Power Tools for Six Sigma
Session G1

Bev Daniels, March 5, 2013

beverly-daniels@idexx.com

ASQ®
The Global Voice of Quality™

NOTICE

While I am providing this material for personal use to members and visitors of the Cove, any reproduction or use of these materials for personal gain without my express written approval is prohibited.

Personal gain includes and is not limited to: republishing on personal web sites, forums or other social media, use for paid consulting or training or use within your organization as training or reference material.

Enumerative *vs.* Analytical Statistics: A Review

ASQ
The Global Voice of Quality™

Enumerative Studies

Experimental design should always be framed to answer a question.

Questions about the **current state** of a process or material call for enumerative studies.

These experiments focus on estimating a population at one point in time.

- acceptance sampling of a lot
- data packages for regulatory agencies
- one inventory cycle count

In these studies, a decision is only made on the material that was sampled. No prediction about the acceptability of future lots or inventory cycle counts can be made.

Analytical Studies

In an analytical study, the experiment is designed to **predict future results**.

- Problem Solving experiments
- Testing on multiple lots of material to qualify a new vendor
- Testing to qualify repeatability of a manufacturing process.

Analytical studies require knowledge of the important variables and repetitive (replicate) testing of those variables.

ASQ

"It's tough to make predictions, especially about the future."
-Yogi Berra

"Don't expect more out of an experiment than you're willing to invest in it."
-Bev

Types of Studies[1]

Enumerative

Descriptive

- Estimation of a finite population or static data set
- Quantifies only the product or process in front of us.
- Has no predictive usefulness for future performance.

Statistics and statistical precision of the estimates have value.

Used to determine what action should be taken on the population under study.

Examples: Census, customer surveys, acceptance sampling,

Critical Structural Element:
Randomness (representativeness) of the sample

Analytic

Predictive

- Understand the causal mechanisms and resulting performance of a system in order to make predictions about future performance
- Used to improve products and processes in the future

Tests of statistical significance are often redundant

Proper structure is more important to our belief in the prediction than statistical estimates of the precision or accuracy of the analysis.

Examples: Y-X problem solving studies, factorial experiments, response surface experiments

Critical Structural Elements:
Independence & **Replication**

ASQ

Structure Reduces Uncertainty[2]

The greatest source of uncertainty in analytic studies is the structure of the study.

The conditions and factors that will change in the future must also change in the study or there can be very little confidence in the prediction of future results.

This uncertainty cannot be statistically quantified.

Conversely, study results that are replicated across the full range of expected variation are far more reliable than any statistical test of significance might indicate.

Structure is more important than statistical formulae.

ASQ

What is Independence?

In layman's terms: The results of each test or trial appear to behave randomly - *i.e*. if you know the value of the first unit you *cannot* predict the value of the next unit.

In Physics terms: The first trial or test does not influence or effect subsequent trials

In Probability terms: The first value or result contains no information about subsequent values or results.

A practical rule is to assume independence *between* samples of the largest component of variation...

What is Replication?

Replication is an independent trial of an experiment. All of the factors must be reset or selected to be different.

This means using different lots of materials, different equipment, different consumables, different operators, performed on different days, *etc*.

Do it again!

Why Replicate?[3]

Replication enables us to have high confidence in future results when all of the factors and conditions have changed (within normal process variation).

Replication enables us to test all alternative theories in the most efficient manner.

Replication enables us to use smaller sample sizes.

Replication is more logistically difficult.

Independence *vs*. Replication

Independent samples are not necessarily replicates of each other.

Replicated subgroups are by definition independent of each other.

Power Tools

Powerful Tools

- **Versatile in their use**; applicable to many diverse situations
- **Easy to use**
- **Intuitively insightful**; do not rely on statistical summaries, p values, *etc*.
- **Compelling in their results**

Power Tools and Kryptonite

Power Tool	Kryptonite
Looking (viewing function & failure)	Tables of Statistics
Graphical Display of Data	Transformations
Multi-Vari	Control Charts, Histograms
MSA – Youden plots	Traditional Gauge R&R
Assembly-Disassembly (ex situ decomposition)	a-priori Fractional Factorials
Full Factorials	OFATs
Matched Pairs (Power Blocking)	Grouped Testing / Traditional Hypothesis Tests
Control Charts	Dashboards and Tables of Numbers

Beverly Daniels
March 5, 2013

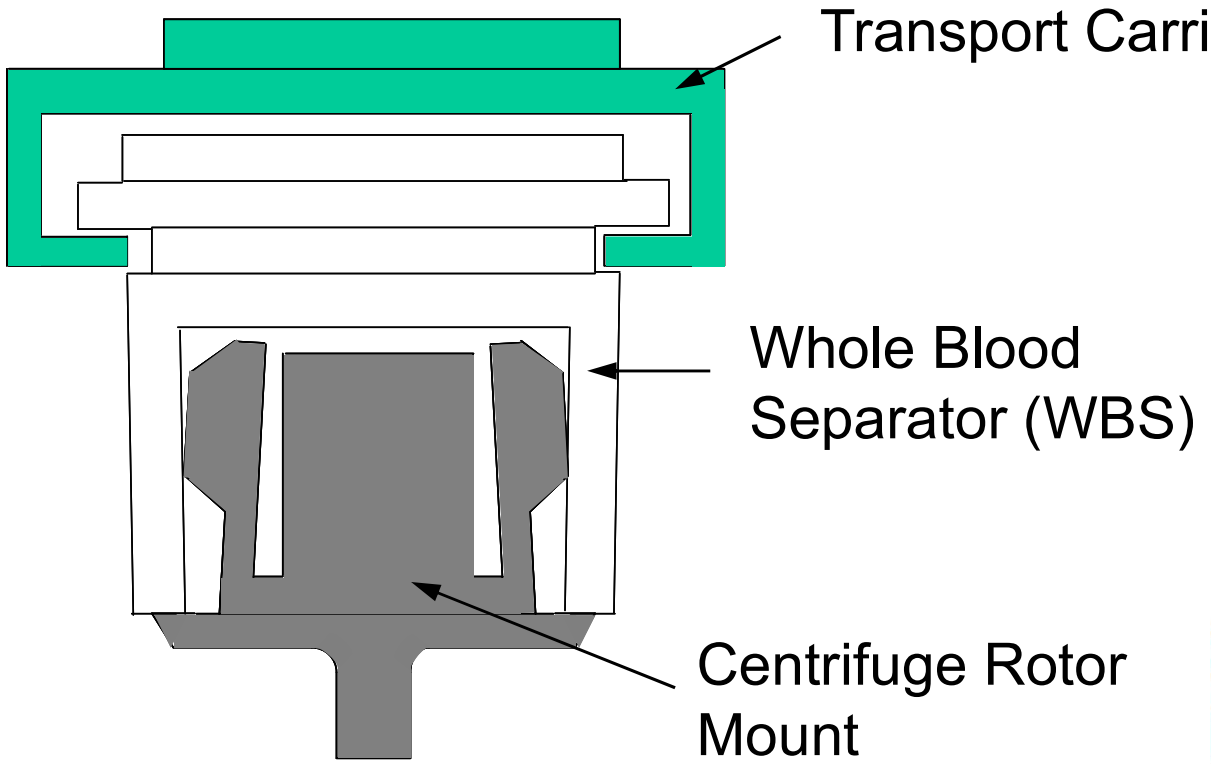
ASQ

Power Tools: Looking

Example 1: Spin Error*

Customers were requesting service on an instrument for a “10502 Error”. Returns for this error had recently escalated from a nuisance level of less than .5% to > 2% of the instrument install base.

A “10502 Error” occurs when the WBS cup doesn’t spin as expected – for the required time and rpm



Beverly Daniels
March 5, 2013

Potential Failure Modes that Would Result in a “Spin Error”

No Spin

Spins but **doesn’t reach target rpm**

Reaches target rpm but **slows down** before it should

Reaches target rpm but **stops** before it should

Each failure mode has different causal mechanisms.

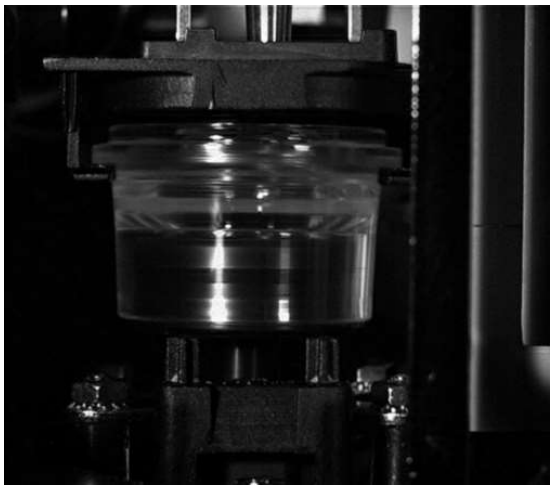
Some causes are unique to a failure mode

Some causes are shared by 2 or more failure modes

ASQ

“Learning to See” Begins with Remembering to **Look**

A high speed camera was used to video the returned instruments and instruments without the error



The video showed the Rotor ‘wobbling’ and the WBS hitting the carrier and “walking off” the rotor. This jams the rotor and stops the spin

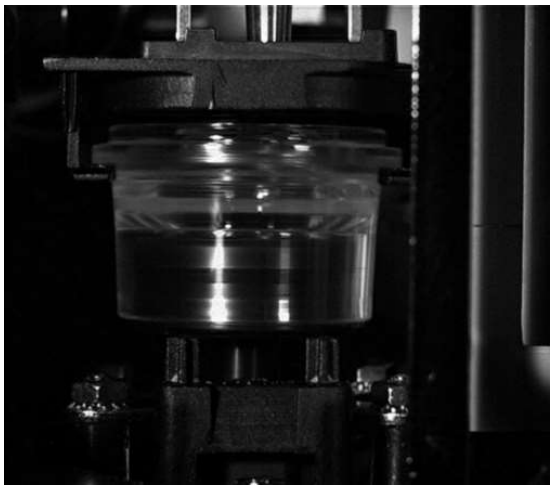
It’s amazing the things you can see when you look
– Yogi Berra

Beverly Daniels
March 5, 2013

ASQ

Structural Decomposition

The second step was to determine if the WBS would come off the rotor chuck by itself or if it needed the carrier to walk itself off the rotor chuck.



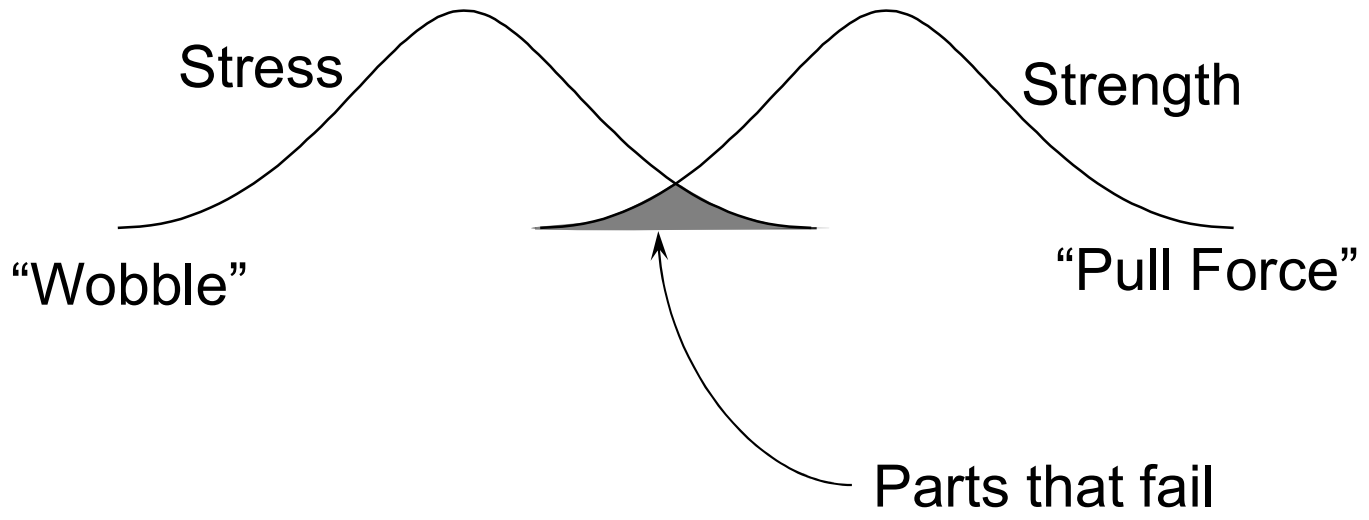
The video showed that the WBS did not come off the rotor even with wobble.

