

How To: Perform a Repeatability and Reproducibility Study for Destructive Tests

Using STATGRAPHICS Centurion

by

Dr. Neil W. Polhemus

October 29, 2008

Introduction

Repeatability and reproducibility studies are designed to quantify the error associated with a measurement system. To determine pure measurement error, separate from any variation amongst the items measured, it is necessary to be able to take more than one measurement of the same item. Unfortunately, some measurement processes result in the destruction of the item being measured, making it impossible to take more than one measurement.

While it may be impossible to repeat the measurement on the same item, it is sometimes possible to identify a set of items alike enough to consider any variation in the measurements of those items representative of measurement error rather than item-to-item variability. This might occur by selecting multiple samples from a single batch or some from other homogenous sampling unit. The observed variation amongst measurements made on the homogenous samples provides an estimate (albeit possibly conservative) of the variation of the measurement process.

Sample Data

The August 2002 edition of Six Sigma Forum Magazine contained an article by Douglas Gorman and Keith M. Bower titled “Measurement System Analysis and Destructive Testing”. They describe an example of a measurement system which measures the impact strength of steel by breaking a metal bar. While multiple measurements could not be made on the same bar, bars constructed from the same ingot were assumed to be homogenous enough to provide an estimate of measurement error.

The data used in that article have been saved in the file named *impact.sfb*. A portion of that data is shown in the table below:

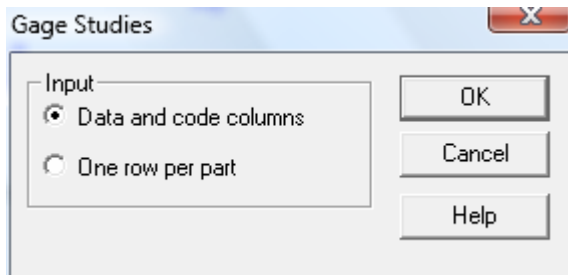
Row	Strength	Operator	Ingot	Sample
1	110.799	1	3	1
2	113.338	1	2	1
3	111.805	1	4	1
4	112.475	1	1	1
5	113.677	1	5	1
6	111.515	2	4	1
7	109.910	2	5	1
8	111.133	2	1	1

9	111.019	2	3	1
10	110.596	2	2	1
11	112.108	3	2	1
12	111.334	3	3	1
13	109.514	3	4	1
14	111.407	3	5	1
15	111.549	3	1	1
...				

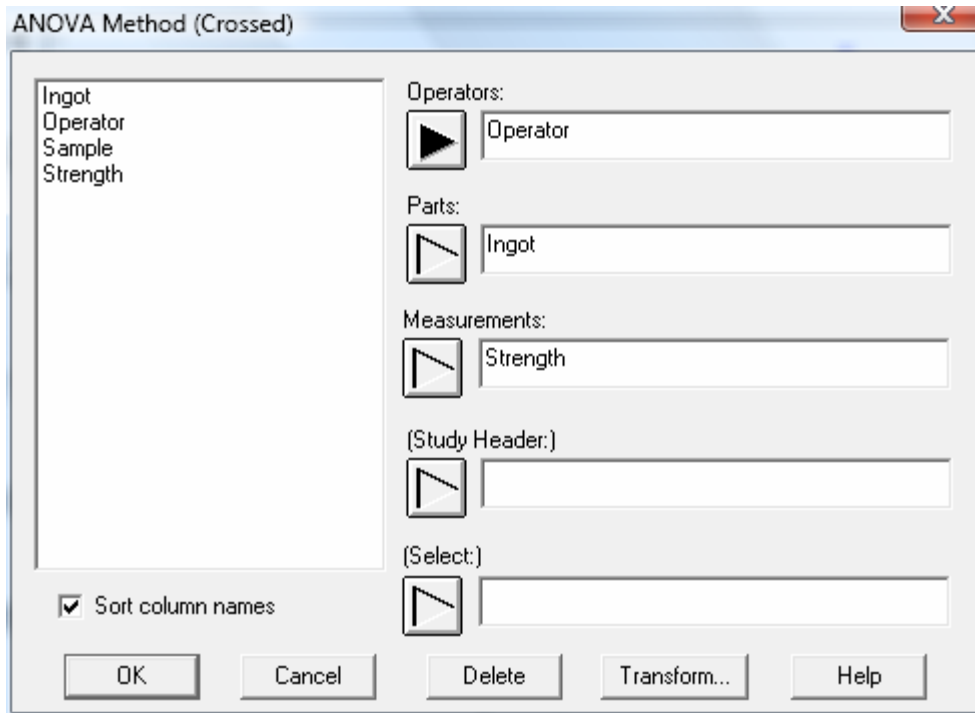
In the study, each of 3 operators measured the strength of 3 bars taken from each of 5 ingots. Since no more than 3 bars could be constructed from each ingot, each operator measured bars from different ingots. Such a sampling plan is said to have *nested* factors, since ingots are nested within operators and bars are nested within ingots.

