illegible]	1	Treated	Large	None	Anneal	Slow	Mechanical	None	[illegible]	

Experiment Goal: determine process to maximize field life of welded castings.

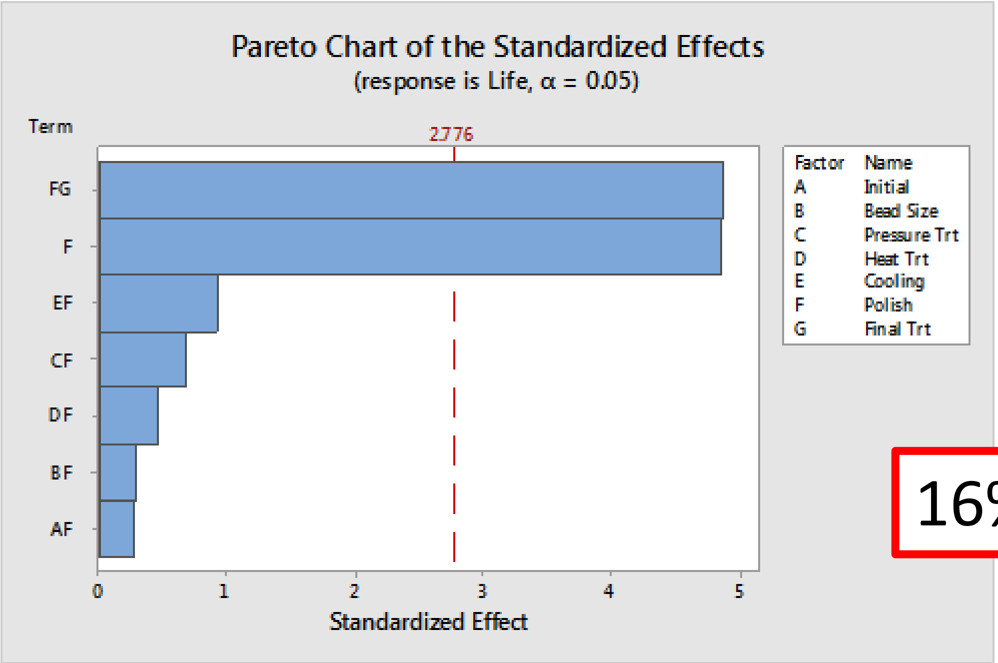
Experiment Plan

- Resolution III, 7 Factor
- Plackett-Burman, 12 runs

Experimental Factors

A. Initial structure; As is - Treated
B. Bead Size; Small - Large
C. Pressure Treat; None - HIP
D. Heat Treat; Anneal – Age
E. Cooling; Slow – Rapid
F. Polish; Chemical – Mechanical
G. Final Treatment; None – Peen

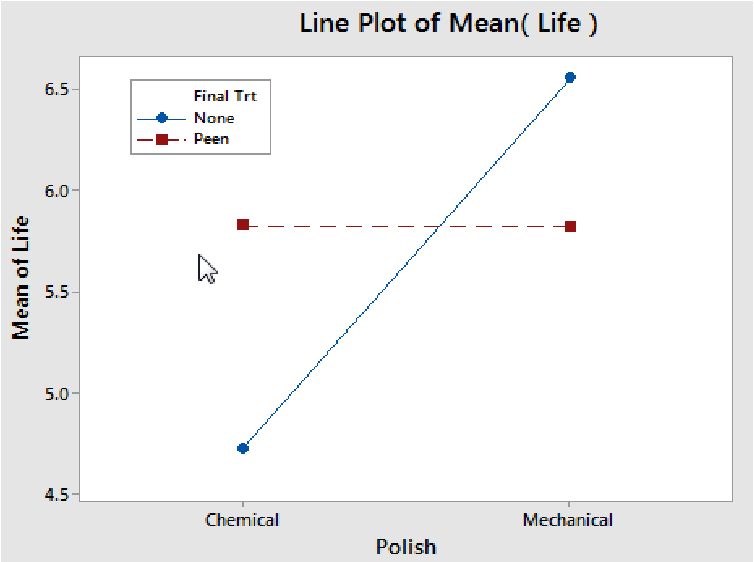
Results of Placket-Burman Casting Experiment



One significant main effect, Polish.

Projection properties of the design allow estimation of two-way interactions with Polish. Polish * FinalTrt is key.

16% Fatigue Life Improvement



Regression Equation in Uncoded Units

Life = 5.73 + 0.458 Polish + 0.459 Polish*Final Trt

Properties of Definitive Screening Designs

- **Confounding –** Main effects are clear. Square terms are partially confounded with two-way interactions. Two-way interactions can not be estimated.

  A
  AA + 0.26 AB - 0.26 AD + 0.08 AE - BC - 0.74 BD ….