ASQ

Wobble and Pull Force

This failure was a classical strength-stress interaction.

The excursion was caused by a sudden increase in stress from the wobble of the assembly.



Beverly Daniels
March 5, 2013

Measuring the effect of Wobble and Pull force

The ‘wobble’ of the rotor chuck was measured as movement in the vertical (Y) direction at the edge of the flange.

The vertical ‘pull’ force to remove the WBS was also measured.

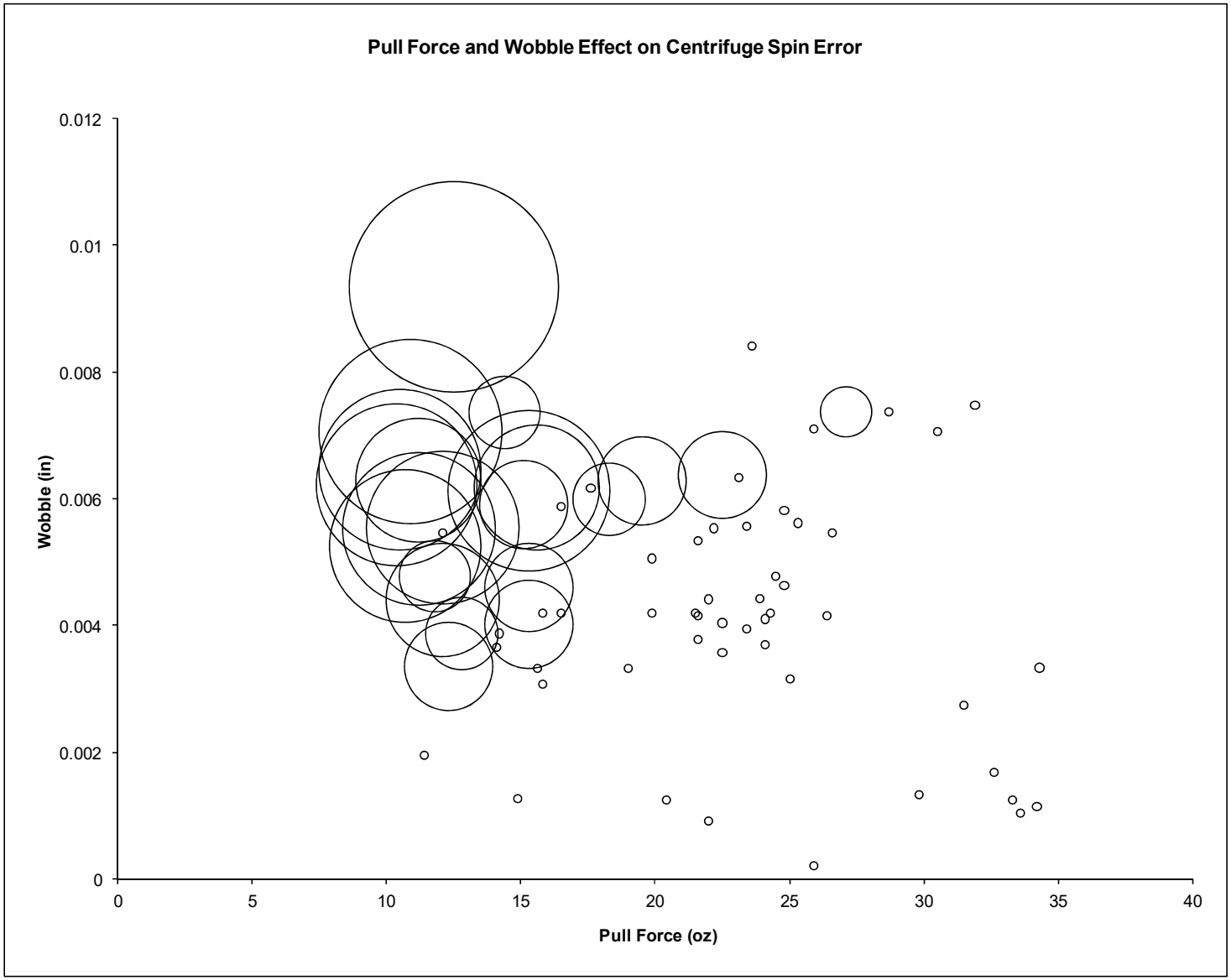
50 Rotor Assemblies were measured for wobble and pull force. They were then run 25 times in an instrument to determine their failure rate.

The wobble and pull force of each Rotor Assembly was plotted on a scatter diagram

Beverly Daniels
March 5, 2013

ASQ

Looking at the Data: A Screening Specification



Beverly Daniels
March 5, 2013

ASQ

Example 2: Chemical Reaction Rate on a Slide*

A small slide is used to detect the presence of a specific chemical in an animal.

A small drop of blood sample is dispensed on a slide

the slide contains agents that will react with the sample

The instrument optically “reads” the slide after a sufficient time has passed.

Some instruments are not accurate in their reads using Control samples.

Beverly Daniels
March 5, 2013

ASQ

It’s Enormous

There are hundreds of potential causes for this failure

- Insufficient sample **dispense**
- Insufficient reactive agent on the **slide**
- Inability of the **instrument** to accurately read a proper reaction
- Interfering substance in the **sample**

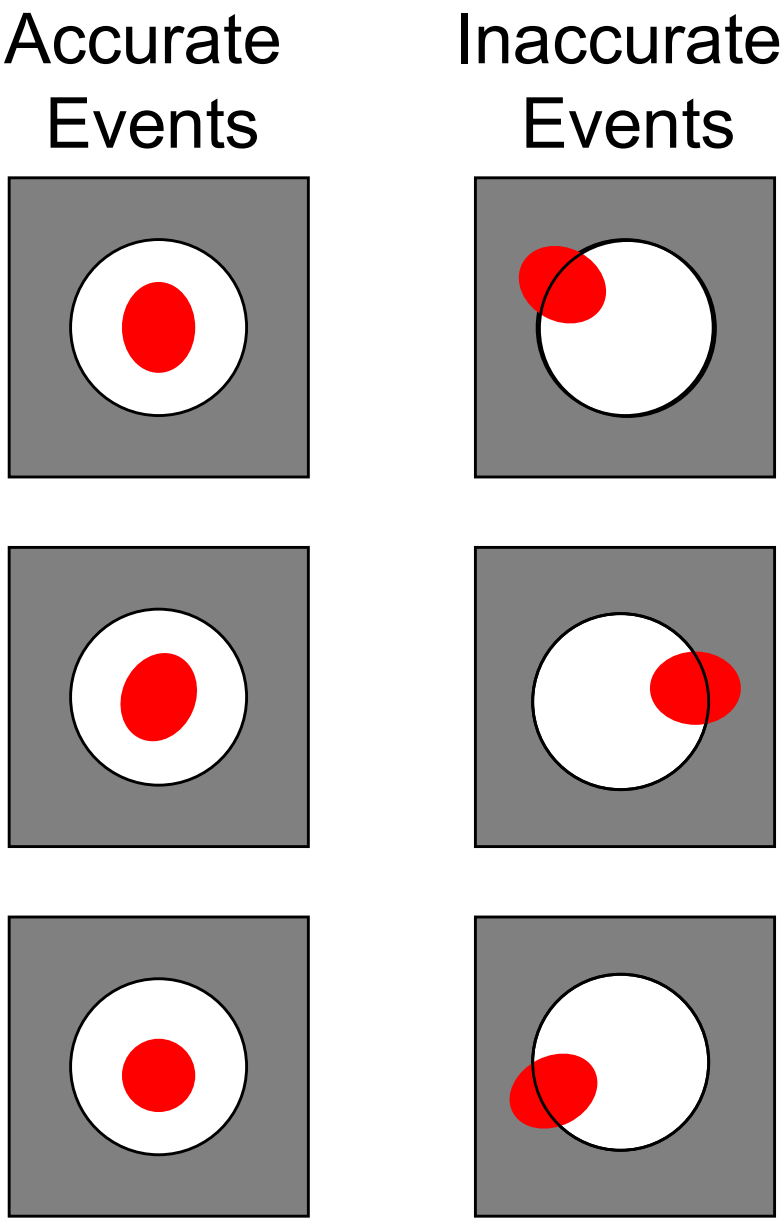
.

.

.

It's Amazing!

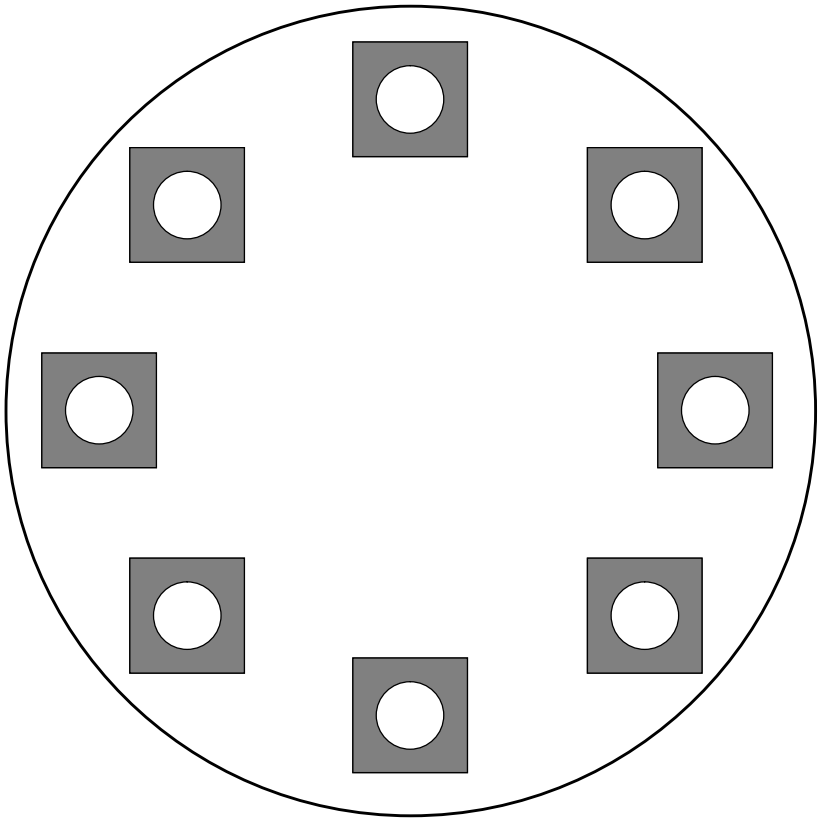
A camera was placed inside the instrument to view accurate events and inaccurate events using 3 'good' instruments and 3 'bad' instruments



Beverly Daniels
March 5, 2013

Verification of Alignment

A ‘pin’ was attached to the dispense nozzle to ‘dimple’ slides to assess the consistency and positioning of the dispense-slide alignment. There are multiple slides on a rotor within each instrument. The rotor moves the slides under the fixed dispense position.

