There is no "I" in "Youden", but there is "You"

Steven P. Bailey, PhD, CMBB

StevenPBailey@Comcast.net

W. J. Youden Memorial Address

61st Annual ASA/ASQ Fall Technical Conference

Philadelphia, PA – Thursday, October 5th, 2017

ASQ

The Global Voice of Quality™

John Cornell's Youden Address (at 36th FTC in 1992)

YOUDEN ADDRESS

BY JOHN A. CORNELL, UNIVERSITY OF FLORIDA

W.J. YOUDEN – THE MAN AND HIS METHODOLOGY

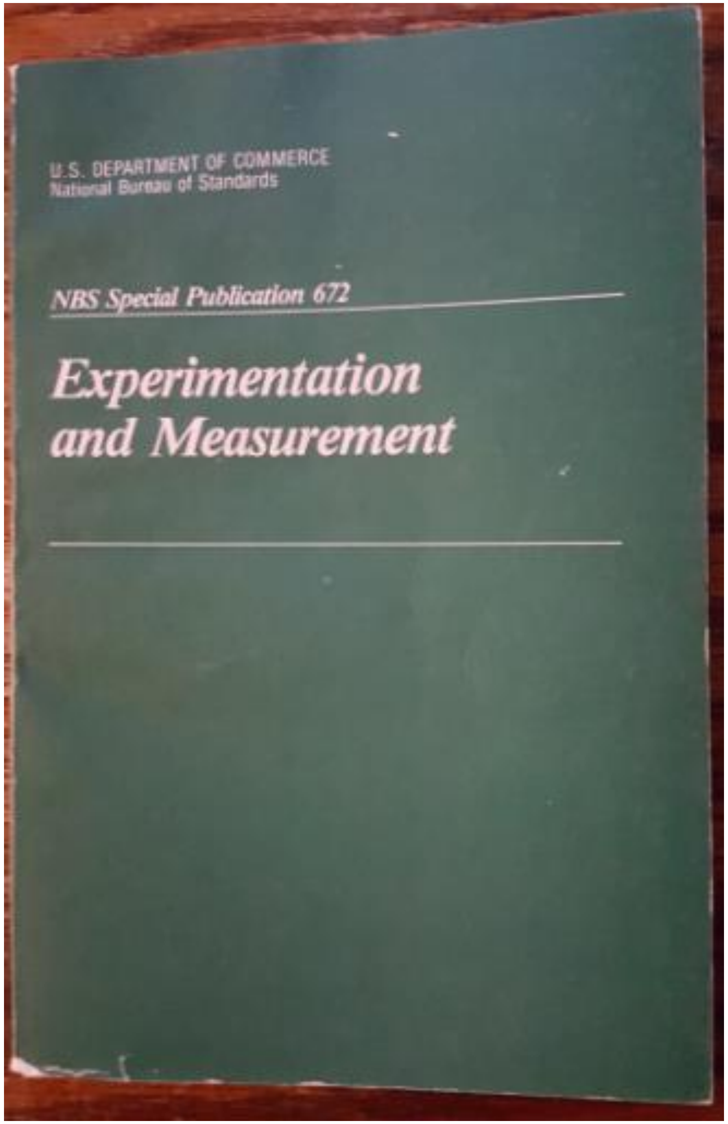
Jack Youden

"Competitiveness Through Continuous Improvement" was the theme of this year's 36th Annual Fall Technical conference and that is why I selected the topic "Jack Youden – The Man and His Methodology" for this address. Jack Youden believed in continuous improvement; he spent his life improving the ways measurements are taken.

A BRIEF BIOGRAPHICAL SKETCH (1900-1971) OF JACK YOUDEN'S LIFE

As with so many others who have contributed much to our profession, Jack Youden began his career, not as a statistician, but rather in a related discipline, as a physical chemist. Born in Townsville, Australia, in 1900, Jack's

Jack Youden's Experimentation and Measurement



This talk will cover:

- A roadmap for Measurement Systems Analysis (MSA) will be presented that has proved useful in guiding MSA studies for six sigma improvement projects.
- A review of DuPont's Strategy of Experimentation that has been taught and used successfully for over 50 years.
- A response surface example with both "design" (or "control") and "environmental" (or "noise") factors, showing how to achieve both "functional" and "robust" products.

Measurement System Analysis (MSA)

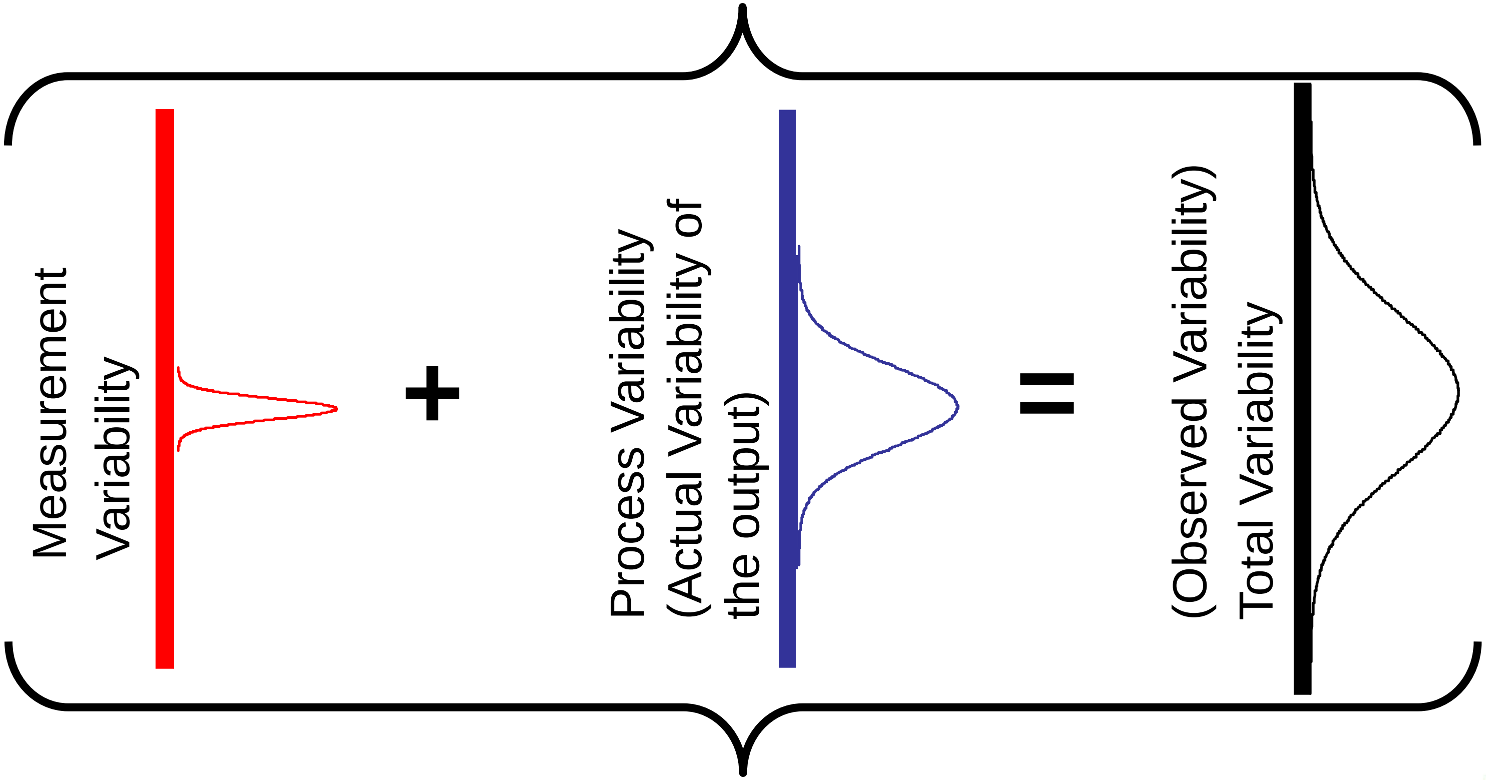
MSA is the process of …

- Identifying potential sources of measurement variation
- Choosing the appropriate analysis tool to quantify variation in the measurement system
- Comparing the extent of measurement variation to what is required for your needs (project)
- Improving the measurement system to reduce variation, if necessary

ASQ

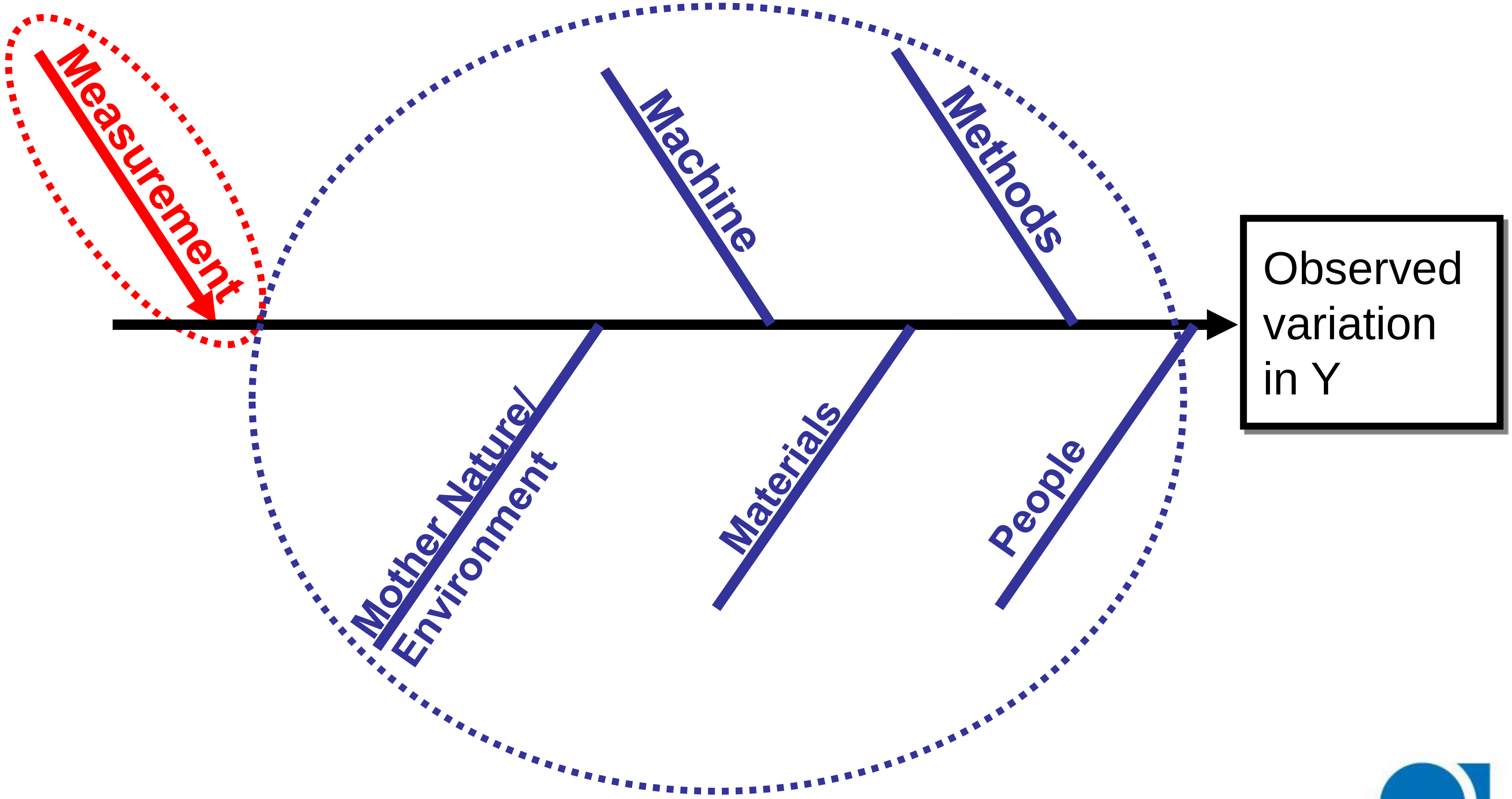
Two Fundamental Questions of MSA

#1: "Is the variation (spread) of the measurement system too large to successfully achieve the objectives?"



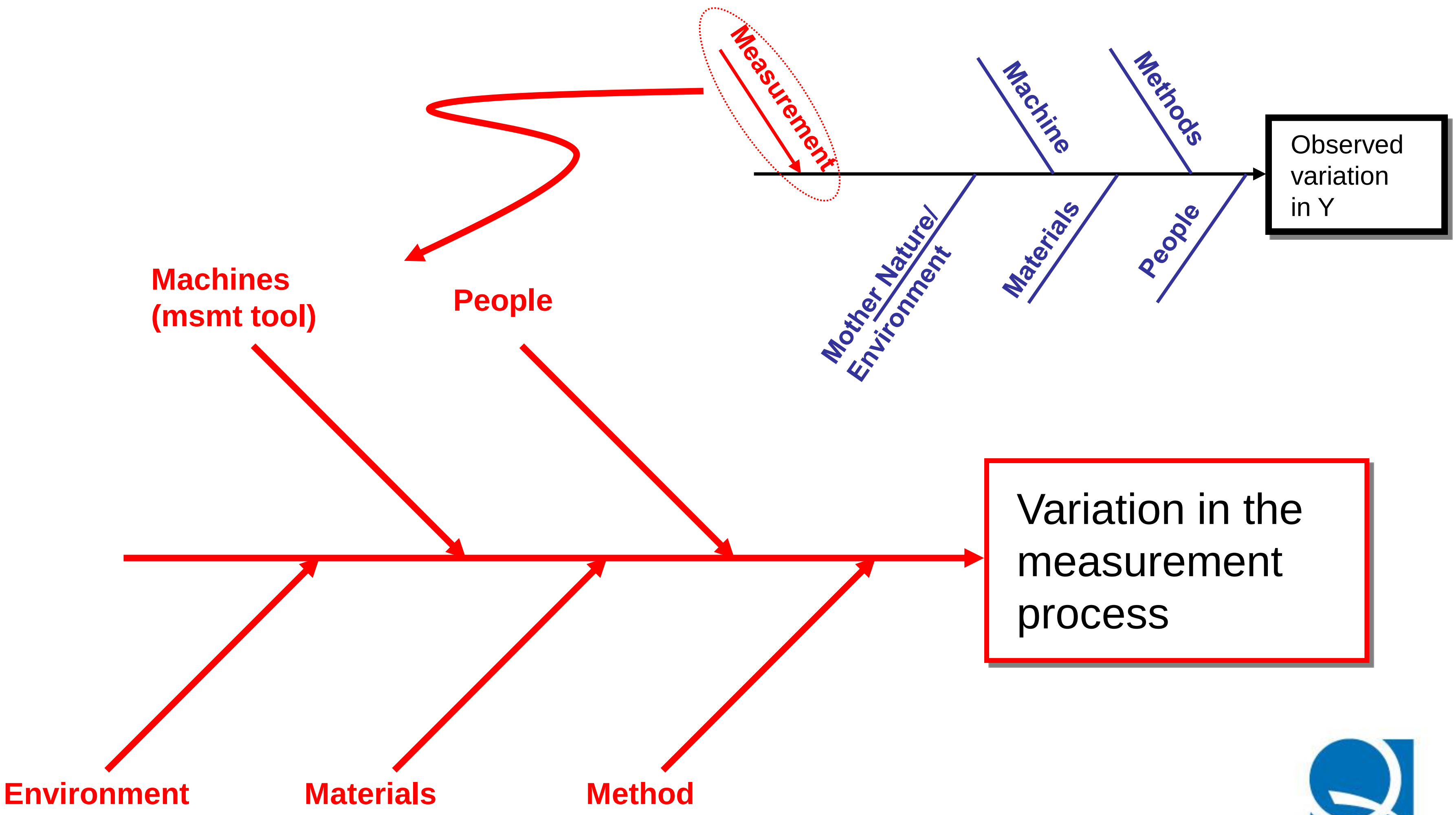
#2: "What must be done to assure that the measurement system is adequate?"