- **Run Efficiency** – 2k + 1 runs for screening k factors, 1 extra run for each categorical factor

- **Power** – mathematically, slightly less than PB but practically much higher due to the fact that main effects are clear.

- **Projection** – different for odd or even k, but in general;
  - ❖ full quadratic model for any 2 or 3 factors

Definitive Screening Design Example

Minitab 7 factor design ***

↓	C1	C2	C3	C4	C5	C6	C7	C8	C9-T	C10	C11	C12
	StdOrder	RunOrder	PtType	Blocks	pH	O2 Flow	Temp	Feed	Media	OD600	%DO	Titer
1	6	1	2	1	6.8	120	41.0	3.5	Synthetic	20	10	1640
2	16	2	1	1	6.8	120	39.5	1.9	Crude	40	10	1536
3	14	3	2	1	6.8	80	42.5	1.9	Synthetic	40	20	1530
4	17	4	0	1	7.0	100	41.0	2.7	Synthetic	30	20	1600
[illegible]	18	5	0	[illegible]	[illegible]	[illegible]	[illegible]	2.7	[illegible]	[illegible]	[illegible]	[illegible]

Lonza

A Novel pDNA Platform

Experiment Goal – improve the yield of a plasmid DNA fermentation process

Experiment Plan

- Definitive Screening Design in 7 factors, 18 runs

Experimental Factors

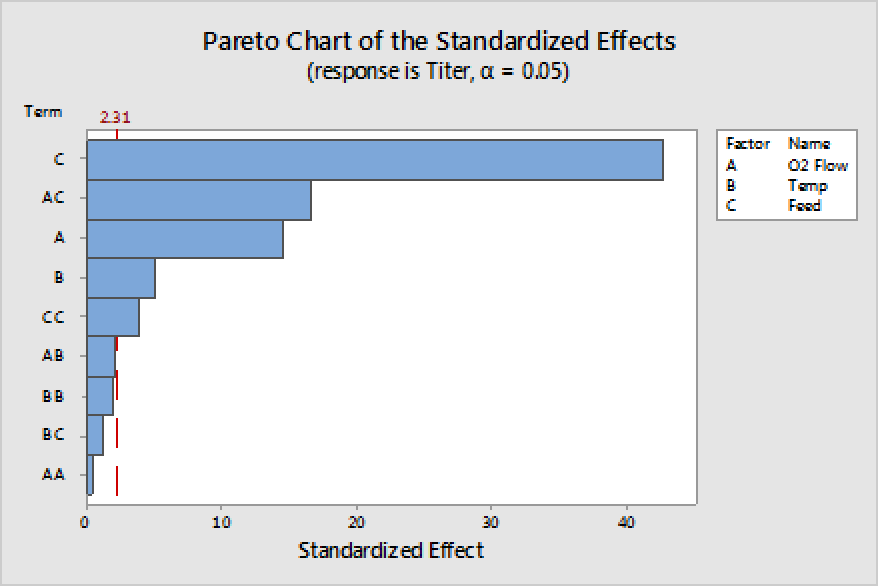
1. pH; 6.8 - 7.2
2. Oxygen Flow; 80 – 120 ml / day
3. Temperature; 39.5 – 42.5º C
4. Feed; 1.9 – 3.5
5. Media; Synthetic - Crude
6. OD600; 20 - 40
7. %DO; 10 – 30%

*Ormek, D., Hamilton, M., Chopra, M., et. al. (2012).

