How to Plan and Analyze a Verification DOE

Before we get started, feel free to download the presentation from our website (no data files with this one.)

http://www.statease.com/webinar.html

Assumed knowledge:

Fractional factorial designs and aliasing

Who We Are

This Month's Webinar

Webinar presented by:
Shari Kraber

With the Stat-Ease, Inc. consulting team

Pat Whitcomb

Mark Anderson

Wayne Adams

2

Learning Objectives
How to Plan and Analyze a Verification DOE

At the conclusion of this session you should be able to:

- Select an appropriate factorial design for a verification DOE.
  - Aliasing (appropriate resolution)
  - Power (appropriate size)
- Interpret the results from a verification design.

3

Transition
How to Plan and Analyze a Verification DOE

Our talk has three parts:

1. **Broad brush description of a factorial design planning process**
2. Illustrate key points via an example
3. Summary

4

Factorial Design Planning Process *(1 of 2)*

1. Identify opportunity and define objective.
2. State objective in terms of measurable responses.
   a. Define the change (Δy) that is important to detect for each response.
   b. Estimate experimental error (σ) for each response.
   c. Use the signal to noise ratio (Δy/σ) to estimate power.
3. Select the input factors to study. (*Remember that the factor levels chosen determine the size of Δy.*)

5

Factorial Design Planning Process *(2 of 2)*

4. Select a design and:
   - Evaluate aliases (fractional factorials and/or blocked designs).
   - Evaluate power (probability of finding an effect of a given size, i.e. Δy/σ).
   - Examine the design layout to ensure all the factor combinations are safe to run and are likely to result in meaningful information (no disasters).

6

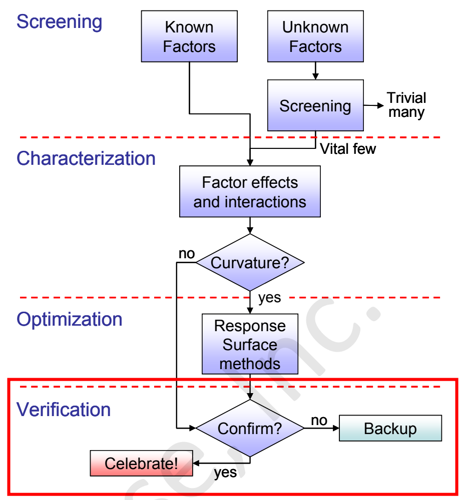
Strategy of Experimentation

What is acceptable aliasing and power depends on the purpose of the design:

Screening:
- Aliasing ≥ Res IV
- Power ≥ 80%

Characterization:
- Aliasing ≥ Res V
- Power ≥ 80%

Verification:
- **Aliasing ≥ Res III**
- **Power ≥ 90%**



7

Selection of Fractional Factorials
Resolution versus Experimental Phase

Screening: Resolution IV designs
- estimate main effects clear of 2fi's
- might get some idea of interactions
- do NOT include factors you know are significant.

Characterization: Resolution V designs
- estimate main effects and 2fi's
- be careful if you block.

Verification: Resolution III designs if power is OK
- minimize number of runs
- basically a go / no-go test.

8

Validation versus Verification

Validation: The requirements for a specific intended use are consistently fulfilled.

- Satisfying external (customer) needs.
- Implies the voice of the customer has been successfully translated into specifications.
- Ensuring "you built the right product."

Verification: Compliance to internal requirements.

- Implies meeting specifications.
- Ensuring "you built the product right."

9

Verification Testing

Strategy: Vary key factors over ranges that are expected to be encountered during normal use.

- Process or product should be "robust" to these normal variations.
- Resolution III design minimizes runs – check power!
- If ANOVA significant, cause unclear, but you will know that additional work is required.

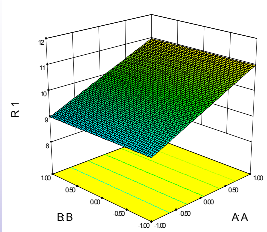
This application of DOE boils down to a go / no-go test.