MSA Evaluates and Minimizes the Extent of Measurement Variation

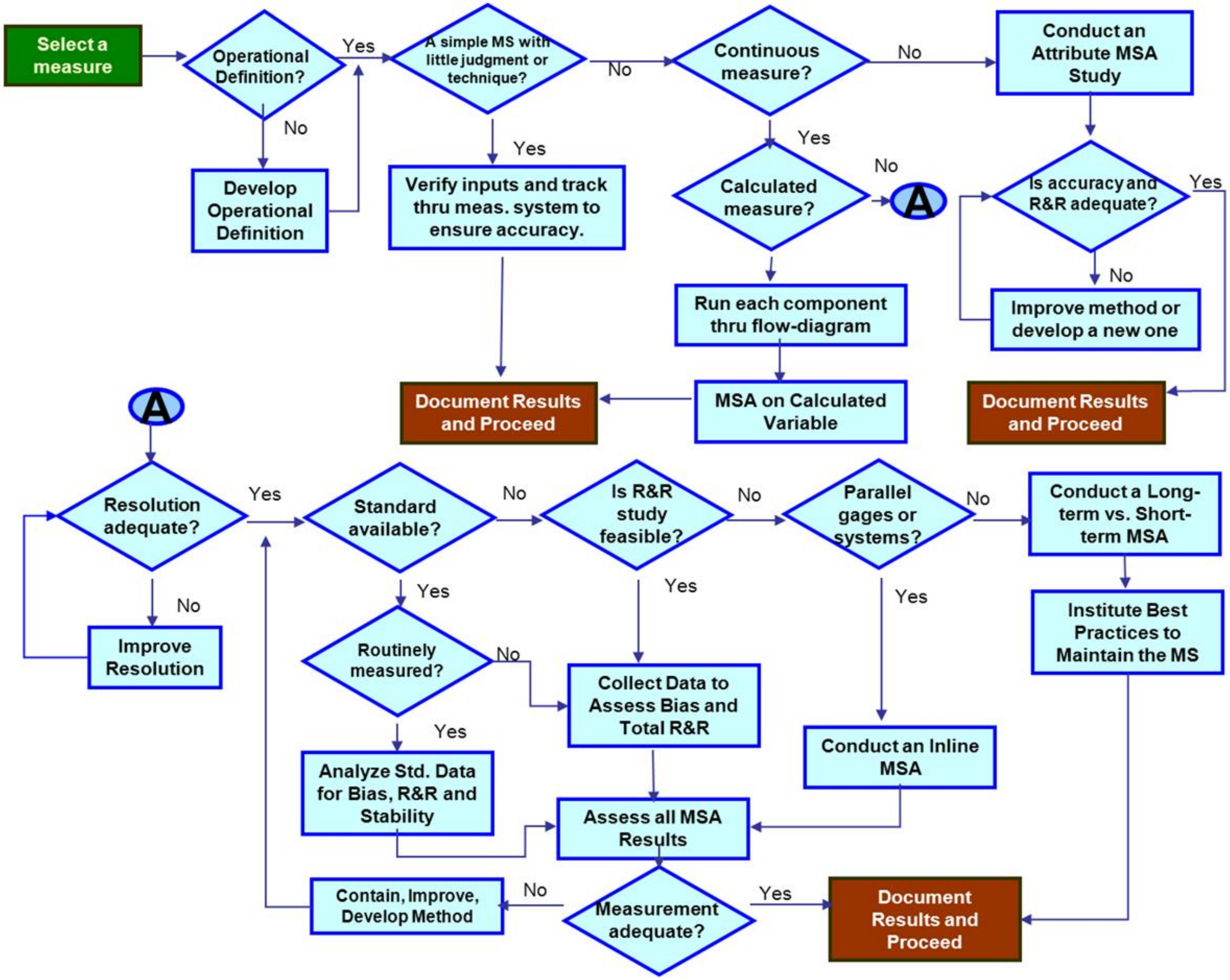


Variation caused by the measurement	+	Variation caused by the process	=	Total variation observed in Y

What Causes Measurement Variation?

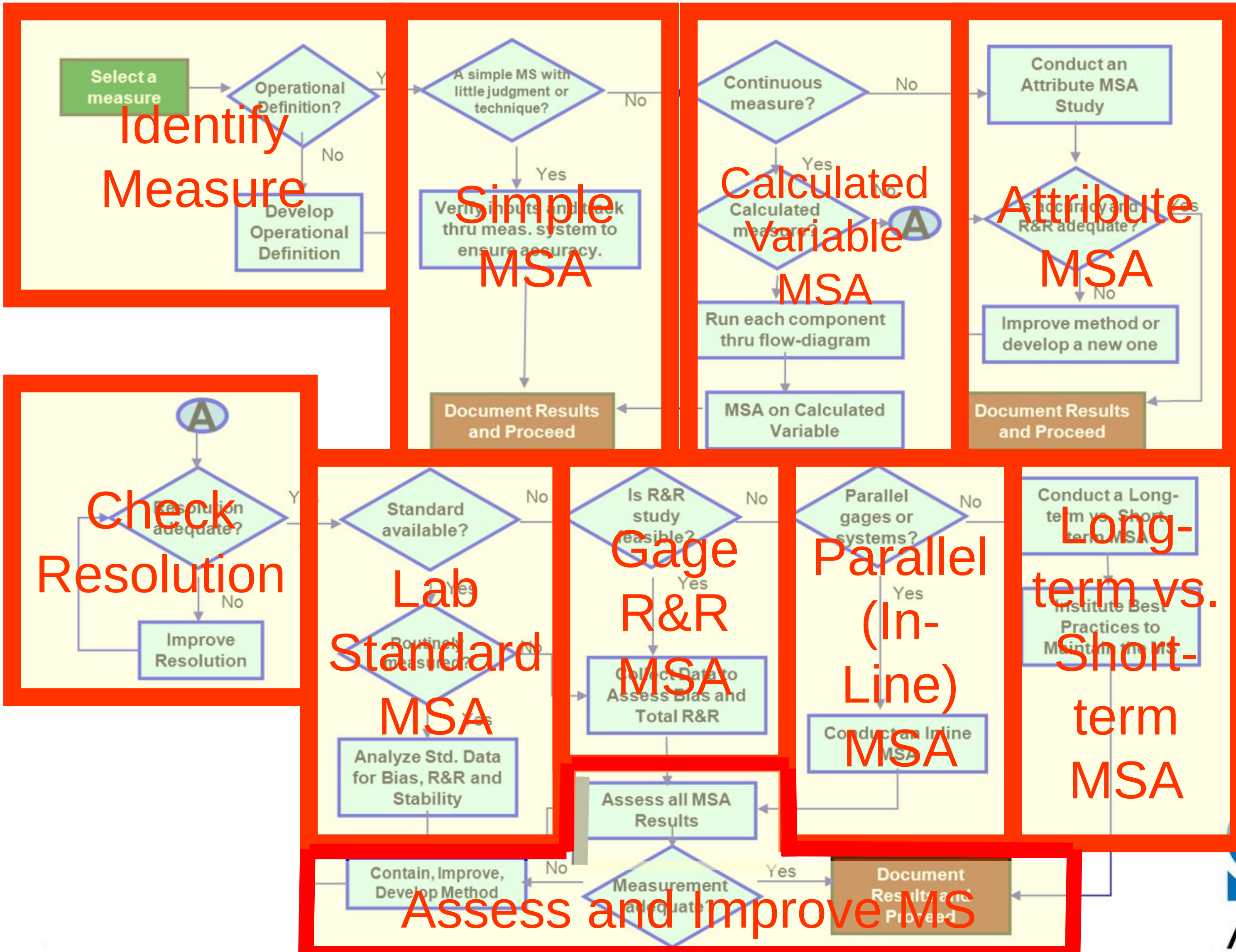


MSA Roadmap



Select a measure
Operational Definition?
Yes
No
Develop Operational Definition
A simple MS with little judgment or technique?
No
Yes
Verify inputs and track thru meas. system to ensure accuracy.
Continuous measure?
No
Yes
Conduct an Attribute MSA Study
Calculated measure?
No
A
Run each component thru flow-diagram
MSA on Calculated Variable
Document Results and Proceed
Is accuracy and R&R adequate?
Yes
No
Improve method or develop a new one
Document Results and Proceed
A
Resolution adequate?
Yes
No
Improve Resolution
Standard available?
No
Yes
Routinely measured?
No
Yes
Analyze Std. Data for Bias, R&R and Stability
Is R&R study feasible?
No
Yes
Collect Data to Assess Bias and Total R&R
Parallel gages or systems?
No
Yes
Conduct an Inline MSA
Conduct a Long-term vs. Short-term MSA
Institute Best Practices to Maintain the MS
Assess all MSA Results
Measurement adequate?
No
Yes
Contain, Improve, Develop Method
Document Results and Proceed

MSA Roadmap


Identify Measure
Select a measure
Operational Definition?
No
Develop Operational Definition
Simple MSA
A simple MS with little judgment or technique?
No
Yes
Document Results and Proceed
Calculated Variable MSA
Continuous measure?
No
Yes
Calculated measure?
Run each component thru flow-diagram
MSA on Calculated Variable
Attribute MSA
Conduct an Attribute MSA Study
R&R adequate?
No
Improve method or develop a new one
Document Results and Proceed
Check Resolution
Improve Resolution
No
Lab Standard MSA
Standard available?
No
Yes
Analyze Std. Data for Bias, R&R and Stability
Gage R&R MSA
Is R&R study feasible?
No
Yes
Assess Bias and Total R&R
Parallel (In-Line) MSA
Parallel gages or systems?
No
Yes
Long-term vs. Short-term MSA
Practices to
Assess all MSA Results
Assess and Improve MS
Contain, Improve, Develop Method
No
Yes
Document

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- If a continuous measure
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok (“ten bucket rule”)
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

Following the MSA Roadmap

- ***First, Identify Measure (and Operational Definition)***
- If a "simple measure", do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- If a continuous measure
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok ("ten bucket rule")
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

ASQ

Activity: Operational Definition

- **Count the following:**
  - (a) number of blue shirts in the room
  - (b) number of tall people in the room
  - (c) number of old people in the room
- **Record your number**
- **Tally numbers on the front chart pad**
- **Discuss the results**

ASQ

“Blue Shirt” Operational Definition

What is a shirt and how can you determine if it is blue?

- A shirt is any garment that covers 70% or more of the torso, above the skirt or pants of the wearer, and the lower extremity of which garment (when hanging freely) falls between 3” and 7” (incl.) below the utmost line of the skirt or pants. If the wearer is wearing neither skirt nor pants, then the garment in question is not a shirt.
- Any shirt so defined will be held to be blue if more than 50% of its outward and visible surface (as worn) is blue in color.
- Any color will be deemed to be blue if it matches any portion of the marked ranges on the color cards provided when both shirt and cards are judged by an inspector medically certified as having passed the U.S.A.F. test for color-blindness.

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- ***If a "simple measure", do Simple MSA***
- If a discrete (attribute) measure, do **Attribute MSA**
- If a continuous measure
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok ("ten bucket rule")
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

Simple Measurements

- **Examples of possible simple measurements**
  - Cycle Times (Hours, Days, Months)
  - Cost to Repair a Piece of Equipment
  - Number of Safety Incidents
  - Number of Complaints
- **What is common among all of these measurements?**
  - Require a good operational definition
  - Little judgment involved in determining the result
  - The potential for variation still exists!
    - Data entry
    - Incorrect formulas
    - Not applying the operational definition

Examples: Simple MSA

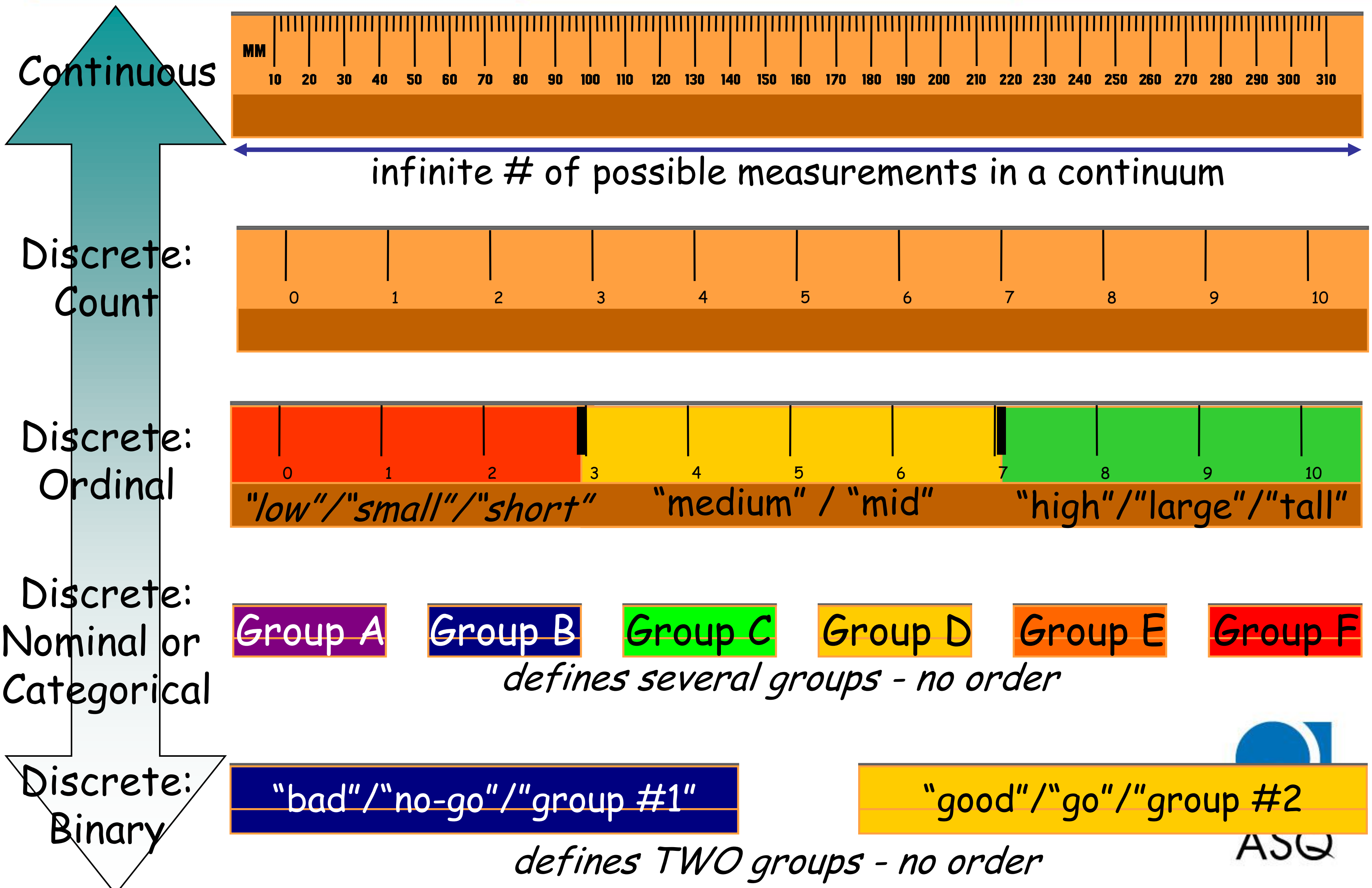
- A project was undertaken to reduce start-up cycle time for a manufacturing process (from 12 hours to 6 hours). The time was automatically calculated in a Distributive Control System (DCS). The project leader manually recorded the beginning and end times for 10 different startups and compared these to those in the system. All times were found to match within 1 minute. This completed the MSA.

- A project to reduce transportation costs tracked expenses for company and rental cars. A sample of 40 trips taken using company cars was compared to expenses reported in the accounting system. 25% of the trips were not found in the system (people were entering these expenses into the wrong location). The system was subsequently changed and additional data was collected to confirm the accuracy. The MSA was considered complete.

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- ***If a discrete (attribute) measure, do Attribute MSA***
- If a continuous measure
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok (“ten bucket rule”)
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

Discrete vs Continuous Data



Analysis of an Attribute MSA

- **For binary data use Assessment Agreement**
- **For nominal data use Assessment Agreement and Kappa**
- **For ordinal data:**
  - Use Assessment Agreement and Kappa if you want to know amount of absolute agreement
  - Use Kendall's coefficients for relative amount of agreement (normally the most useful)

Analysis Interpretation*

Stoplight Color	Assessment Agreement	Kappa	Kendall's	Decision	Action
RED (not acceptable)	<70%	<0.7	<0.7	Measurement system incapable	Improve measurement system before proceeding with project.
YELLOW (acceptable)	70-90%	0.7-0.9	0.7-0.9	Measurement system moderately capable	Consider improving measurement system while proceeding with project.
GREEN (preferred)	>90%	>0.9	>0.9	Measurement system capable	Measurement system adequate, proceed with project.

*Guidelines only; specific situations may suggest tighter or more relaxed requirements than these

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- ***If a continuous measure***
    - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
    - ***Check that Resolution is ok (“ten bucket rule”)***
    - If lab data sufficient, do **Lab Standard MSA**
    - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
    - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
    - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

Evaluating Measurement Resolution

10-Bucket Rule for Evaluating Measurement Resolution

Does the measurement tool work on a scale that gives at least 10 subdivisions within the range of interest?

Example: Process wait times often exceed 15 minutes, and we want to eventually get them down to 10 minutes or less.

- What level of resolution is required for measuring wait times?
- Which of the following have adequate resolution for measuring wait times?
  - Digital stopwatch (measures to nearest 0.01 sec)
  - Digital wall clock (displays HH:MM:SS time)
  - Analog wall clock (usual clock face)
  - Sundial

ASQ

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- ***If a continuous measure***
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok (“ten bucket rule”)
  - ***If lab data sufficient, do Lab Standard MSA***
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

Continuous MSA Roadmap

Resolution? → Use the 10-bucket Rule

OK

Bias?

OK

Stability?

OK

Linearity?

OK

"Precision" (Repeatability & Reproducibility?)

MSA evaluates the Performance of the Measurement System by evaluating each of these

Then

If unacceptably high measurement variation is found, corrective action is taken to reduce to acceptable levels

ASQ

Use of Standards

- **Standards are run in most laboratories to maintain control of test methods**
- **On-line instruments may also have standards data available**
- **A look at this data will often provide information on stability, bias and total R&R**
- **If more than one standard is available then linearity may also be tested from the standards data**
- **Seriously consider if and how this will represent typical variation. Issues include:**
  - Samples are not blind
  - Different operators may test the standard vs. run the routine tests

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- ***If a continuous measure***
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok (“ten bucket rule”)
  - If lab data sufficient, do **Lab Standard MSA**
  - ***If crossed, nested or expanded Gage R&R feasible, do Gage R&R MSA***
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

Gage R&R Study Design

- **Select items from which to take measurements**
  - Items should represent typical range observed from the process
  - Recommend at least 10 samples
- **Select at least two operators to take measurements**
  - Preferred: use as many different operators as time will allow
  - Use only personnel that will normally make these measurements
- **Each operator measures each part at least two times**
  - df = (#items) * (#operators) * (#repeats per operator - 1) at least 30
  - Restrictions lead to **"nested"** rather than **"crossed"** analyses
- **Consider "expanded" data collection structures**
  - On process generating the "items" – process DOE factors
  - On process generating the "measurements" – eg, **multiple instruments (within lab) or multiple laboratories**

AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics

What is Statistics? Donate Join Login

ABOUT MEMBERSHIP EDUCATION PUBLICATIONS MEETINGS POLICY & ADVOCACY YOUR CAREER SEARCH

W. J. Youden Award in Interlaboratory Testing

No winner chosen for 2017.