ASQ

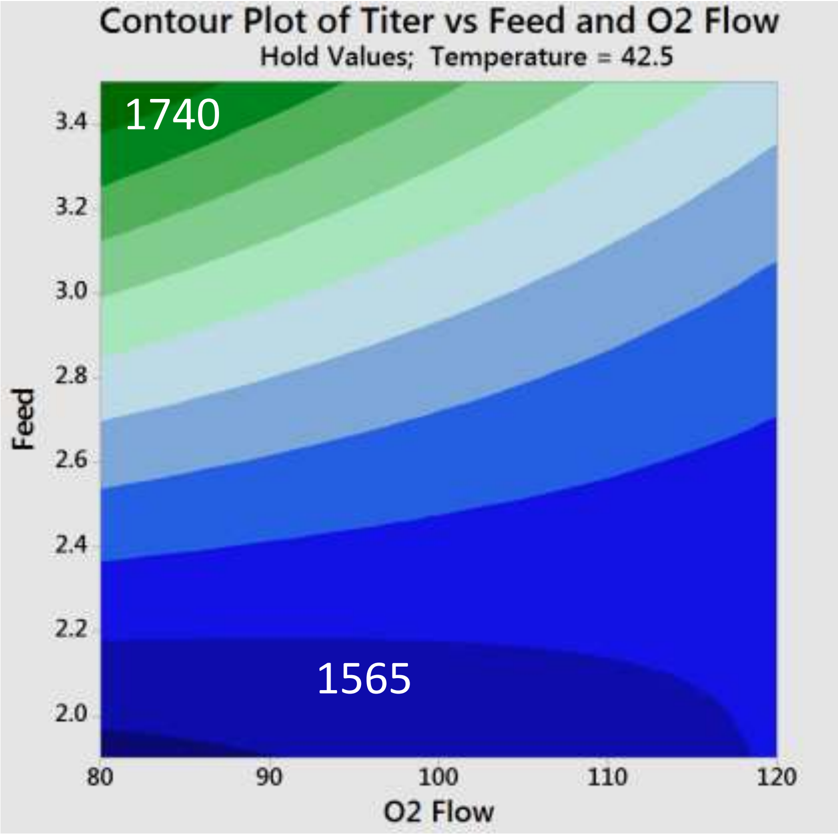
Data Experts

Results of Definitive Screening Example



Experiment Plan

- Definitive Screening Design in 7 factors, 18 runs



Results

- O_2 Flow, Temperature and Feed identified as major variables
- Model reduced to a 3 factor quadratic design
- (1740 – 1565) / 1653 = 11%

11% Yield Improvement from low to high

When Hobby and Expertise Combine

ReCafe at Windswept Farm

Organic Sustainable Farm to Table Cafe

Green Bean Seedling Screening Experiments

Experiment Goal – determine which factors have the strongest effect on germination rate (%) , average days to germination and plant growth.

Experiment Plans

- Resolution IV – 2^{7-2} Fractional Factorial, 32 runs
- Plackett-Burman, 12 runs
- Definitive Screening Design, 18 runs

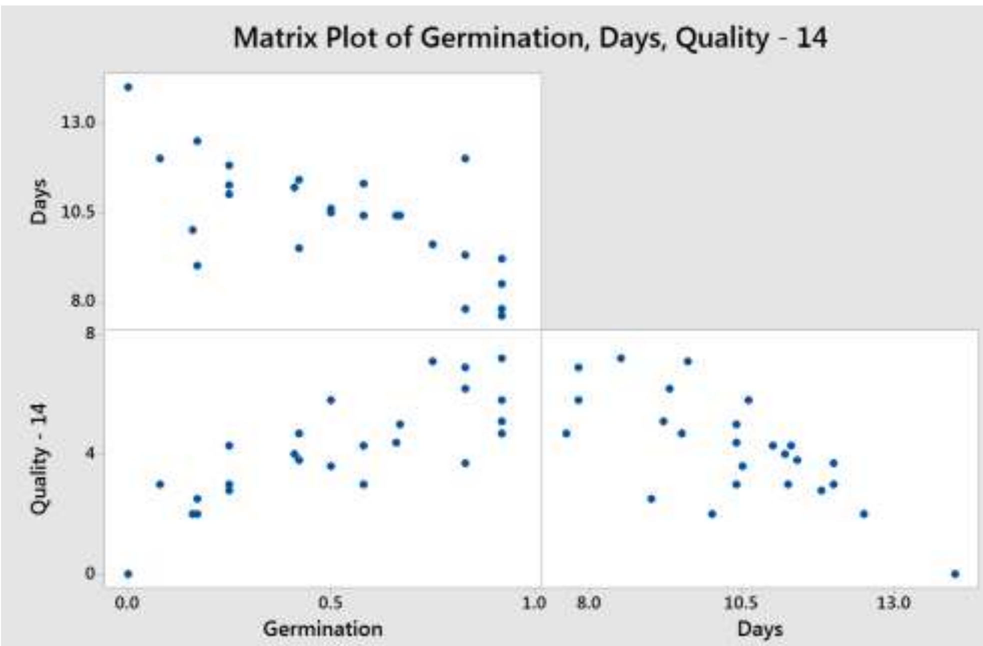
Experimental Factors

1. Seed Depth; ¾” – 1.5”
2. Watering; 1 / day – 1 / 2 day
3. Miracle Grow; None – Twice
4. Compost; None – 100% of soil
5. Vermiculite; None – 33% of soil
6. Seed Variety; Tendergreen – Blue Lake
7. Sunshine; Shade – Sun

Green Bean Experiments – Score Response

Green Bean Responses are Correlated

- Germination Rate (%)
- Days to Germination (avg days)
- Plant Quality after 14 days



