StatEase 40 YEARS

Make the most from every experiment![SM]

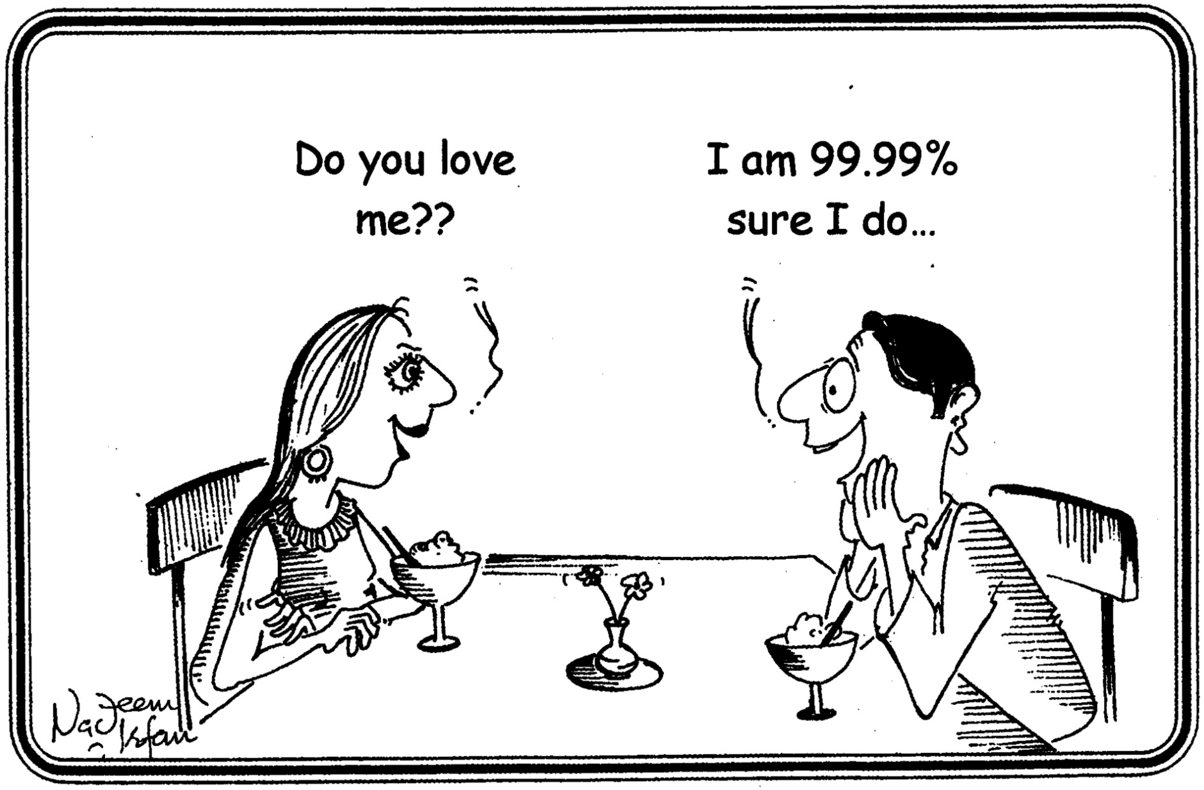
INDUSTRY, ENGINEERING & MANAGEMENT SYSTEMS CONFERENCE

Multifactor Design of Experiments (DOE) for Rapid Process Improvement

By Mark J. Anderson, PE, CQE, MBA
Engineering Consultant
Stat-Ease, Inc.
mark@statease.com

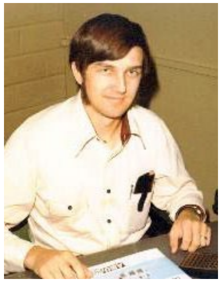
Trying Not to Go Too Far with Statistics

**Meaning never having to say you are certain*



Talking Points

Based on 50+ years of DOE experience

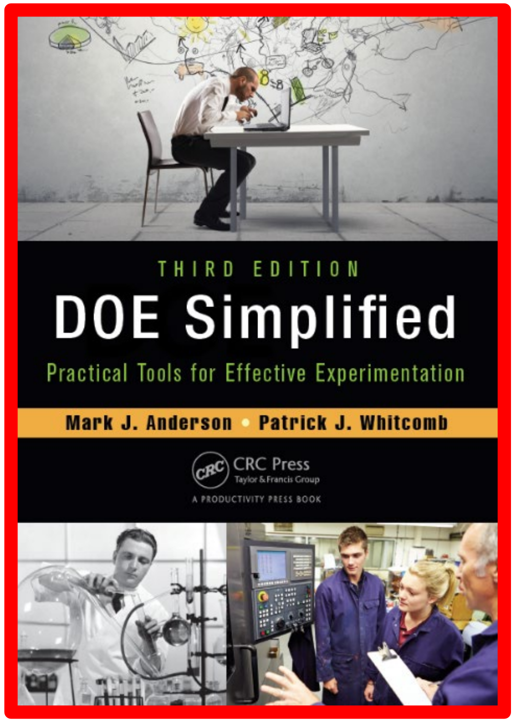


- Multifactor testing versus one-factor-at-a-time (OFAT)
- The evolution of DOE going back to early 1900's
- Case studies that lay out core tools
- Industrial DOE success stories

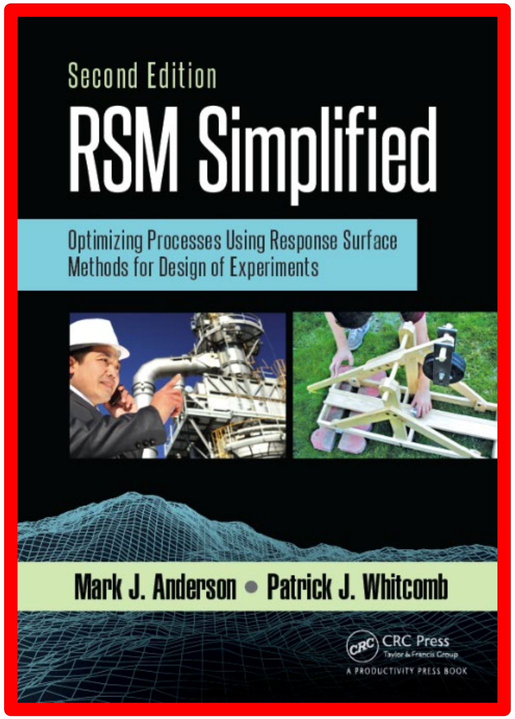
My Payback: Practical Paperbacks*

KISMIF: Keeping it simple and making it fun!