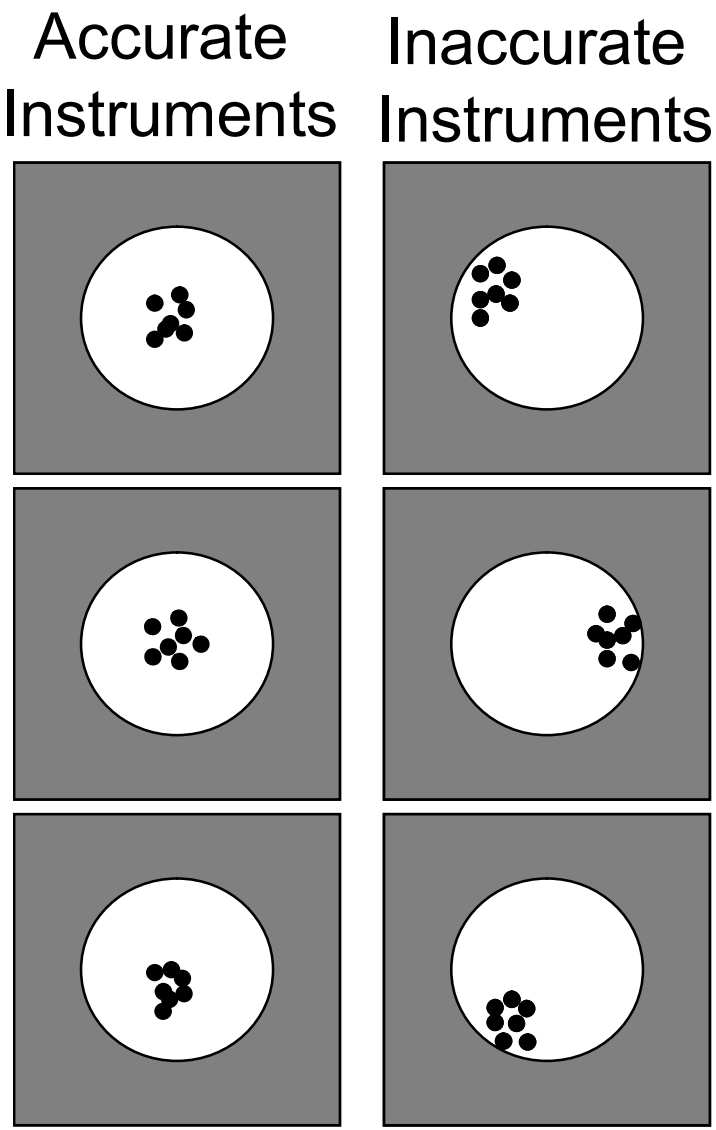


Beverly Daniels
March 5, 2013

Seeing Alignment

The dispense and slide positions are not aligned in the ‘bad’ instruments and are aligned in the ‘good’ instruments.

This graphical display is a form of multi-vari known as a measles chart

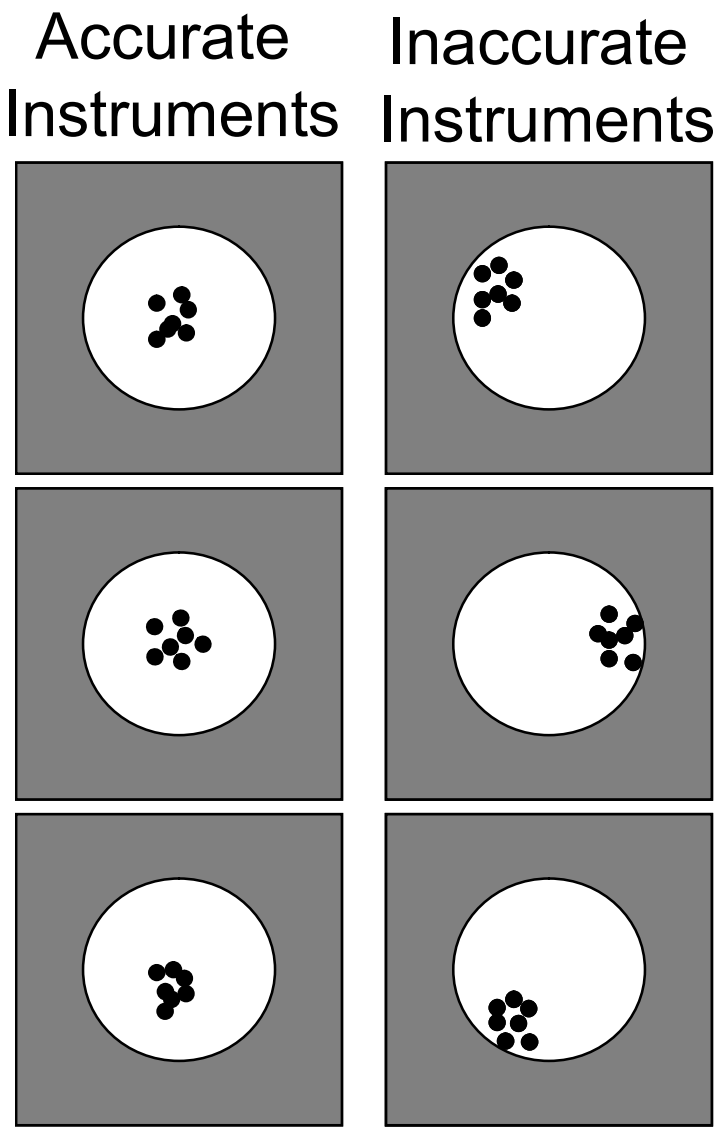


Beverly Daniels
March 5, 2013

ASQ

Quiz

What is the next diagnostic question?



Beverly Daniels
March 5, 2013

ASQ

Power Tools: Multi-Vari

Intro to Multi-Vari

The Multi-Vari Chart is a method for detecting where & when the largest variation in Y is occurring.

It is a graphical method that is very objective in it's approach to determining the significant X's: it looks at all of them, *known and unknown.*

The chart is a graph of the Y characteristic plotted *in time sequence* in groupings known as **components of variation**.

The grouping that produces the largest change in Y is the category of variation that contains the significant X. **It is most appropriate for manufacturing or fabrication type processes.** It is not typically the best choice for field failures or final assembly functional test failures...

Beverly Daniels
March 5, 2013

ASQ

Stratified Time Sequence Sampling

- To perform a multi-vari, it is necessary to identify Y data in *time ordered sequence* by when the Y characteristic is *created* not measured.
- **The components of variation are identified simply as categories of input factors (grouped Xs) that can create variation in the Y independently from other categories.**
- The Y data is not summarized, individual data points are analyzed as a set. No averages, no ranges, no control charts.
- No random samples: all positional and sequential identification must be maintained.
  This is also called a stratified sample.
- Each Component of Variation is sampled 3 times

ASQ

The “Big 3” Categories of Variation

Within Piece: Typically 2 - 4 locations within the piece would be measured. Conditions to look for include: taper, out-of-round, bowing, etc.

Piece to Piece: A minimum of 3 sequential pieces must be measured. More can be measured if there is reason to believe that the X's vary in a somewhat slower manner.

Time to Time: This would most likely include vendor batch changes, environmental changes such as (external to the equipment or ambient) temperature, humidity, barometric pressure, etc.

Nested Categories of Variation

Operator to Operator: Again be sure that operator identification is maintained for the Y data.

Shift to Shift: If your process runs more than one shift, you should maintain the key shift data: at a minimum this would be starting shift (time) and ending shift (time).

Machine to Machine: The Y data for each machine must be collected separately and plotted independently with the same sample plan.

Process Specific Categories of Variation

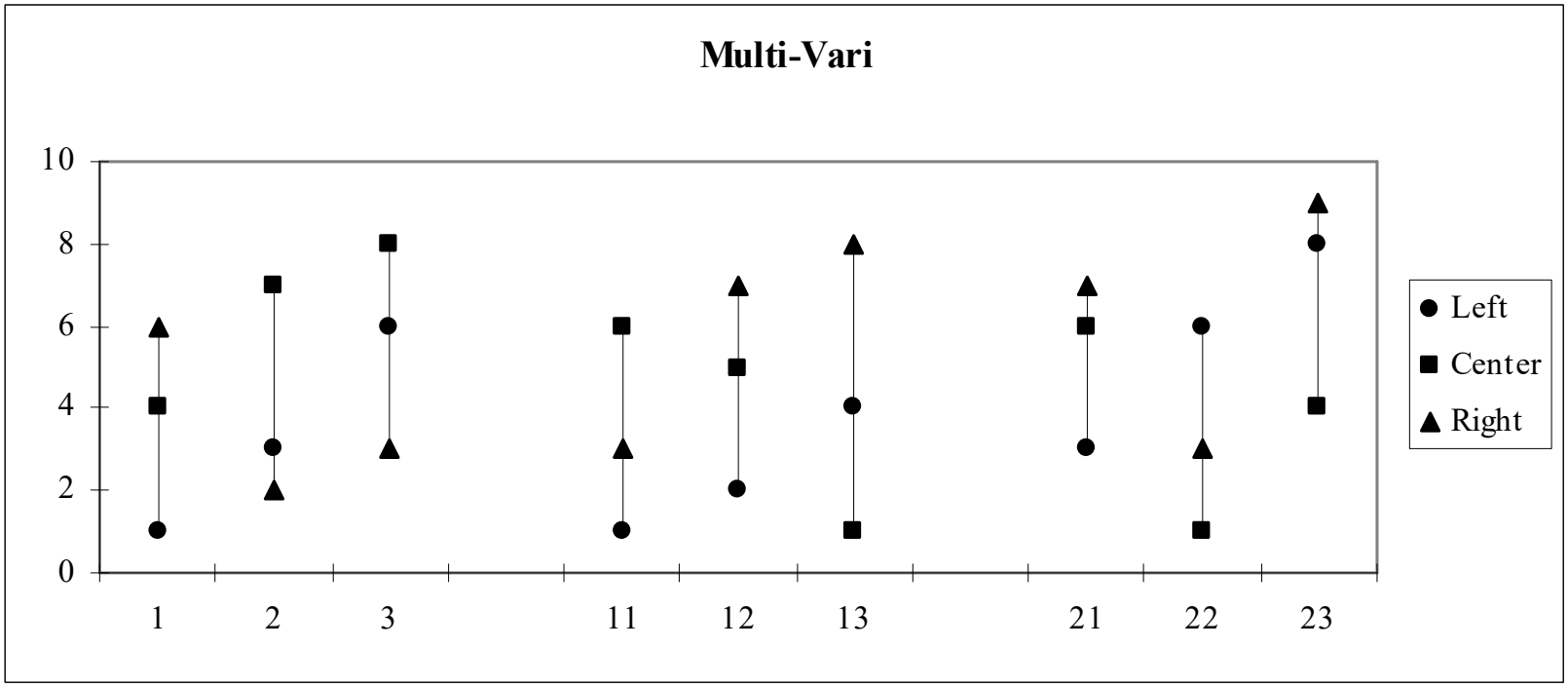
Within Fixture: In the case of a batch turn or drilling or other similar operation with a circular fixture, each piece in the fixture should be measured to determine Within Fixture type variation such as runout or particular fixture stations that may vary significantly from each other.

Mold to Mold or Cavity to Cavity: Data must be taken for each mold or cavity and positional tags maintained.

Batch to Batch: If your batch process has fixed positions or molds or cavities, etc., it essential to maintain positional tags. If your batch process is a random grouping of units then a small random sample of 3 units from each batch would be sufficient.

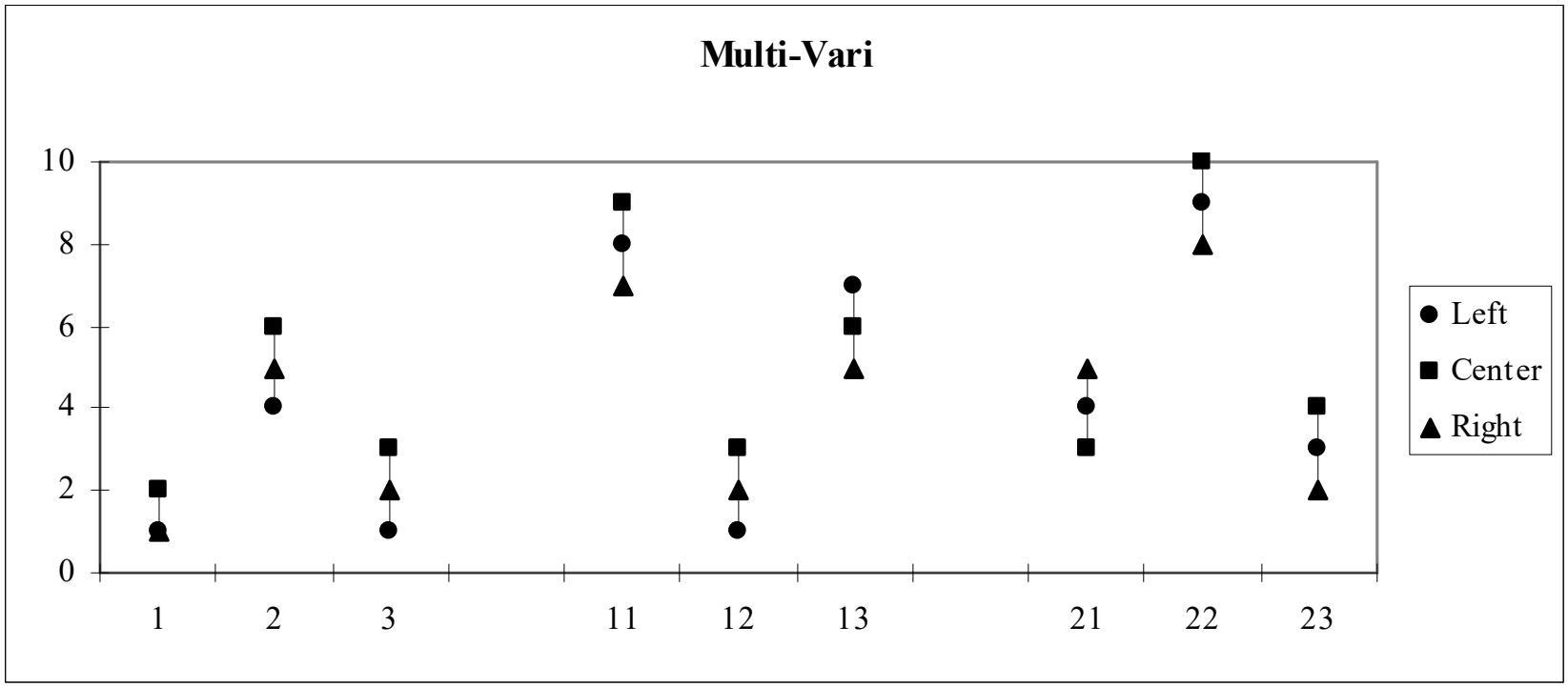
Cycle to Cycle: This is seen by looking at consecutive parts from the same cavity, mold, or fixture position.