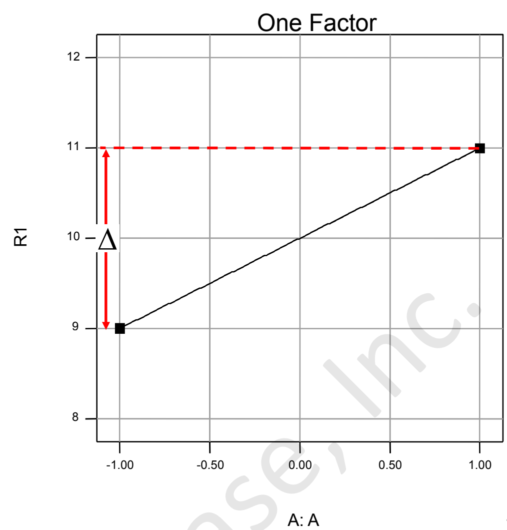
For verification we hope none of the factor effects are large!

10

Factorial Design – Power

2^3 Full Factorial Δ=2 and σ=1





11

What is Power?

No Factor Effect; H_0: Δ = 0

Confidence = $(1-\alpha)$*100%

Power = $(1-\beta)$*100%

Desire power to be high (at least 90% for verification) for an effect size of interest, Δ.

ANOVA says...

Truth is...	Retain H_0	Reject H_0
No Effect	OK	Type I error (alpha rate) *False alarm*
Effect	Type II error (beta rate) *Failure to detect*	OK

12

Factorial Design – Power

2^3 Full Factorial Δ=2 and σ=1

Leave Sigma and Delta fields blank to skip power calculation.

Responses: 1 (1 to 999) Edit model for power...

Name	Units	Diff. to detect Delta("Signal")	Est. Std. Dev. Sigma("Noise")	Delta/Sigma (Signal/Noise Ratio)
R1		2	1	2

Power is reported at a 5.0% alpha level to detect the specified signal/noise ratio.

Signal (delta) = 2.00 Noise (sigma) = 1.00 Signal/Noise (delta/sigma) = 2.00

A

65.7 %

Since power is low, we may choose to replicate if the desired detectable effect and the standard deviation are fixed.

13

Factorial Design – Power

Two Replicates of 2^3 Full Factorial Δ=2 and σ=1

Leave Sigma and Delta fields blank to skip power calculation.

Responses: 1 (1 to 999) Edit model for power...

Name	Units	Diff. to detect Delta("Signal")	Est. Std. Dev. Sigma("Noise")	Delta/Sigma (Signal/Noise Ratio)
R1		2	1	2

Power is reported at a 5.0% alpha level to detect the specified signal/noise ratio.

Signal (delta) = 2.00 Noise (sigma) = 1.00 Signal/Noise (delta/sigma) = 2.00

A

96.0 %

14

Power
The probability of finding an effect!

Power depends on:

- The size of the difference Δ:
  the larger the difference the higher the power.
- The size of the experimental error σ:
  the smaller σ the higher the power.
- The α risk chosen:
  the larger α the higher the power.
- Choosing a design appropriate to the problem:
  more orthogonal and larger designs have more power.
- The number of replicates:
  the more runs the higher the power.

15

Transition
How to Plan and Analyze a Verification DOE

Our talk has three parts:

1. Broad brush description of a factorial design planning process
2. **Illustrate key points via an example**
3. Summary

16

Viscosity Verification Test

Conduct a verification test on a laboratory assay for measuring viscosity.