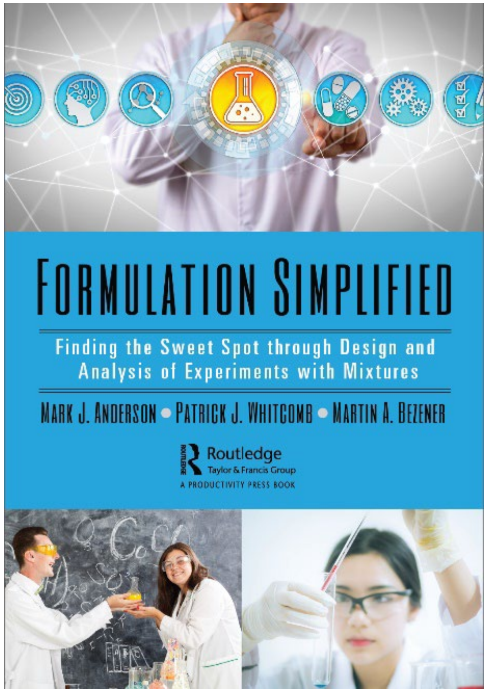
3rd edition



2nd edition



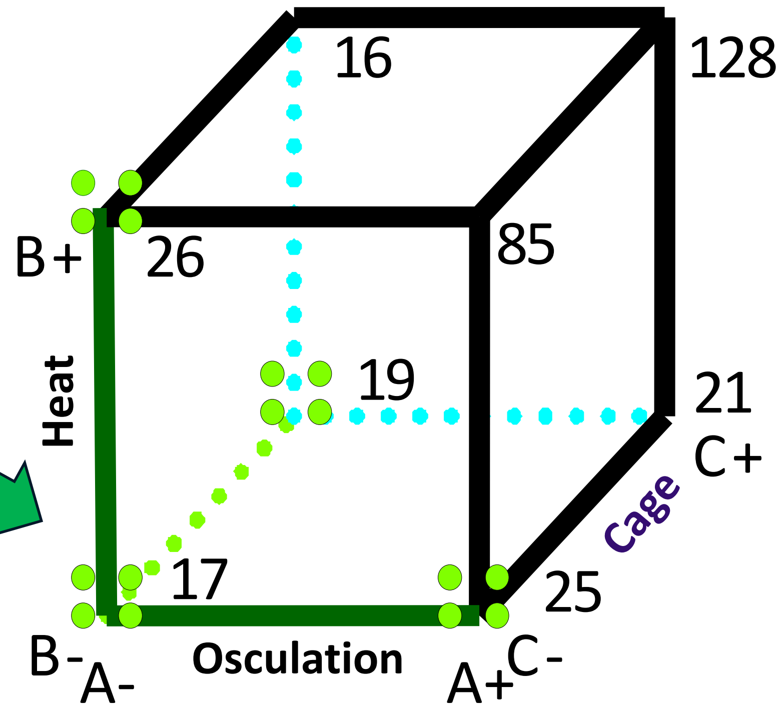
1st edition



Multi-Factorial (VS OFAT)

*Life from accelerated bearing test**

*Start point for One Factor at a Time **(OFAT)***



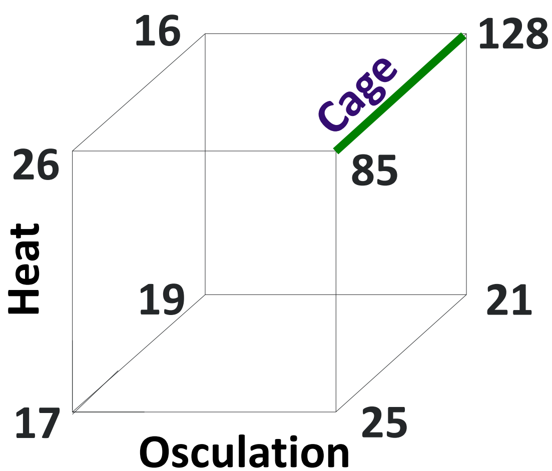
Relative efficiency = 16/8

⇒ 2 to 1!

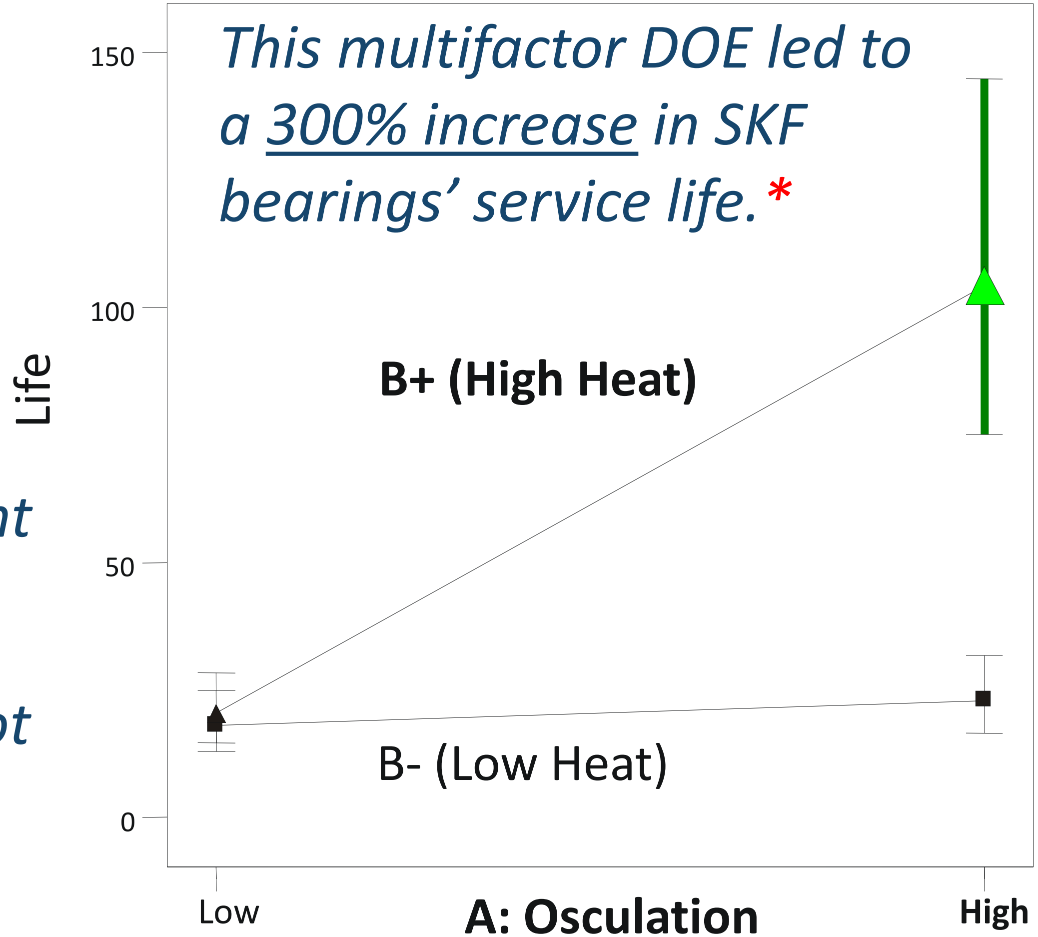
*Hellstrand, C., "The necessity of modern quality improvement and some experience with its implementation in the manufacture of rolling bearings", *Philosophical Transactions of the Royal Society of London A* 327, 529-537, 1989.

Interaction Plot

The happy ending!



PS: The least significant difference (LSD) bar widens at high-life setting. Thus, Cage not significant. But worth changing to plastic!

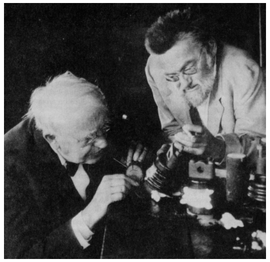


* "Breaking the Boundaries," *Design Engineering*, Feb 2000, pp 37-38.