If Within Piece is the Largest Category



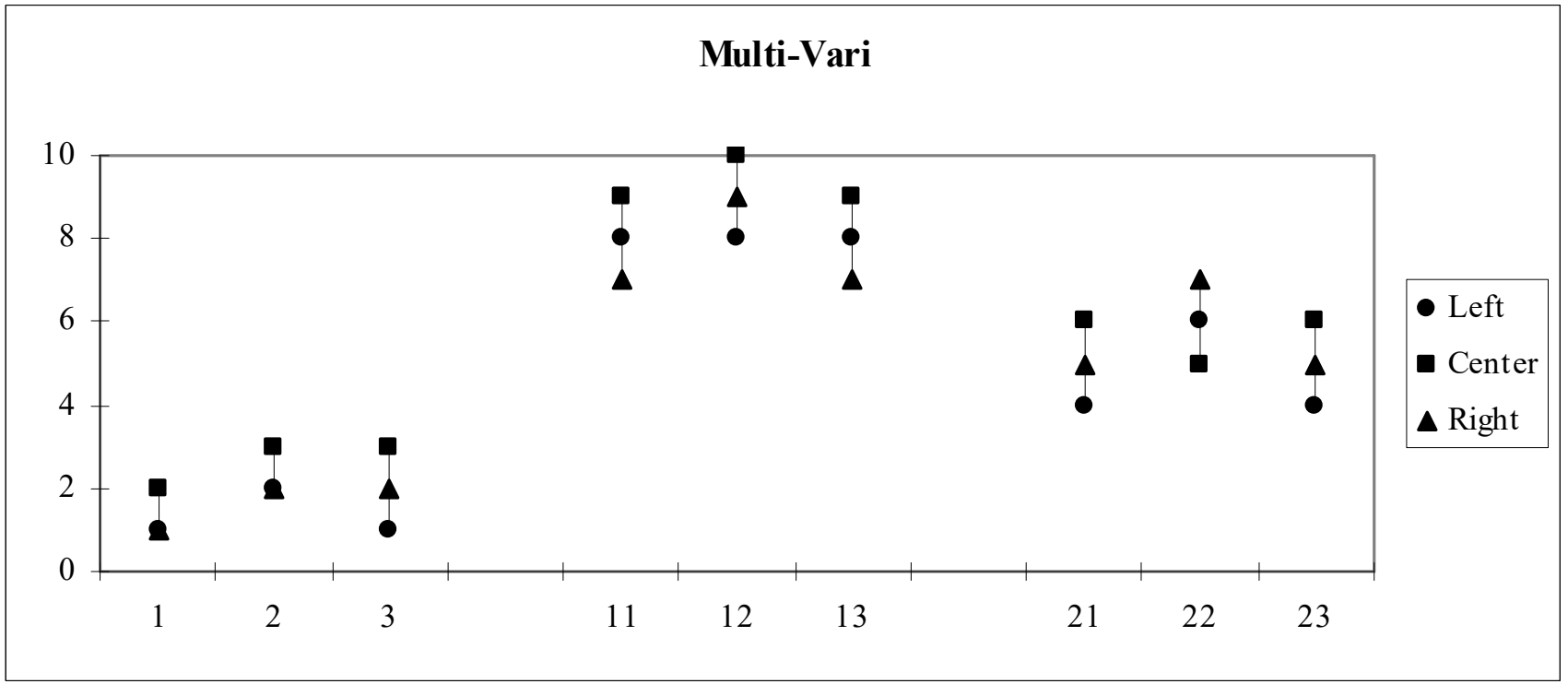
Beverly Daniels
March 5, 2013

If Piece to Piece is the Largest Category



Beverly Daniels
March 5, 2013

If Time to Time is the Largest Category



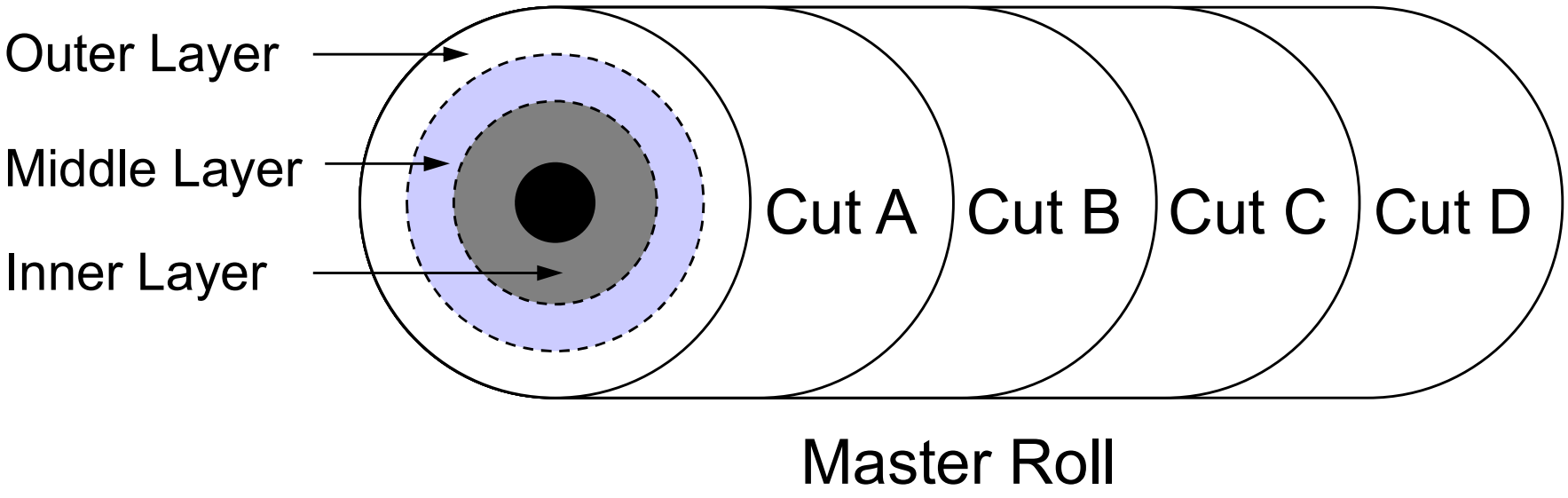
Beverly Daniels
March 5, 2013

Example 3: Sticky Layer Coefficient of Friction*

A functional failure has been traced back to the kinetic coefficient of friction of a 'sticky' layer on a foil backing.

The foil is a commodity part produced in large rolls.

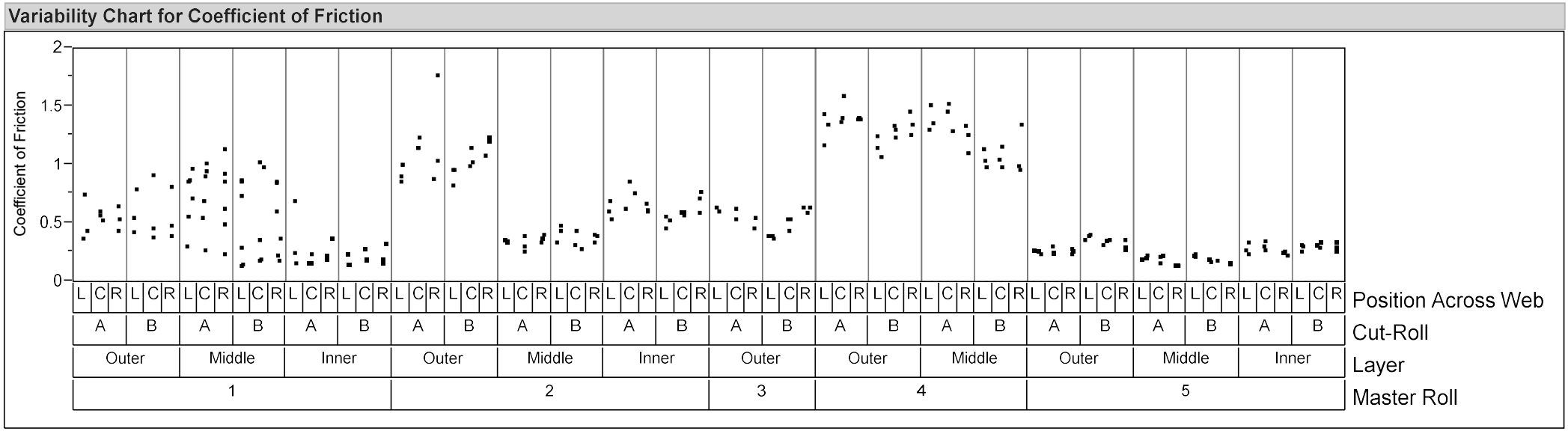
The master rolls are divided into sub-rolls using the following protocol:



Beverly Daniels
March 5, 2013

Looking at the Contribution From Across the Web

3 subgroups were taken from each sub-roll at the beginning, middle and end of the roll. 3 samples taken for each subgroup across the web at the left, center and right positions across the web



The largest component of variation is the Master Roll

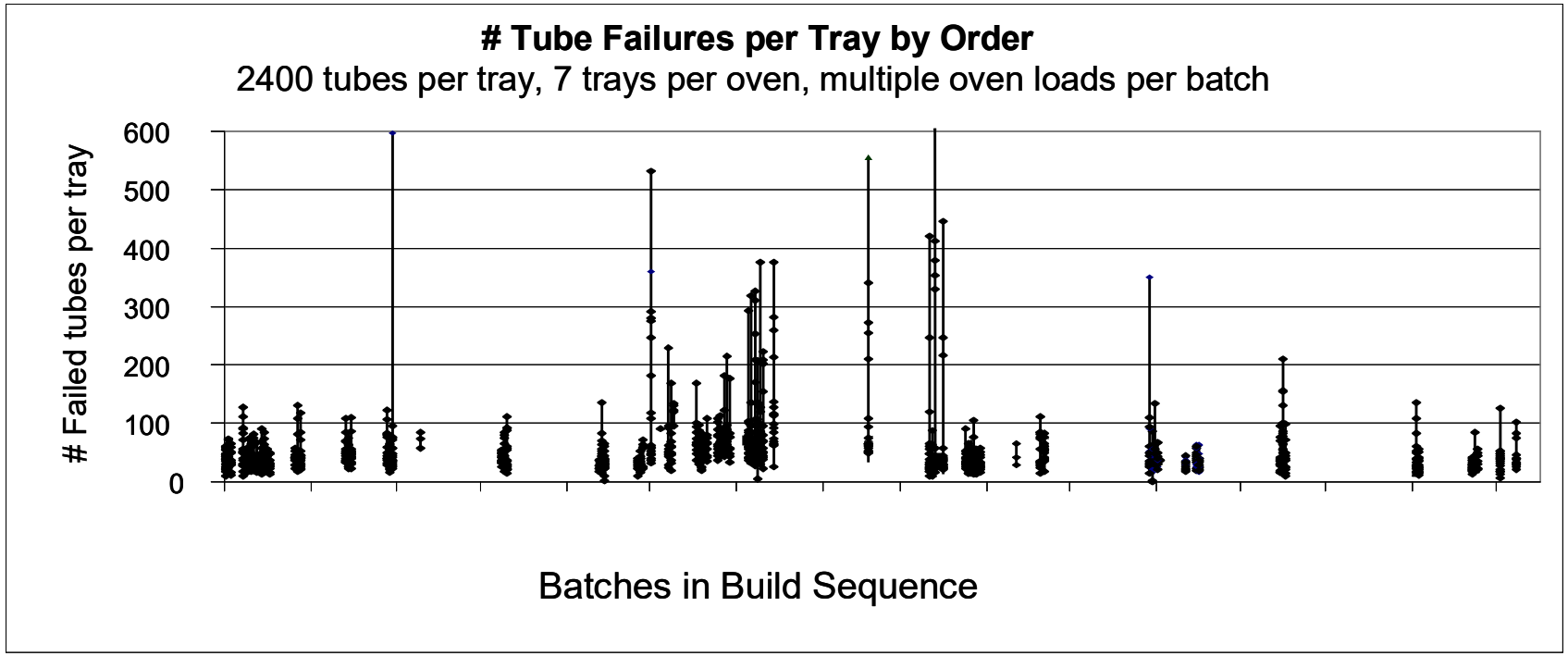
The primary causal mechanism of high CoF is a factor that changes from Master Roll to Roll

Beverly Daniels
March 5, 2013

ASQ

Example 4: Freeze Drying Failures*

The following chart shows the number of failed tubes per tray for each batch. (Each dot is the number of failed tubes in a tray).) The batches are plotted in time sequence.