About the Award

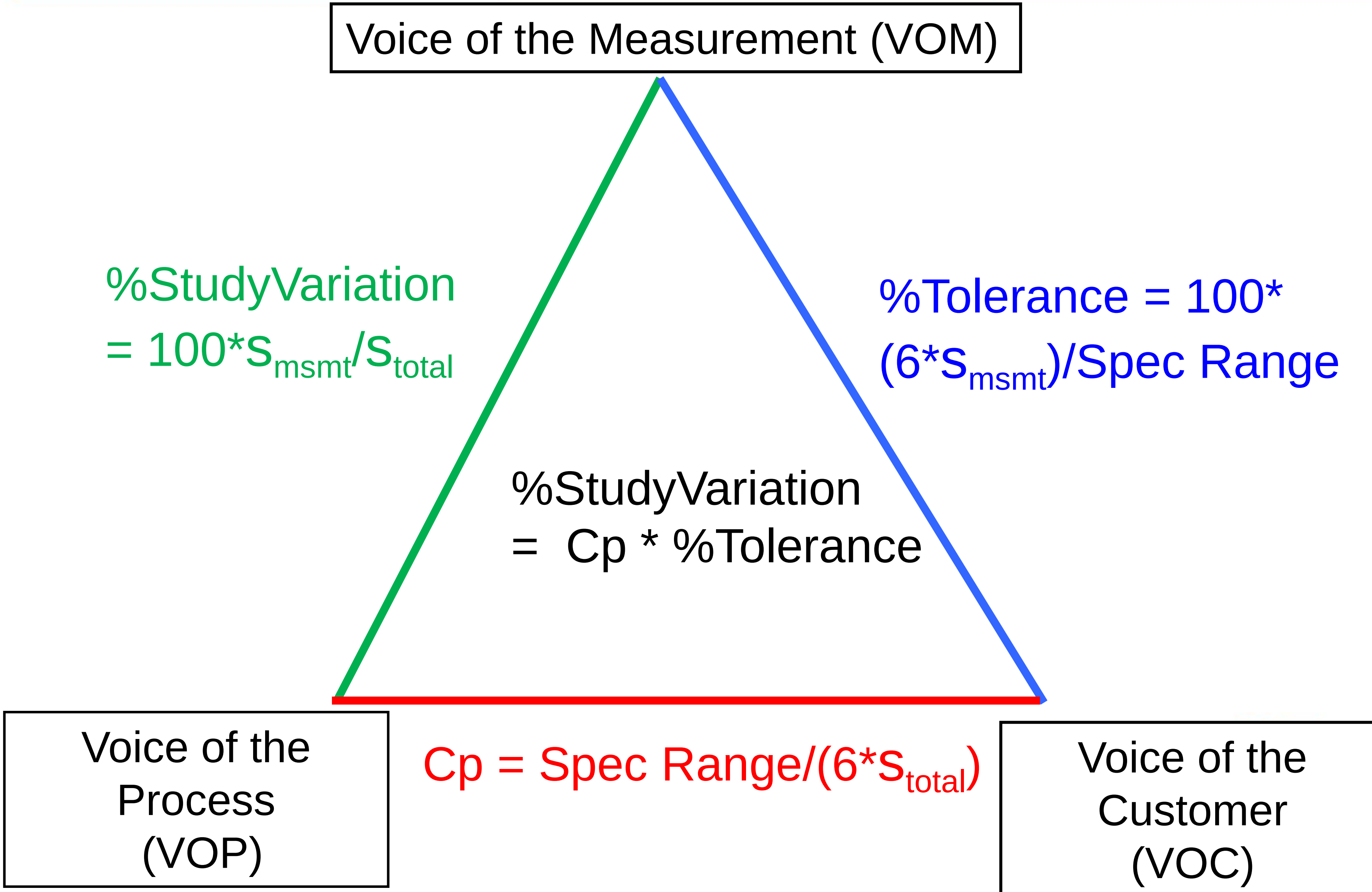
The W. J. Youden Award in Interlaboratory Testing was established in 1985 to recognize the authors of publications that make outstanding contributions to the design and/or analysis of interlaboratory tests or describe ingenious approaches to the planning and evaluation of data from such tests. Award recipients are presented with an engraved award and $1,000, which is divided evenly among the recipients. The award is presented annually if, in the opinion of the awards committee, an eligible and worthy publication is

Recent Award Recipients

2015: Alexander Franks, Gábor Csárdia, D. Allan Drummond, and Edoardo M. Airoldi
2014: Yunda Huang, Ying Huang, Shuying Sue Li, Felicity Zoe Moodie, and Steven Self
2013: Lane F. Burgette and Jerome P. Reiter
2012: David Dunson and Garritt L. Page
2011: Ryan Browne, Jock MacKay, and Stefan Steiner

ASQ

Linking Voice of Measurement



Analysis Interpretation*

Stoplight Color	%Study Variation	%Tolerance	Decision	Action
RED (not acceptable)	>30%	>30%	Measurement system incapable	Improve measurement system before proceeding with project.
YELLOW (acceptable)	10-30%	10-30%	Measurement system moderately capable	Consider improving measurement system while proceeding with project.
GREEN (preferred)	<10%	<10%	Measurement system capable	Measurement system adequate, proceed with project.

*Guidelines only; specific situations may suggest tighter or more relaxed requirements than these

ASQ

Following the MSA Roadmap

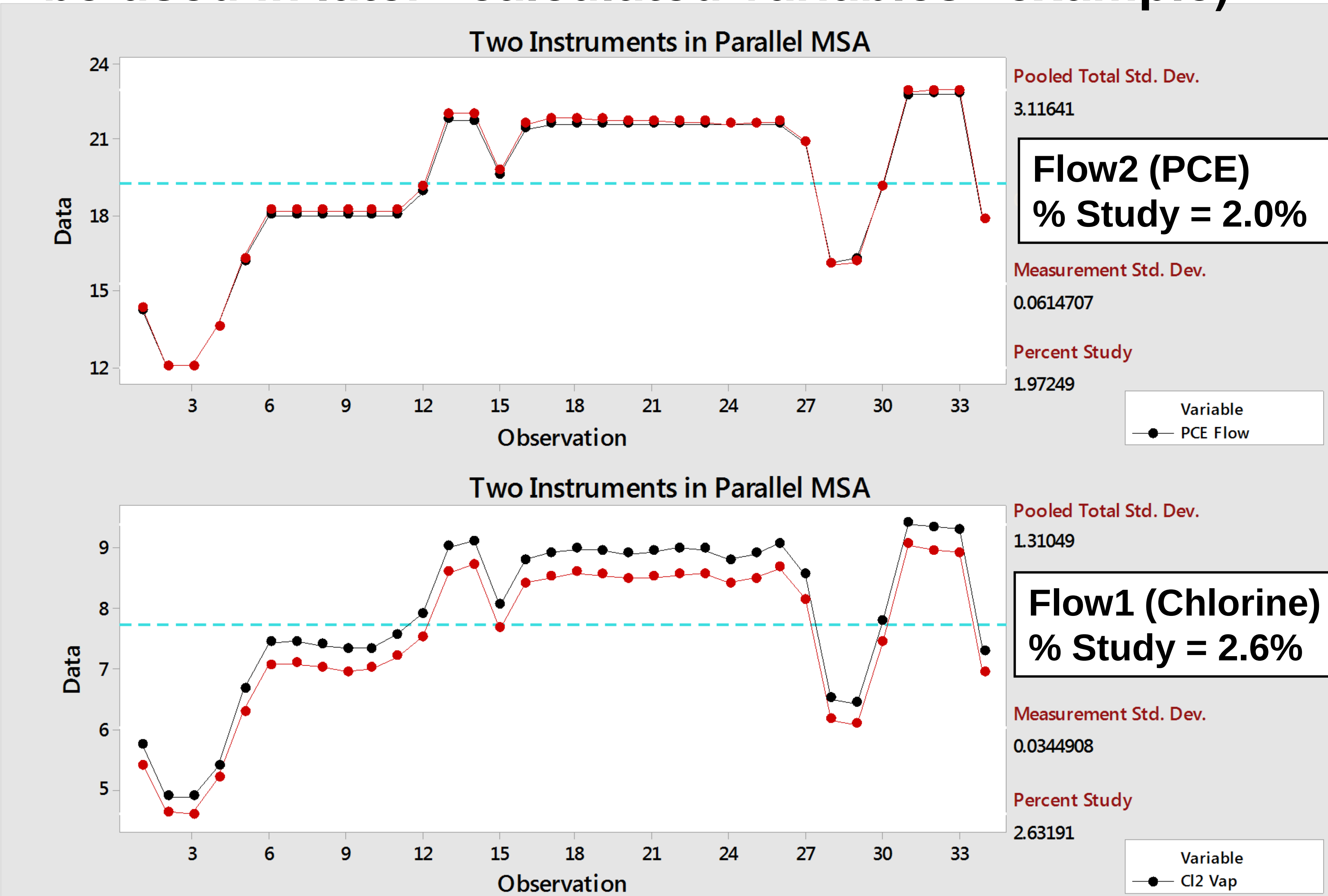
- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- ***If a continuous measure***
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok (“ten bucket rule”)
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - ***If parallel or in-line instruments, do Parallel (In-Line) MSA***
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

Example: Two instruments in-line or on parallel lines

- **Design**
  - Two instruments, either in series or in parallel
  - We can make measurements on both at the same time so that we can assume that they are seeing the same “sample”
  - Collect enough data to allow all potential sources of measurement variation
- **Calculations**
  - Measurement variation is calculated from the differences between the pairs of readings taken by the two instruments at a given time (after subtracting out instruernt bias)
  - Process variation calculated individually from each instrument’s readings and then pooled between the two instruments.

ASQ

Two examples from same process (with results to be used in later "calculated variables" example)



Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- ***If a continuous measure***
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok (“ten bucket rule”)
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - ***As last resort, do Long-Term vs. Short-Term MSA***
- Finally, **Assess and Improve MS** (if needed) and document

ASQ

Long-term vs. Short-term MSA

- **Data Requirements**
  - Short-term variability dominated by measurement variation
    - Sampling frequency short relative to process inertia
  - Long-term variability indicative of the normal range of the data
- **Analysis**
  - Use moving-range-based calculations
  - Alternatively, variance components analysis (requires specific grouping)

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a "simple measure", do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- ***If a continuous measure***
  - ***If calculated from several continuous variables, do MSA on each component and then do Calculated Variable MSA***
  - Check that **Resolution** is ok ("ten bucket rule")
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- Finally, **Assess and Improve MS** (if needed) and document

ASQ

How Does Error Propagate?

From Physical Chemistry for independent variables:

For the equation $Y = f(x_1,x_2,\ldots x_n)$, error propagates as

$$(dY)^2 = (\partial f/\partial x_1)^2(dx_1)^2 + (\partial f/\partial x_2)^2(dx_2)^2 + \ldots(\partial f/\partial x_n)^2(dx_n)^2$$

or written another way

$$(\sigma_Y)^2 = (\partial f/\partial x_1)^2(\sigma_{x_1})^2 + (\partial f/\partial x_2)^2(\sigma_{x_2})^2 + \ldots(\partial f/\partial x_n)^2(\sigma_{x_n})^2$$

References:
"Theory of Error", by Yardley Beers or any Physical Chemistry Lab Text
Chapter 9 of "Statistics for Experimenters" by Box, Hunter and Hunter

Common Approaches to Propagation of Error (POE)

- **By calculation**
  - Only feasible for simple calculations, or if you like doing the calculus!
- **By setting up a simulation (monte carlo)**
  - Set up individual columns representing the measurement error in each X and perform the Y calculation to simulate how the error propagates

Activity: In-Line and Calculated Variable MSA

- **A project has been undertaken to improve yield of a chemical process. Yield is calculated based on two flow meters using the following formula.**

$$\text{Yield} = (\text{Flow}_1/\text{Flow}_2) \times 2.339$$

- **Flow_1 and Flow_2 are measured via on-line instruments each with parallel flow meters. The data represents samples taken at the same time from each instrument. Both Flow_1 and Flow_2 instruments are expected to be close to equal in measurement variation.**

Activity: In-Line and Calculated Variable MSA Continued

From R&R and process studies we have:

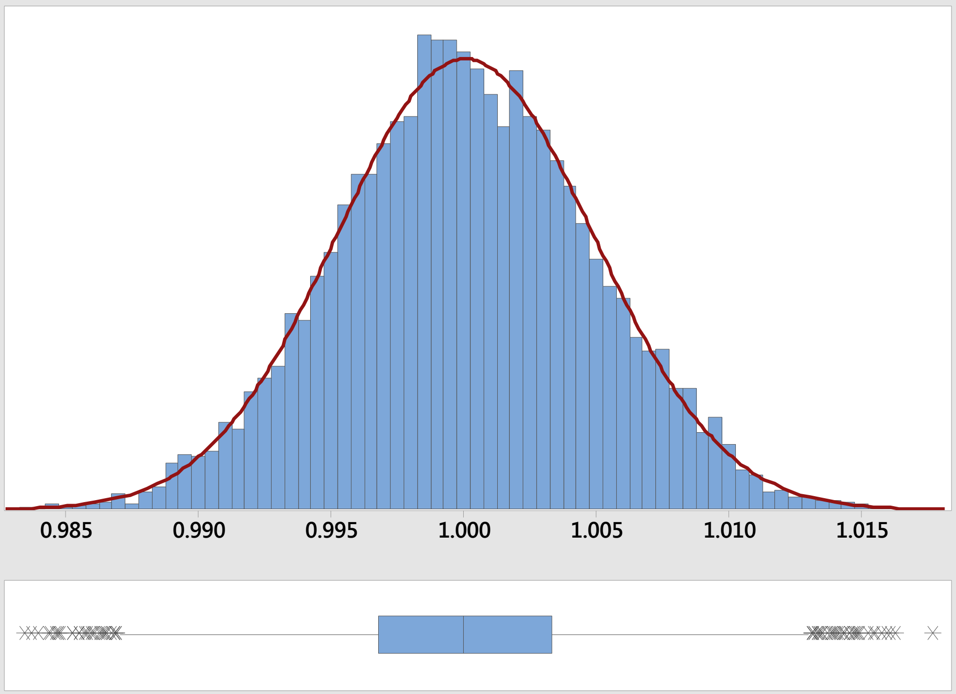
	Average	R&R StdDev	Process StdDev
Flow 1	8.85	0.0345	1.331
Flow 2	20.7	0.0615	3.092
Yield	0.96067		0.0139

Work in teams to determine

1. **Measurement %study variation for $Flow_1$ and $Flow_2$**
2. **Measurement %study and % tolerance for Yield**
   -- Determine at $Flow_1$=8.85 and $Flow_2$=20.7
   -- Note Yield Specifications: LSL = 0.85 to USL = 1.15
3. **Does this measurement system appear to be adequate?**

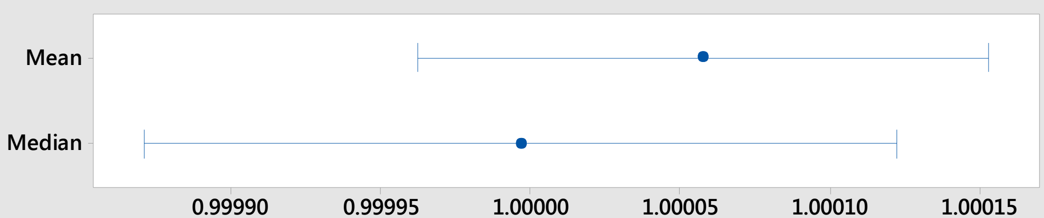
From Simulation: Yield Measurement StdDev is 0.0049

Summary Report for Y



Anderson-Darling Normality Test	
A-Squared	0.52
P-Value	0.188
Mean	1.0001
StDev	0.0049
Variance	0.0000
Skewness	0.0127857
Kurtosis	-0.0498914
N	10000
Minimum	0.9834
1st Quartile	0.9968
Median	1.0000
3rd Quartile	1.0033
Maximum	1.0177
95% Confidence Interval for Mean	
1.0000	1.0002
95% Confidence Interval for Median	
0.9999	1.0001
95% Confidence Interval for StDev	
0.0048	0.0049

95% Confidence Intervals



Answer: In-Line and Calculated Variable MSA

1. **Flow 1 %Study Variation = 0.0345/1.331 × 100% = 2.63%**

   Flow 2 %Study Variation = 0.0615/3.092 × 100% = 1.97%

2. **Yield % Study Variation and % Tolerance**

$$\sigma_Y = \sqrt{(\partial Y / \partial F_1)^2 \sigma_{F_1}^2 + (\partial Y / \partial F_2)^2 \sigma_{F_2}^2}$$

$$\sigma_Y = \sqrt{(2.339 / F_2)^2 \sigma_{F_1}^2 + (-2.339 \times F_1 / F_2{}^2)^2 \sigma_{F_2}^2}$$