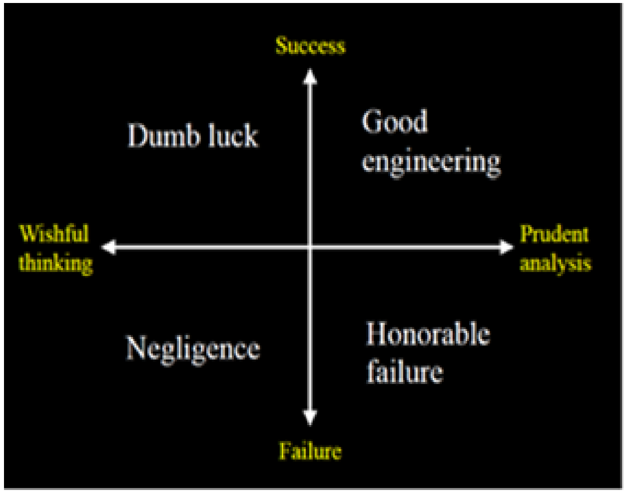
Before Statistics and Multifactor Testing

How industrial experimenters succeeded

1. Scientific method: Commonly attributed to Francis Bacon in the 17th century, stemming from Aristotle in mid-300s BC.
2. Persistence: Edison's 1% inspiration and 99% perspiration.
3. Good engineering: Edison's protégé Charles Steinmetz once charged $1000 to GE for knowing which part to investigate on an electrical device, $1 for the chalk mark and $999 for knowing where to put it.



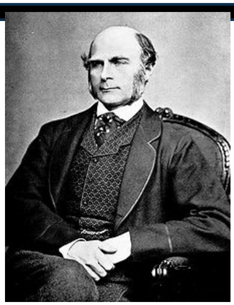
4. "Dumb luck"!



Source: "Beyond Probability, A pragmatic approach to uncertainty quantification in engineering," Scott Ferson, NASA Statistical Engineering Symposium, Williamsburg, Virginia, 4 May, 2011

The Beginning of Statistical Methods

Regression of happenstance data (1/2)



Regression analysis, invented in the late 19th century by Francis Galton (pictured),* connects the responses (Y's) to the input factors (X's) via mathematical models of the form: $\hat{Y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 \ldots. + \beta_k X_k + \varepsilon$

where k is the number of factors and ε represents error.

*"Regression towards mediocrity in hereditary stature". *The Journal of the Anthropological Institute of Great Britain and Ireland* (1886). 15: 246–263

"Engineers are quite comfortable these days - in fact, far too comfortable – with results from the blackest of black boxes: neural nets, genetic algorithms, data mining, and the like." [e.g., AI]

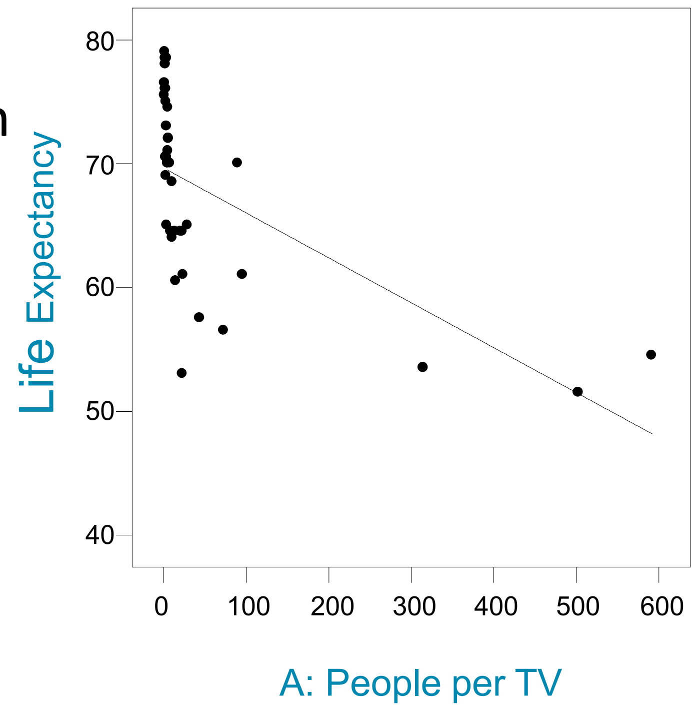
- Russell Lenth (Professor of Statistics, University of Iowa)

The Beginning of Statistical Methods

Regression of happenstance data (2/2)

A Cal Poly stats prof observed* that life expectancy in various countries varies with the number of people per television (TV). This solves our problems replacing obsolete devices: Ship them to developing nations so these poor TV-deprived people can live longer! ;)

*Allan Rossman, “Televisions, Physicians, and Life Expectancy.” *Journal of Statistics Education* 2, no. 2 (1994).



The Beginning of Statistical Methods

Simple comparative experiments

More than a century ago, William Sealy Gossett, a chemist at Guiness Brewery, developed a statistical method called the "t-test" to determine when the soft-resin content (desirable for stout) in hop flowers differed significantly from the brewery's standard.*

This is a simple comparative experiment on one factor at a time (OFAT). It is still widely used of sensory and other evaluations.

*(Published in 1908 under the pseudonym "Student".)

"He possessed a wickedly fertile imagination and more energy and focus than a St. Bernard in a snowstorm."

– Stephen Ziliak

A Very Small Dose of Stat Detail

One-factor comparison via t-tests

Legal judgment: Innocent until proven guilty.

Hypothesis test: Same until proven different.

H_0("null"): $\mu_1 = \mu_2$ (samples from same population)

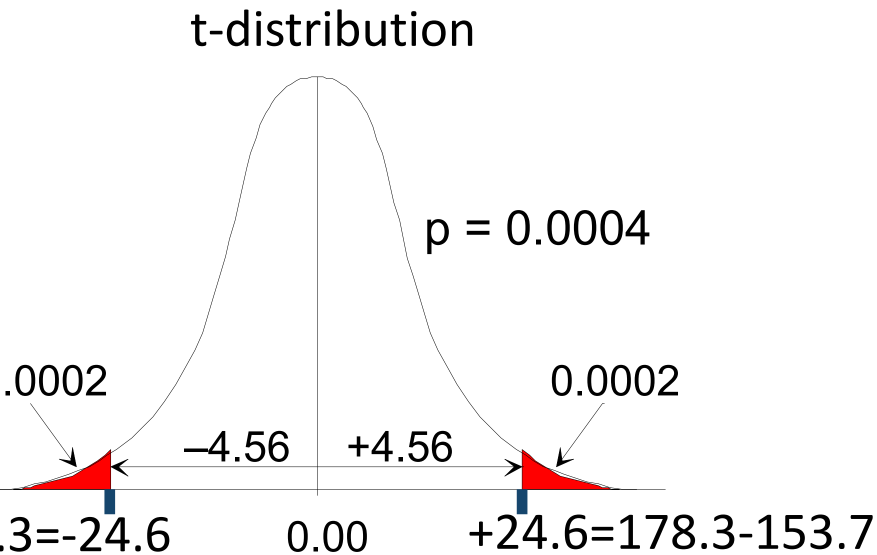
H_1("alternative"): $\mu_1 \neq \mu_2$ (samples from different populations)

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{s_{\bar{Y}_1 - \bar{Y}_2}}$$

$$t = \frac{\text{difference between averages}}{\text{standard deviation of difference}}$$