Data Input

To analyze the data, select the *Gage Studies – Variable Data - ANOVA Method* from the main menu. If you are using the classic menu, this is located under the *SPC* heading. If you are using the Six Sigma menu, it is located under *Measure*. When the first dialog box appears, indicate that the data have been entered in “Data and Code Columns” (rather than putting all measurements for the same ingot in a single row):

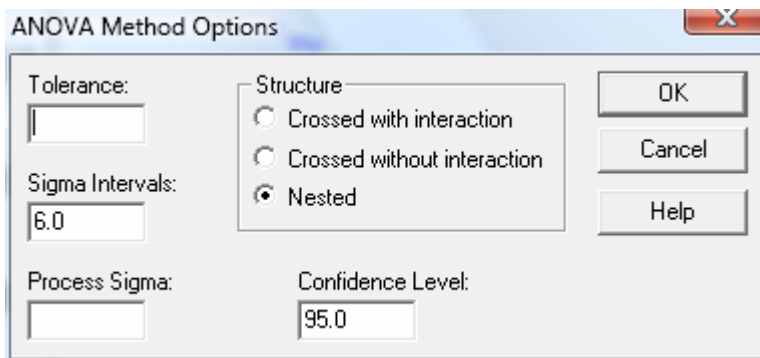


Then complete the second data input box as shown below:



Notice that each ingot plays the role of a single part.