At F_1 = 8.85 and F_2 = 20.7

$$\sigma_Y = \sqrt{(2.339/20.7)^2 (0.0345)^2 + (-2.339 \times 8.85/(20.7)^2)^2 (0.0615)^2}$$

$$= \sqrt{(0.01277 \times 0.00119) + (0.002334 \times 0.003782)}$$

= 0.0049

%Study Var. = 0.0049 / 0.0139 × 100% = 35.3%

%Tolerance = 6 × 0.0049 / (1.15-0.85) × 100%

= 0.0294/0.30 × 100% = 9.8%

Summary: VOM vs. VOP vs. VOC for In-Line and Calculated MSA Activity

Voice of the Measurement (VOM)

%Study Variation
$= 100 \times s_{msmt} / s_{total}$
$= 35\%$

%Tolerance
$= 100 \times (6 \times s_{msmt})$/Spec. Range
$= 10\%$

%Study Variation
= C_P × %Tolerance
or
35% = (3.5) × (10%)

Voice of the Process (VOP)

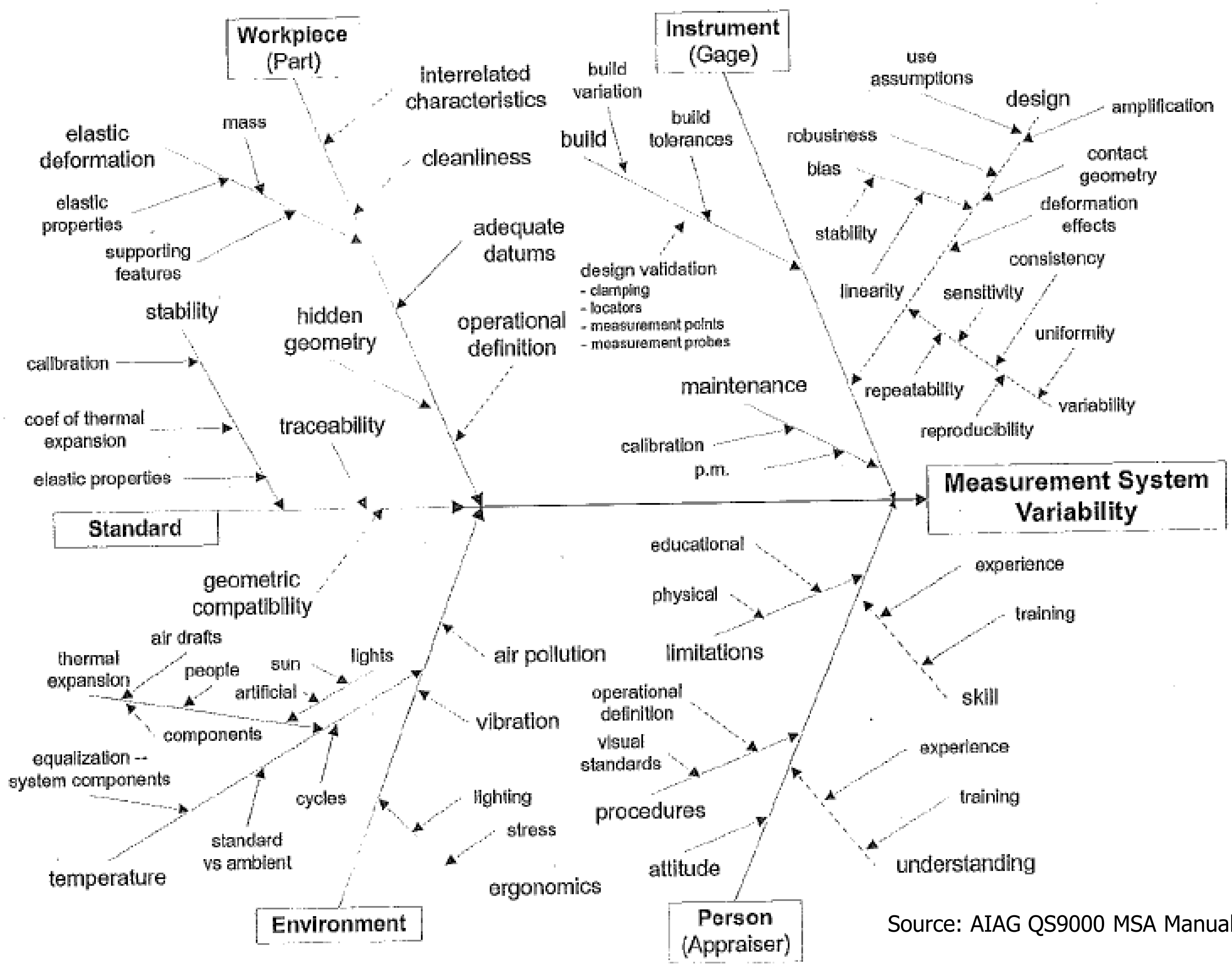
C_P = Spec. Range/$(6 \times s_{total})$
$= 3.5$

Voice of the Customer (VOC)

Following the MSA Roadmap

- First, **Identify Measure** (and Operational Definition)
- If a “simple measure”, do **Simple MSA**
- If a discrete (attribute) measure, do **Attribute MSA**
- If a continuous measure
  - If calculated from several continuous variables, do MSA on each component and then do **Calculated Variable MSA**
  - Check that **Resolution** is ok (“ten bucket rule”)
  - If lab data sufficient, do **Lab Standard MSA**
  - If crossed, nested or expanded Gage R&R feasible, do **Gage R&R MSA**
  - If parallel or in-line instruments, do **Parallel (In-Line) MSA**
  - As last resort, do **Long-Term vs. Short-Term MSA**
- ***Finally, Assess and Improve MS (if needed) and document***

Improving the Measurement System



Source: AIAG QS9000 MSA Manual

Improving Measurement Variation By Averaging

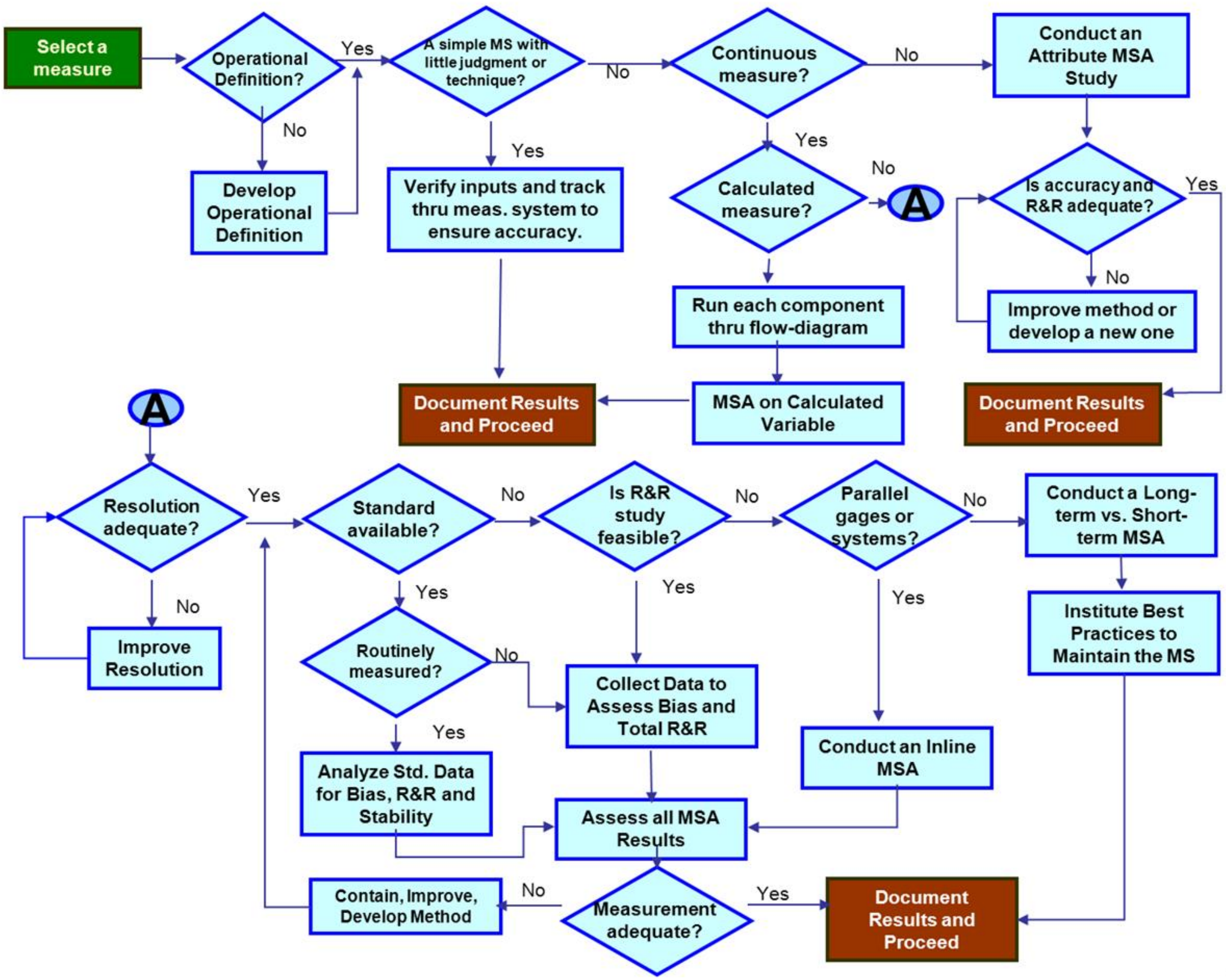
- **One option for improving the measurement variation that is always available is to report the average of multiple results. This can be considered when the measurement is already performing near its best or improvement of the measurement is cost prohibitive.**
- **By taking 2 measurements each time (1 operator) and averaging**

$$\sigma^2_{R\&R} = \frac{\sigma^2_{repeatability}}{2} + \sigma^2_{reproducibility}$$

- **By taking those same 2 measurements using 2 operators**

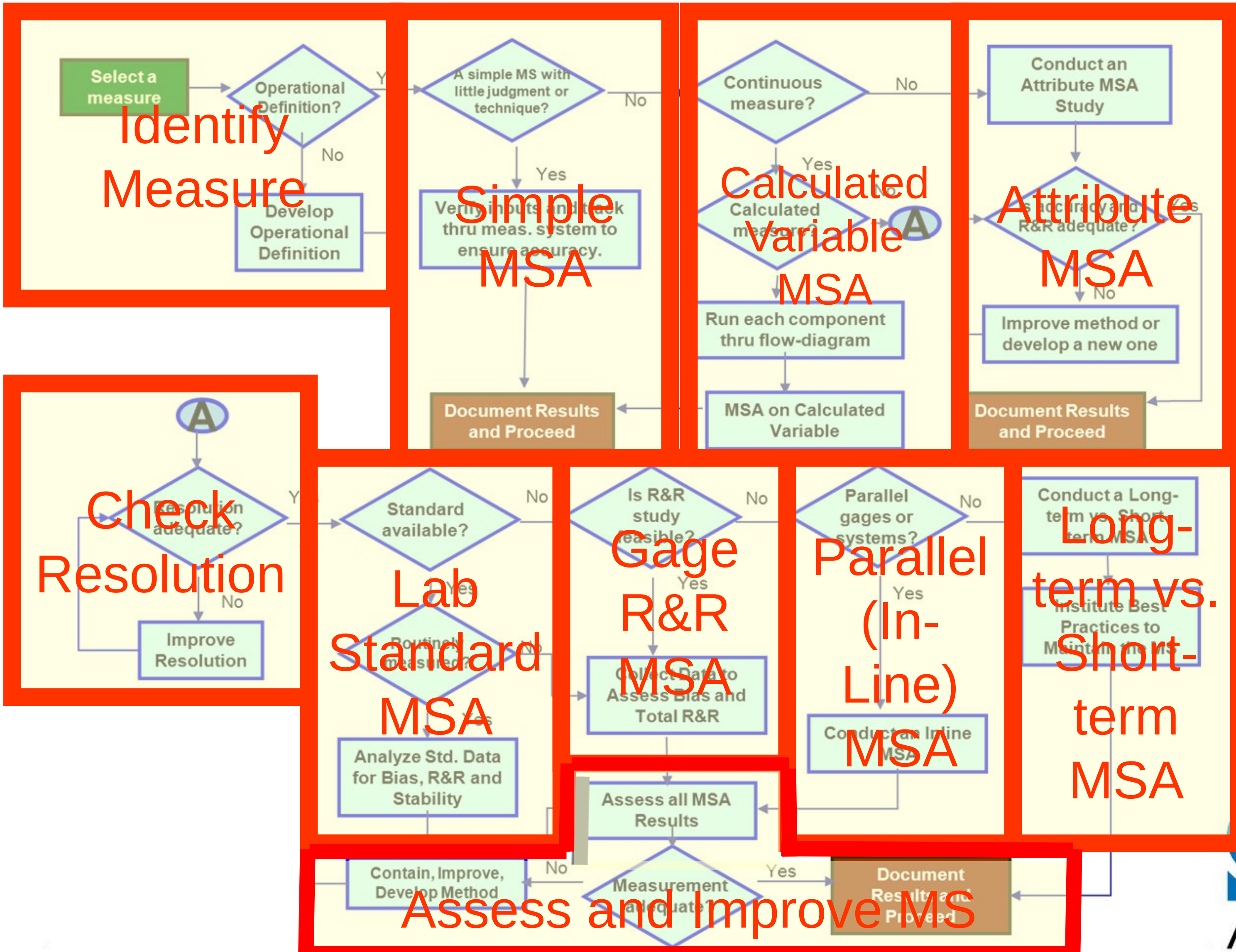
$$\sigma^2_{R\&R} = \frac{\sigma^2_{repeatability} + \sigma^2_{reproducibility}}{2}$$

MSA Roadmap



Select a measure
Operational Definition?
Yes
No
Develop Operational Definition
A simple MS with little judgment or technique?
No
Yes
Verify inputs and track thru meas. system to ensure accuracy.
Document Results and Proceed
Continuous measure?
No
Conduct an Attribute MSA Study
Yes
Calculated measure?
No
A
Run each component thru flow-diagram
MSA on Calculated Variable
Is accuracy and R&R adequate?
Yes
No
Improve method or develop a new one
Document Results and Proceed
A
Resolution adequate?
Yes
No
Improve Resolution
Standard available?
No
Yes
Routinely measured?
No
Yes
Analyze Std. Data for Bias, R&R and Stability
Is R&R study feasible?
No
Yes
Collect Data to Assess Bias and Total R&R
Parallel gages or systems?
No
Yes
Conduct an Inline MSA
Conduct a Long-term vs. Short-term MSA
Institute Best Practices to Maintain the MS
Assess all MSA Results
Measurement adequate?
No
Contain, Improve, Develop Method
Yes
Document Results and Proceed

MSA Roadmap



ASQ

References

- AIAG Measurement Systems Analysis, Reference Manual, 3rd ed. (2003). Automotive Industry Action Group, Southfield, MI.
- Bailey, S. and Johnson, L.A. (2012). "Implementing an Expanded Gage R&R (Repeatability and Reproducibility) Study," ASQ World Conference on Quality and Improvement, Anaheim, CA, May 21 3-4pm.
- Burdick, R.K., Borror, C.M., Montgomery, D.C. (2003). "A Review of Methods for Measurement Systems Capability Analysis," *Journal of Quality Technology,* Vol 35, No. 4, p342.
- Trusting Measurement Results in the Chemical and Process Industries (2001). ASQ Chemical and Process Industries Division Chemical Interest Committee, Milwaukee, WI.