This multi-vari chart shows more than a simple time series of the yields (some trays yield well others do not) but **it doesn't provide strong clues as to causality**.

Beverly Daniels
March 5, 2013

ASQ

Filling & Capping the Tubes: Operator Groupings

Each tray has 2400 tubes

One to three operators can be utilized to fill and cap the tubes prior to freeze drying

There are 7 trays per oven and multiple ovens are used per batch

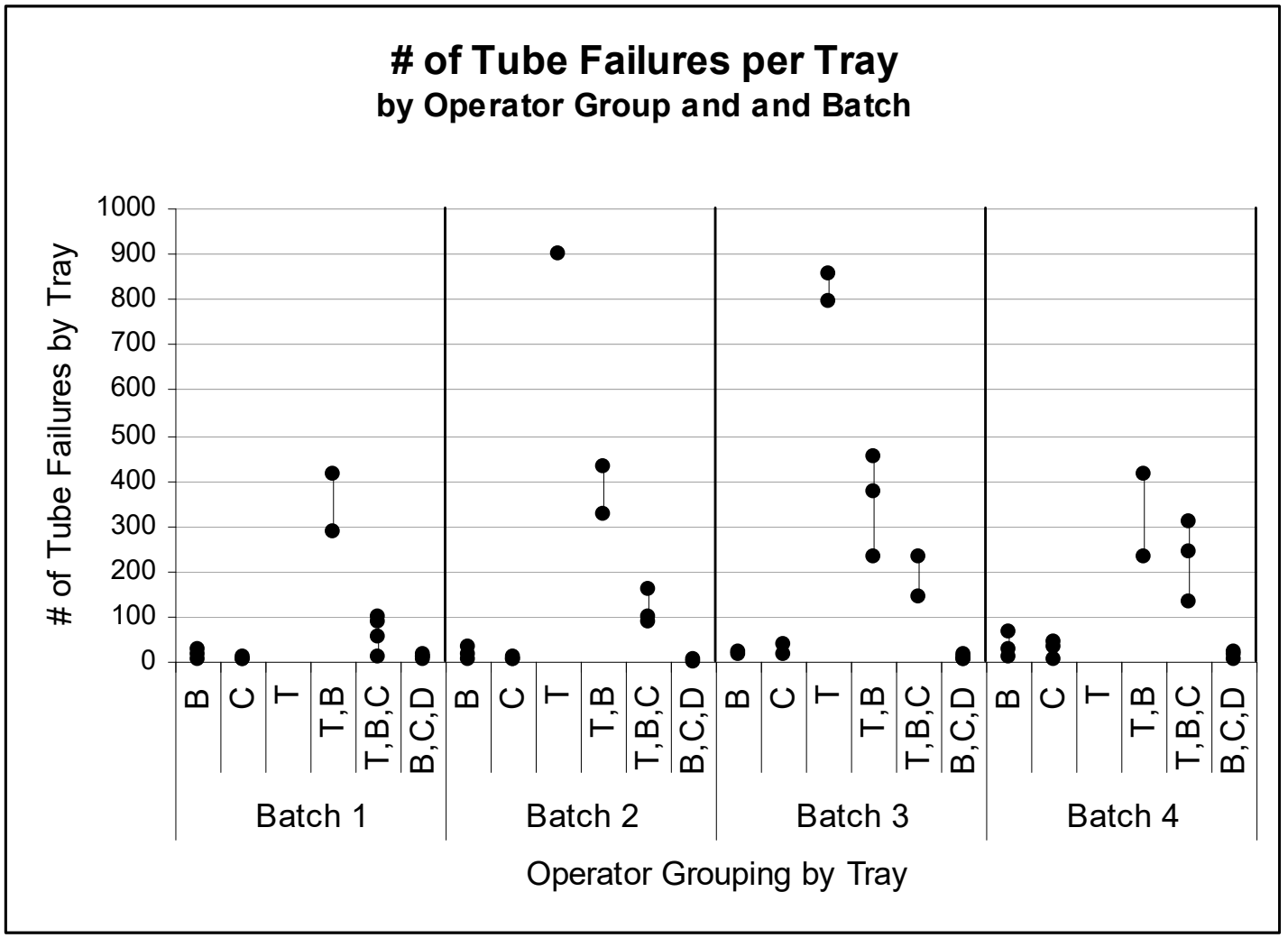
The production records were analyzed to determine which operators were assigned to each tray and the tube failures for each tray were plotted by operator group.

There is no way to know how many tubes each operator filled and capped in a multi operator group.

Beverly Daniels
March 5, 2013

ASQ

Multi-Vari Break Down by Operator



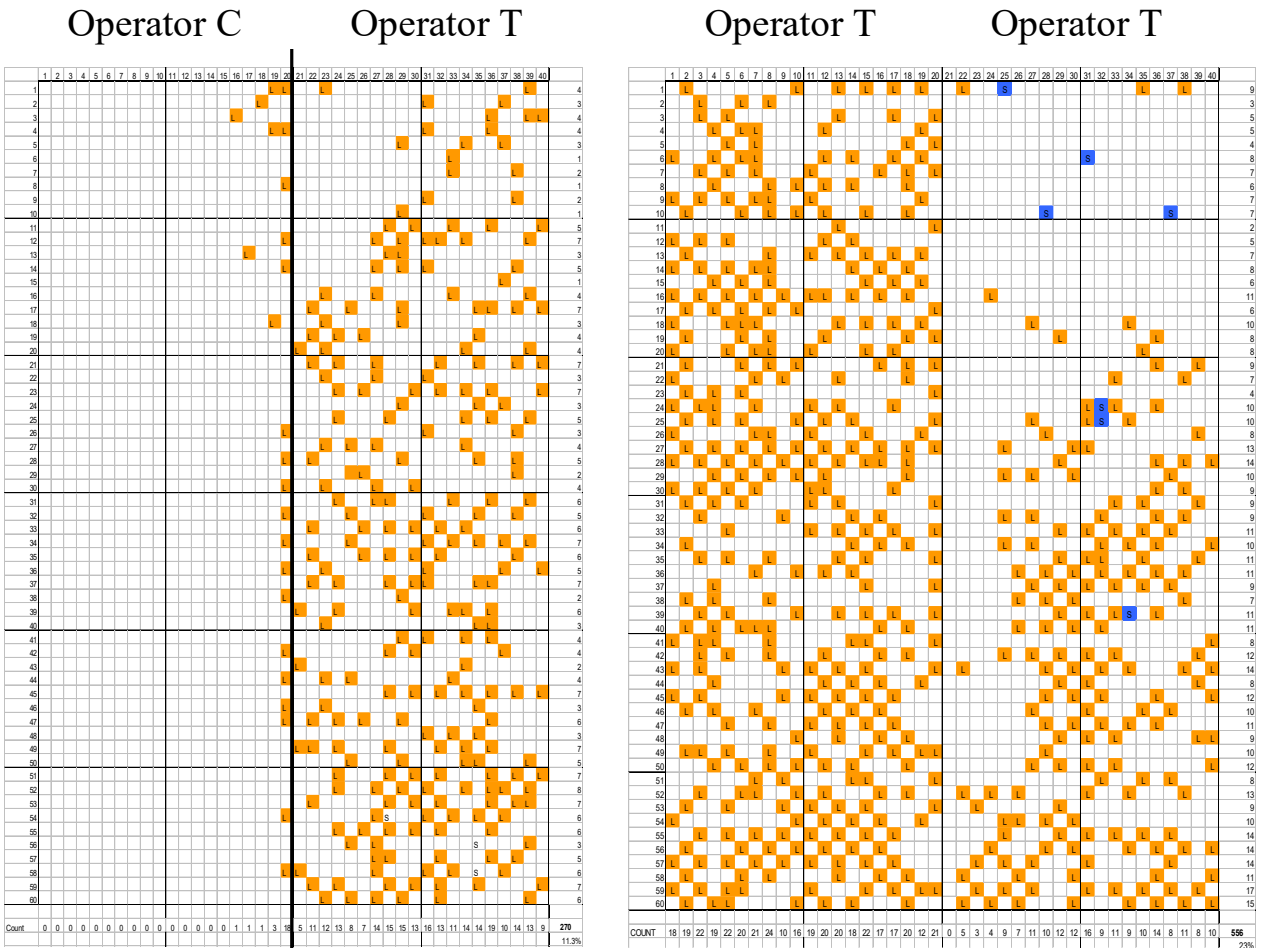
Almost every time operator ‘T’ is involved there are elevated tube failures

Beverly Daniels
March 5, 2013

ASQ

Final Breakdown: Measles Chart

Tube failures were mapped for multiple operator/tray combinations



The maps verify that Operator T has a high level of failures AND that they occur in a non-random pattern

Beverly Daniels
March 5, 2013

Power Tools: MSA
Not Your Father's Gage R&R

MSA Sampling Plan[4,5]

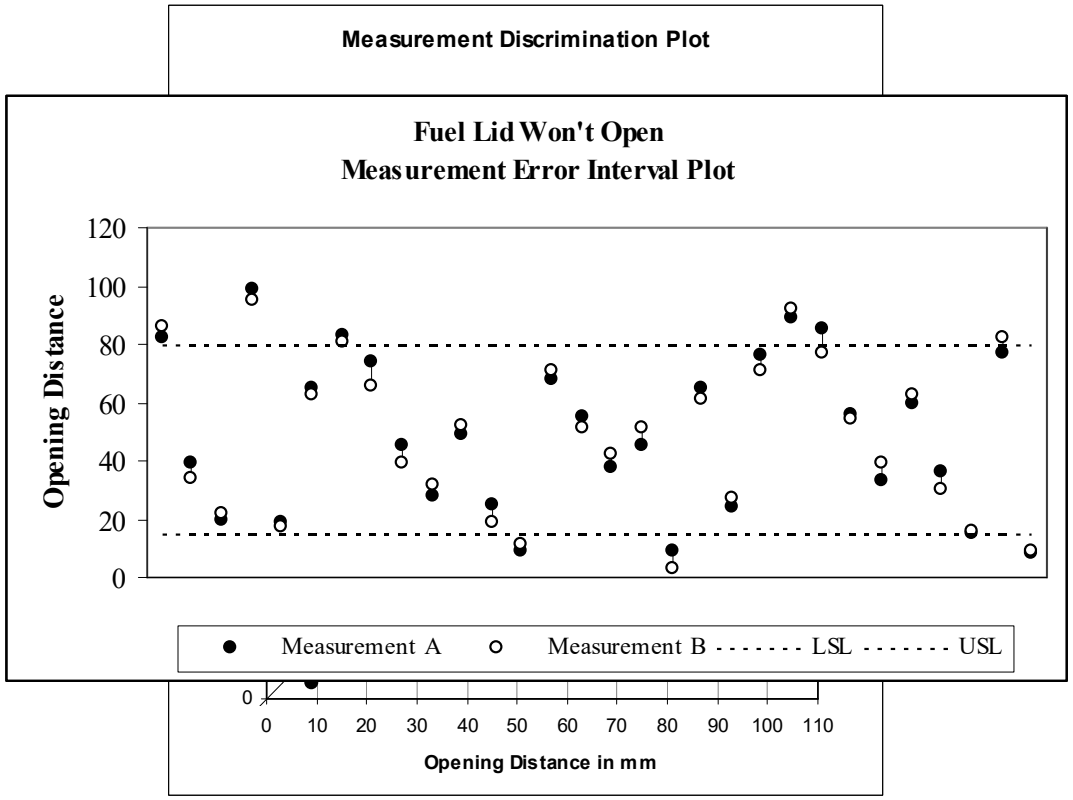
- Use 30 parts, measured twice.
- The study is intended to estimate the standard deviation of measurement error (the variation in repeated measures)
- The increase in the ‘accuracy’ of an individual standard deviation does not increase substantially when increasing the number of repeated measures from 2 to 3.
- Estimates of standard deviation are more ‘accurate’ with many small subgroups.
- 30 subgroups of 2 repeated measures provides the same number of data points as 10 subgroups of 3 repeated measures, but with better ‘accuracy’

