Gage Repeatability and Reproducibility in Rockwell Hardness Testers

Bill O’Neill*
Instron
Norwood, Mass.

Gage repeatability and reproducibility, commonly known as a GR&R, is a method used in statistical process control (SPC) to measure the precision and variation present in a measuring device and the subsequent effectiveness of the instrument to be used as a measuring tool. Repeatability is related to the ability of the gage itself to give consistent results (precision) under a repeated test of the same manner. Reproducibility is the ability of the gage to provide repeated results regardless of the operator performing the test (variation among operators).

In Rockwell hardness testing, a GR&R study is a valuable measure of how accurate the tester is performing and how much variation the tester may be contributing to the overall process. While GR&R evaluations in Rockwell testing are not achieved without some complications, once these are understood and accounted for, GR&R assessment provides useful process data. Many types of gages, such as micrometers and calipers, can be GR&R performance tested relatively easily and without sample variation (using a gage block of a known value to make repeated measurements on the same block and calculating the repeatability). However, a GR&R evaluation on a Rockwell tester poses inherent difficulties. For example, as a Rockwell test is never performed in the same location, a primary obstacle is the variation in the test sample. As a result, as no material is completely uniform in hardness, a variation is introduced. This phenomenon is typically accounted for through introducing process tolerances to the mathematical analysis when performing the GR&R calculations.

The purpose of performing a GR&R study is to determine how much of the process tolerance is being used up by variation in a hardness testing instrument (known as equipment variation or repeatability), as well as between operators (known as appraiser variation or reproducibility). When the combination of these sources (repeatability and reproducibility, or R&R) becomes a significant portion of the process tolerance, you cannot be sure whether you are measuring the hardness of a part or simply generating random numbers with the Rockwell tester machine. For SPC to work effectively, the combined variation should be less than 10% of the process tolerance (<10% GR&R). Rockwell hardness testing machines with a GR&R between 10 and 30% may be acceptable on an interim basis; machines with a GR&R of greater than 30% should not be used for SPC.

Understanding a Tester’s Repeatability and Reproducibility

Most users do not have quantitative knowledge of how well their testers perform. The repeatability and reproducibility of a tester are frequently determined by performing a GR&R study. By doing GR&R studies periodically, it is easy to establish and monitor the performance of an instrument. A typical GR&R study can quickly establish the short- and long-term performance of a tester including operator influence.

A full, or long-method, GR&R study involves ten different test blocks and three operators, each making three tests on each block. The results of 90 tests will indicate what part of the samples’ tolerance will be used up by the inaccuracy of the tester. A short, or mini, GR&R involves ten different test blocks with three tests on each by a single operator for a total of 30 tests. Comparisons on Rockwell testers between the two methods show negligible differences in the final result.

The Test Process and Calculation

Due to the previously mentioned variation in materials, particularly on production parts, GR&R testing should be performed on standardized test blocks to reduce the material variation influence as much as possible. Test blocks, by design, are manufactured to be as uniform as possible, thereby making them the best material on which to perform the GR&R test. In a full GR&R, each of three operators should perform a single hardness test on each of the ten test block sequentially to complete the first run. To compensate for the random nonuniformity inherent in a test block, the operator must ensure that each set of indents on each block is repeated in as close proximity as possible to each other (grouped or radially). The process is repeated for three full runs.

The GR&R calculation essentially is the
comparison of the combination of machine and operator variation with the process tolerance. If the variation is low or the process tolerance wide by comparison, then the \% GR&R will also be low. Conversely, if the variation is high or the process tolerance narrow by comparison, the \% GR&R will be high. In the analysis, the ranges of readings for each operator are calculated and the average range is generated for each. In addition, the average test value is also determined for each operator. The data are used to generate the full GR&R result.

The process tolerance aspect of the calculation is relatively simple. It is plugged in directly from the engineering specification for the part’s hardness (for example, a required hardness of 42-48 HRC has a tolerance of 6 points). Note that the calculation of GR&R is only relevant in the context of process tolerance. For example, comparing machine and operator variation to a test block tolerance is not meaningful, as it says nothing about the machine’s suitability to measure real parts. Test block tolerances are for ensuring the accuracy of a machine, not its repeatability.