There are 7 key factors and the factor ranges represent normal laboratory variation:

	NAME	UNITS	LOW	HIGH
A	Sample Prep		Short	Long
B	Moisture		Volume	Weight
C	Speed	rpm	1100	1300
D	Mixing time	minutes	2	4
E	Rest time	minutes	30	60
F	Spindle		Type 1	Type 2
G	Cover		Absent	Present

17

Viscosity Verification Test

Factorial Design Planning Process *(page 1 of 2)*

1. Identify opportunity and define objective.
   Verify that these 7 critical parameters have no effect on the response when varied in these ranges.
2. State objective in terms of measurable responses.
   Response = Viscosity in mPa-sec.
   a. Define the change (Δy) that is important to detect for each response. Δ = 35 mPa-sec
   b. Estimate experimental error (σ) for each response.
      σ = 20 mPa-sec
   a. Use the signal to noise ratio (Δ/σ = 1.75) to estimate power.

18

Viscosity Verification Test
Factorial Design Planning Process *(page 2 of 2)*

3. Select the input factors to study. (*Remember that the factor levels chosen determine the size of Δ.*) *done*
4. Select a design and:
   - Evaluate aliases (fractional factorials and/or blocked designs)
   - Evaluate power (for verification desire power ≥ 90% an effect of interest) *1 main effect*
   - Examine the design layout to ensure all the factor combinations are safe to run and are likely to result in meaningful information (no disasters)

Try a 2^{7-4} factorial

19

Viscosity Verification Test
Check Aliasing of 2^{7-4}

Is the aliasing in the 2^{7-4} (8-run) design acceptable?

[Est. Terms] Aliased Terms
[Intercept] = Intercept +ABD +ACE +AFG +BCF +BEG +CDG +DEF
[A] = A +BD +CE +FG +BCG +BEF +CDF +DEG
[B] = B +AD +CF +EG +ACG +AEF +CDE +DFG
[C] = C +AE +BF +DG +ABG +ADF +BDE +EFG
[D] = D +AB +CG +EF +ACF +AEG +BCE +BFG
[E] = E +AC +BG +DF +ABF +ADG +BCD +CFG
[F] = F +AG +BC +DE +ABE +ACD +BDG +CEG
[G] = G +AF +BE +CD +ABC +ADE +BDF +CEF

Resolution III is acceptable in a verification test, as long as the product/process passes the verification test.

20

Viscosity Verification Test
Check Power of 2^{7-4}

Does the 2^{7-4} (8-run) design have sufficient power?

Power at 5 % alpha level
Signal (delta) = 35.00
Noise (sigma) = 20.00
Signal/Noise (delta/sigma) = 1.75
Power = 54.6 %

Not sufficient power! Should we replicate?

21

Viscosity Verification Test
Inadequate Power

Should a fractional factorial be replicated?

No – Generally it makes more sense to use a larger fraction, rather than doing a replicate.

- Replicating will only add power.
- Adding another fraction will
  - add power
  - provide the ability to estimate more terms
  - test more combinations of the factors

Try a 2^{7-3} factorial

22

Viscosity Verification Test
Check Aliasing of 2^{7-3}

Aliasing in the 2^{7-3} is better (*res IV*) than in the 2^{7-4} (*res III*)

[A] = A + BCE + BFG + CDG + DEF
[B] = B + ACE + AFG + CDF + DEG
[C] = C + ABE + ADG + BDF + EFG
[D] = D + ACG + AEF + BCF + BEG
[E] = E + ABC + ADF + BDG + CFG
[F] = F + ABG + ADE + BCD + CEG
[G] = G + ABF + ACD + BDE + CEF
[AB] = AB + CE + FG
[AC] = AC + BE + DG
[AD] = AD + CG + EF
[AE] = AE + BC + DF
[AF] = AF + BG + DE
[AG] = AG + BF + CD
[BD] = BD + CF + EG

23

Viscosity Verification Test
Check Power of 2^{7-3}

Does the 2^{7-3} (16-run) design have sufficient power?

Power at 5 % alpha level
- Signal (delta) = 35.00
- Noise (sigma) = 20.00
- Signal/Noise (delta/sigma) = 1.75
- **Power = 90.6 %**

YES! Let's continue with this design.