ASQ

The Youden Plot[6]

The first and second measurements are plotted on a “square scatter diagram”

Fuel Lid Opening Distance
Steel Ruler; cm

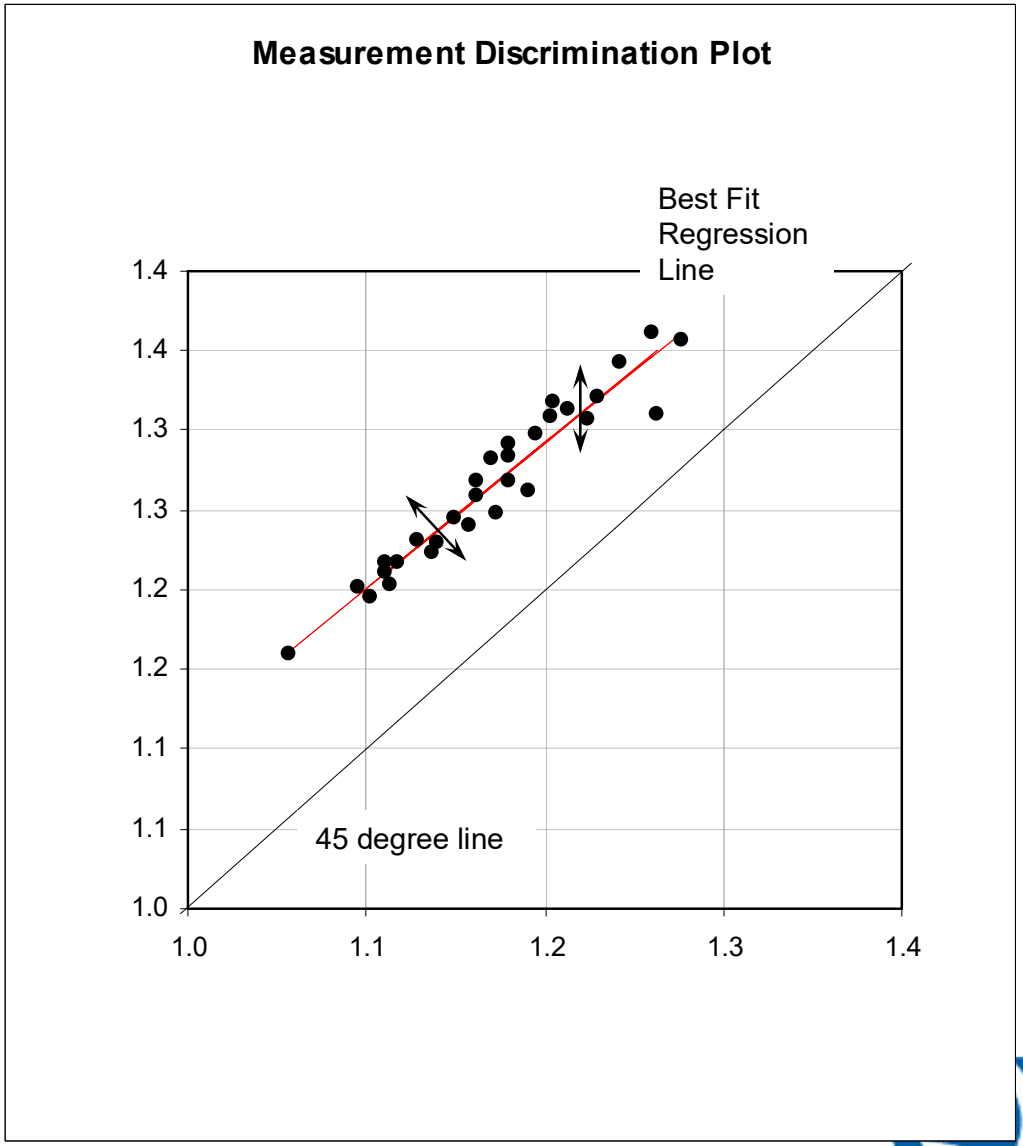


MSA Analysis Does Not Use Regression

Measurement error is the scatter in the direction that is perpendicular to a 45 degree line

Error in a regression is in the vertical direction: it is the variation in Y due to all other factors besides the X axis factor…

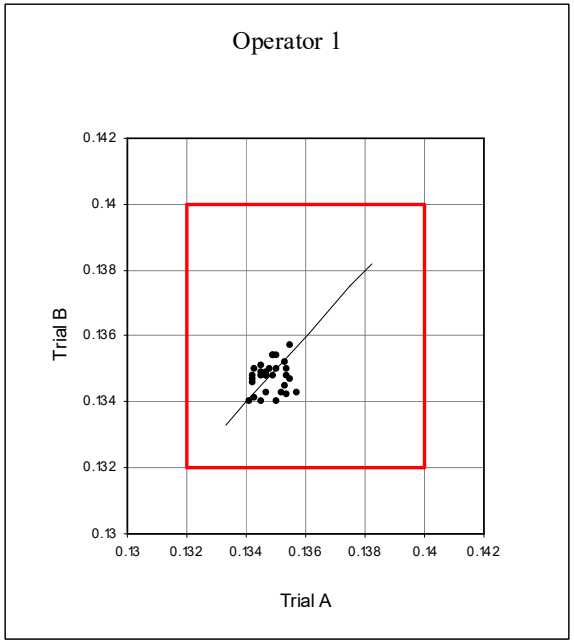
If regression is used to display measurement error, the regression line is the best fit line of the data points so that bias will not be visually detected.

