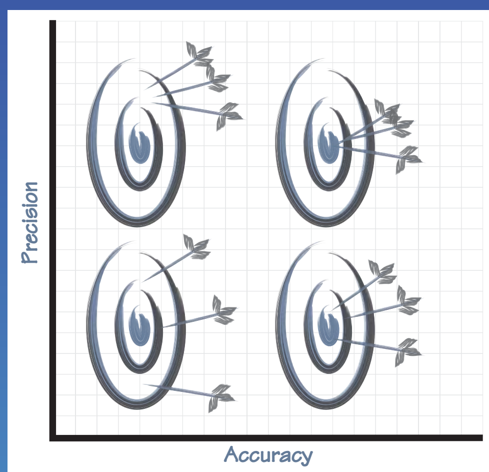
NIST Recommended Practice Guide

Computing Uncertainty for Charpy Impact Machine Test Results



J. D. Splett
H. K. Iyer
C.-M. Wang
C. N. McCowan

NIST

Special Publication 960-18

Special Publication 960-18

NIST Recommended Practice Guide: Computing Uncertainty for Charpy Impact Machine Test Results

J. D. Splett
H. K. Iyer
C.-M. Wang
Statistical Engineering Division
Information Technology Laboratory

C. N. McCowan
Materials Reliability Division
Materials Science and Engineering Laboratory

National Institute of Standards and Technology
Boulder, Colorado 80305

October 2008

U.S. Department of Commerce
Carlos M. Gutierrez, Secretary

National Institute of Standards and Technology
Patrick D. Gallagher, Deputy Director

National Institute of Standards and Technology Special Publication 960-18
Natl. Inst. Stand. Technol. Spec. Publ. 960-18, 34 pages (October 2008)
CODEN: NSPUE2

U. S. GOVERNMENT PRINTING OFFICE
Washington: 2008

For sale by the Superintendent of Documents, U. S. Government Printing Office
Internet: bookstore.gpo.gov – Phone: (202) 512-1800 – Fax: (202) 512-2250
Mail: Stop SSOP, Washington, DC 20402-0001

Contents

NIST Recommended Practice Guide: Computing Uncertainty for Charpy Impact Machine Test Results

J. D. Splett,[1] C. N. McCowan,[2] H. K. Iyer,[1] and C.-M. Wang[1]

National Institute of Standards and Technology
325 Broadway
Boulder, CO 80305

This recommended practice guide demonstrates how to determine the uncertainty associated with mean absorbed energy of specimens tested on a Charpy impact machine. We assume that the Charpy machine has successfully met the requirements for both direct and indirect verification as described in the ASTM E 23, Standard Test Methods for Notched Bar Impact Testing of Metallic Materials. We follow the recommendations and procedures in the "Guide to the Expression of Uncertainty in Measurement" for computing uncertainty. We assume the reader is somewhat familiar with the Charpy machine verification program administered by the National Institute of Standards and Technology.

Keywords: absorbed energy; Charpy V-notch; impact test; pendulum impact test; uncertainty; verification testing

1. Introduction

The absorbed energy of a test material, measured using a Charpy impact machine, is often reported as the mean absorbed energy of a set of specimens. However, the sample mean does not account for known sources of bias, including machine bias, which can be substantial. We address the estimation of a test result for the case in which the test result is corrected for known biases and the case in which it is not. It is left to the user's discretion whether or not to correct a test result.

Computing the reported test result is straightforward; however, computing the uncertainty associated with the test result requires more consideration. The purpose of this document is to clarify the concept of uncertainty and to provide Charpy laboratories with a procedure for computing the uncertainty of a test result.

---

1 Statistical Engineering Division, Information Technology Laboratory
2 Materials Reliability Division, Materials Science and Engineering Laboratory

Before valid Charpy measurements can be made in the laboratory, the machine needs to pass both direct and indirect verification tests as specified in ASTM E 23 [1]. Even if a Charpy machine has passed the indirect verification test, it is likely that results for the verification specimens differ from the certified value. This difference can be used to quantify machine bias. Thus, the indirect verification results and the certified reference value (along with their uncertainties) play key roles in the calculation of uncertainty of a test result.