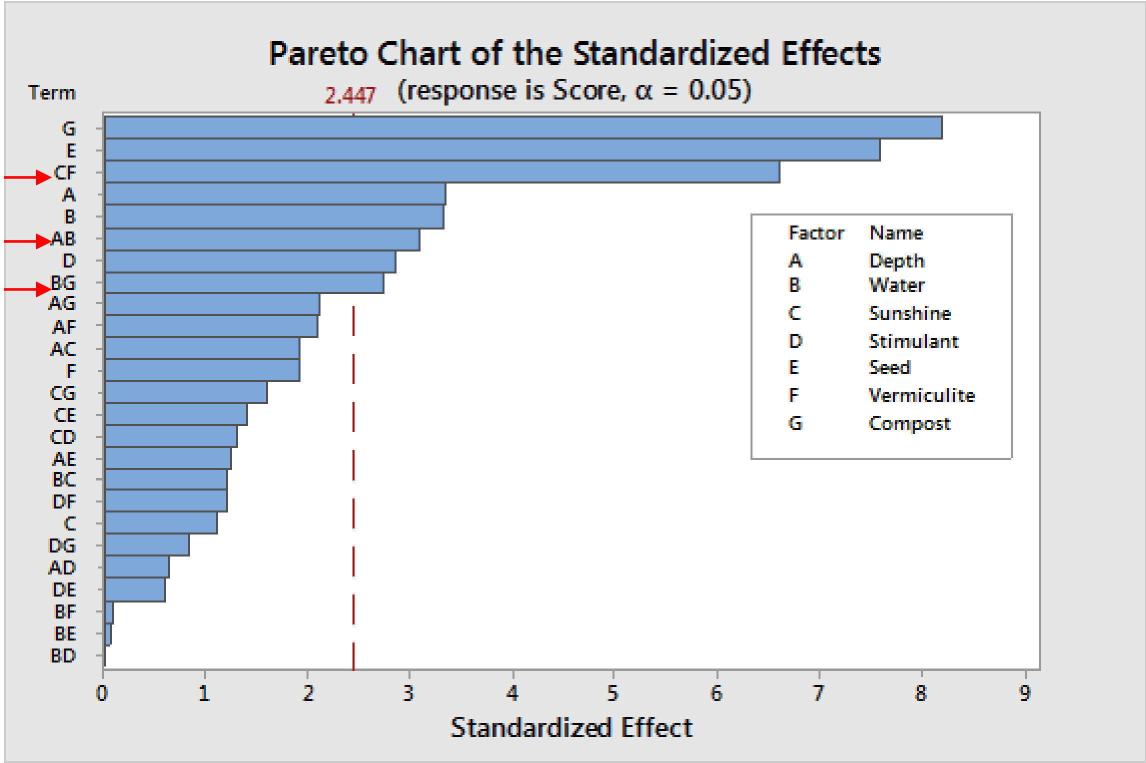
Combine to one response called Score using Principal Components

Eigenanalysis of the Correlation Matrix

Eigenvalue	2.7268	0.1706	0.1026
Proportion	0.909	0.057	0.034
Cumulative	0.909	0.966	1.000

About 97% of the variation in the responses is captured using the first two components

Green Bean Results – Res IV Fractional Factorial



Five main effects and three interactions are significant.

Two-way interactions are confounded with each other. What is the alias structure for the significant interactions?

CF – EG + ABD ?