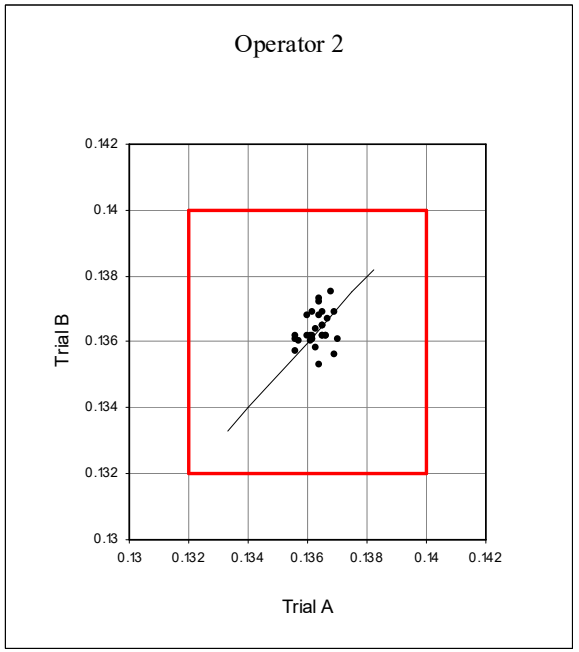


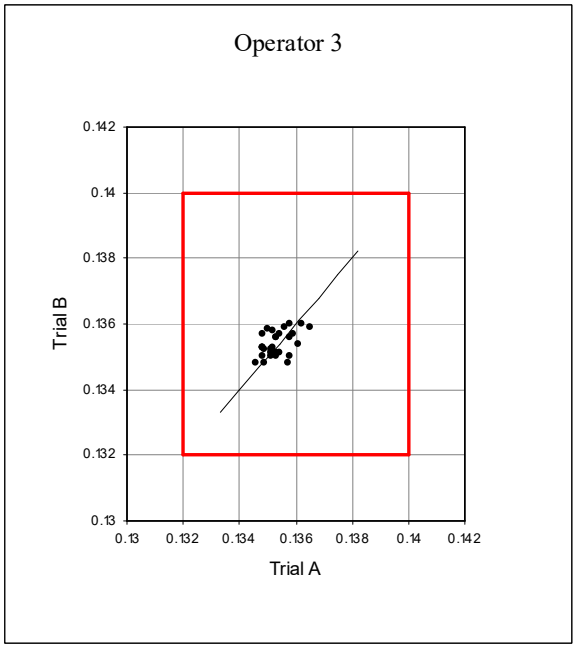
Beverly Daniels
March 5, 2013

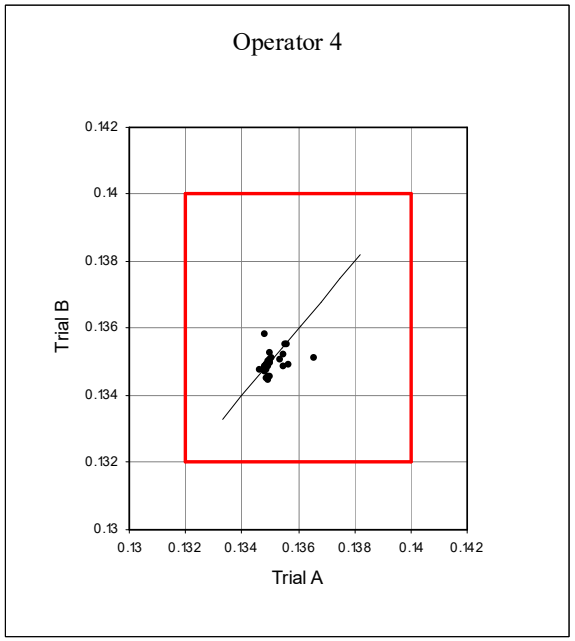
ASQ

Comparing Multiple Operators



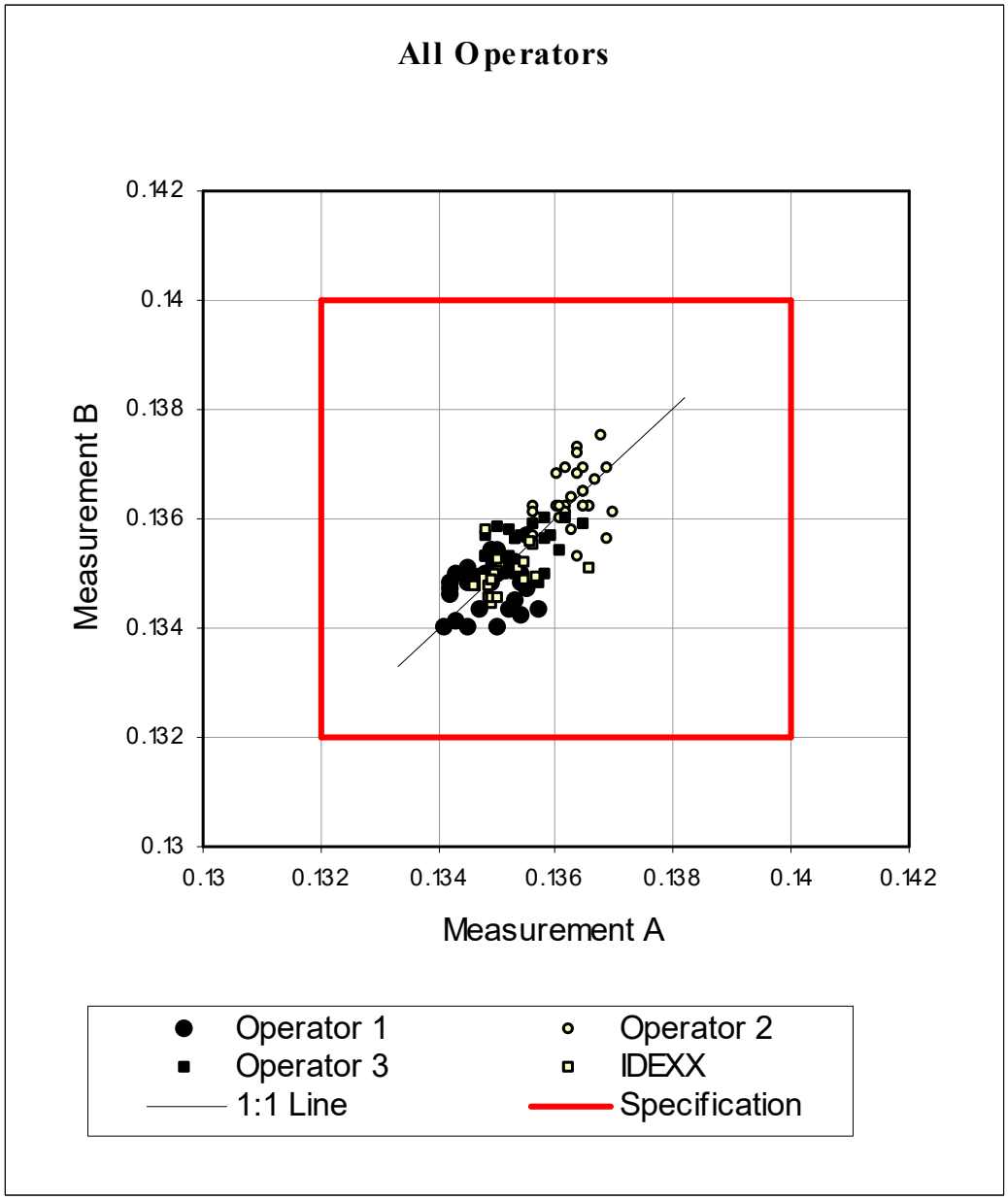






Beverly Daniels
March 5, 2013

Comparing Operators Together



Beverly Daniels
March 5, 2013

Functional Reproducibility

Functional tests pose a special challenge since the exact conditions of the test can never be repeated.

Repeatability is established thru multiple measurements in a very short run; test conditions vary only minimally

Reproducibility is established over multiple days, operators, tester set-ups, etc.; test conditions vary a lot.

Example 5: Over Recovery*

A Hematology instrument is experiencing an unacceptably large delta from the known sample value for a critical CBC characteristic.

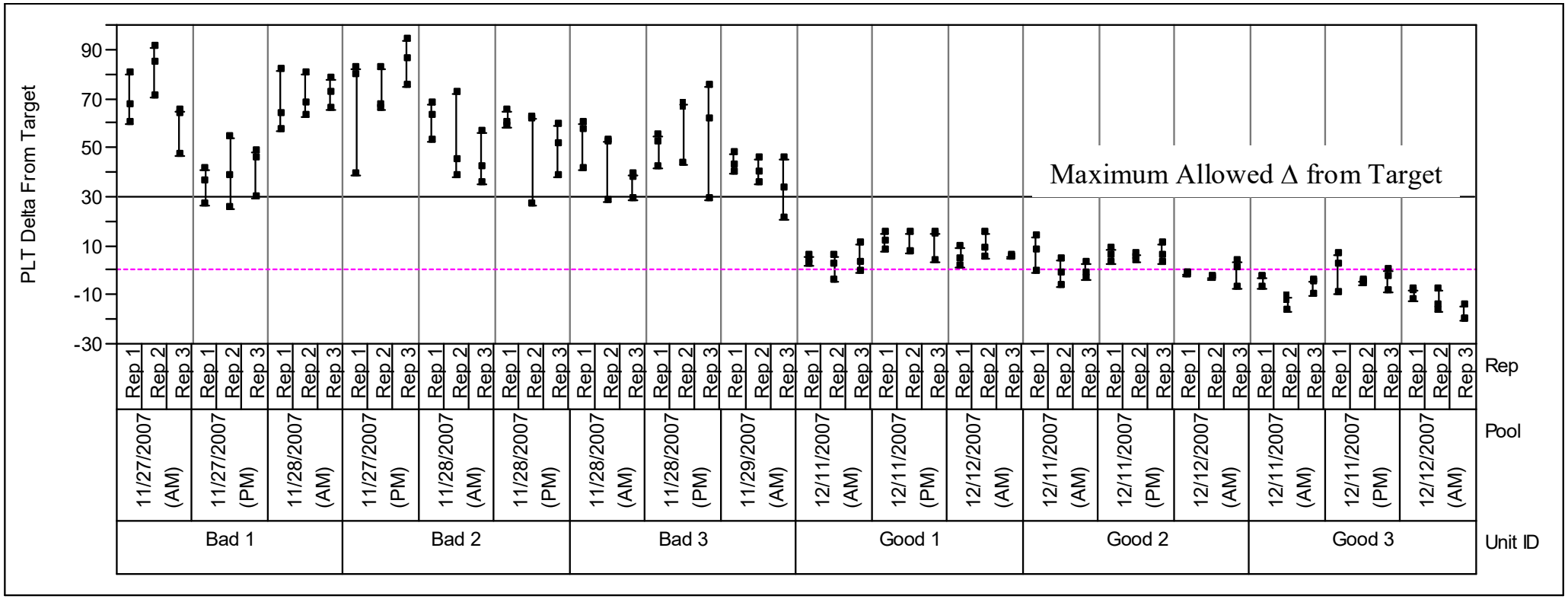
Repeatability: The within blood sample variability was tested using 3 good and 3 bad instruments. A total of nine runs were conducted on each blood sample, in three sets of three runs, on three separate samples. The 3 samples were tested without resetting the Instrument.

Reproducibility: each blood sample is further broken down into three sets of three runs.

During the test protocol a prime was performed after every three runs to mimic how the instruments are tested in the calibration process (in sets of 3 runs for QC blood samples).

This was done to evaluate the within time and time to time variation.

Over Recovery MSA Results

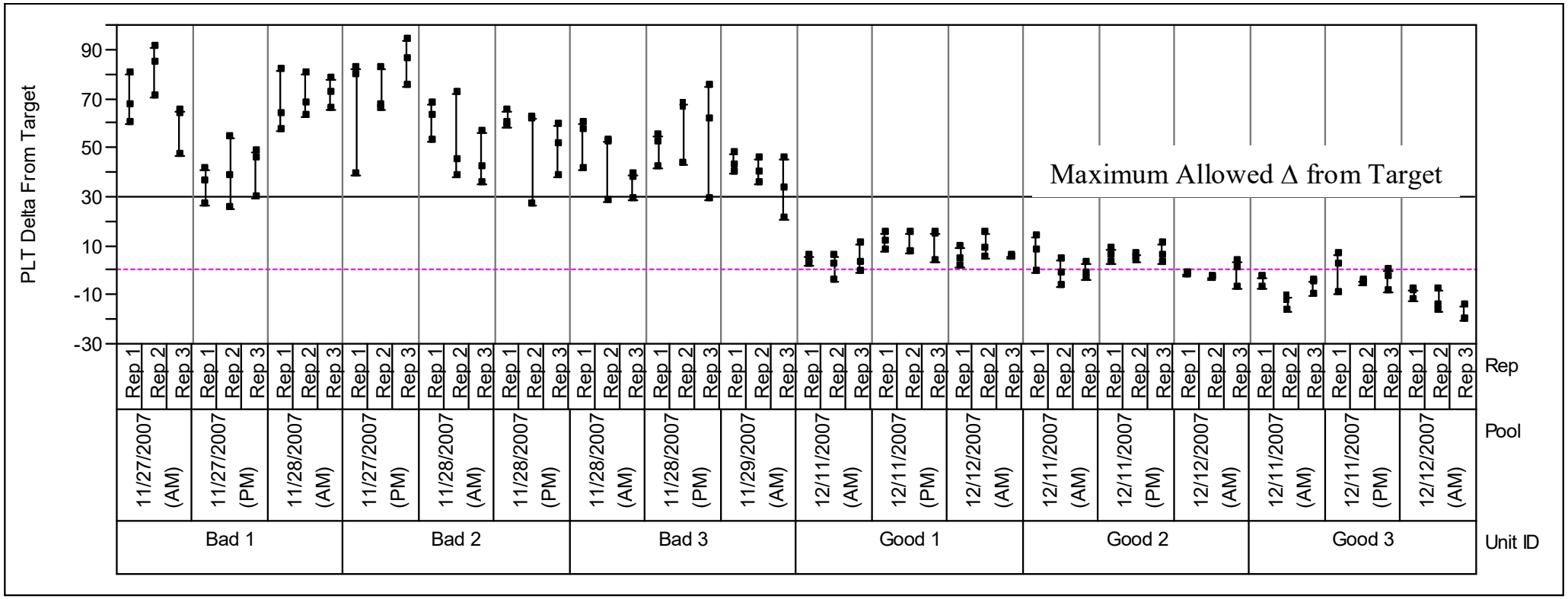


All Instruments indicate good repeatability & reproducibility.

Notice that the variation for bad instruments is much larger than the variation for good instruments.

Beverly Daniels
March 5, 2013

When OK is Good Enough



While there was decent separation of the results from good to bad, the data could only be separated into two distinct categories. **Therefore each experiment required 3 runs of the same sample.**

Beverly Daniels
March 5, 2013

ASQ

Power Tools: Assembly-Disassembly

Assembly-Disassembly[7,8]

An equal number of ‘good’ assemblies and ‘bad’ assemblies are collected.

The major components and parts are identified

Each assembly is disassembled and reassembled twice, measuring the characteristic of interest after each reassembly. This separates the assembly process from the components.

Components and parts are then switched and the characteristic of interest is measured again. The parts are returned to their original assemblies.

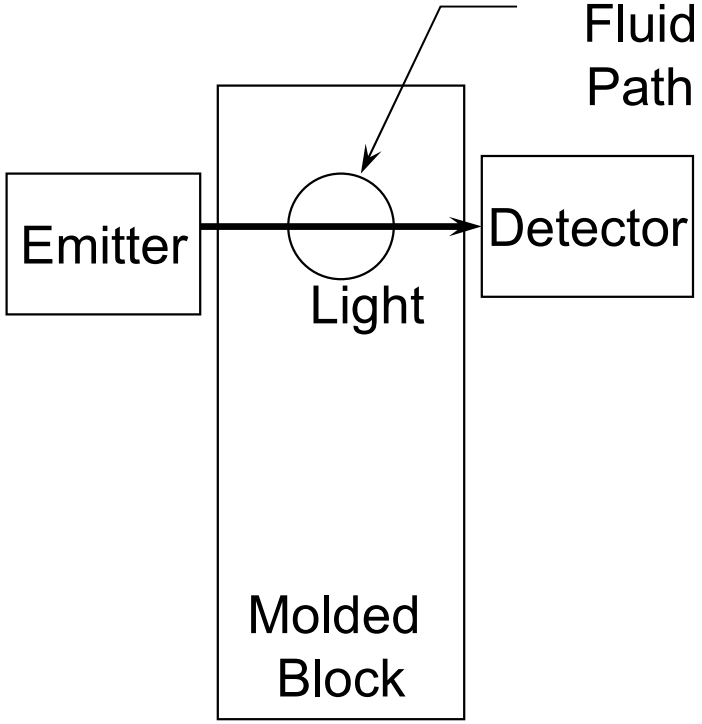
The process is completed until ‘good’ assemblies become ‘bad’ and ‘bad’ assemblies become ‘good’.

Example 6: Optical Block*

A fluid is passed thru a small hole in a translucent block of molded material. An emitter transmits light through the fluid path and a detector on the opposite side detects the amount of light that passes thru the fluid path (and the block).

An algorithm translates the difference between the emitted light and the detected light and calculates a particle count in the fluid.