We express uncertainty according to the accepted criteria described in the "Guide to the Expression of Uncertainty in Measurement," or GUM [2], taking into consideration both random and systematic sources of error. The procedure we recommend for computing uncertainty is very general and can accommodate any number of random or systematic error sources including the following:

Anvils and supports	Material inhomogeneity
Center of percussion	Operator
Center of strike	Potential energy
Friction loss	Repeatability
Height of pendulum fall	Scale accuracy
Impact velocity	Test temperature

The uncertainty contributions from individual error sources can be estimated if they are identified as significant, but generally these errors are assumed to be minimized by adjustments made to the machine during direct verification and by following the standard test procedure. As will become apparent, the calculation of uncertainty is greatly (and often) simplified by assuming that direct verification contributions are zero, and only contributions from indirect verification are considered. This is a widely accepted approach to the calculation of uncertainty for Charpy impact tests, and is used in standards such as ISO 148-1 [3]. We present more detail here, because understanding the individual contributions to uncertainty, and how to quantify them, leads to better control of the test. We encourage the users to consider these, and other relevant details.

We present an example in Section 2 that provides instructions for calculating the uncertainty of a test result. Section 3 provides details regarding the Type B evaluation of errors, Section 4 addresses the computation of machine bias, Section 5 discusses direct verification sources of error, Section 6 addresses temperature measurement errors, Section 7 provides some information about expanded uncertainty, and Section 8 gives some example uncertainty calculations. Complete details regarding the justification of the uncertainty procedures are given in Appendix A.

2. Uncertainty of a Test Result

In this section, we provide details for computing the uncertainty of a test result within the context of an example. A Charpy laboratory will typically compute the sample mean and sample standard deviation of n specimens of the test material using the following two equations:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} , \tag{1}$$

$$s = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{n-1}} . \tag{2}$$

The degrees of freedom (df) associated with the sample standard deviation, s, are $n-1$. It is important to note that s includes *all* sources of random error, including machine variability, material variability, and the typical variations expected when following the standard test procedure. The individual components of the random error cannot be estimated separately in the case of destructive impact testing (multiple measurements on the same specimen are needed to do this). In addition, random errors (unlike systematic errors) do *not* remain constant during the measurement of n specimens, so these errors do not result in a bias.

The data given in Table 1 are used to illustrate the calculations needed to assess the uncertainty associated with the result for our example. Table 1 lists test results and summary statistics for $n = 5$ observations of a particular test material measured at 80 °C.

Suppose we are also given the values in Table 2. (We provide details for computing the quantities in Table 2 in subsequent sections.) Our best estimate of the machine bias is $\hat{b}$, which is defined as the difference between the verification result for the test machine and the certified value of the verification specimens. Systematic errors due to all other factors that are not already included in the machine bias are denoted by $\hat{e}_{\text{systematic}}$. The values $u(\hat{b})$ and $u(\hat{e}_{\text{systematic}})$ are the uncertainties associated with $\hat{b}$ and $\hat{e}_{\text{systematic}}$, while df_b and df_e, represent degrees of freedom for $u(\hat{b})$ and $u(\hat{e}_{\text{systematic}})$.

Table 1. Measurement results for a test material.

Observed data, J	Summary statistics
58.0	$n = 5$
62.0	$\bar{y} = 57.6$ J
54.0	$s = 3.6$ J
54.0	df = 4
60.0	

Table 2. Example quantities required to compute uncertainty of a test result.

	Machine bias	Systematic error
Estimate	$\hat{b} = -4.2\ \text{J}$	$\hat{e}_{\text{systematic}} = 3.0\ \text{J}$
Uncertainty	$u(\hat{b}) = 2.8\ \text{J}$	$u(\hat{e}_{\text{systematic}}) = 0.6\ \text{J}$
Degrees of freedom	$\text{df}_\text{b} = 84$	$\text{df}_\text{e} = 2$

To compute the uncertainty of the test result, we first compute a mean of the test material that is corrected for machine bias and all other systematic effects:

$$\bar{y}_{\text{corrected}} = \bar{y} - \hat{b} - \hat{e}_{\text{systematic}} \ . \tag{3}$$

Substituting the values from Tables 1 and 2 into Eq. (3) gives

$$\bar{y}_{\text{corrected}} = 57.6\ \text{J} - (-4.2\ \text{J}) - (3.0\ \text{J}) = 58.8\ \text{J} \ .$$