Comparisons via t-Test

Case study: Stat-Ease bowling contest

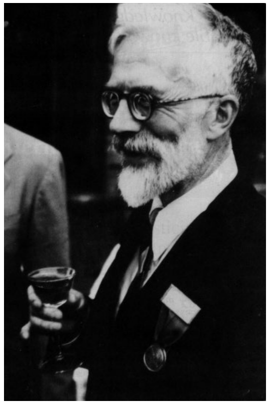


Run	Pat	Mark
1	160	165
2	150	180
3	140	170
4	167	185
5	157	195
6	148	175
Avg	153.7	178.3

t = 4.56 standard deviations between means, so by two-tailed test (Pat-Mark or Mark-Pat) p = 0.0004, thus with >99.9% confidence Mark is the better bowler. ☺

Fisher: Inventor of Modern-Day Statistics
and multilevel, multifactor experiment designs

"Personally, the writer prefers to set a low standard of significance at the 5 per cent point and ignore entirely all results which fail to reach this level. A scientific fact should be regarded as experimentally established only if a properly designed experiment rarely fails to give this level of significance."



-Sir Ronald Fisher
"The Arrangement of Field Experiments,"
The Journal of the Ministry of Agriculture, 1926, 33, 504.

Little known fact:

When Fisher invented DOE at Rothamsted Experimental Station in England, computations were done by 'calculators' – mathematical adepts, mainly female.

Example of Fisher's pioneering work:

A randomized, replicated, blocked DOE (1/3)

In a landmark field trial on barley in Minnesota, agronomists guided by Fisher grew 5 varieties (M, S, V, T, P) at 5 ag stations in 1931 and 1932.

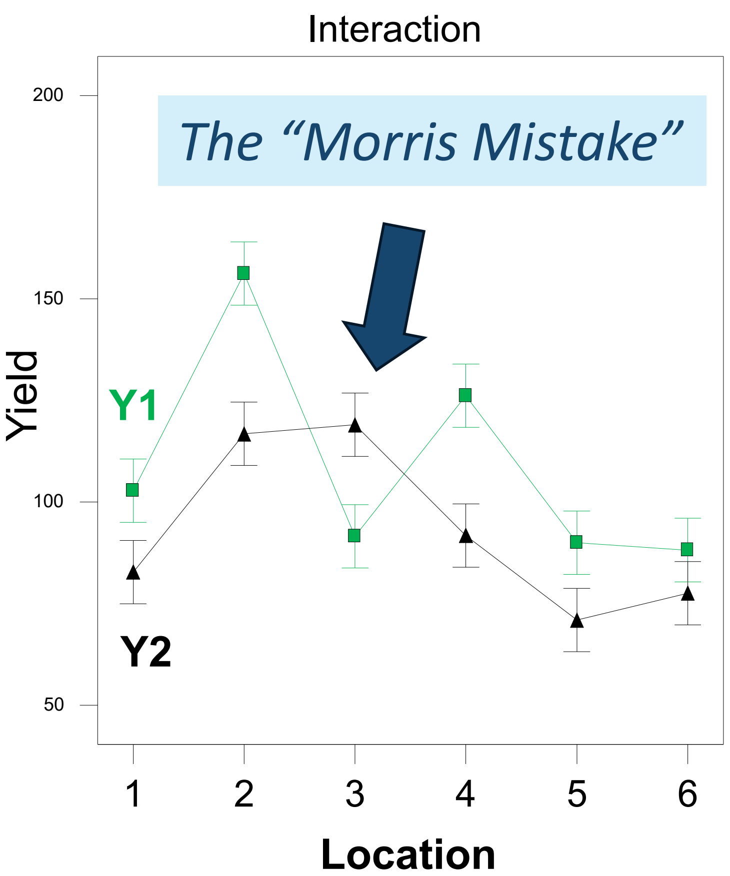
Which variety stands out? (Hint: See Graphs!)

Location	Year	M	S	V	T	P
1	1	81	105	120	110	98
	2	81	82	80	87	84
2	1	147	145	151	192	146
	2	100	116	112	148	108
3	1	82	77	78	131	90
	2	103	105	117	140	130
4	1	120	121	124	141	125
	2	99	62	96	126	76
5	1	99	89	69	89	104
	2	66	50	97	62	80
6	1	87	77	79	102	96
	2	68	67	67	92	94

Example of Fisher's pioneering work:
A randomized, replicated, blocked DOE (2/3)

In a book called *Visualizing Data* (Hobart Press, 1993) William S. Cleveland suggests that the experimenters* reversed the numbers year-by-year in their report for location 3 (Morris, MN). It is hard to see in the raw data, but obvious when graphed with varieties averaged. The 'take home' message:

One picture = 1000 numbers!

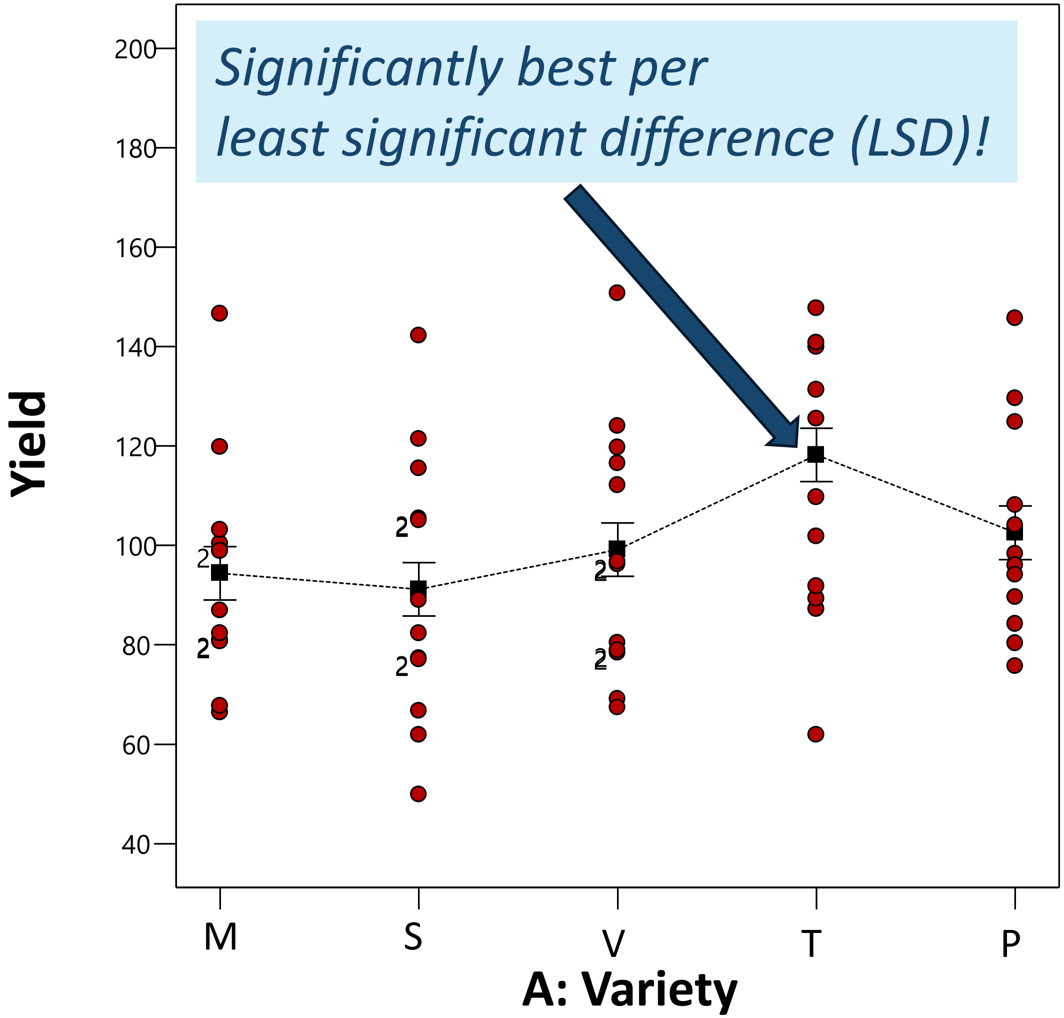


*(Immer, et al, *Journal of Agronomy*, 26, 403-419, 1934).