AB + CDF + DEG Clear

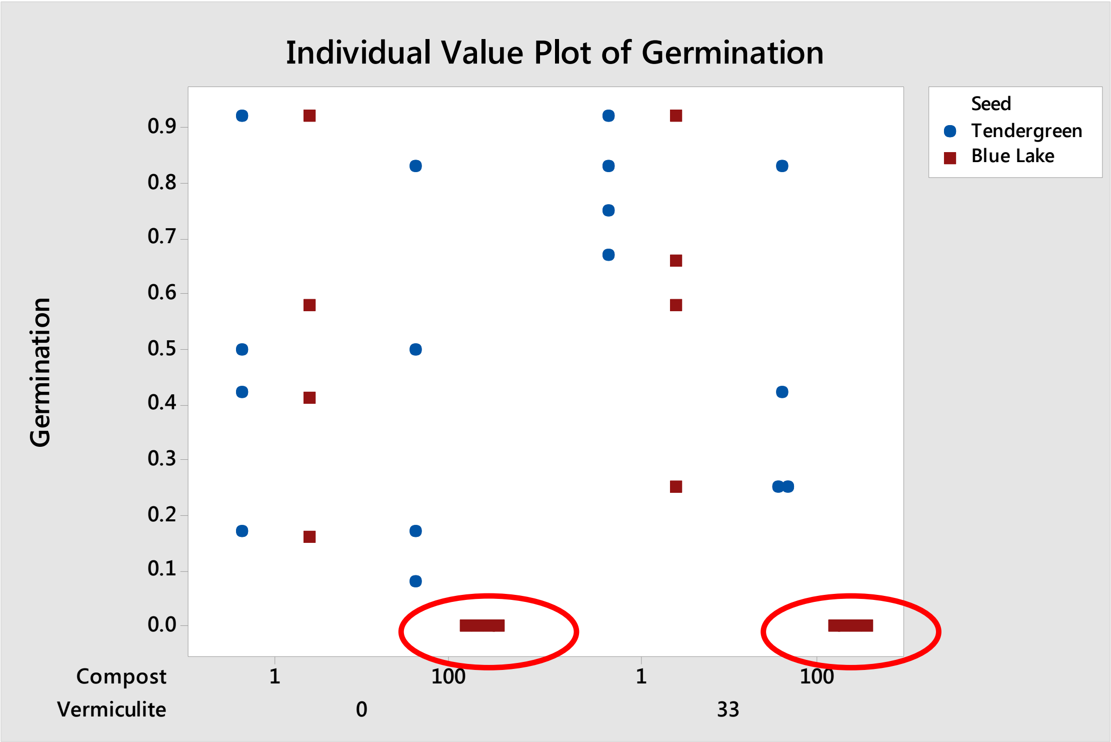
BG - ADE Clear

Sunshine * Vermiculite

or

Seed * Compost

Green Bean Results – Res IV Fractional Factorial



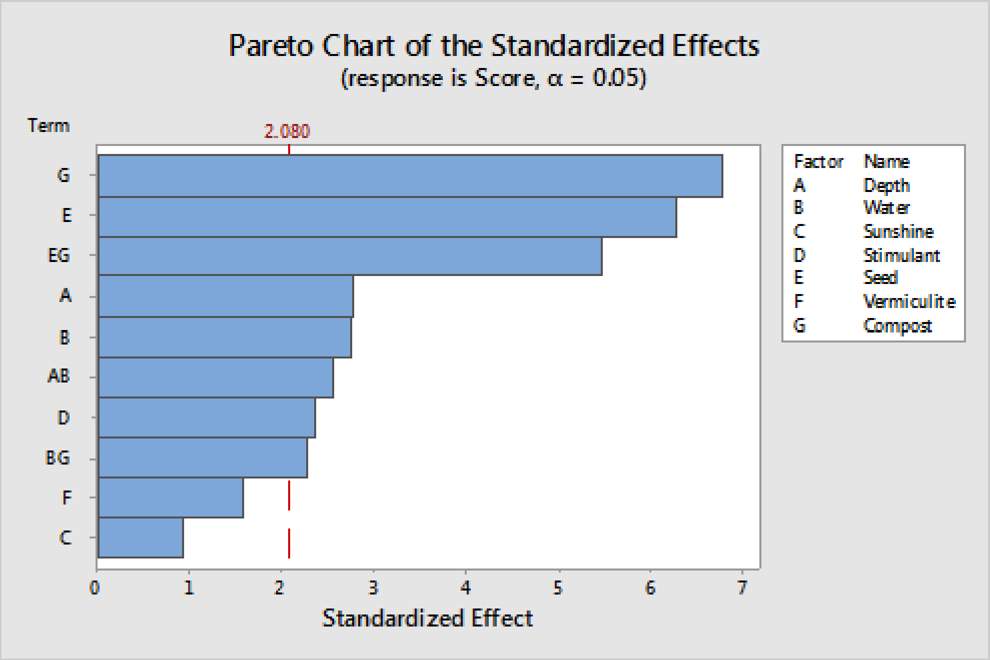
Seed * Compost Interaction

The Blue Lake seed variety clearly had problems in the 100% compost soil

Next Step

Reduce the model with the EG (Seed* Compost) interaction, AB, DG and main effects.