Next, we calculate the uncertainty of the corrected mean, $\bar{y}_{\text{corrected}}$. Assuming all of the terms in $\bar{y}_{\text{corrected}}$ are independent, the combined standard uncertainty of $\bar{y}_{\text{corrected}}$ is

$$u(\bar{y}_{\text{corrected}}) = \sqrt{\frac{s^2}{n} + u^2(\hat{b}) + u^2(\hat{e}_{\text{systematic}})} \ . \tag{4}$$

Substituting the appropriate values from Tables 1 and 2 into Eq. (4) gives

$$u(\bar{y}_{\text{corrected}}) = \sqrt{\frac{(3.6\ \text{J})^2}{5} + (2.8\ \text{J})^2 + (0.6\ \text{J})^2} = 3.3\ \text{J} \ .$$

Typically, the standard uncertainty is multiplied by a coverage factor that expands the uncertainty to form an “uncertainty” interval about the measurement result. The interval is expected to encompass a large fraction of possible values of the result. Thus, the expanded uncertainty is defined as the combined standard uncertainty multiplied by a coverage factor. The coverage factor is often set equal to two for simplicity, but this approximation can be problematic, so it is recommended that the degrees of freedom be used to obtain the appropriate coverage factor.

An uncertainty interval with $100(1-\alpha)$ % coverage probability (α is 0.05 for 95 % coverage) is given by

$$\bar{y}_{\text{corrected}} \pm t_{1-\alpha/2,\,\text{df}_{\text{eff}}} \cdot u(\bar{y}_{\text{corrected}}) \;, \qquad (5)$$

where $t_{1-\alpha/2,\,\text{df}_{\text{eff}}}$ is found in a *t*-table (see Appendix C). The degrees of freedom associated with $u(\bar{y}_{\text{corrected}})$,

$$\text{df}_{\text{eff}} = \frac{u^4(\bar{y}_{\text{corrected}})}{\frac{1}{\text{df}}\left(\frac{s^2}{n}\right)^2 + \frac{u^4(\hat{b})}{\text{df}_b} + \frac{u^4(\hat{e}_{\text{systematic}})}{\text{df}_e}} \qquad (6)$$

are determined from the Welch-Satterthwaite approximation as described in the GUM. [2]

Substituting appropriate values from Tables 1 and 2 into Eq. (6) gives

$$\text{df}_{\text{eff}} = \frac{(3.3)^4}{\frac{1}{4}\left(\frac{(3.6)^2}{5}\right)^2 + \frac{(2.8)^4}{84} + \frac{(0.6)^4}{2}} = 47.9 \;,$$

which rounds down to 47. Using a *t*-table we get a coverage factor of $t_{0.975,47} = 2.012$. Thus, a 95 % interval for our example is

$$\begin{gathered} 58.8\ \text{J} \pm t_{0.975,47} \cdot 3.3\ \text{J} \\ 58.8\ \text{J} \pm 2.012 \cdot 3.3\ \text{J} \\ 58.8\ \text{J} \pm 6.6\ \text{J}. \end{gathered}$$

The expanded uncertainty U is 6.6 J, indicating that 95 % of possible measurement results lie within the uncertainty interval (52.2 J, 65.4 J). If we compute the interval based on the *uncorrected* value, we can express our uncertainty interval as

$$\begin{gathered} (\bar{y} \pm U) - (\hat{b} + \hat{e}_{\text{systematic}}) \\ (57.6\ \text{J} \pm 6.6\ \text{J}) - (-4.2\ \text{J} + 3.0\ \text{J}) \\ (51.0\ \text{J},\ 64.2\ \text{J}) + 1.2\ \text{J} \end{gathered}$$

and just report the uncorrected interval (51.0 J, 64.2 J) along with the correction (1.2 J). The decision to report a corrected test result is left to the user. However, if the corrected test result is reported, we recommend that the report clearly state how the correction was computed and include pertinent information such as the magnitude and sign of the correction, the test standard used, and the source of the indirect verification specimens tested.

The remainder of this document is dedicated to providing additional details regarding the computation of individual components needed to compute the uncertainty of a Charpy test result.