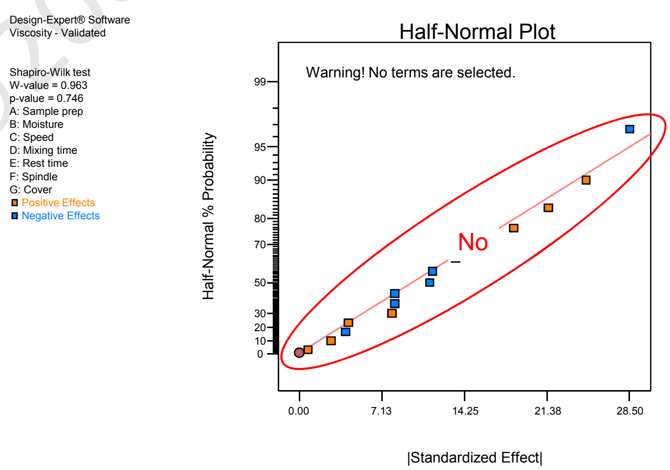
24

Viscosity Verification Test
Four Potential Outcomes

1. No significant effects and acceptable variability; successful verification. ☺
2. No significant effects and unacceptable (large) variability; does not verify. ☹
3. Statistically significant effect that is not practically important; qualified verification. 😐
4. Statistically significant effect that is practically important; does not verify. ☹

25

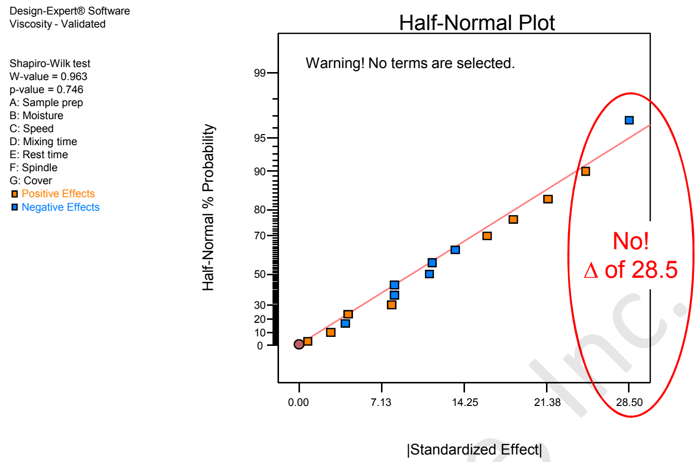
Viscosity Verification Test
1st Outcome: Verified



Is anything statistically significant?

26

Viscosity Verification Test
1st Outcome: Verified



Is anything large enough to be practically important?

27

Viscosity Verification Test
1st Outcome: Verified

Conclusion:

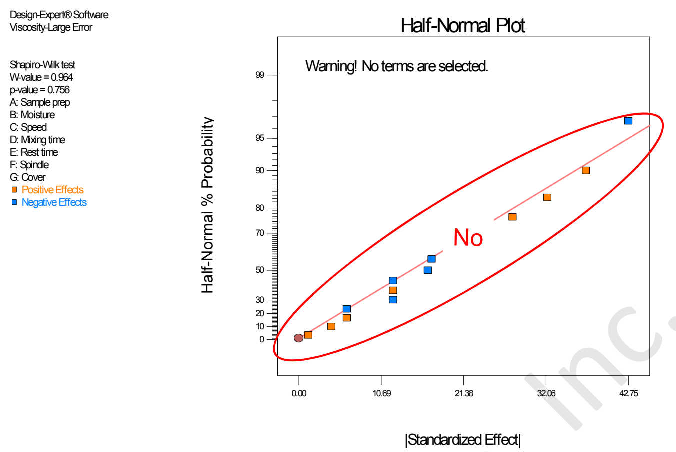
The variation of these 7 factors in their designated ranges did not cause a change in the response that was as large or larger than our specified effect (Δ=35 mPa-sec)

Success! Verification confirmed.