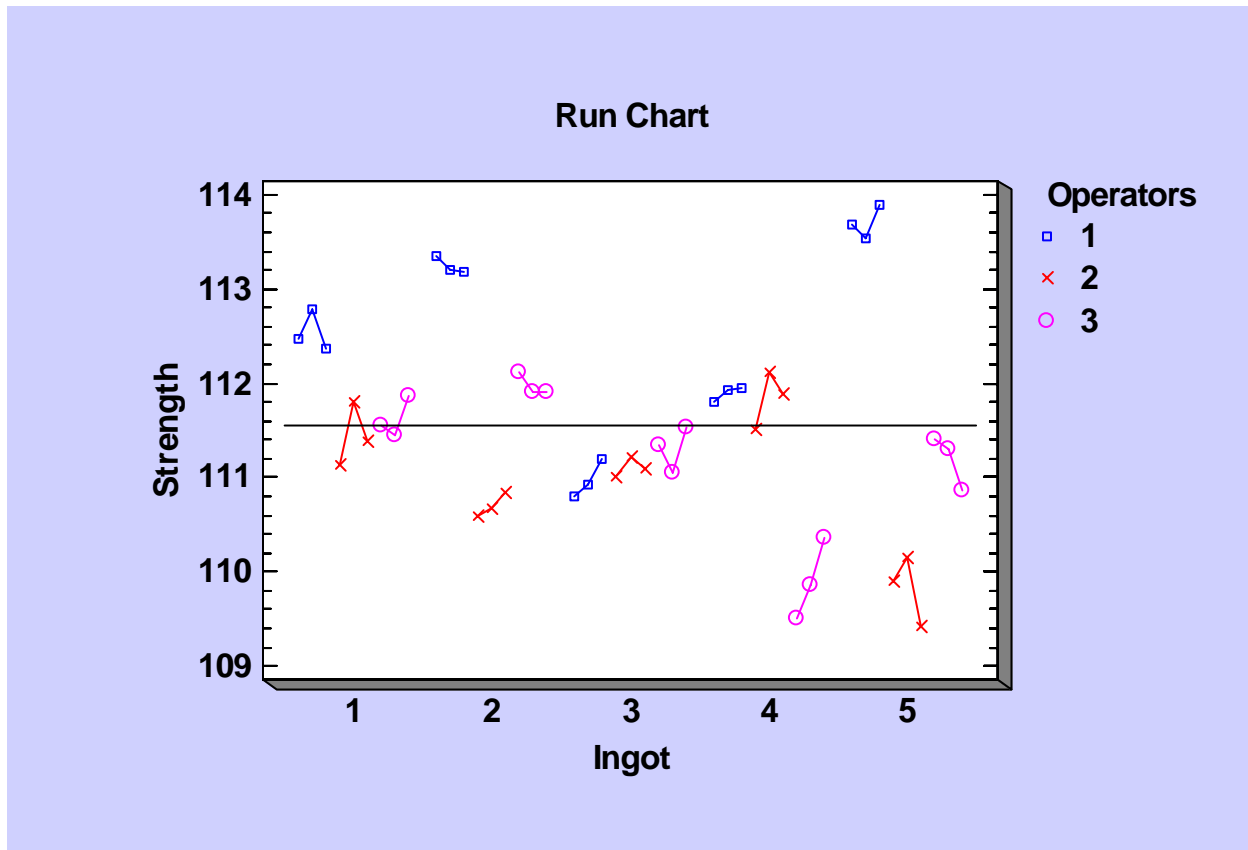
Before examining the results, it is important to press the right mouse button and select *Analysis Options*:



Notice the field labeled “Structure”. Since each operator measured samples from different ingots, the *Nested* radio button must be checked. Had it been possible to construct 9 bars from each ingot, so that each operator measured bars from the same 5 ingots, then a *Crossed* analysis would have been performed. In the later case, one has the option of including a operator-by-part interaction, which would allow the differences between operators to vary from ingot to ingot.

Run Chart

It is always helpful to plot the data before examining the numerical results. In the *Run Chart* below, the measurements made on the 3 bars from each of the 15 ingots are displayed:



Notice the relatively small variation amongst bars within ingots, relative to the differences between ingots. This supports the assumption that samples taken from the same ingot are homogenous.

Analysis Summary

In a nested R&R study, the overall variation in the data is assumed to consist of 3 variance components (repeatability, reproducibility, and part-to-part) which combine in the following manner:

$$\hat{\sigma}_{total} = \sqrt{\hat{\sigma}_{repeat}^2 + \hat{\sigma}_{repro}^2 + \hat{\sigma}_{parts}^2}$$

The STATGRAPHICS *Analysis Summary* table shows estimates of each of these components:

Gage R&R - ANOVA Method - Strength

Operators: Operator
Parts: Ingot
Measurements: Strength
ANOVA: nested

3 operators 5 parts 3 trials

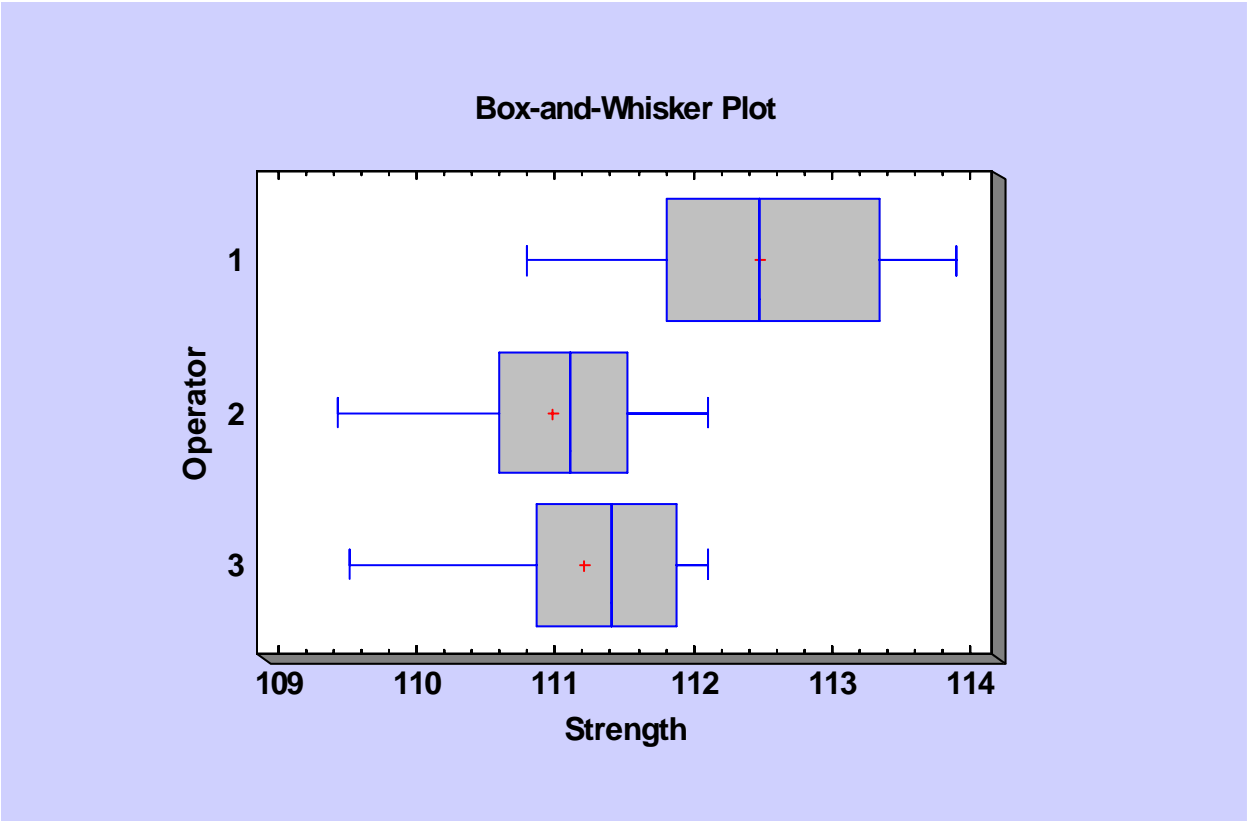
Gage Repeatability and Reproducibility Report

<i>Measurement</i>	<i>Estimated</i>	<i>Percent</i>	<i>Estimated</i>	<i>Percent</i>	<i>Percent</i>
<i>Unit</i>	<i>Sigma</i>	<i>Total Variation</i>	<i>Variance</i>	<i>Contribution</i>	<i>of R&R</i>
Repeatability	0.242463	21.141	0.0587883	4.4694	10.81
Reproducibility	0.696612	60.7393	0.485268	36.8927	89.19
R & R	0.737602	64.3133	0.544056	41.3621	100.00
Parts	0.878234	76.5754	0.771295	58.6379	
Total Variation	1.14689	100.0	1.31535		

Of particular interest is the column labeled *Percent Contribution*, which shows the percentage of the total variance attributable to each component. For the sample data, repeatability contributes only 4.5% to the total variance of the results, implying that the measurements are very repeatable when the same operator makes the multiple measurements. However, the reproducibility contribution (36.9%) is quite large compared to the actual differences between the ingots (58.6%). This would indicate a problem with the measurement process, in that the operators appear to be quite different. If desired, an ANOVA table can be requested from the list of tabular options:

<i>Source</i>	<i>Sum of Squares</i>	<i>Df</i>	<i>Mean Square</i>	<i>F-Ratio</i>	<i>P-Value</i>
Operators	19.3034	2	9.65169	4.07	0.0448
Parts	28.4721	12	2.37267	40.36	0.0000
Residual	1.76365	30	0.0587883		
Total	49.5391	44			