3. Type B Uncertainty Evaluation

Typically, direct estimates of systematic errors based on actual measurements are difficult to obtain and even harder to quantify because the required data are not generally available. In such cases, uncertainties due to systematic errors are estimated based on past experience, engineering knowledge, information from published literature, and so on. An uncertainty evaluation that does not involve actual measurements is called a Type B uncertainty evaluation. Type A uncertainty evaluations are based on data obtained under repeatability conditions. Type B uncertainty evaluations can be associated with either random or systematic errors, but are most commonly used with systematic errors.

Type B uncertainty evaluations utilize assumptions regarding distributions of errors. For example, instrument manufacturer's specifications can be thought of as limits to a rectangular distribution. From this, the standard uncertainty associated with measurements by that instrument can be deduced. These types of uncertainties can be highly subjective, but are sometimes useful.

The following example (also shown in Section B.7) illustrates how to use a manufacturer's specification for a Type B uncertainty evaluation. Suppose r is the random error in the Charpy machine scale mechanism and $\pm\Delta r$ represents the manufacturer's specified error bounds of the measurement instrument. Assuming that the error can be anywhere within the $\pm\Delta r$ bounds, a rectangular distribution is used to describe the distribution of possible biases, and in this case bounds are already expressed in the proper units (joules). The standard uncertainty of r is

$$u(r) = \frac{\Delta r}{\sqrt{3}}.$$

A rectangular distribution is often used in the absence of specific information about the error distribution; however, other distributions can be used if more is known about the errors. (See Reference [2] for details regarding Type B uncertainty evaluations.) It is also necessary to provide an estimate of degrees of freedom for each uncertainty component. We will assume $\mathrm{df}_r = \infty$, which implies that we know $u(r)$ exactly. The GUM provides a method for assigning a df value to Type B estimates of uncertainty, which will be demonstrated shortly.

In the previous scale-error example, the distribution of possible errors was defined by the interval $(-r, r)$, which is centered on zero. Sometimes the distribution of a systematic error is centered on a value other than zero, resulting in a nonzero systematic error estimate. For example, an operator might be consistently reading the scale too high, so that the distribution of errors is described by a rectangular distribution defined by (a,b), where a and b are both greater than zero ($0 < a < b$). In this case, the estimated systematic error is $(a+b)/2$ and the associated standard uncertainty is $(b-a)/2\sqrt{3}$.

There are also systematic errors associated with the test procedures that can be approximated using a Type B uncertainty evaluation. Suppose an operator notices that the lengths of fractured specimen halves are uneven and determines that the specimens were all impacted off-center (striker impact is not aligned with the notch). In addition, the operator knows that the 1 mm to 2 mm offsets observed for the broken specimens result in an increase in the absorbed energy between 2 J and 4 J based on extensive experience with this particular material. To estimate the systematic error and its uncertainty, we assume that the 2 J and 4 J limits to error represent bounds of a rectangular distribution so that

$$\hat{e}_{\text{systematic}} = \frac{2\text{ J} + 4\text{ J}}{2} = 3\text{ J} \quad \text{and} \quad u(\hat{e}_{\text{systematic}}) = \frac{4\text{ J} - 2\text{ J}}{2\sqrt{3}} = 0.6\text{ J}.$$

To determine the degrees of freedom associated with $u(\hat{e}_{\text{systematic}})$, we employ a useful relationship from the GUM (Eq. (G.3)). In general,

$$\text{df} = \frac{1}{2}\left[\frac{\Delta u}{u}\right]^{-2},$$

where the quantity in square brackets represents the *relative* uncertainty, or the uncertainty of the uncertainty. In our example, we judge the uncertainty of $u(\hat{e}_{\text{systematic}})$ to be 50 % or 0.50, so that

$$\text{df}_{\text{e}} = \frac{1}{2}[0.50]^{-2} = 2.$$

In general, the degrees of freedom provide information regarding the quality of the uncertainty estimate. For Type A uncertainty evaluations, the degrees of freedom provide an *objective* measure of quality, while degrees of freedom associated with Type B uncertainty evaluations provide a *subjective* measure of quality.