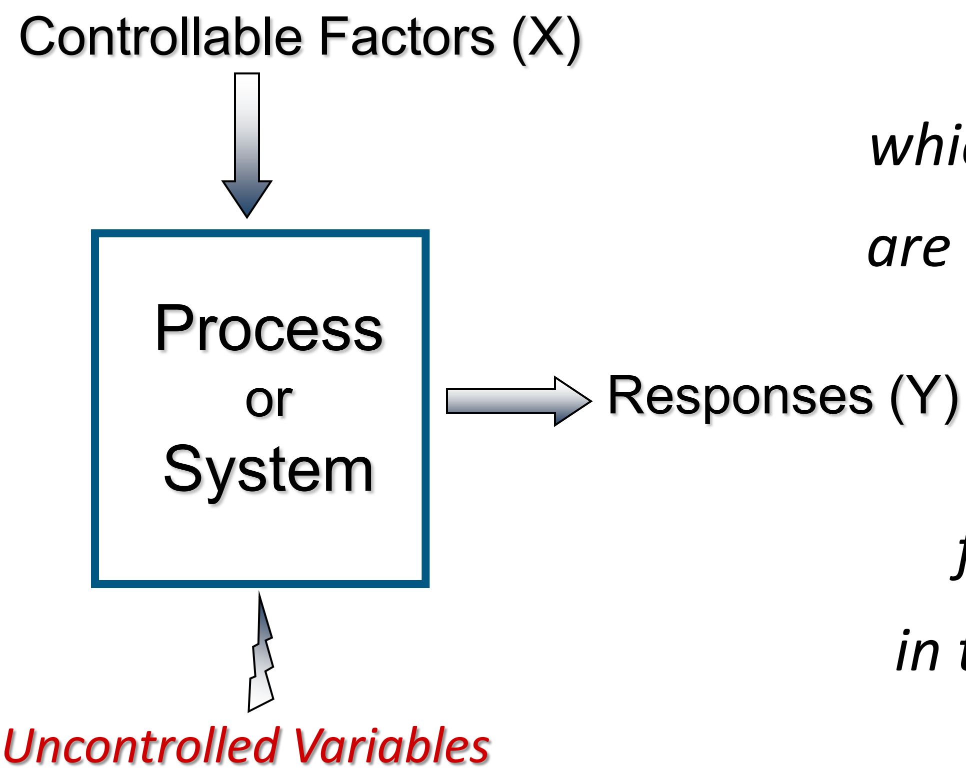
Example of Fisher's pioneering work:
A randomized, replicated, blocked DOE (3/3)

Key to Varieties:

- M: Manchuria
- S: Svansota
- V: Velvet
- T: Trebi
- P: Peatland



DOE Works on Any Process or System

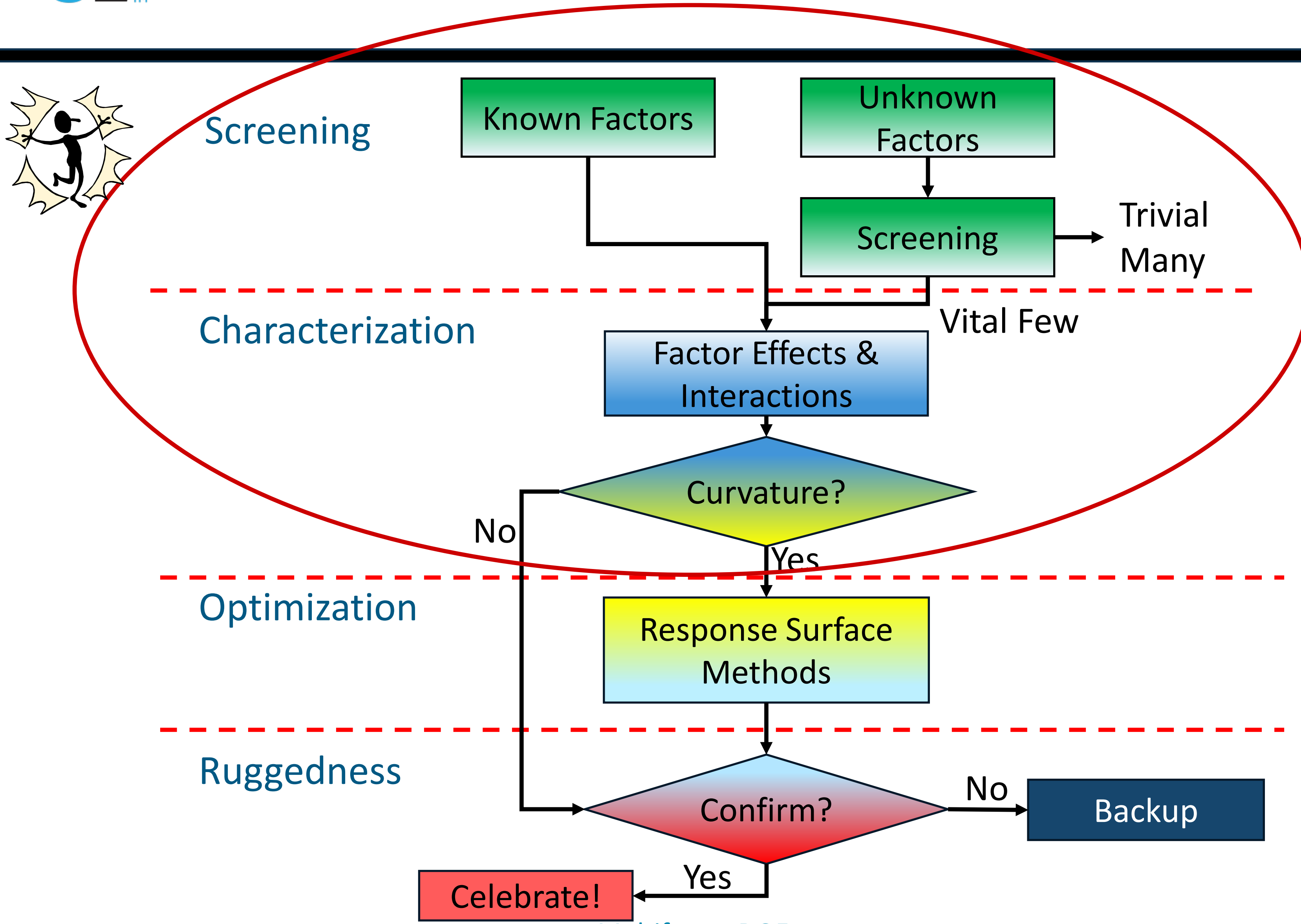


DOE is:

"A series of tests, in which purposeful changes are made to input factors,

to identify causes for significant changes in the output responses."