ASQ

Coauthors & Acknowledgements

DuPont MSA Roadmap & Toolkit

- Pat DeFeo, DuPont Safety and Construction
- Steve Larson, Sam Houston State University
- Andy Rabbani (deceased), Momentive
- Dave Schussler, Cigna

Other Acknowledgements

- Stephanie DeHart, Eastman
- Jen Van Mullekom, Virginia Tech

Strategy of Experimentation (SOE) History

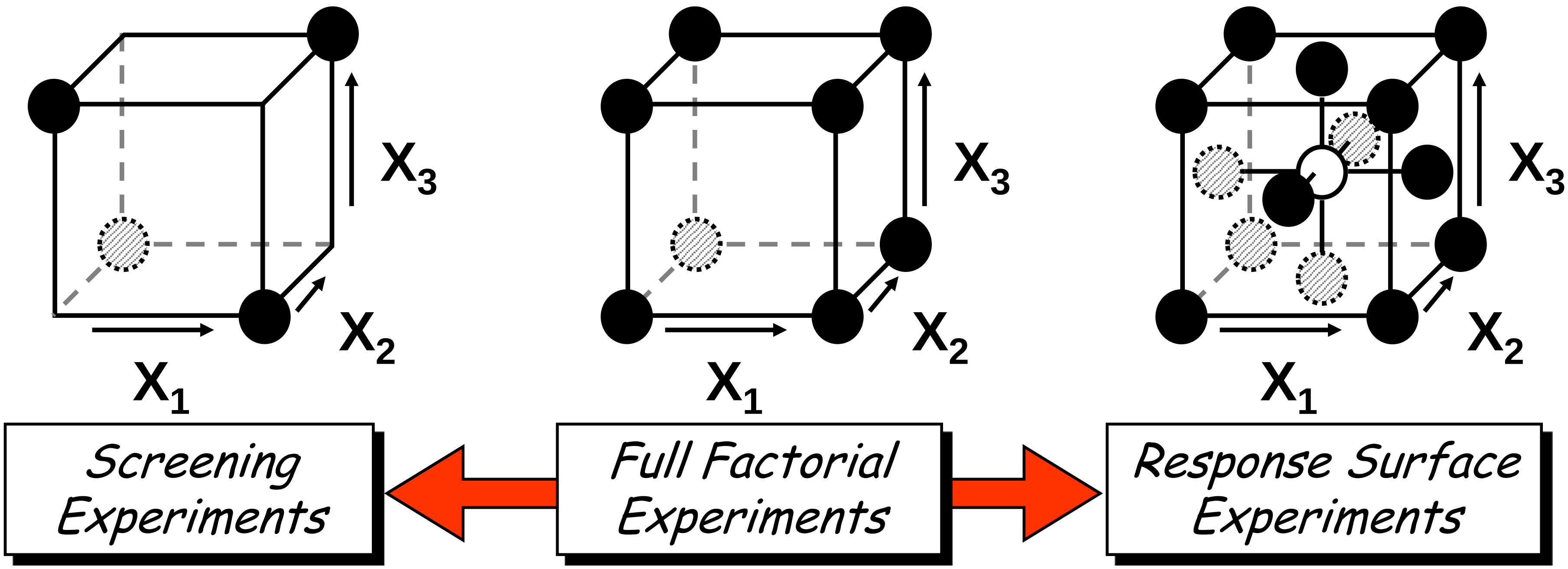
Year	DOE Theory	Area	Software	Course	Audience
1920s	Agriculture split plot experiments (Fisher & Yates)				
1930s					
1940s	Plackett-Burman designs				
1950s	Response surface methods (Box et al.)	R&D	Hand calculations Univac		
1960s	Mixture designs (Scheffe)	R&D MFG		SOE (1964)	Internal offering
1970s	Design optimality and computer-aided designs Conjoint analysis		Univac programs developed internally	SOE SOFD added	Internal & External offering (Began selling course externally in 1974)
1980s	Robust parameter design	R&D, MFG Agriculture	RS/Discover (VAX), Minitab® (VAX), ECHIP (PC)	Last major content update	
1990s	Industrial split plot designs	R&D, MFG, Agriculture Tech. Sales/Marketing	Design Expert (PC), JMP® (PC), Minitab® (PC)	Software updates	Internal & DuPont Customers
2000s	Computer experiments	R&D, MFG, Agriculture, Tech. Sales/Marketing	Minitab®, JMP®, SAS®	SOE & SOEFD course DOE	

Milestones In Designed Experiments
From "Design Of Experiments Makes A Comeback",
Chemical & Engineering News, April 1, 2013 Issue - Vol. 91 Issue 13

1. Ronald A. Fisher (1890 – 1962)
2. George E. P. Box (1919 – 2013)
3. **"In the 1970s, DuPont's Quality Management & Technology Center trains DuPont employees on DOE and offers the training to other companies. The company continues the service into the 1990s."**
4. FDA Quality by Design (2011)

Evolution of the Experimental Environment

Full Factorials as Building Blocks for Screening and Response Surface Experiments



Over 40,000 students internally and externally trained in DuPont's Strategy of Experimentation (SOE)!

ASQ

1990 FTC Shewall Award-Winning Papers

II 10:30 to 12:00	**Experimental Design** *Dubious Observations in 2^n Factorial Experiments* Peter W.M. John, University of Texas *Supersaturated Designs in Quality Improvement* Joseph G. Voelkel, Rochester Institute of Technology **Moderator:**	**Characterizing Variability** *Assessing the Measurement Process: Can I Find the Forest Through the Trees?* Thomas J. Boardman, Colorado State University *Understanding Process and Measurement Variability* John T. Herman, Du Pont **Moderator:**	**Tutorial** *Optimization and Variation Reduction* Wayne A. Taylor, Baxter Healthcare Corporation **Moderator:**
12:15 to 1:45	**LUNCHEON**	*Title* Speaker: Michael B. Emery, Du Pont, Vice President of Engineering Presiding: James M. Lucas, Du Pont, Chairman ASQC-CPID	
III 2:00 to 3:30	**Robust Design** *Understanding Robust Design, Loss Functions, PERMIAs, and Signal to Noise Ratios* Thomas J. Lorenzen and Miguel A. Villalobos, General Motors Research Laboratories *Giving Your Response Surface a Robust Workout* Steven P. Bailey, Kenneth A. Chatto, William H. Fellner, and Charles G. Pfeifer, Du Pont	**JQT Session** *Some SPC Methods for Autocorrelated Data* Douglas C. Montgomery and Christina Mastangello, Arizona State University *Statistical Tolerancing Based on Consumer's Risk Considerations* Robert G. Easterling et al., Sandia National Labs **Session Organizer and**	**Tutorial** *Using Reviews To Improve Your Process for Planning and Implementation* Spencer B. Graves, Hewlett Packard Co. Casey Collett, Goal/QPC **Moderator: Beth [illegible]**

GIVING YOUR RESPONSE SURFACE A ROBUST WORKOUT

Steven P. Bailey
Du Pont Company
Newark, DE 19714-6091

Kenneth A. Chatto
Consultant
Lugoff, SC 29078

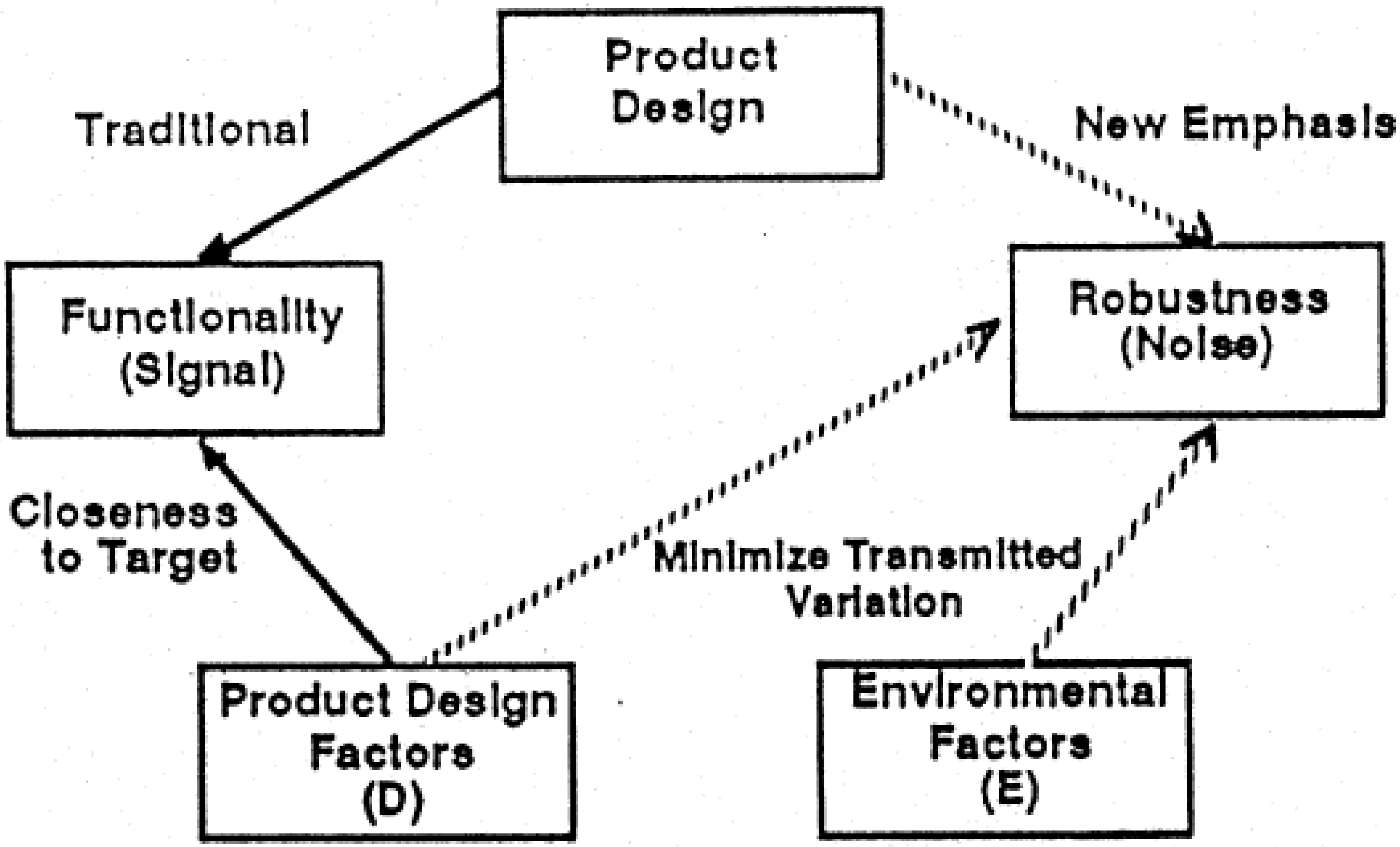
William H. Fellner
Du Pont Company
Newark, DE 19714-6091

Charles G. Pfeifer
Du Pont Company
Newark, DE 19714-6091

ABSTRACT

In this paper we demonstrate how classical experimental design methodology can be used to develop quality products that have characteristics important to the customer (functionality) while minimizing the effects of uncontrolled sources of variation during manufacture or consumption (robustness). In particular, we describe an iterative process for extracting information from a response surface equation that enables the practitioner to be confidently in control of the analysis. Several new techniques are introduced which simplify interpretation of traditional analyses found in commercial software. The approach described is based on lessons learned from three decades of experience within Du Pont and offers some important advantages over methods attributed to Taguchi.

Figure 1
Product Design: Functionality and Robustness



Designing a Circuit

$$\text{Current}(Y) = \frac{V}{\sqrt{R^2 + (2\pi f L)^2}}$$

where V = Voltage

R = Resistance

f = frequency

L = Inductance

Example adapted from Taguchi, Genichi. "The Development of Quality Engineering" in *The American Supplier Institute Journal* 1, No. 1 (Fall 1988).

Circuit Design Goal

To: **Choose "nominal" values for the design factors (Resistance and Inductance)**

In a Way That: **Keeps the property (Current) as close as possible to the desired aim of 10 Amps and**

Minimizes transmitted variability due to variation in environmental factors (Voltage and Frequency) and deviations (tolerances) in the design factors (Resistance and Inductance) from the nominal settings.

So That: **The circuit design is functional and robust.**

Circuit Design Factors and Response

Factors	Name	Low (1)	Middle (0)	High (+)	Units	Tolerances
V	Voltage	90	100	110	V	±10 (Full Range)
R	Resistance	0.5	5.0	9.5	Ohms	± 0.5
f	Frequency	50	55	60	Hz	± 5 (Full Range)
L	Inductance	0.01	0.02	0.03	H	± 0.002

Property	Name	Aim	Units	Spec Range
Y	Current	10	Amps	± 2

Data for Taguchi's Analysis

R = 9.5 and L = 0.010 is Taguchi's preferred design.
Taguchi used signal-to-noise (S/N) ratio
but we will use mean squared error (MSE).

R	L	NMin	NMax	Avg	StDev	MSE
0.5	0.010	19.43	43.77	31.60	17.21	614.55
0.5	0.020	10.77	19.45	15.11	6.14	44.97
0.5	0.030	7.43	12.51	9.97	3.59	6.43
5.0	0.010	12.64	21.34	16.99	6.15	67.79
5.0	0.020	9.04	15.22	12.13	4.37	14.09
5.0	0.030	6.79	11.13	8.96	3.07	5.80
9.5	0.010	8.20	11.77	9.99	2.53	3.19
9.5	0.020	6.93	10.35	8.64	2.42	4.78
9.5	0.030	5.74	8.74	7.24	2.12	9.85

Taguchi's Analysis and Solution

<u>Inner Array</u>: Used 3x3 full factorial design for (R, L).

- R = 0.5, 5.0, 9.5
- L = 0.010, 0,020, 0.030

<u>Outer Array</u>: Used "compounded noise" approach, noting (from the equation) where the extreme values for the response (Current) occur in tolerance space:

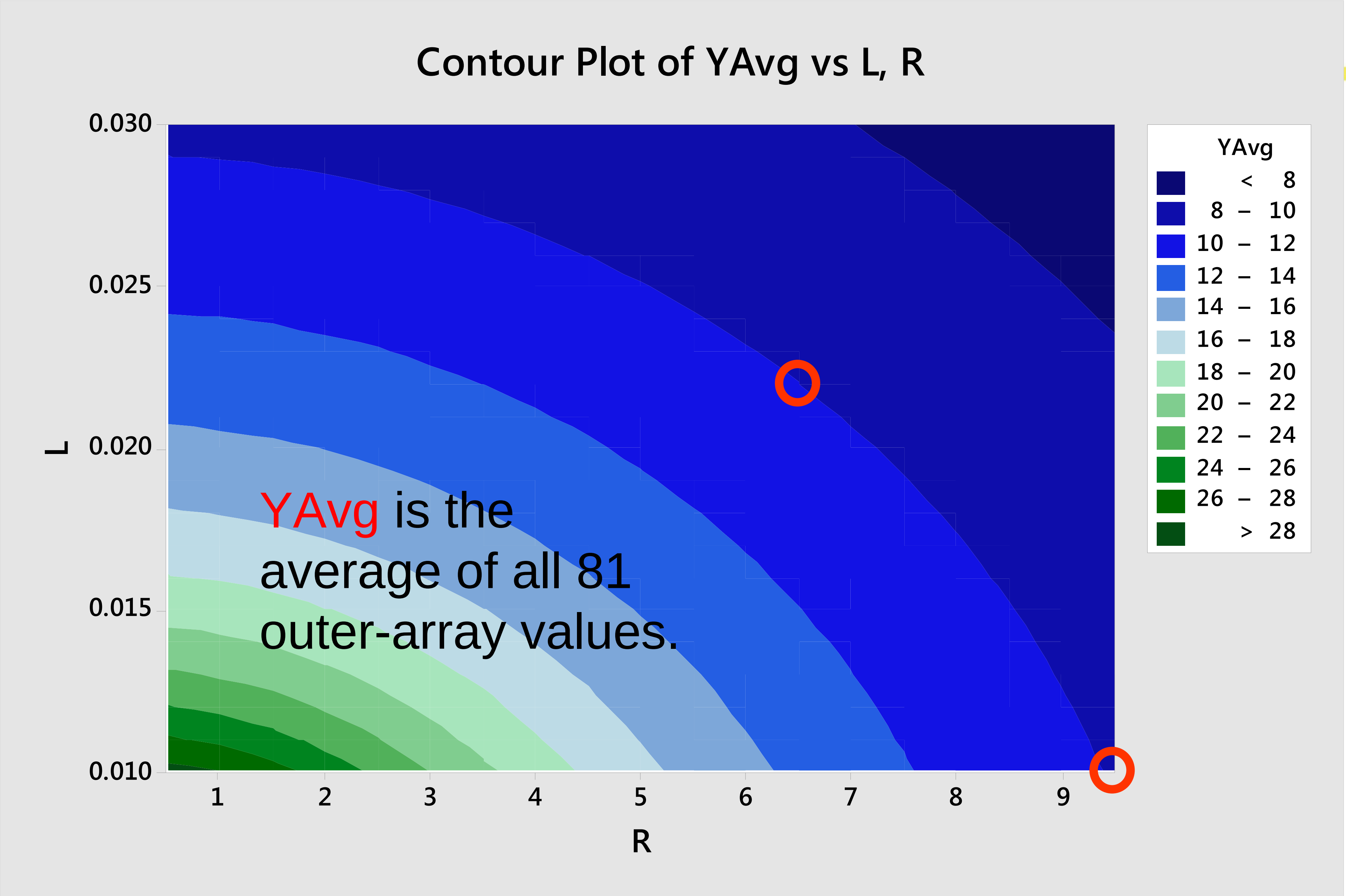
- Minimum (NMin) at V = 90, f = 60, R+0,5 L+0.002
- Maximum (NMax) at V = 110, f = 50, R-0.5, L-0.00

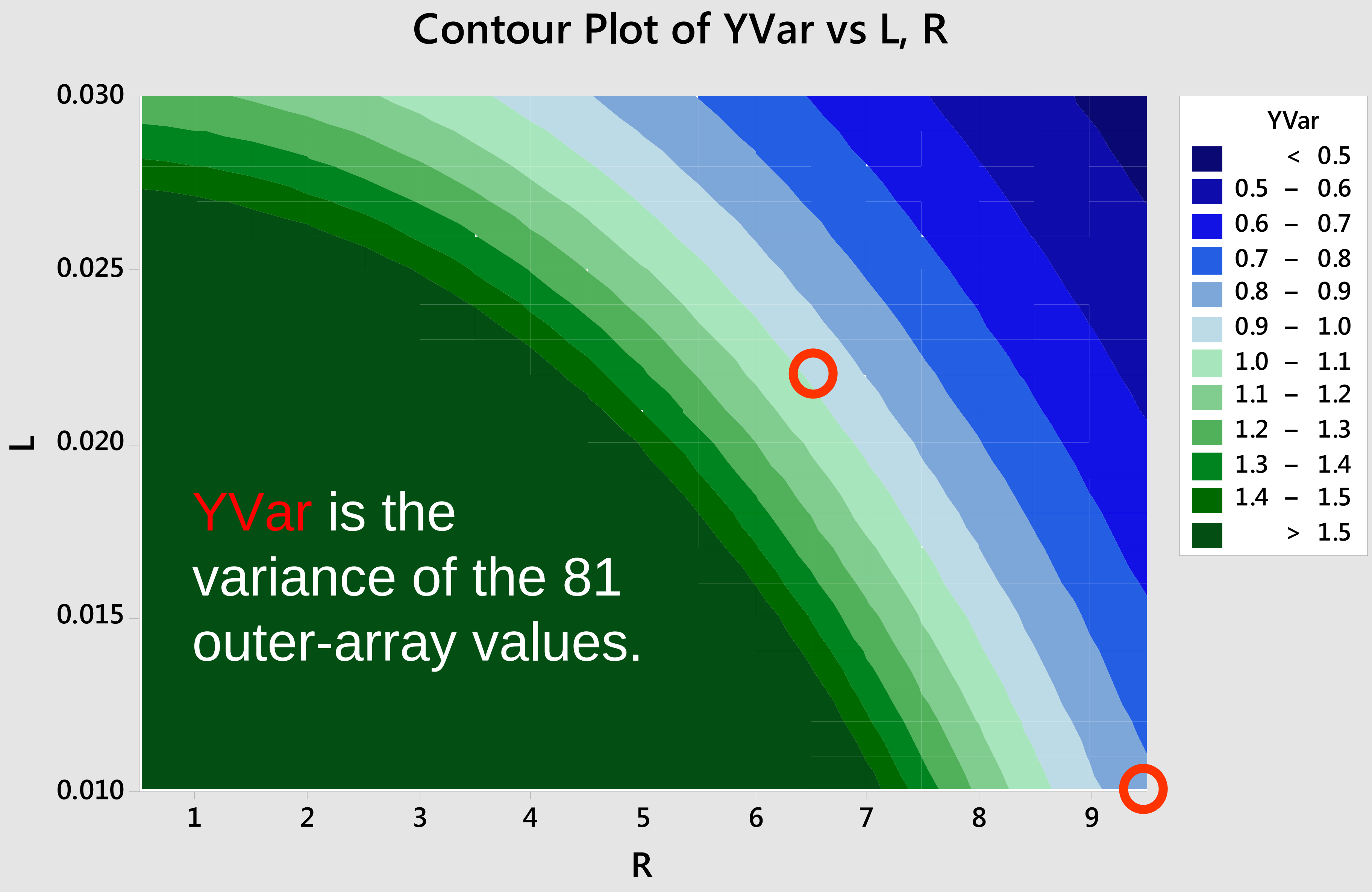
<u>Conclusion</u>: Based on summary statistics (on previous chart), R = 9.5 and L = 0.010 is the preferred design.