☺

28

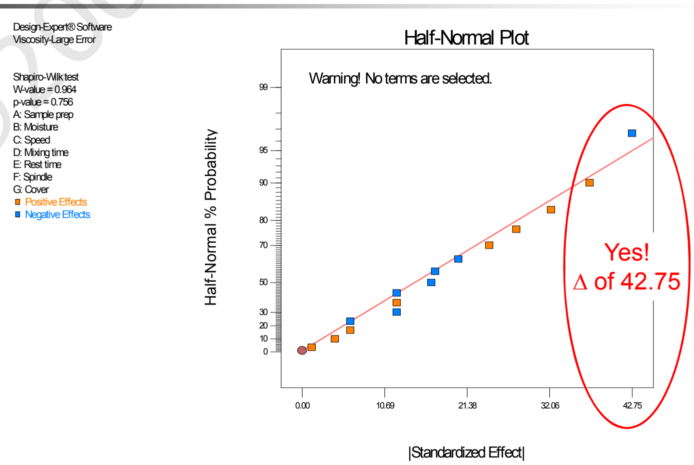
Viscosity Verification Test
2nd Outcome: Large Error



Is anything statistically significant?

29

Viscosity Verification Test
2nd Outcome: Large Error



Is anything large enough to be practically important?

30

Viscosity Verification Test
2nd Outcome: Large Error

Conclusion:

Although there are no significant effects, the variability in this system is larger than our cut off limit of (Δ=35 mPa-sec).

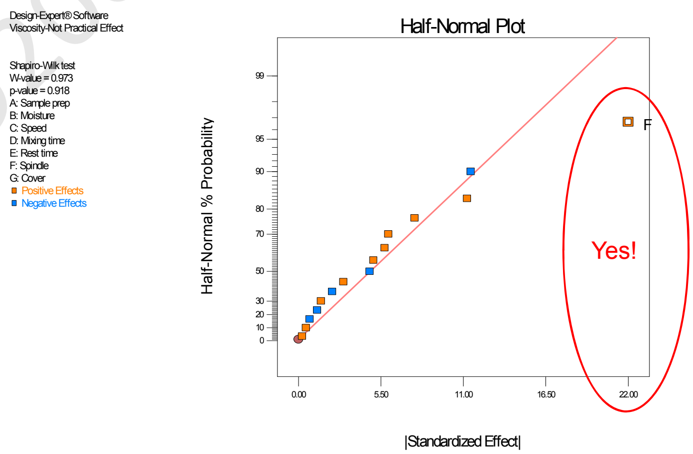
The standard deviation was large, but this variability cannot be attributed to any factor effect.

Verification Failure!

☹

31

Viscosity Verification Test
3rd Outcome: Significant – Not Important



Is anything statistically significant?

32

Viscosity Verification Test

3rd Outcome: Significant – Not Important

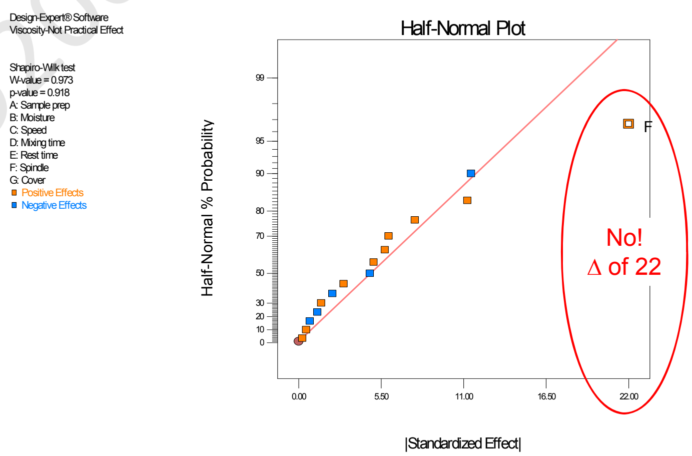
ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	1936.00	1	1936.00	14.92	0.0017
F-Spindle	*1936.00*	*1*	*1936.00*	*14.92*	*0.0017*
Residual	*1817.00*	*14*	*129.79*		
Cor Total	*3753.00*	*15*			

33

Viscosity Verification Test

3rd Outcome: Significant – Not Important



Is anything large enough to be practically important?