Strategy of Experimentation



Screening/Characterization

Purpose: Quickly sift through a large number of potential factors to discard the trivial many. Then follow-up with an experiment that focuses on the vital few.

Tool: Two-level factorial designs:

1. Fractional for resolving main effects in minimal runs.
2. Full (or less fractional) to resolve two-factor interactions.

Screening/Characterization

Case study

Factor	Name	Low (–)	High (+)
A	Prime pumps	3	5
B	Pulls at choke	3	5
C	Gas at choke	0	100%
D	Final choke	0	50%
E	Gas for start	0	100%

Primer bulb

Choke control

Full and Fractional Two-Level Designs

Green light this design that cuts runs in half at high resolution for possible two-factor interactions.

Factors

Runs	2	3	4	5	6	7	8	9	10
4	2^{2}	2^{3-1}_{III}							
8		2^{3}	2^{4-1}_{IV}	2^{5-2}_{III}	2^{6-3}_{III}	2^{7-4}_{III}			
16			2^{4}	2^{5-1}_{V}	2^{6-2}_{IV}	2^{7-3}_{IV}	2^{8-4}_{IV}	2^{9-5}_{III}	2^{10-6}_{III}
32				2^{5}	2^{6-1}_{VI}	2^{7-2}_{IV}	2^{8-3}_{IV}	2^{9-4}_{IV}	2^{10-5}_{IV}
64					2^{6}	2^{7-1}_{VII}	2^{8-2}_{V}	2^{9-3}_{IV}	2^{10-4}_{IV}

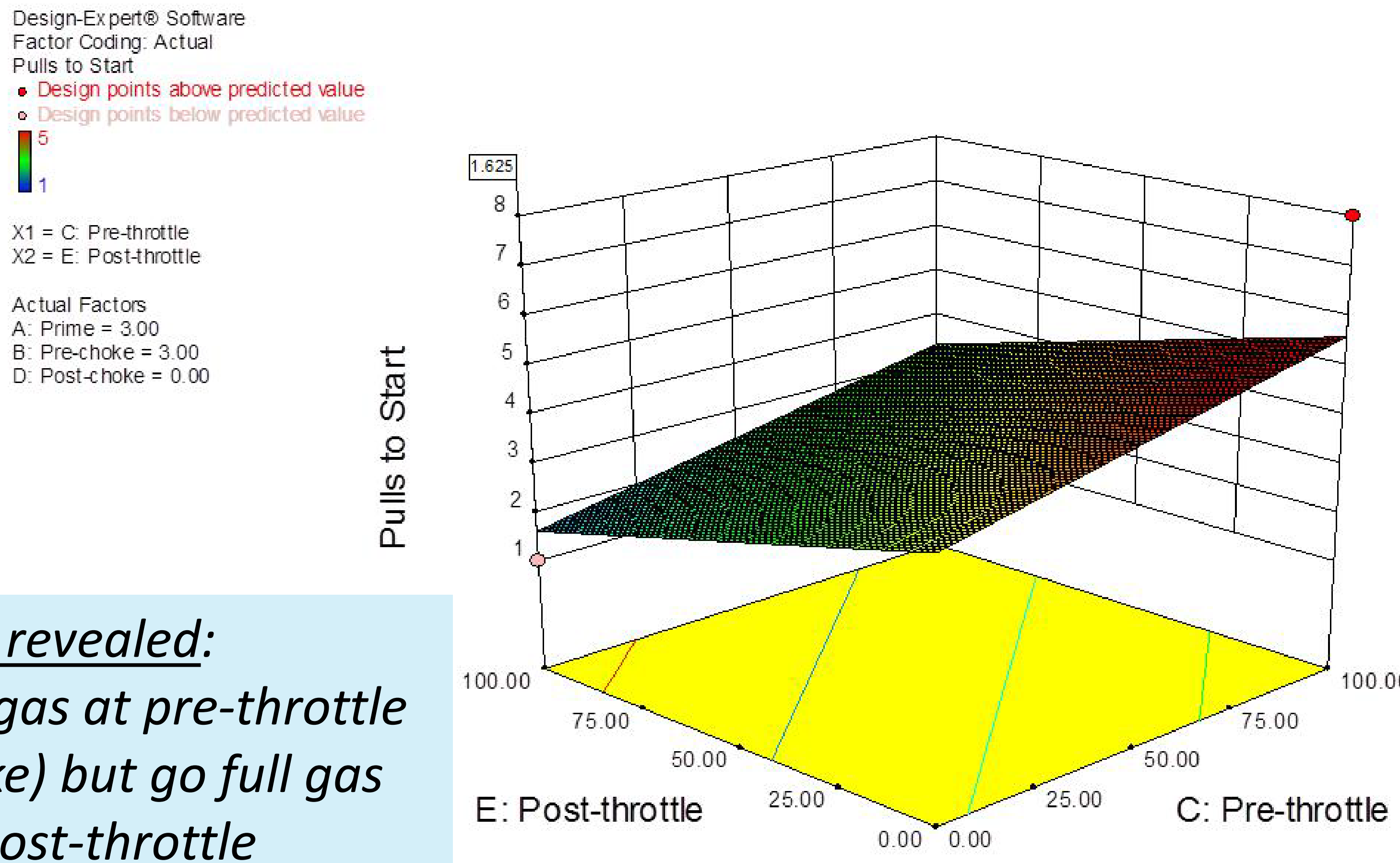
Experiment-design template

Std	A	B	C	D	*E*	Pulls
1	–	–	–	–	+	1
2	+	–	–	–	–	4
3	–	+	–	–	–	4
4	+	+	–	–	+	2
5	–	–	+	–	–	8
6	+	–	+	–	+	2
7	–	+	+	–	+	3
8	+	+	+	–	–	5
9	–	–	–	+	–	3
10	+	–	–	+	+	1
11	–	+	–	+	+	3
12	+	+	–	+	–	4
13	–	–	+	+	+	3
14	+	–	+	+	–	4
15	–	+	+	+	–	6
16	+	+	+	+	+	5

This factor deliberately aliased to cut runs by half!

[E]=E+ABCD

Ladies and Gentlemen: Start Your Engines!



<u>The trick revealed</u>:
No (0%) gas at pre-throttle (full choke) but go full gas (100%) post-throttle (starting engine).