Since the P-value for *Operators* is less than 0.05, there are statistically significant differences between operators at the 5% significance level. These differences can be easily seen by selecting the *Box-and-Whisker Plot* from the list of graphs available in the *Gage R&R* procedure:

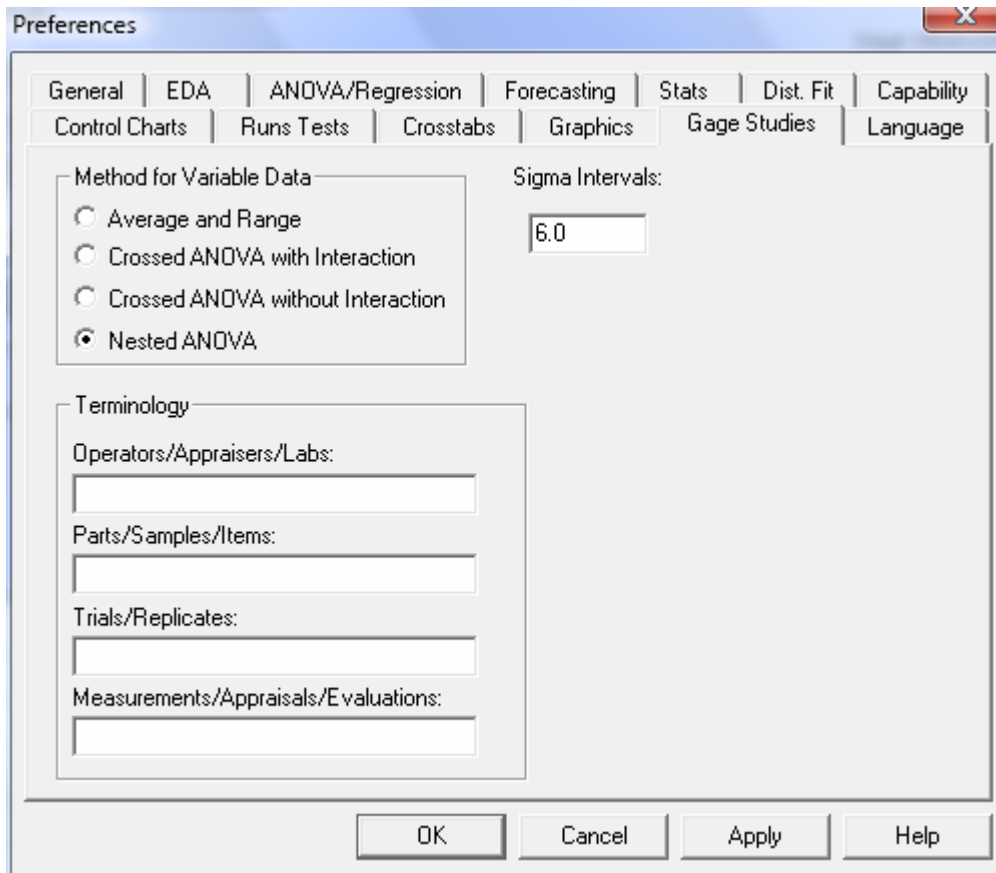


It is evident that Operator 1 tends to obtain higher measurements than the other two operators. Whether these measurements are better or worse would need to be determined from a different type of study.

SnapStat

The same analysis could be performed using the Gage R&R *SnapStat*.