34

Viscosity Verification Test

3rd Outcome: Significant – Not Important

Conclusion:

Although varying these 7 factors across their designated ranges did produce a significant effect, the effect detected did not cause a change in the response that was as large as our specified cut off (Δ=35 mPa-sec)

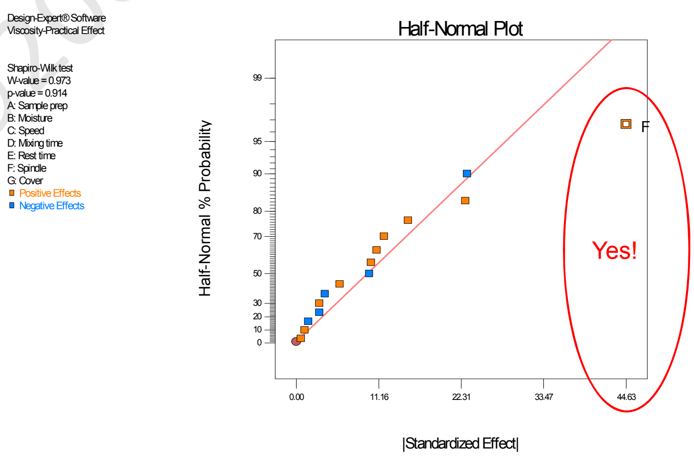
Success! Qualified Verification

😐

35

Viscosity Verification Test

4th Outcome: Significant – Important Effect



Is anything statistically significant?

36

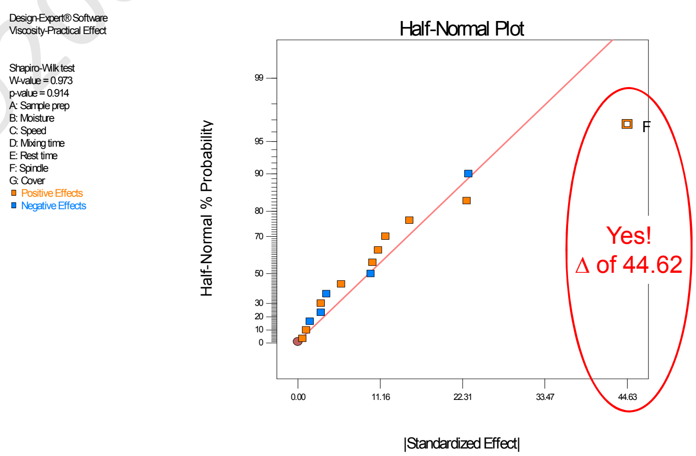
Viscosity Verification Test
4th Outcome: Significant – Important Effect

ANOVA for selected factorial model
Analysis of variance table [Partial sum of squares - Type III]

Source	Sum of Squares	df	Mean Square	F Value	p-value Prob > F
Model	7965.56	1	7965.56	15.32	0.0016
F-Spindle	7965.56	1	7965.56	15.32	0.0016
Residual	7277.87	14	519.85		
Cor Total	15243.44	15			

37

Viscosity Verification Test
4th Outcome: Significant – Important Effect



Is anything large enough to be practically important?

38

Viscosity Verification Test

4^{th} Outcome: Significant – Important Effect

Conclusion:

A statistically significant effect detected that caused a change in the response that was larger (44.62) than our specified cut off (Δ=35 mPa-sec)

Verification Failure!