Minimum-Run Designs (up to 50 factors)

Considerable savings over standard fractions

Characterization

Factors	Std Res V	MR5*
6	32	22
7	64	30
8	64	38
9	128	46
10	128	56
11	128	68
12	256	80
13	256	92
14	256	106

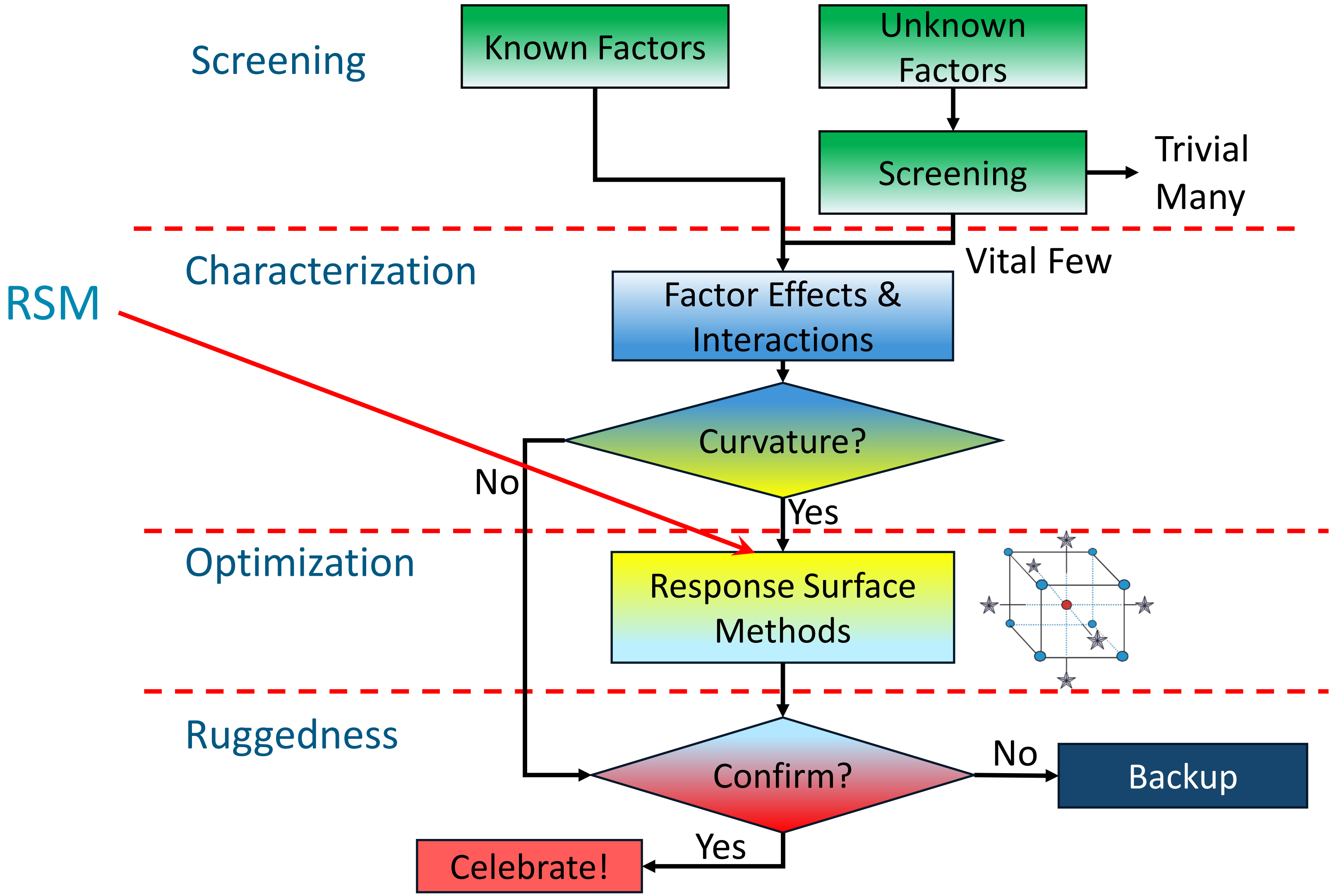
Screening

Factors	Std Res IV	MR4**
9	32	18
10	32	20
11	32	22
12	32	24
13	32	26
14	32	28
15	32	24
16	32	26
17	64	28

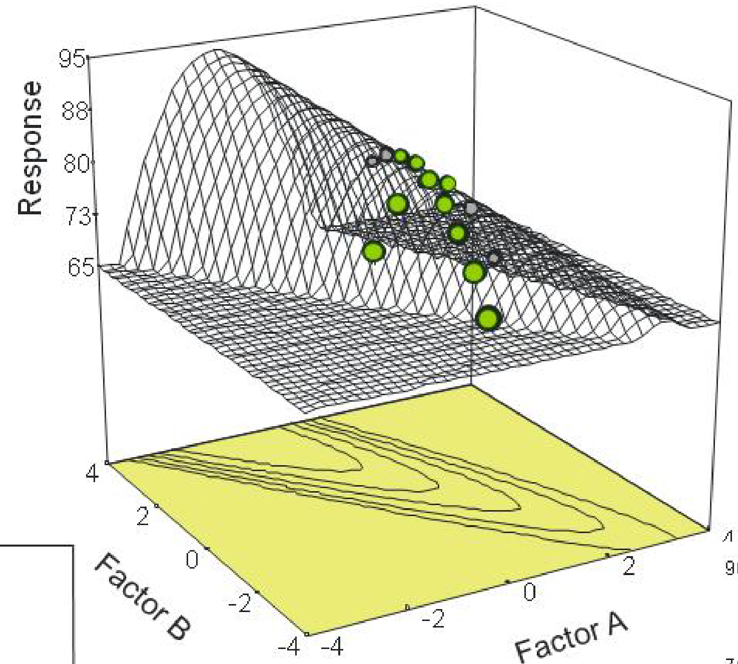
* Oehlert & Whitcomb, "Small, Efficient, Equireplicated Resolution V Fractions of 2^k designs …", Fall Technical Conference, 2002.

** Anderson & Whitcomb, "Screening Process Factors In the Presence of Interactions," Annual Quality Congress, American Society of Quality, Toronto, 2004.

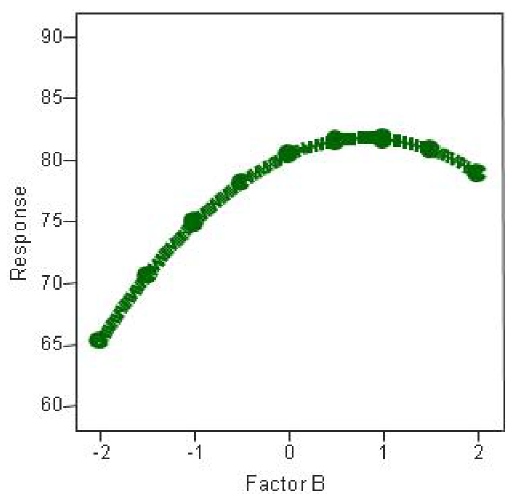
Strategy of Experimentation



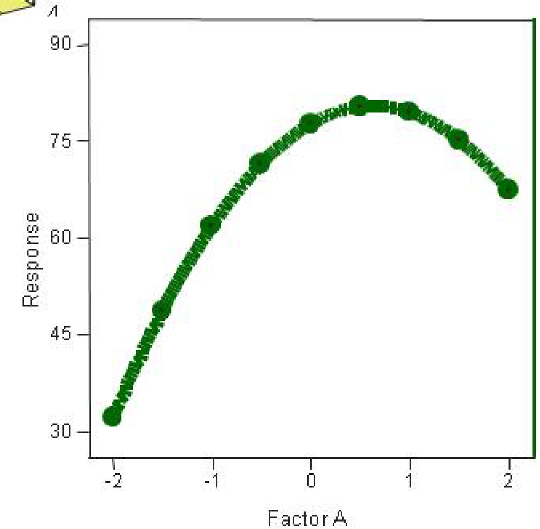
RSM vs OFAT



Response
95
88
80
73
65
Factor B
4
2
0
-2
-4
Factor A
-4
-2
0
2



Response
90
85
80
75
70
65
60
-2
-1
0
1
2
Factor B



Response
90
75
60
45
30
-2
-1
0
1
2
Factor A

RSM: Flowchart

Subject Matter Knowledge
(Plus, Factorial Screening)