We can also combine several sources of systematic error to determine $\hat{e}_{\text{systematic}}$ and its uncertainty. For example, suppose we would like to combine three independent sources of systematic error: friction loss, potential energy, and impact velocity, so that

$$\hat{e}_{\text{systematic}} = \hat{D} + \hat{E} + \hat{v}.$$

Then the combined standard uncertainty of $\hat{e}_{\text{systematic}}$ is

$$u(\hat{e}_{\text{systematic}}) = \sqrt{u^2(\hat{D}) + u^2(\hat{E}) + u^2(\hat{v})} \ ,$$

with effective degrees of freedom from the Welch-Satterthwaite approximation,

$$\text{df}_{\text{e}} = \frac{u^4(\hat{e}_{\text{systematic}})}{\dfrac{u^4(\hat{D})}{\text{df}_{\text{D}}} + \dfrac{u^4(\hat{E})}{\text{df}_{\text{E}}} + \dfrac{u^4(\hat{v})}{\text{df}_{\text{v}}}} \ .$$

This type of procedure can be applied to any number of independent systematic errors.

4. Machine Bias

To estimate the machine bias, we assume that the machine bias for the material under test is the same as the machine bias based on the indirect verification. This is an important assumption that allows us to estimate machine bias for all test materials. We use the results of an indirect verification test and the associated reference value for our best estimate of machine bias,

$$\hat{b} = \bar{V} - \hat{\delta}_{\text{systematic}} - R \quad , \tag{7}$$

where

$$\bar{V} = \frac{\sum_{i=1}^{n_V} V_i}{n_V} \tag{8}$$

is the sample mean absorbed energy from the indirect verification test based on $n_V = 5$ test results, $\hat{\delta}_{\text{systematic}}$ represents errors due to all systematic effects associated with indirect verification test, and R represents the certified reference value for the batch of verification specimens.

To illustrate the computation of machine bias and its uncertainty, we will return to the example from Section 2. Table 3 lists quantities provided by the National Institute of Standards and Technology (NIST) with the high-energy verification test specimens that were used for the most recent high-energy indirect verification of the Charpy machine of interest.

We use the high-energy indirect verification test results because the nominal value of the absorbed energy of the test material is closest to the high-energy verification material.

The uncertainty associated with the certified verification specimens ($u(R)$) is provided by NIST with the results of the indirect verification test (or by request). Table 4 displays the indirect verification data that were observed when the verification set was broken on the machine of interest.

Table 3. Information provided by NIST for high energy verification specimens.

Reference value, R	109.9 J
Reference value standard uncertainty, $u(R)$	2.6 J
Degrees of freedom, df_R	102

Table 4. High energy indirect verification test results.

Verification set data, J	Summary statistics
108.0	$n_V = 5$
104.0	$\bar{V} = 106.2$ J
109.0	$S_V = 2.3$ J
106.0	$\mathrm{df}_V = 4$
104.0	

Table 5. Systematic error associated with the indirect verification.

Estimate, $\hat{\delta}_{\text{systematic}}$	0.5 J
Standard uncertainty, $u(\hat{\delta})$	0.2 J
Degrees of freedom, df_δ	10

The "V" subscript is used to distinguish the indirect verification results from the test material results. The sample standard deviation associated with the indirect verification specimens (S_V) is calculated as s was calculated previously in Section 2,

$$S_V = \sqrt{\frac{\sum_{i=1}^{n_V}(V_i - \bar{V})^2}{n_V - 1}} \quad . \tag{9}$$

As was the case for s, S_V also includes *all* sources of random error related to both machine variability and material variability, and the individual contribution of errors cannot be determined.

Suppose we are given , $\hat{\delta}_{\text{systematic}}$ its associated uncertainty, and degrees of freedom, as shown in Table 5. We will not elaborate on the origin of the systematic error in Table 5; however, the same general procedures used to estimate $\hat{e}_{\text{systematic}}$, discussed in detail in Section 3, can also be used to estimate $\hat{\delta}_{\text{systemati}}$.
Although $\hat{\delta}_{\text{systematic}}$ = 0.5 J in this illustrative example, typically
is assumed to be zero because errors that are well understood and could be corrected for are minimized during direct verification of the machine. So, neglecting contributions to the bias from $\hat{\delta}_{\text{systematic}}$, the estimated machine bias is calculated as the difference between the mean of the specimen tested in the indirect verification test and the certified value of the specimens tested. For our example, in which $\hat{\delta}_{\text{systematic}}$ is not assumed to be zero, the machine bias is

$$\hat{b} = 106.2 \text{ J} - 0.5 \text{ J} - 109.9 \text{ J} = -4.2 \text{ J} .$$