☹

39

Viscosity Verification Test

4^{th} Outcome: Check Aliasing

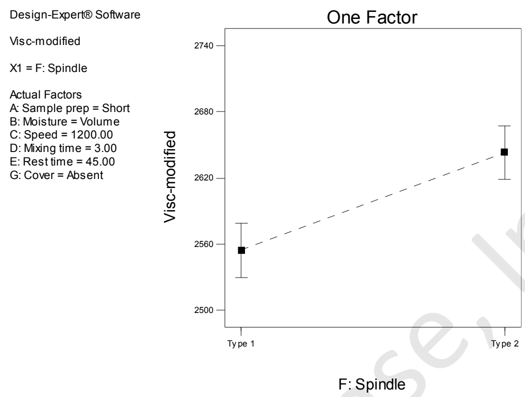
Look at the alias structure:

[A] = A + BCE + BFG + CDG + DEF
[B] = B + ACE + AFG + CDF + DEG
[C] = C + ABE + ADG + BDF + EFG
[D] = D + ACG + AEF + BCF + BEG
[E] = E + ABC + ADF + BDG + CFG
[F] = F + ABG + ADE + BCD + CEG
[G] = G + ABF + ACD + BDE + CEF
[AB] = AB + CE + FG
[AC] = AC + BE + DG
[AD] = AD + CG + EF
[AE] = AE + BC + DF
[AF] = AF + BG + DE
[AG] = AG + BF + CD
[BD] = BD + CF + EG

40

Viscosity Verification Test
4th Outcome: Explore Likely Causes

F – Spindle Type causes a statistically significant change in the Viscosity that is larger than the stated Δ, therefore the process is not verified.



41

Verification

What next?

If the process is not verified, additional decisions must be made (e.g. standardize on one spindle type) or more work must be done to control and/or find the influencing factors.

42

Transition
How to Plan and Analyze a Verification DOE

Our talk has three parts:

1. Broad brush description of a factorial design planning process
2. Illustrate key points via an example
3. **Summary**

43

Factorial Design Planning Process *(1 of 2)*

1. Identify opportunity and define objective.
2. State objective in terms of measurable responses.
   a. Define the change (Δy) that is important to detect for each response.
   b. Estimate experimental error (σ) for each response.
   c. Use the signal to noise ratio ($\Delta y/\sigma$) to estimate power.
3. Select the input factors to study. (*Remember that the factor levels chosen determine the size of Δy.*)

44

Factorial Design Planning Process *(2 of 2)*

4. Select a design and:
   - Evaluate aliases (fractional factorials and/or blocked designs).
   - Evaluate power (probability of finding an effect of a given size, i.e. $\Delta y/\sigma$).
   - Examine the design layout to ensure all the factor combinations are safe to run and are likely to result in meaningful information (no disasters).

45

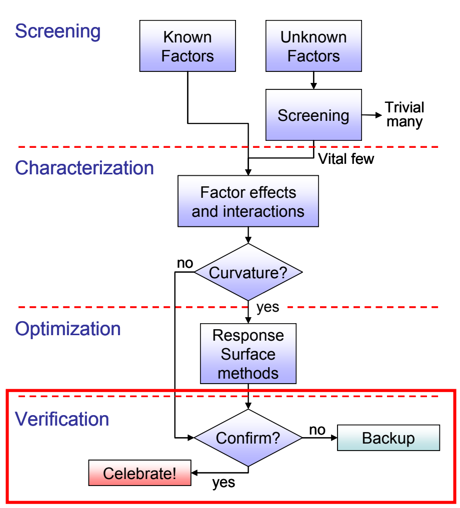
Factorial Design Planning Process

What is acceptable aliasing and power depends on the purpose of the design:

Screening:
- Aliasing ≥ Res IV
- Power ≥ 80%

Characterization:
- Aliasing ≥ Res V
- Power ≥ 80%

Verification:
- **Aliasing ≥ Res III**
- **Power ≥ 90%**



46

Learning Objectives
How to Plan and Analyze a Verification DOE

At the conclusion of this session you should be able to:

- Select an appropriate factorial design for a verification DOE.
  - Aliasing (appropriate resolution)
  - Power (appropriate size)
- Interpret the results from a verification design.

47

How to get help

- ❑ Search publications posted at www.statease.com
- ❑ E-mail statHelp@statease.com for answers from Stat-Ease's staff of statistical consultants
- ❑ Call 612.378.9449 and ask for "statistical help"

Thanks for attending!

48