Green Bean Results – Res IV Fractional Factorial

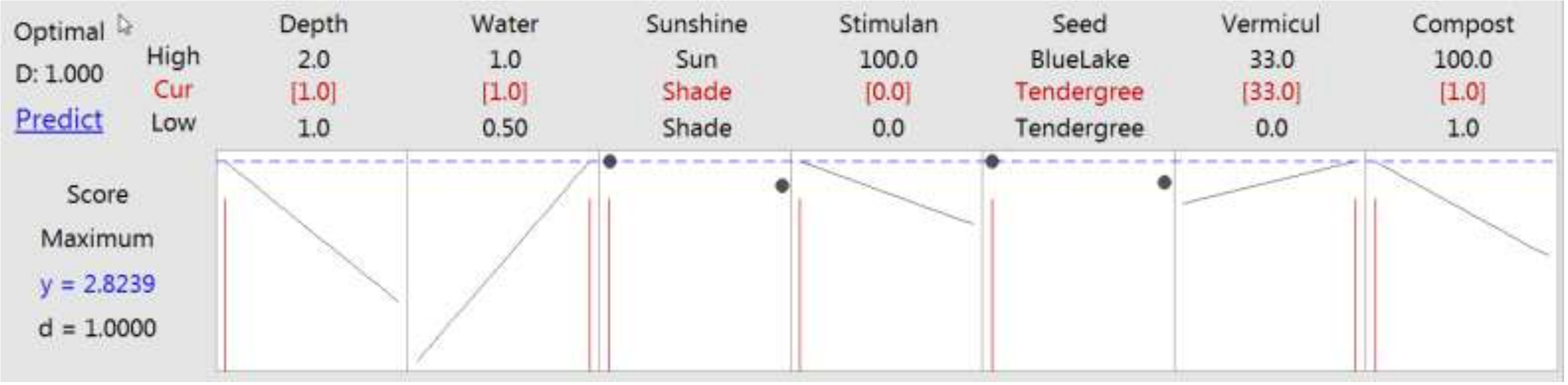


Five main effects and three interactions are significant.

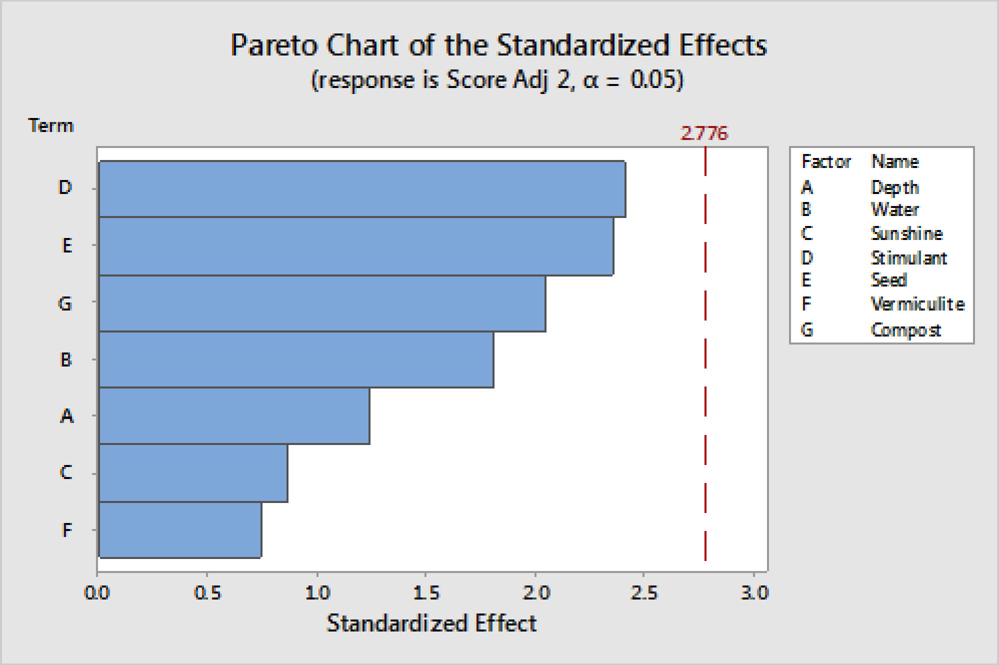
<u>Heredity Principle</u>
Two-way interactions are most likely found with significant main effects.

A, B, D, E and G are important as is EG, AB and BG.

We will look more closely at Water * Depth (AB) later.



Green Bean Results – Plackett-Burman



No main effects are significant.

Due to low power, some would select D, E and G as important.

Stimulant (D) is found to be the largest effect, but these main effect estimates are biased by partial confounding with interactions.

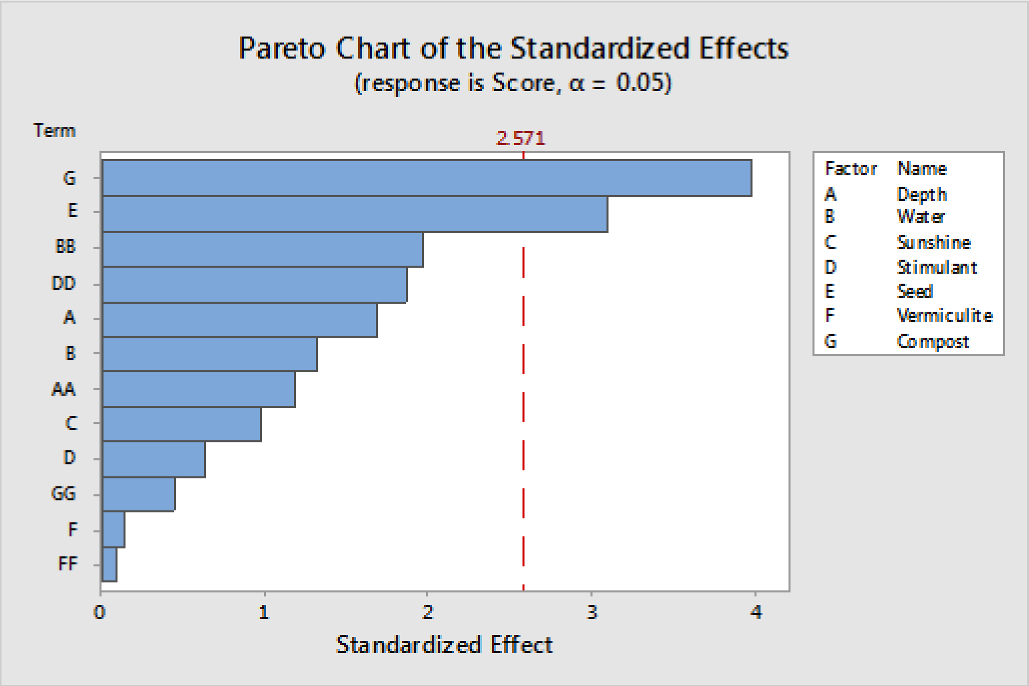
Good:

- A,B,D,E,G in the top 5
- C, F identified as least important
- Only 12 runs

Not Good:

- Interactions are not found
- Main effect ranking is biased
- Low power

Green Bean Results – Definitive Screening Design



Two main effects are significant.

Due to low power, some would select BB and DD as important.

BB + 0.44 AB - 1.06 AC + 0.85 AD - 1.06 AE + 0.85 AF + 0.85 AG + 0.06 BC + 0.15 BD + 0.06 BE + 0.15 BF + 0.15 BG + 1.06 CD - 1.06 CE + 1.24 CF + 0.65 CG + 1.24 DE - 0.85 DF - 0.85 DG + 0.65 EF + 1.06 EG - 0.85 FG

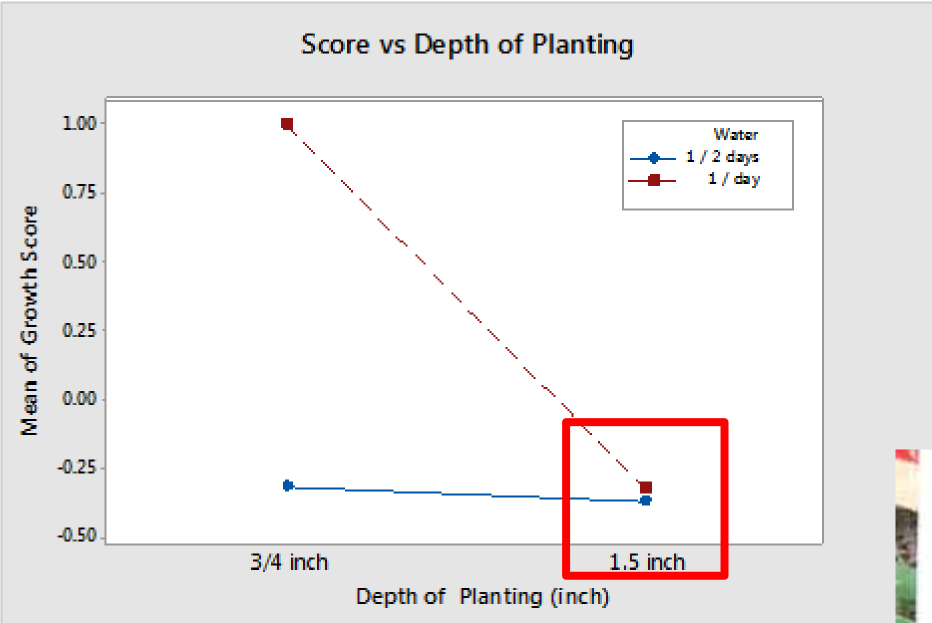
Good:

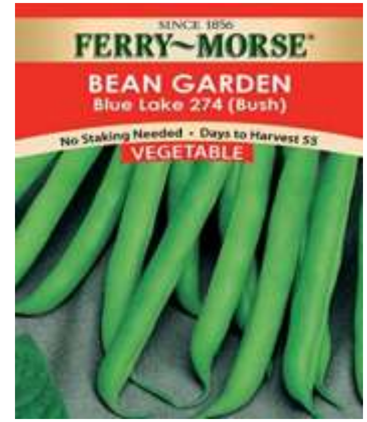
- A,B,E,G in the top effects
- C, F identified as less important
- Only 18 runs