The calculations for variation are somewhat vague; they simply convert average range values and operator differences into an approximation for six sigma (six times the standard deviation for all the data). Six sigma is the statistical description for a machine’s total variation. Figure 1 shows standard deviation differences on a test block between different performance levels. Assuming the machine is varying in a normal manner, six sigma says that over 99% of all tests done on a given block (or set of ten blocks in the case of GR&R) will fall within this region. It is also, in a sense, the uncertainty of the machine at that hardness level, meaning that for a given reading: the actual hardness value could be up to plus or minus three sigma away.

**Making Sense of the Results**

Given that well trained operators and/or automatic machines yield negligible operator (appraiser) variation, it is easiest to understand the final calculation for \% GR&R in the context of only the machine (equipment) variation. This calculation simply divides the six sigma approximation (average range for all tests multiplied by a statistical constant) by the process tolerance and multiplies by 100. A 10% or less GR&R requires that six sigma (total machine variation) be 10% or less of the process tolerance. Figure 2 demonstrates typical GR&R results from various types of Rockwell testers including analog and digital deadweight, as well as closed loop. The analog instruments yield the poorest GR&R performance, while a closed loop tester with highly regulated force application yields the best performance.

Some Rockwell testers have features that contribute to high performing gage repeatability and reproducibility, and some manufacturers ensure that every tester will undergo a GR&R evaluation and will not be shipped unless it has a GR&R performance equal to or better than a stated percentage. Proof of the study is typically provided in a GR&R certificate issued with each tester.

**Contributing factors to high GR&R performance include:**

- Closed-loop load control: load cell provides means of feedback through the load cell and indenter (wear and friction from mechanical parts are compensated for within the loop). Closed loop control provides the most accurate form of force application.

- Attaching the penetrator to the load cell so that errors from friction are eliminated.

- Design where the force measuring device and the depth measuring scale are directly in line with the indenter in a single-axis arrangement.

- Elimination of elevating screws (source of nonrecoverable deflection)

- Using digital readout as opposed to analog, and increasing resolution to .01.

GR&R is a useful overall measure of a Rockwell tester’s performance, but it is not a replacement for the recommended regular block/indirect verification process. Daily indirect performance verification of the testing instrument is important; the scales being used should be verified using standardized test blocks or coupons. If possible, it is recommended that the system is verified with each scale change and at each shift start-up. Blocks should be selected that are in the approximate range of the material being tested and used only on the calibrated side. Five total readings should
be made in the verification process; the measured values must fall within the tolerance stated on block and the block certificate. If verification fails, the machine should be removed from operation until the appropriate adjustments or repairs are made. Maintenance and authorized verification of the instrument is imperative to the continued smooth operation and assurance that the system meets the precision requirements of a Rockwell test. ASTM recommends annual maintenance and verification of a Rockwell tester and more frequent verification under heavy use or extreme conditions. Verification should be performed by an accredited verifying agency, and the report should follow and reference ASTM E18 Rockwell test method.

Summary

Gage repeatability and reproducibly is a useful, informative tool to assess the performance level of a Rockwell hardness tester. Hardness testing is regularly used as an important, revealing process in materials testing, quality control, and acceptance and performance of materials, so the generation of accurate data is critical. Testers verify heat treatment, structural integrity, and component quality to determine if a material has the properties necessary to ensure that the materials used in the items we use every day contribute to a well-engineered, efficient, and safe world. Proper technique, procedure, and standards, in addition to ensuring a well performing instrument, will greatly contribute to the accuracy and usefulness of Rockwell testing.

For more information: Bill O’Neill, Wilson Hardness, an Instron company, 825 University Ave., Norwood, MA 02062; tel: 800/695-4273; e-mail: bill_oneill@instron.com; www.wilsonghardness.com.