↓

Vital Few Factors (x's)

↓

Process → Measured Response(s) (y(s))

↓

Fitting*

↓

Polynomial Model

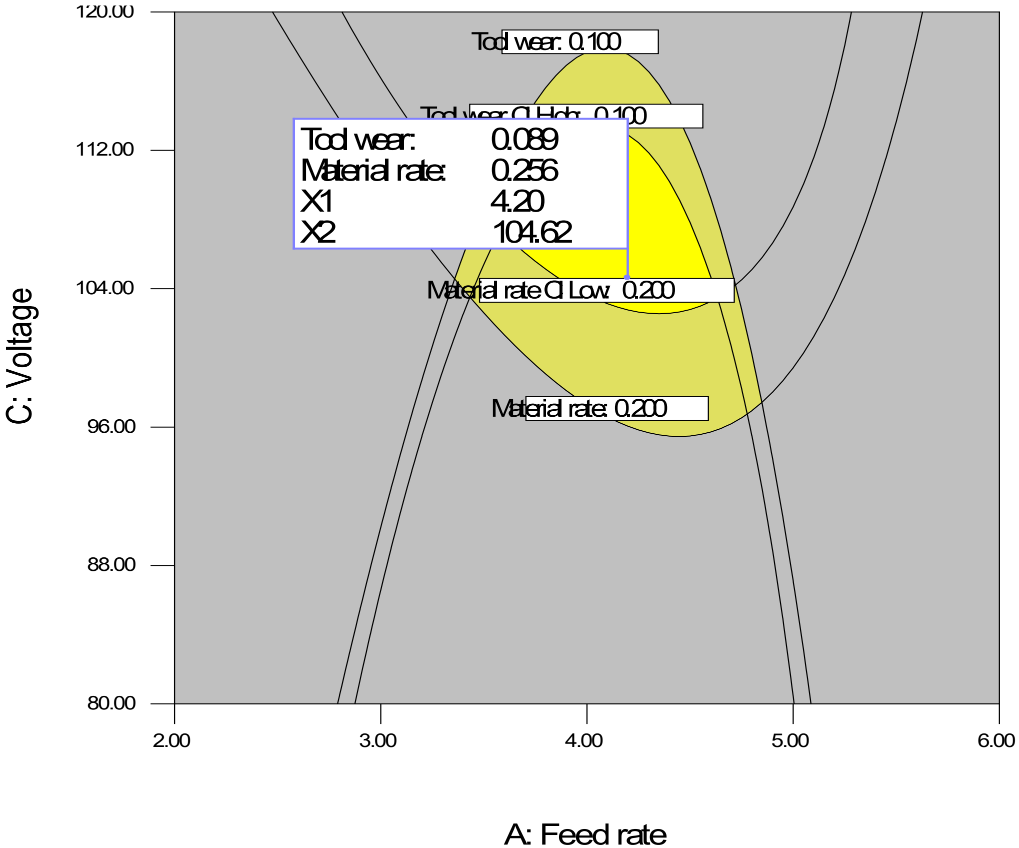
↓

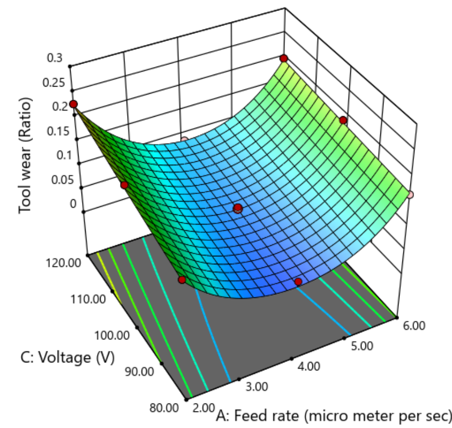
Response Surface

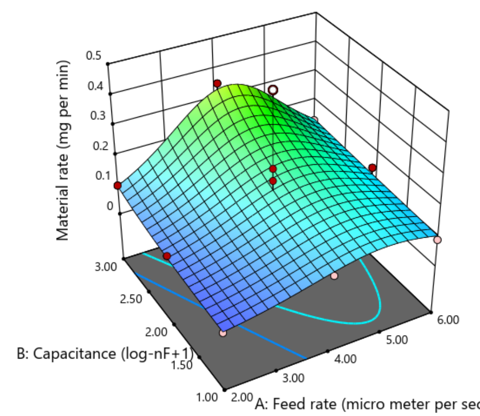
"All models are wrong, but some are useful." - George Box

Multiple Response Optimization

Electrodischarge milling (EDM)







"Numerical optimization via desirability investigation of machining parameters for the multiple-response optimization of micro electrodischarge milling," Mehfuz & Ali, *International Journal of Advanced Manufacturing Technology* (2009) 43:264–275

Importance of Achieving Robust Settings

Sent: Monday, April 7, 2025 5:37 AM

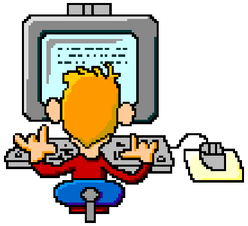
Hi Mark, Thank you for the follow up, I have successfully installed the software and ran through the RSM examples last night. I am setting up a new manufacturing facility with a range of procedures for which proper DOE will be essential. I am mostly trying to find optimum settings for machines, but also wanting to find robust setting ranges that can adjust for tool wear etc.

Tom ________, Product Engineering Manager

Good DOE Tools Vital to Success

A short list of recommended software

- General packages that include DOE:
  - JMP: Aimed at scientists & engineers, but, due to tremendous array of features, best suited to master statisticians, IMO
  - Minitab: Go to for Six Sigma practitioners and quality engineers, well suited for manufacturing QA and QC, e.g., via SPC
- Dedicated to DOE:
  - Modde: Focused on regulated (pharma, food) and other process industries, developed by Swedish chemometricians, most popular in Europe
  - Stat-Ease 360 or Design-Expert: Aimed at industrial R&D, powerful, yet easy to use, very compatible with JMP or Minitab—many companies license both: *Make the most from every experiment*



*Demo w **SE360 v23** if time allows: Milling*
Rebuild, re-open,
analyze: note transforms, 1 potential outlier
*optimize, overlay A vs C with CI's**
**missing CI for R2--a bug*

Conclusion

- **Trim out the OFAT!**

By making use of multifactor design of experiments (DOE) starting with simple two-level factorials and graduating to response surface methods (RSM) for processes, you will greatly accelerate product development and process optimization.

statistics made easy®

Make the most from every experiment! SM

Inspired by "Multifactor DOE for Rapid Process Improvement"

Mark J. Anderson, Engineering Consultant

Stat-Ease, Inc., Minneapolis, MN

mark@statease.com

www.linkedin.com/in/markstat/

Please email questions!

Let's connect!