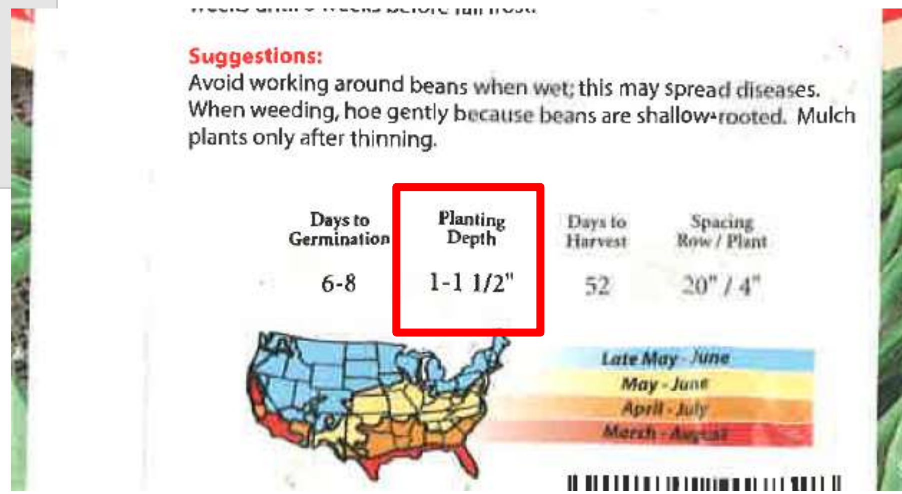
Not Good:

- Square and interaction effects are partially confounded.
- Interactions not identified.

Never Underestimate your Ability to Learn







" I told them the 1 ½ " was wrong!
We use 1" max depth. "

Erica - Technical Support
McKenzie Seed Company

Comparing the Three Design Approaches

- **Plackett-Burman –** Lowest number of runs, appropriate when cost is critical. Works well when the sparsity of effects principle applies and interactions are few and small.

- **Resolution IV Fractional Factorial** – useful when characterization is a goal as well as screening. Best choice when interactions are anticipated. Can require higher run count.

- **Definitive Screening Designs** – Medium run count. Main effects are clear so excellent for screening. Useful for biological and chemical systems where quadratic effects might be important.

Tips for Using any Screening Design

- **Remove known factors –** to discover important variables, leave known factors out. This reduces the size and complexity of the screening experiment. A follow-up experiment will study new and known factors in detail.
- **Use conservative set-points –** one factor set-point that moves the process "off the cliff" will derail 40 - 50% of your runs. Balance this with the "be bold approach" to increase power. Mee, Robert, W. (2009).
- **Multiple center points is not the best idea** – this is common because it allows error estimation and a test for curvature. However, I find additional factorial design points are more powerful and provide the same value.
- **When is the turnip out of blood? –** " If the analysis finds several incompatible models, it is a strong indication that the information provided by the data and design are limited and no analysis method can distinguish between them". Wu, J. F. and Hamada, M. (2012).
- **Advance model fitting techniques can be misleading** – modern computer techniques can find best fit models based on statistical criteria that make no sense based on first principles. Adams, Wayne. (2016).

Learning Objectives

- Understand the strengths and weaknesses of definitive screening, fractional factorial, and Plackett-Burman designs for screening input variables to your process.
- Determine the trade-offs between run efficiency, power, and confounding for definitive screening, fraction factorial, and Plackett-Burman screening designs.
- Compare the performance of these three screening designs when executed on the same seven-factor variable space.
- Apply the concepts of run efficiency, power, and confounding to select the screening experiment design applicable to your process scenario.

References

Adams, Wayne, (2016) "Definitive Screening Designs: Characteristics and Analysis Methods" , ASQ Statistics Division Webinar. https://www.youtube.com/watch?v=vHy2fknHxHc

Jones,B., Nachtsheim, C., " A Class of Three-Level Designs for Definitive Screening in the Presence of Second-Order Effects," January, 2011, *Journal of Quality Technology.*

Mee, Robert, W. (2009). *A Comprehensive Guide to Factorial Two-Level Experimentation.* Springer Science and Business Media, London – New York, p182.

Montgomery, D. C. (2013). *Design and Analysis of Experiments*, 7th ed. Hoboken, NJ: John Wiley & Sons.

Ormek, D., Hamilton, M., Chopra, M., (2012). Characterization of a Biomanufacturing Fermentation Process Using and New DOE Approach: Definitive Screening Designs, *BioProcess International*, http://www.bioprocessintl.com/wp-content/uploads/2012/11/Expression___Produc_185693a.jpg

Plackett, R.L. and Burman, J.P., "The Design of Optimum Multifactorial Experiments", *Biometrika* , 1946 , 33(4), p305-325.

Tamhane A. C. (2009). *Statistical Analysis of Designed Experiments; Theory and Application,* Hoboken, NJ: John Wiley & Sons, p319 - 321.

Wu J. C. F., and Hamada M. S., (2009). *Experiments: Planning, Analysis and Optimization, 2nd Edition.* Hoboken, NJ: John Wiley & Sons, p362.

The Global Voice of Quality®

Questions?

Lou Johnson – Email: lou@jmcdataexperts.com

Datasets in Minitab 18 at www.dataexperts.com