<u>Note</u>: Other (R, L) combinations in the design space (besides the 9 in the inner array) were not considered.

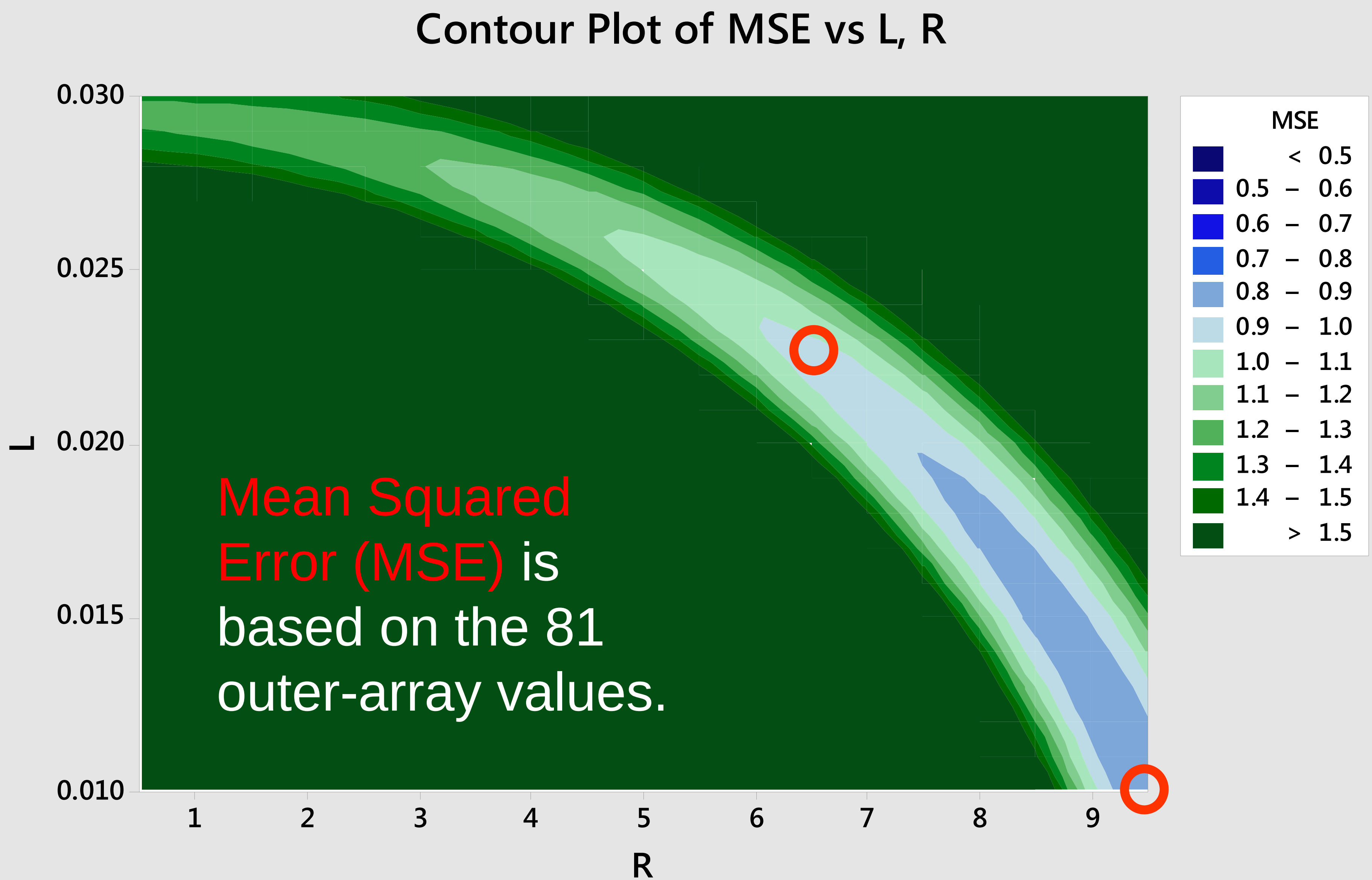
<u>Our approach </u>will look at the whole (R, L) design space, and a full 3x3x3x3 = 81 outer array for the "noise space".

ASQ


Contour Plot of YAvg vs L, R
YAvg is the average of all 81 outer-array values.
L
0.030
0.025
0.020
0.015
0.010
R
1
2
3
4
5
6
7
8
9
YAvg
< 8
8 – 10
10 – 12
12 – 14
14 – 16
16 – 18
18 – 20
20 – 22
22 – 24
24 – 26
26 – 28
> 28



Contour Plot of YVar vs L, R
0.030
0.025
0.020
0.015
0.010
L
1
2
3
4
5
6
7
8
9
R
YVar
< 0.5
0.5 – 0.6
0.6 – 0.7
0.7 – 0.8
0.8 – 0.9
0.9 – 1.0
1.0 – 1.1
1.1 – 1.2
1.2 – 1.3
1.3 – 1.4
1.4 – 1.5
> 1.5
YVar is the variance of the 81 outer-array values.


Contour Plot of MSE vs L, R
0.030
0.025
0.020
0.015
0.010
L
1
2
3
4
5
6
7
8
9
R
MSE
< 0.5
0.5 – 0.6
0.6 – 0.7
0.7 – 0.8
0.8 – 0.9
0.9 – 1.0
1.0 – 1.1
1.1 – 1.2
1.2 – 1.3
1.3 – 1.4
1.4 – 1.5
> 1.5
Mean Squared Error (MSE) is based on the 81 outer-array values.

Circuit Design Experiment

- ✓ **Generate data using an experimental design: 4-factor face-centered central composite design**

How this was done for this example

$$Y = \frac{V}{\sqrt{R^2 + (2\pi f L)^2}} + \text{experimental error } (\sigma = 1)$$

- ✓ **Fit a polynomial model to the data.**
- ✓ **Examine the regression summaries and contour plots.**

Figure 5
Steps 4,5 - Calculate Signal and Noise and Plot Contours

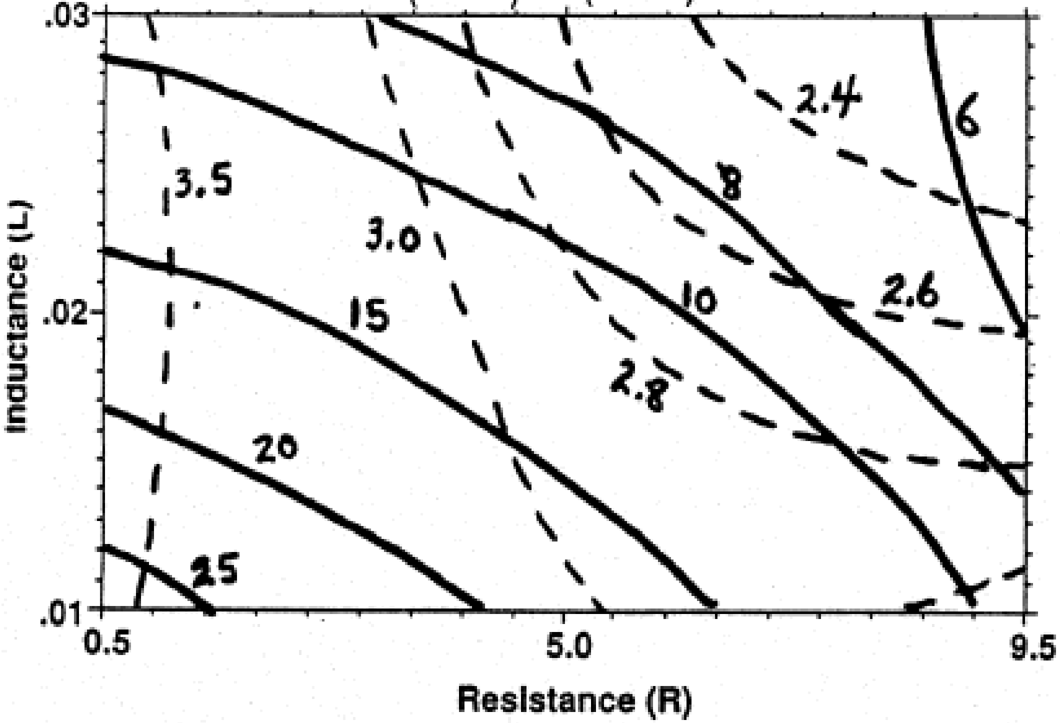
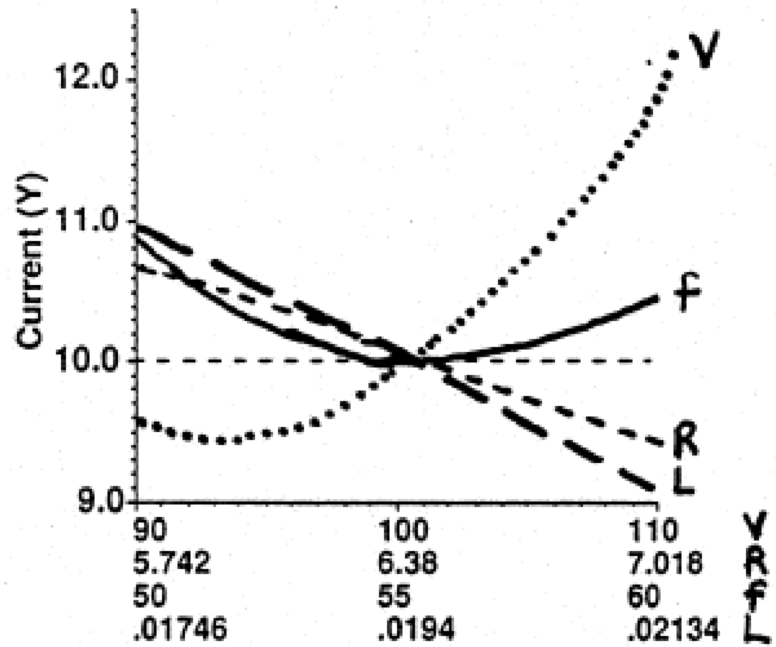
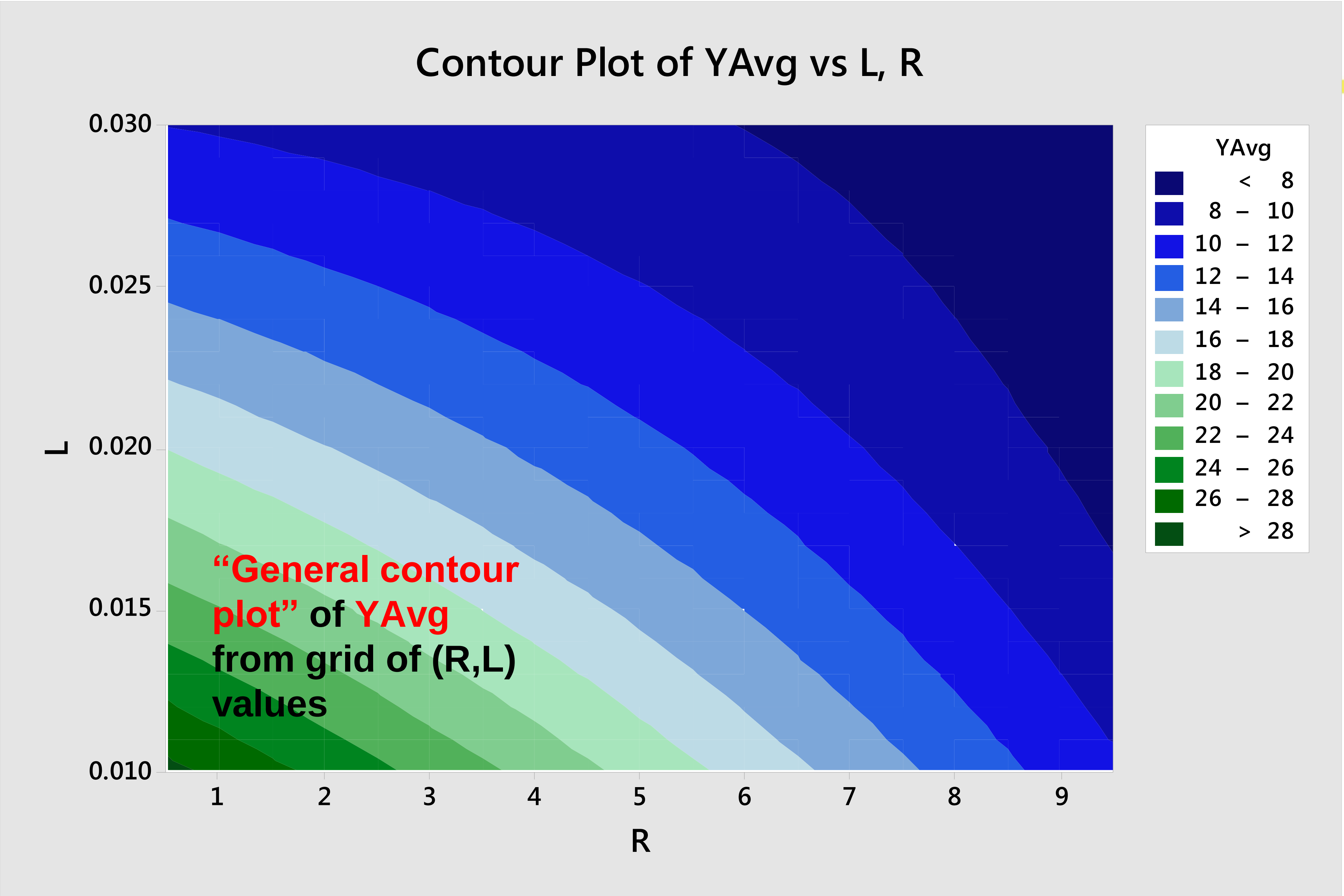
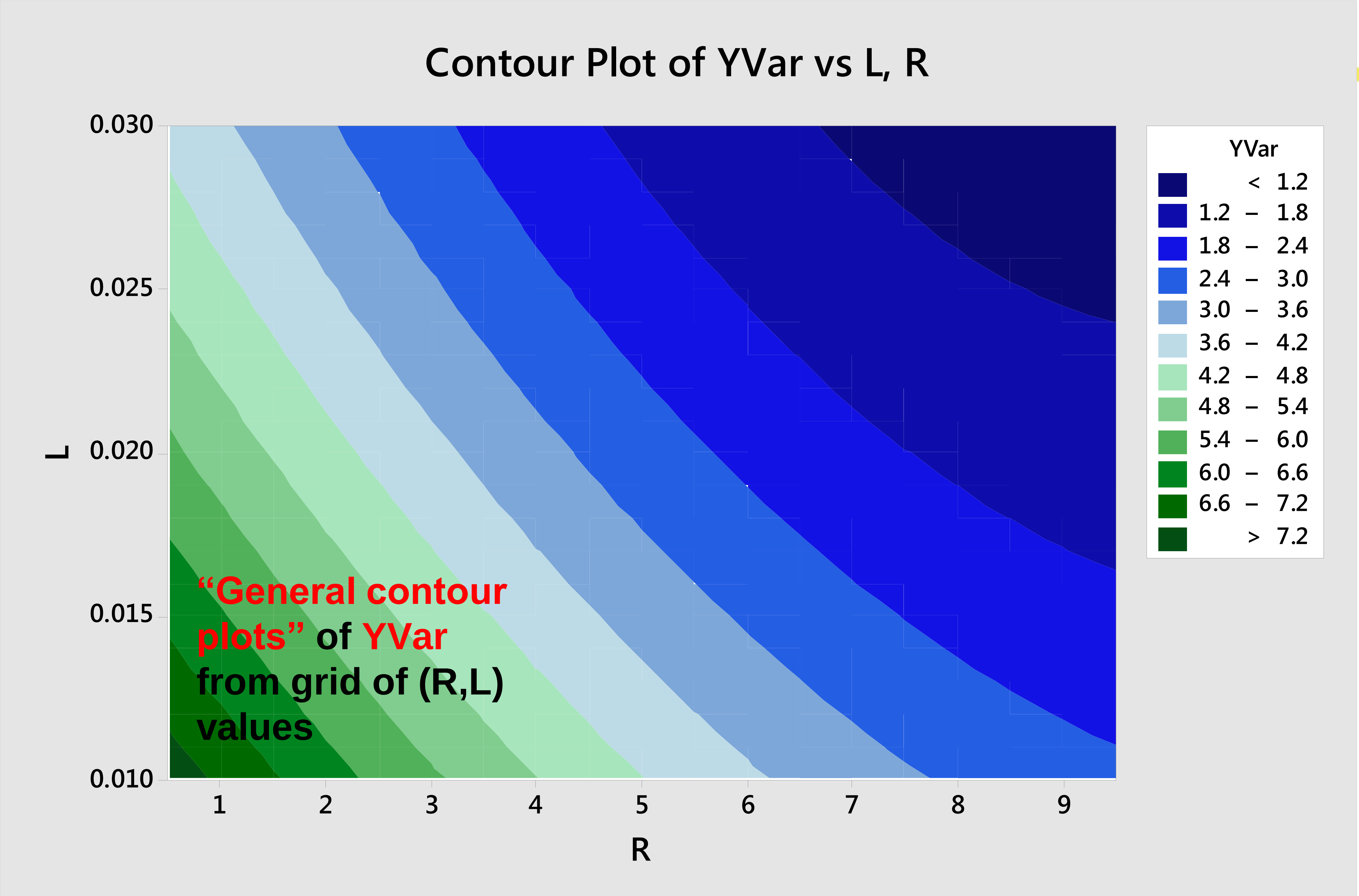
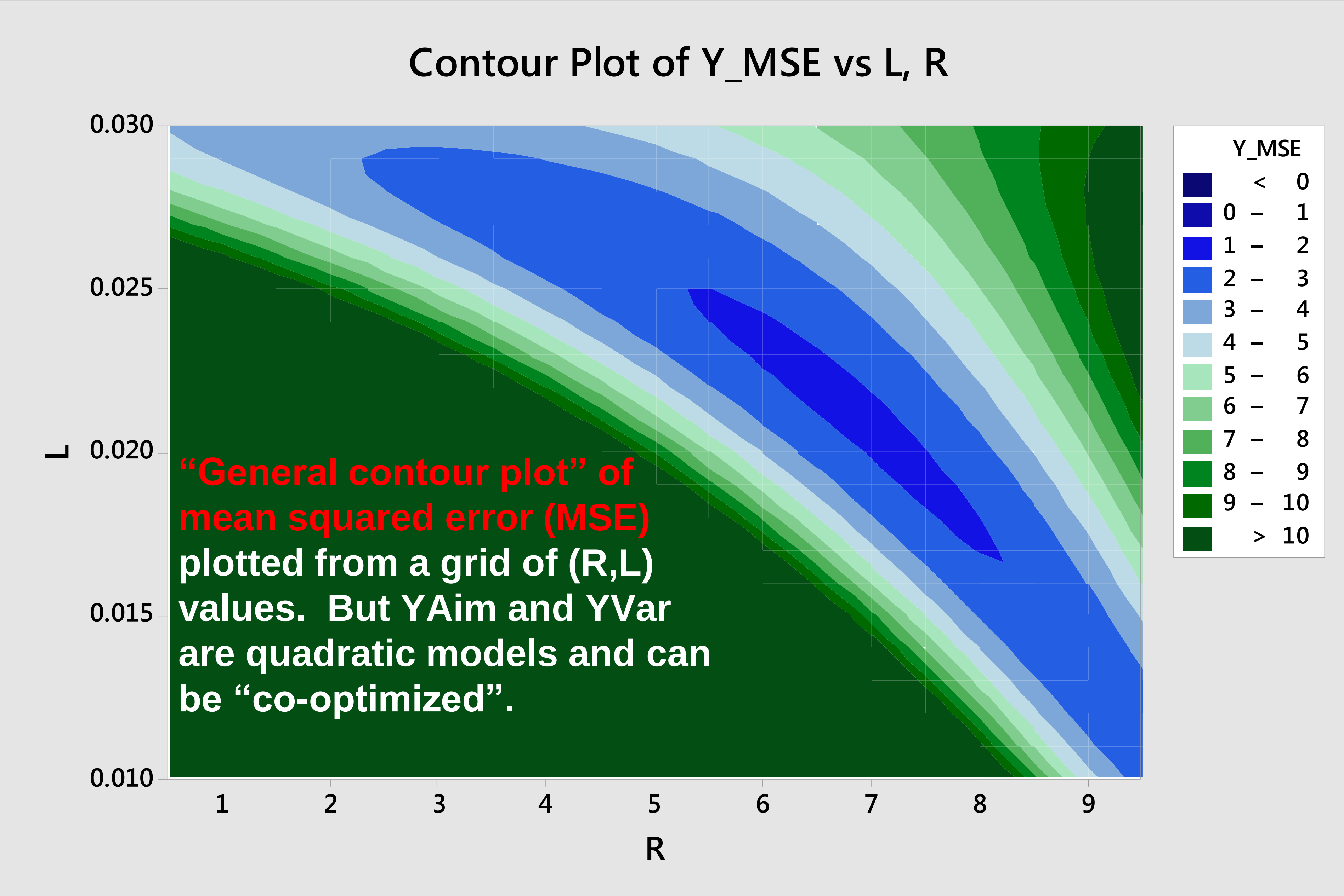