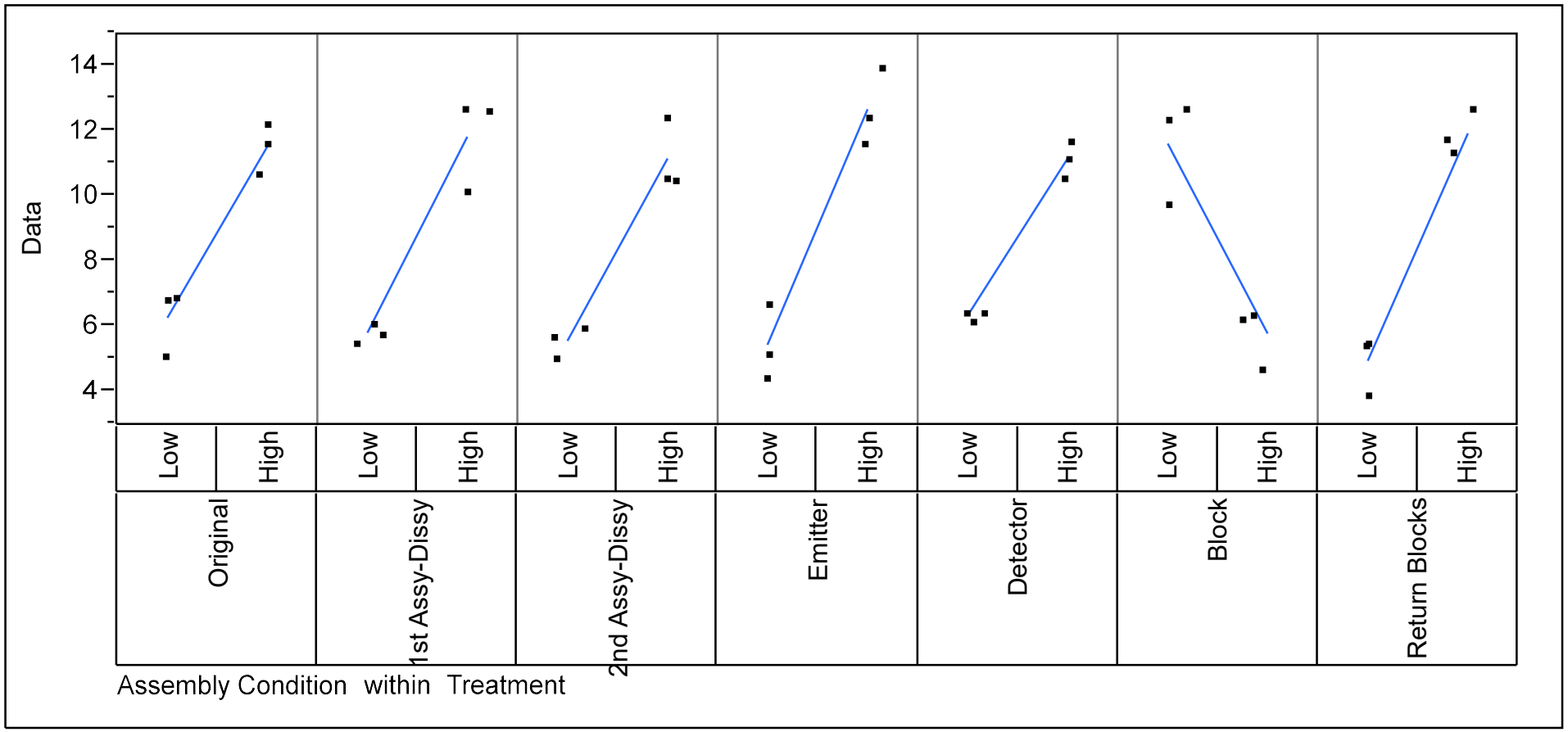
At some point, the behavior of the assembly changes so that the counts are not correct when compared to a known standard…they are too low.



Beverly Daniels
March 5, 2013

Assembly-Disassembly Results

3 low units and 3 high units are selected.



Beverly Daniels
March 5, 2013

Versatility of the Tool

Assembly-Disassembly can be used for many types "assemblies"

- Chemical products
- Hybrid electrical/mechanical/chemical assemblies
- Different production lines

Example 7: Slow Flow*

An Assembly that flows sample fluid in a forward direction and then ‘washes’ the sample fluid back to it’s point of origin is flowing too slowly.

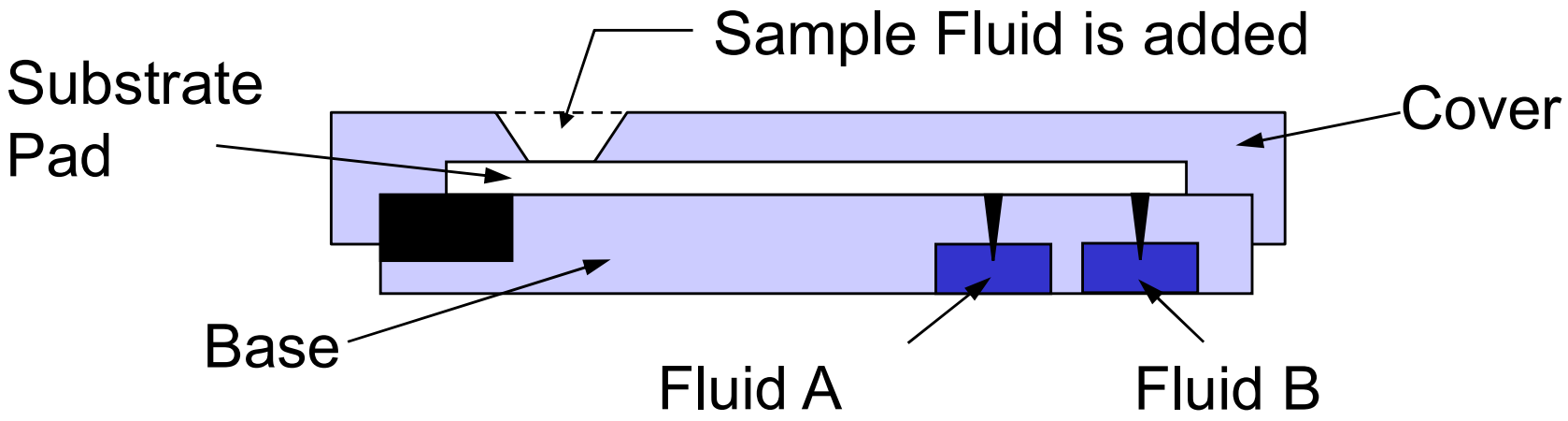
The assembly is comprised of

- A substrate pad through which the fluid flows
- The sample fluid is comprised of two different liquids
- A base that holds the substrate pad and the wash fluid
- A cover
- A couple of other smaller components within the base

This assembly can only be used once

Beverly Daniels
March 5, 2013

Assembly Diagram



Beverly Daniels
March 5, 2013

Assembly-Disassembly

Two lots of these devices were collected

One lot had a very slow flow rate

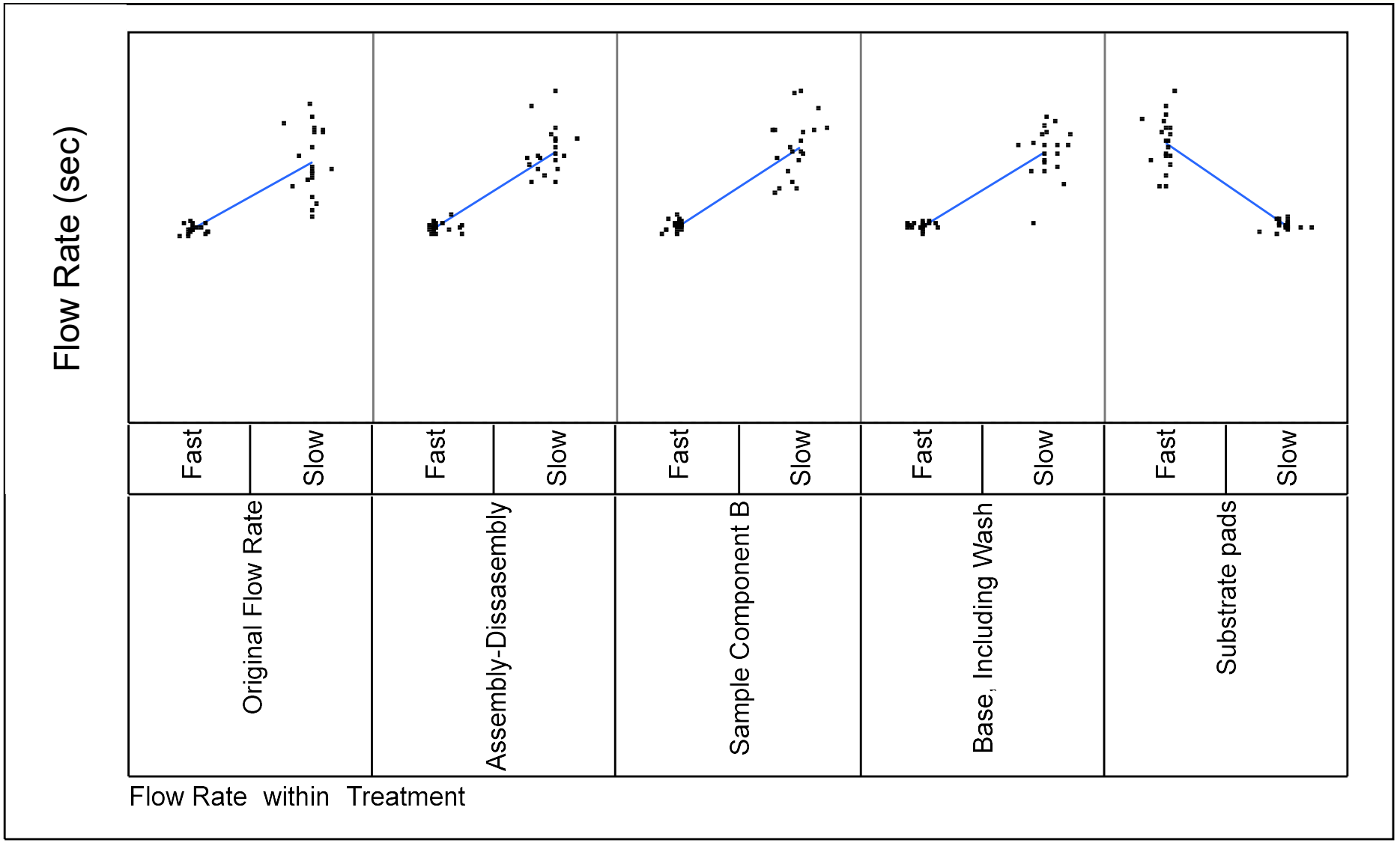
The other lot had a fairly fast flow rate

The sample size for each treatment was 20 units from each lot.

Since one of the components in the sample fluid came in lot sizes only large enough to test 50 units, a new lot was used for every treatment.

The experiment began by disassembling the 20 units for the fast lot and 20 units for the slow lot and then testing for flow time.

Results



The causal mechanism lies in the substrate pads

Beverly Daniels
March 5, 2013

ASQ

References and Bibliography

1 Deming, W. Edwards, “On Probability as a Basis for Action”, *The American Statistician*, 1975, Vol. 29, No. 4, pp146-152

2 Moen, Ronald D., Nolan, Thomas, W., Provost, Lloyd P., “Quality Improvement through Planned Experimentation”, 2nd Edition, McGraw-Hill, 1999

3 Kida, Thomas, “Don’t Believe Everything You Think, Prometheus Books, 2006

4 Donald J Wheeler, Craig Award Paper, “Problems With Gauge R&R Studies”, 46th Annual Quality Congress, May 1992, Nashville TN, pp. 179-185.

5 Donald S. Ermer and Robin Yang E-Hok, “Reliable data is an Important Commodity”, *The Standard*, ASQ Measurement Society Newsletter, Winter 1997, pp. 15-30.

6 Youden, William John, “Graphical Diagnosis of Interlaboratory Test Results”, *Industrial Quality Control*, May 1959, Vol. 15, No. 11

7 Ott, Ellis R., “A Production Experiment with Mechanical Assemblies*”, Industrial Quality Control*, Vol. 9, No. 6, May 1953, pp. 124-130

8 Bhote, Keki R. and Bhote, Adi K., World Class Quality - Using Design of Experiments to Make it Happen, Second Edition, AMACOM, 2000

9 Steiner, Steven and Mackay, Jock, “Statistical Engineering: An Algorithm for Reducing Variation in Manufacturing Processes”, American Society for Quality, March 2005

ASQ

*Acknowledgements

“Spin Error” project led by Kurt Brink, Black Belt

“Accurate Chemistry Read” project led by Everett Hall, Black Belt candidate

“Sticky Foil” project led by Dave White, Black Belt candidate

“Freeze Drying Failures” project led by Bob Cashman, Black Belt

“Over Recovery” project led by Angela Haiss, Black Belt

“Optical Block” project led by Luis Rodriguez, Black Belt

“Slow Flow” project led by Sharon Tynan, Black Belt