Step 1: First, select *Edit – Preferences* from the main STATGRAPHICS menu. Go to the *Gage Studies* tab and select the *Nested ANOVA*:



Step 2: Next, select *SnapStats!! – Gage Studies* from the main menu and complete the two data input dialog boxes as before. This will create a single page of output as shown below:

SnapStat: Gage R&R - ANOVA Method (Nested)

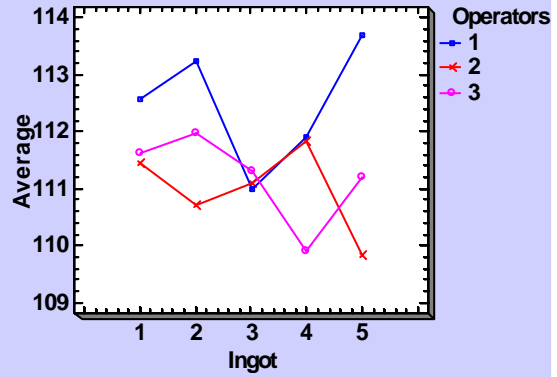
Data variable: Strength

	Est. Sigma	% TV	% Contrib.
Repeat	0.242463	21.141	4.4694
Reprod	0.696612	60.7393	36.8927
R&R	0.737602	64.3133	41.3621
Parts	0.878234	76.5754	58.6379
Total	1.14689	100	

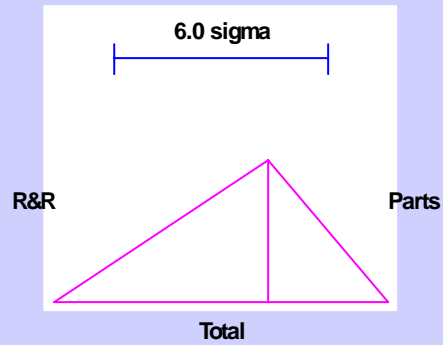
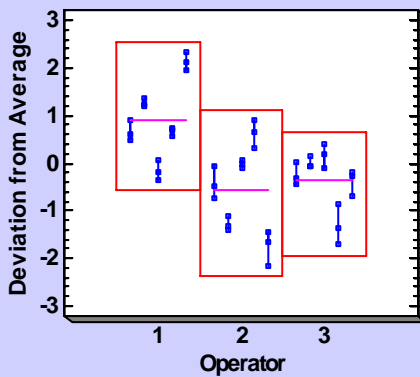
	% R&R	6.0 Sigma	% Tolerance
Repeat	10.81	1.45478	
Reprod	89.19	4.17967	
R&R	100.00	4.42561	
Parts		5.2694	
Total		6.88132	

Number of distinct categories (ndc): 1

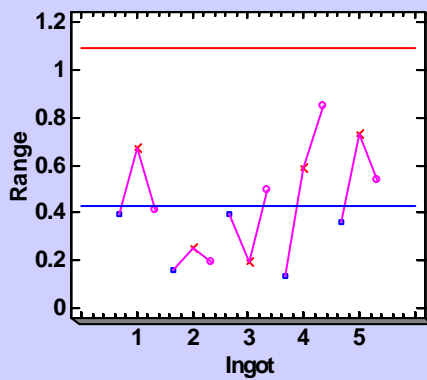
Gage Measurements by Operator



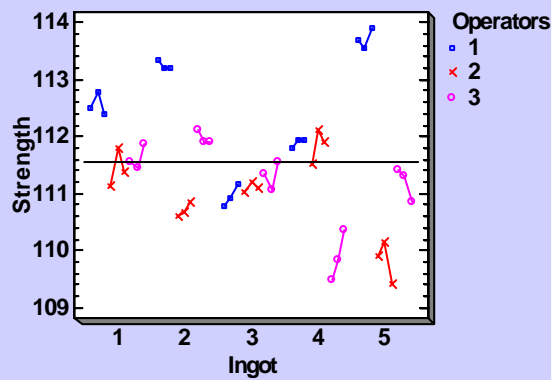
R&R Plot for Strength



Range by Ingots



Run Chart



This page combines the numerical and graphical output in a single display.

Conclusion

When multiple measurements cannot be made on a single item, it is often possible to select items that are similar enough that whatever differences are measured can be assumed to be solely (or at least primarily) measurement error. Since any real differences between homogenous items will tend to inflate the estimated repeatability of the measurement process, the analysis provides a conservative estimate of the ability of the measurement process to distinguish amongst items. Depending on whether operators measure items from within the same homogenous collection or from different collections, either a crossed or nested ANOVA is required to analyze the results.