Figure 6
Step 9 - Perturbation Plot at R=6.38, L=.019




Contour Plot of YAvg vs L, R
0.030
0.025
0.020
0.015
0.010
L
1
2
3
4
5
6
7
8
9
R
YAvg
< 8
8 – 10
10 – 12
12 – 14
14 – 16
16 – 18
18 – 20
20 – 22
22 – 24
24 – 26
26 – 28
> 28
"General contour plot" of YAvg from grid of (R,L) values


Contour Plot of YVar vs L, R
0.030
0.025
0.020
0.015
0.010
L
1
2
3
4
5
6
7
8
9
R
YVar
< 1.2
1.2 – 1.8
1.8 – 2.4
2.4 – 3.0
3.0 – 3.6
3.6 – 4.2
4.2 – 4.8
4.8 – 5.4
5.4 – 6.0
6.0 – 6.6
6.6 – 7.2
> 7.2
"General contour plots" of YVar from grid of (R,L) values



"General contour plot" of mean squared error (MSE) plotted from a grid of (R,L) values. But YAim and YVar are quadratic models and can be "co-optimized".

Robust Design Elements and Options

- Transfer function (including “Control” and “Noise” factors)
  - True equation
  - Full quadratic model (from response surface DOE)
  - Reduced quadratic model
- Noise generation
  - Propagation of error (POE), using +/- 3 StDev tolerances
  - “Outer-array-like” calculations (span tolerance range)
  - Monte Carlo via (normal) distribution
- Optimization criterion
  - Minimize Variance subject to On-Aim Constraint
  - Minimize Mean Square Error (MSE)
  - Maximize Cpk/Ppk or minimize “Out of Spec” (so tolerances on Y needed)

Noise by Propagation of Error

The noise or transmitted variance in Y, V(Y) from the factors (Xs) can be estimated using the propagation of error (POE) relationship.

$$\mathbf{V(Y)} = \left(\frac{d\mathbf{Y}}{d\mathbf{X_1}}\right)^{\mathbf{2}} \sigma^{\mathbf{2}}_{\mathbf{X_1}} + \left(\frac{d\mathbf{Y}}{d\mathbf{X_2}}\right)^{\mathbf{2}} \sigma^{\mathbf{2}}_{\mathbf{X_2}} + \mathbf{....} \left(\frac{d\mathbf{Y}}{d\mathbf{X_k}}\right)^{\mathbf{2}} \sigma^{\mathbf{2}}_{\mathbf{X_k}}$$

Propagation of Error

- ✓ First-order derivatives from a second-order (quadratic) response surface are themselves first-order (linear) equations! (See below for three-factor example of this.)
- ✓ Thus the V(Y) equation on the previous chart is in fact a second-order (quadratic) response surface itself!

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_{12} X_1 X_2 + b_{13} X_1 X_3 + b_{23} X_2 X_3 + b_{11} X_1^2 + b_{22} X_2^2 + b_{33} X_3^2$$

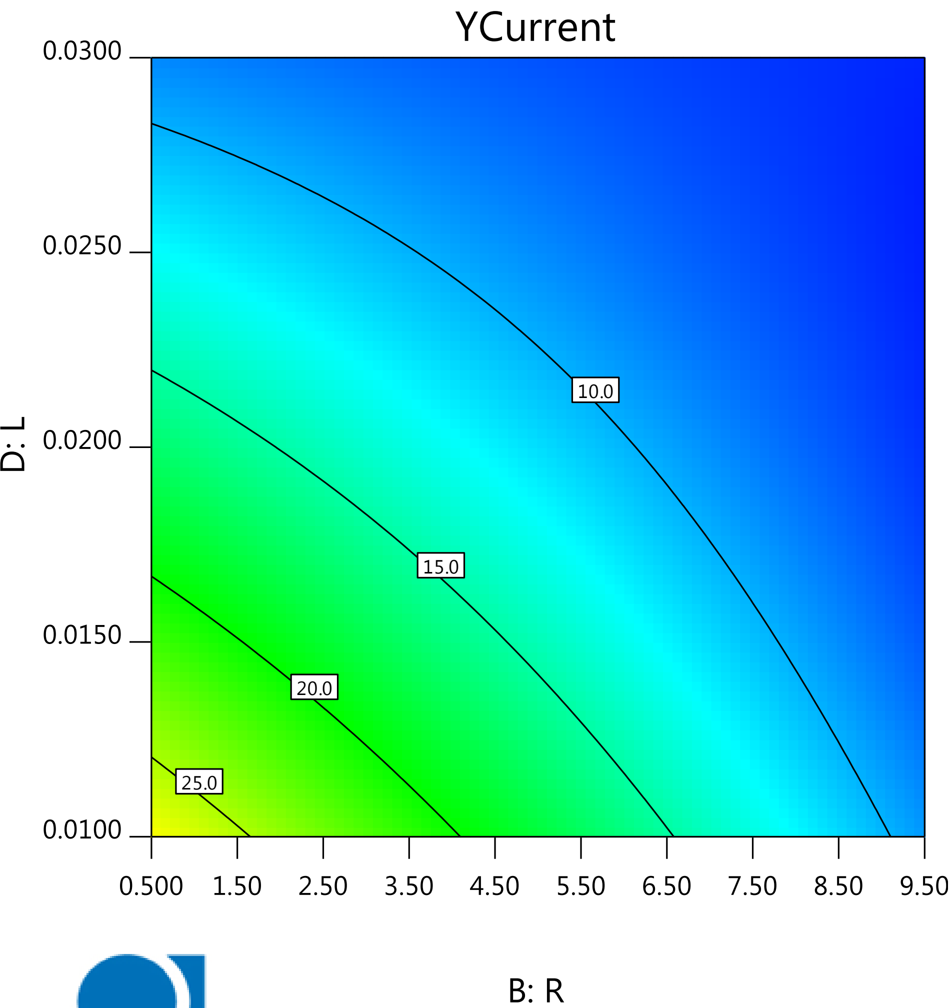
$$\frac{dY}{dX_1} = b_1 + 2b_{11} X_1 + b_{12} X_2 + b_{13} X_3$$

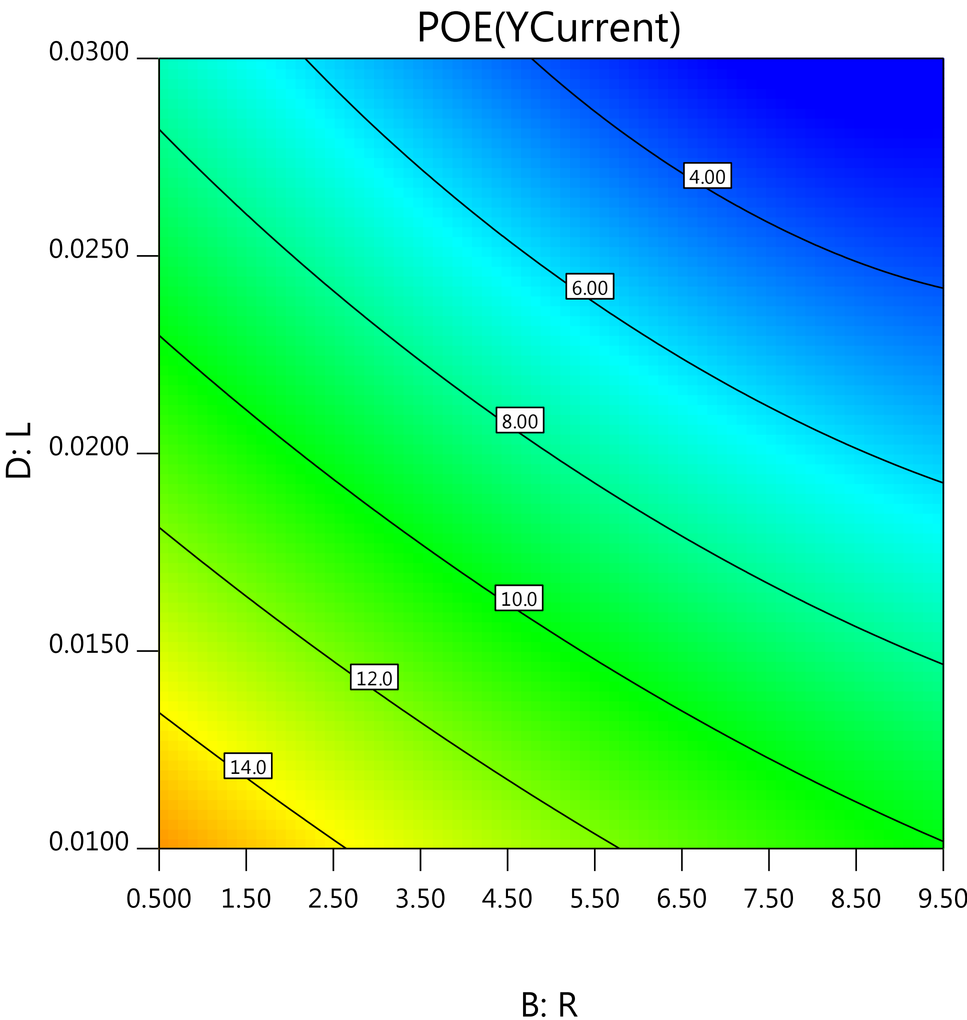
$$\frac{dY}{dX_2} = b_2 + b_{12} X_1 + 2b_{22} X_2 + b_{23} X_3$$

$$\frac{dY}{dX_3} = b_3 + b_{13} X_1 + b_{23} X_2 + 2b_{33} X_3$$

ASQ

POE for quadratic models have been available in Stat-Ease's Design Expert since the mid-1990's



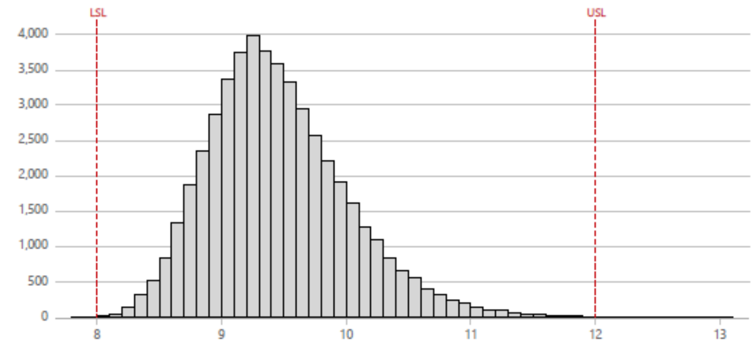


"Parameter Optimization" of a Ppk-like metric is available for any model in Minitab's Companion

At right are monte carlo results using the quadratic response surface model at optimum (R=6.92, L=0.019)

Parameter Optimization Results

Current



Using the new input settings, the simulation indicates that you can expect 0.12 % of the *Current* values to fall outside of the specification limits. This corresponds to a Cpk of 1.04. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

1.04

% Out of Spec

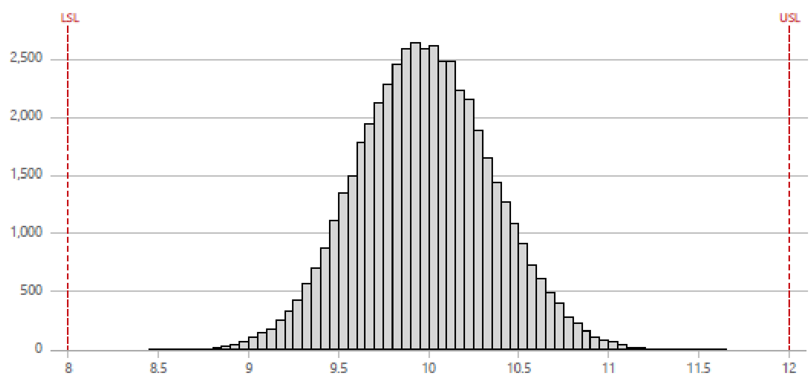
0.12 %

Summary Statistics

N	50,000
Mean	9.4604
Standard Deviation	0.575946

Parameter Optimization Results

Current



Using the new input settings, the simulation indicates that you can expect 0.00 % of the *Current* values to fall outside of the specification limits. This corresponds to a Cpk of 1.74. A generally accepted minimum value of Cpk is 1.33.

Process Performance (Cpk)

1.74

% Out of Spec

0.00 %

Summary Statistics

N	50,000
Mean	9.9791
Standard Deviation	0.377903

At left are the results from the true model (R=9.40, L=0.010)

Coauthors & Acknowledgements

Giving Your Response Surface a Robust Workout (Experimentation for Robust Product Design)

- Bill Fellner, retired
- Ken Chatto, retired
- Chuck Pfeifer, retired

Software Acknowledgements

- Design Expert – Martin Bezener, Pat Whitcomb
- JMP – Brad Jones
- Minitab – Doug Gorman, Jenn Atlas

DOE Big Picture

Equation

y = f(Xs) + e

Words

Output equals Function of Inputs plus Variation

Considerations

- Continuous vs. Discrete
- "As Is" vs. Transformed vs. "Link Function"
- Theoretical vs. Empirical
- Simple vs. Complex
- Continuous vs. Discrete
- Many vs. Few
- Screening vs. Optimization
- Bias Error vs. Random Error
- Bias Error → Blocking and Randomization
- Random Error → Overt (Pure) and Hidden Replication

Degrees of Freedom

Number of rows of data (n) = Number of parameters to estimate (p) + Number of additional treatment combinations for checking lack of fit (l) + Amount of pure replication (r)

Question!

Historical Data Mining (HDM) vs Designed Experiment (DOE)

A final word from Jack Youden (from 50 years ago)

"There is one particular role I am determined that statistics should not have. AOAC must not serve as a playground for statisticians to exhibit their special skills at the price of bewildering the chemist. There is an important reason for insisting on simple and intuitively acceptable statistical techniques. Presentation of evidence before a court, or to a producer whose product is rejected, will be more convincing if it is understandable."

J. Youden (1967)

AOAC = Association of Official Analytical Chemists

There is no “I” in “Youden”, but there is “You”

Thank you for listening!

The Global Voice of Quality™

JMP can also perform similar analyses (per note from Brad Jones below)

- The interface to the profiler allows for simulating from the prediction model allowing noise factors to vary according to a specified distribution. The user can set the number of simulations and the output is the defect rate. The interface also supports running a designed simulation experiment that alters the nominal settings of each factor to find the settings that minimize the predicted defect rate.
- Alternatively, and more simply, you can output the prediction equation as a formula and use the profiler in the graphics menu while indicating which factors are noise factors in the launch dialog. The optimization then finds the settings of the control factors that simultaneously minimize the magnitudes of the first derivative of the prediction equation with respect to the noise factors and match the target response. Of course this is a multiple response optimization, so we use a (user customizable) utility function and maximize that.

Outline for talk (as appeared in program)

Experiences, approaches, and examples of Measurement Systems Analysis (MSA) and Design of Experiments (DOE) will be shared, based on decades of application in the chemical and process industries.

- First, a roadmap for MSA will be presented that has proved useful in guiding MSA studies for six sigma projects.
- Then the Strategy of Experimentation that has been taught and used successfully in DuPont for over 50 years will be reviewed
- The importance of including both “design” (or “control”) and “environmental” (or “noise”) factors in these studies to achieve both “functional” and “robust” products will be illustrated.
- The role of custom (algorithmic or optimal) designs and Definite Screening Designs (the “new tool on the block”) in this strategy will be discussed.
- Finally, some comments on “Big Data” and the combined power of both “Historical Data Mining” and DOE will be shared.

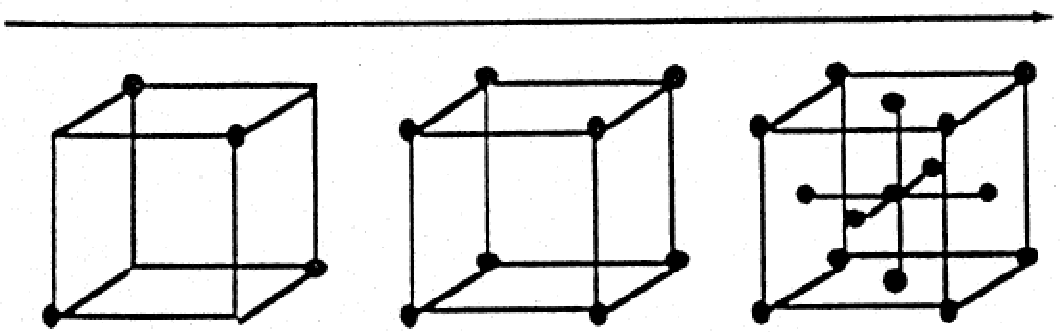
Outline for talk (condensed version)

Experiences, approaches, and examples of Measurement Systems Analysis (MSA) and Design of Experiments (DOE) will be shared, based on decades of application in the chemical and process industries.

- A roadmap for MSA will be presented that has proved useful in guiding MSA studies for six sigma projects in DuPont.
- Then these DOE topics will be briefly covered:
  - **DuPont's 50+ years of Strategy of Experimentation**
  - **An example with both "design" (or "control") and "environmental" (or "noise") factors and how to achieve both "functional" and "robust" products.**
  - **Examples of custom (algorithmic or optimal) designs**
  - **Definite Screening Designs (DSDs, the "new tool on the block") and their "not so definitive" Analysis**
  - **"Big Data" and the combined power of both "Historical Data Mining" and DOE.**

Figure 2

Classical DOX Approach: "Evolution of the Experimental Environment"



Screening

- Plackett-Burman
- Fractional Factorial

Interaction

- Factorial
- Fractional Factorial

Optimization

- Central Composite
- Box-Behnken

Figure 3

Tolerance Region Illustration

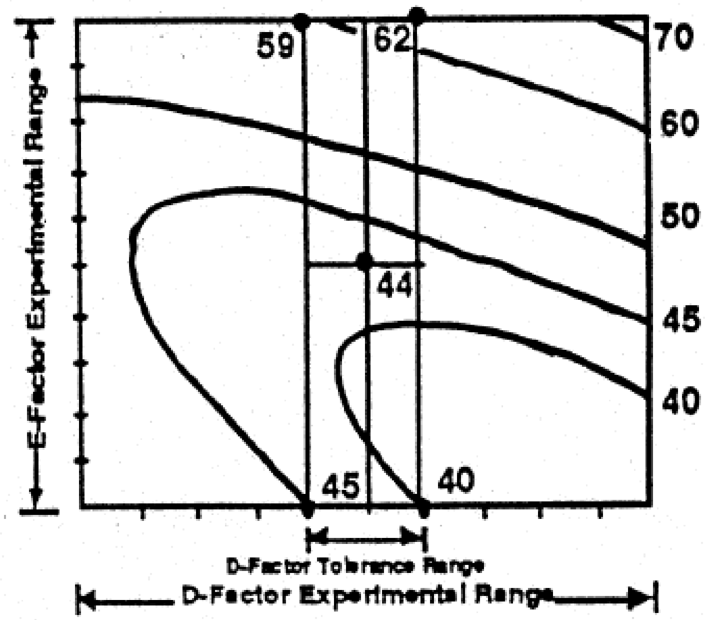
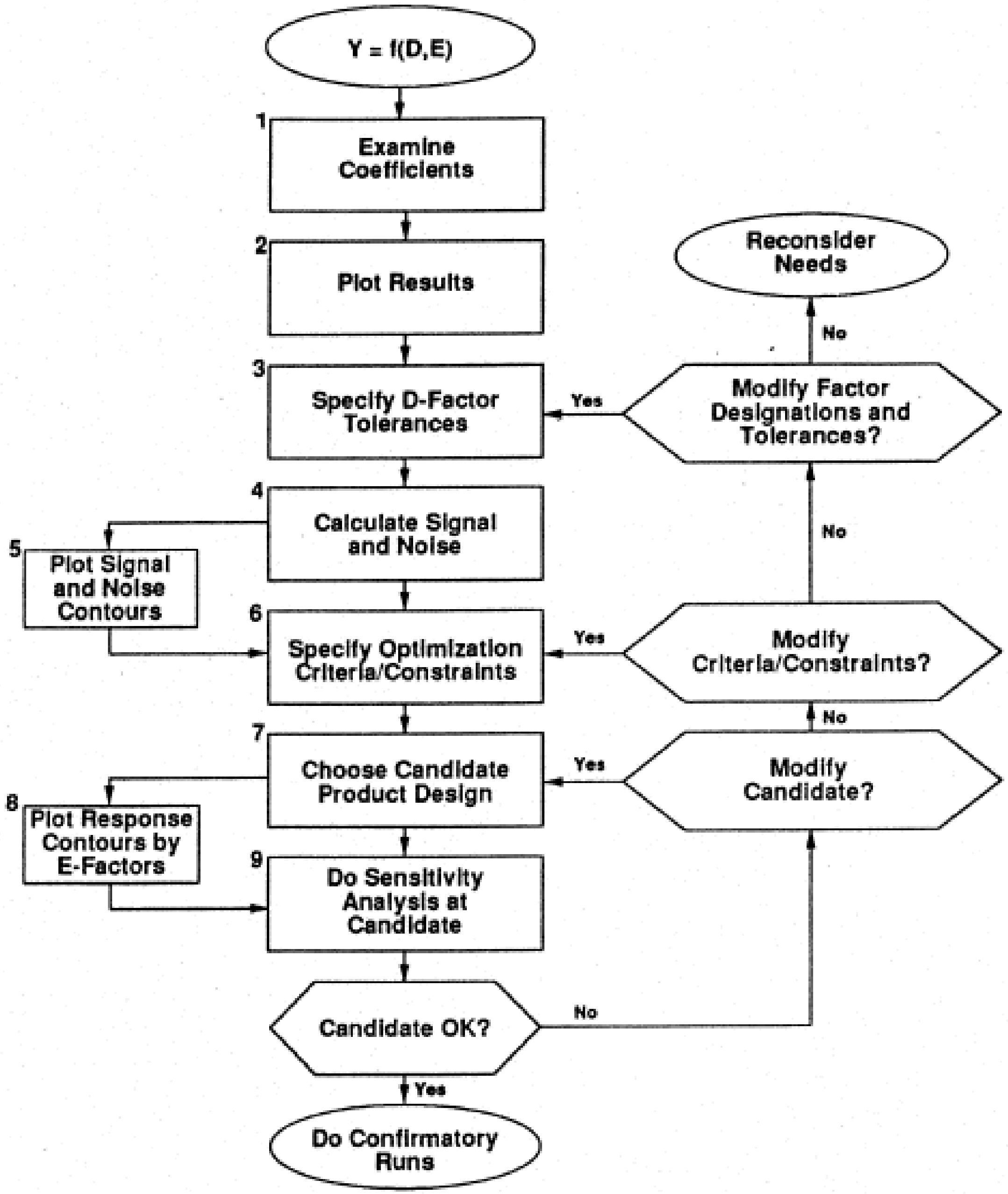


Figure 4
Exercising the Response Surface Equation(s)



Factor	Symbol	Type	Units	Low	Middle	High
Voltage	V	E	Volts	90	100	110
Resistance	R	D	Ohms	.5	5.0	9.5
Frequency	f	E	Hertz	50	55	60
Inductance	L	D	Henrys	.01	.02	.03

Two factors, Voltage and Frequency, are considered environmental factors. The product design goal is to achieve a Current value of 10 while minimizing noise.

The Current data were simulated by adding a random normal error with mean zero and standard deviation one (ampere) to the known theoretical model, $Y = V/[(R^2 + (2\pi fL)^2]^{1/2}$. A regression analysis was performed by fitting a full quadratic model to the data. Examination of the coefficients (Step 1) provides immediate insight into whether noise varies significantly within the product design (D-factor) space. Only when significant interactions among D-factors, at least one of which has a nonzero tolerance, and between D and E-factors exist, can we expect to have an influence on noise through the specification of a product design. In this example, the two design factors, Resistance and Inductance, interact significantly, as do the two environmental factors, Voltage and Frequency; also, Resistance and Frequency interact ($p \leq .10$).

ASQ

A final word from Jack Youden (from 50 years ago)

"When compiling a subject index for the NBS Special Publication 300, Vol. 1, a volume that contains 15 of Youden's publications, I had great difficulty in finding terms in Youden's writing to include in the index, but no trouble at all for the other authors. Youden took great pains to avoid words that needed to be technically defined."

"Youden's handwriting was extremely neat and it seems that his first draft was his final draft (after thinking about what to write for a long period of time). Enclosed is an example, ..."

Harry Ku

The example to which Dr. Ku refers is a handwritten manuscript by Jack two years after retiring from NBS entitled "The Role of Statistics in Regulatory Work". At the end of the introductory paragraph where Jack sets the stage for discussing the role of statistics in regulatory work is the following excerpt:

"There is one particular role I am determined that statistics should not have. AOAC must not serve as a playground for statisticians to exhibit their special skills at the price of bewildering the chemist. There is an important reason for insisting on simple and intuitively acceptable statistical techniques. Presentation of evidence before a court, or to a producer whose product is rejected, will be more convincing if it is understandable."

J. Youden (1967)

William John Youden, 1900–1971

William John (Jack) Youden died suddenly from a heart attack on Wednesday, March 31, 1971, in Washington, D. C. Internationally famous for his contributions to mathematical statistics and for his outstanding ability and "missionary" zeal in communicating statistical techniques to those concerned with experimentation. Dr. Youden and his wisdom and friendship will be missed by many in the experimental field as well as by those associated with him in the statistical profession.

Dr. Youden was born in Townsville, Australia, on April 12, 1900. Two years later his father returned to his own birthplace, Dover, England, with his wife and young son and the three resided there from April 1902 to June 1907. During these years a sister, Dora Alice, and brother, Harry, were born. In 1907 the family of five set out for America and entered the United States through the Port of New York in July 1907. They lived for a while at Ivoryton, Conn., and at Niagara Falls, N.Y., where Jack attended the local public schools, and then moved to Rochester, N.Y., in 1916 for Jack's senior year of high school.

The years 1917–1921 were spent at the University of Rochester, except for one brief interruption to serve his new country as a private in the U. S. Army, October 15 to December 12, 1918. At the University of Rochester Jack was elected to Phi Beta Kappa and was awarded a B. S. in chemical engineering in June 1921. The following academic year, 1921–22, he continued at the University of Rochester as an instructor in chemistry, then went to Columbia University as a graduate fellow in chemistry, earning an M. A. in 1923 and a Ph. D. in 1924.

Immediately following receipt of his doctorate, Youden joined the staff of the Boyce Thompson Institute for Plant Research in Yonkers, N. Y., as a physical chemist. He continued with the Institute in this capacity with two short leaves of absence and one three-year stint as an operations analyst with the Army Air Force, until he joined the National Bureau of Standards in May 1948.

This article is based primarily on material assembled by Churchill Eisenhart in support of Youden's nomination for the Wilks Medal.

Dr. Youden was often heard telling a "client" in consultation on statistical aspects of experimentation or the audience at one of his well-attended lectures on statistical methodology that he was "a chemist," implying, it would appear, that he was really **not** a statistician. Youden may have been **all** chemist for his first seven years at the Boyce Thompson Institute, but by September 1931 the transition from chemist to statistician was underway. The first evidence of this change may be found in his paper entitled "A Nomogram for Use in Connection with Gutzeit Arsenic Determinations on Apples" (1931). During the academic year 1931–32 he commuted on his own volition from Yonkers to Morningside Heights in New York City to attend Professor Harold Hotelling's lectures on "Statistical Inference" at Columbia University. He was on his way to becoming an expert on statistical aspects of experimentation. From then on he became more and more of a statistician, but his laboratory experience was always to remain a treasured asset enabling him to communicate with scientists on their own ground.

The paper that made his name a laboratory, if not a household, word was published in early 1937: "Use of Incomplete Block Replications in Estimating Tobacco Mosaic Virus." Here he gave examples and illustrated the application of a new class of symmetrical balanced incomplete block designs that possessed the characteristic "double control" of Latin square designs, without the restriction that **the number of replications** of each "treatment" (or "variety") **must equal the number of "treatments"** (or "varieties"). This paper and its new designs led to Dr. Youden's obtaining a Rockefeller Fellowship that enabled him to take his first "leave of absence" from Boyce Thompson, devoting the academic year 1937–38 to further work in the field of experiment design under the direction of R. A. Fisher himself at the Galton Laboratory, University College, London. Youden's new rectangular experiment designs, termed "Youden Squares" by Fisher and Yates in the introduction to the first edition of their *Statistical Tables for Biological, Agricultural and Medical Research* (1938), were found immediately to be of broad utility in biological and medical research generally; to be applicable but of less value in agricultural field trials; and, with the coming of