Assuming independent input quantities, the standard uncertainty of the machine bias is

$$u(\hat{b}) = \sqrt{\frac{S_V^2}{n_V} + u^2(\hat{\delta}_{\text{systematic}}) + u^2(R)} \qquad (10)$$

Substituting the appropriate values from Tables 3 through 5 into Eq. (10) provides the following estimate of the standard uncertainty of the machine bias:

$$u(\hat{b}) = \sqrt{\frac{(2.3 \text{ J})^2}{5} + (0.2 \text{ J})^2 + (2.6 \text{ J})^2} = 2.8 \text{ J} .$$

The degrees of freedom associated with the uncertainty estimate,

$$\mathrm{df_b} = \frac{u^4(\hat{b})}{\frac{1}{\mathrm{df_V}}\left(\frac{S_V^2}{n_V}\right)^2 + \frac{u^4(\hat{\delta}_{\text{systematic}})}{\mathrm{df_\delta}} + \frac{u^4(R)}{\mathrm{df_R}}} \qquad (11)$$

are determined from the Welch-Satterthwaite approximation. In our example, the degrees of freedom are

$$df_b = \frac{(2.8)^4}{\frac{1}{4}\left(\frac{(2.3)^2}{5}\right)^2 + \frac{(0.2)^4}{10} + \frac{(2.6)^4}{102}} = 84.4,$$

which rounds down to 84.

In the examples presented here, the "bias compared to what?" issues are clear. Machines verifying to ASTM E 23 requirements are all compared with a single target for impact energy, defined by ASTM E 23. However, when considering the performance of an ASTM E 23 machine to machines not tested under ASTM E 23 requirements, the comparison is less direct because bias can exist between the various verification systems used around the world (multiple certified values for absorbed energy). We encourage the users to understand this issue, and how it might affect them. Users should also know that the various national measurement institutes distributing impact verification specimens are working to minimize biases among them, and make the quantification of bias for impact testing more transparent to users around the world.

5. Direct Verification

Direct-verification uncertainty sources are related to physical properties of the Charpy machine including: anvil and supports, center of strike, potential energy, impact velocity, center of percussion, friction loss, and scale accuracy. With the possible exception of friction loss, all direct verification sources of uncertainty are Type B evaluations. We provide information regarding calculation of the individual sources of direct verification uncertainty in Appendix B. While it is relatively easy to compute each individual source of uncertainty, it is difficult to quantify the uncertainty components in terms of the effect on Charpy measurements in joules.

The recognized sources of uncertainty for our problem are minimized during the direct verification of an impact machine and by following the standard test procedure. So, it is general practice to estimate the uncertainty of impact tests from the results of indirect verifications and the variations associated with repeat measurements on the material being tested. However, it is also of interest, and part of the exercise in calculating uncertainty, to better understand your machine and process so that it might be better controlled and quantified. It is left up to individual laboratories to identify and include the appropriate uncertainty sources.

Although it is common for laboratories to ignore the uncertainty due to direct-verification bias, it is important to acknowledge the potential for error due to these sources. Thus, it is informative for laboratories to document their reasons for either including or excluding direct verification sources of error. If possible, the uncertainty associated with direct verification should be re-examined each time the machine is verified directly.

6. Temperature

Although systematic error due to temperature probably exists to some extent for all Charpy measurements, it is difficult to quantify the sign (direction) and magnitude of the error. Thus, we typically assume the estimated error is zero, but there is some uncertainty associated with the estimate. This section outlines a procedure that can be used to estimate the uncertainty due to systematic temperature errors.

The uncertainty due to temperature does not depend on machine properties; however, it is highly dependent on the material being tested. For example, steels undergo a transition in fracture behavior from brittle to ductile with increasing temperatures. Supplemental data can be collected for a particular steel of interest, and used to estimate the uncertainty associated with temperature. If later measurements are taken in stable regions defined by the lower shelf or upper shelf (Figure 1), then the uncertainty associated with temperature is probably negligible. However, the uncertainty due to temperature can be significant if measurements are being taken in the transition region of the curve.

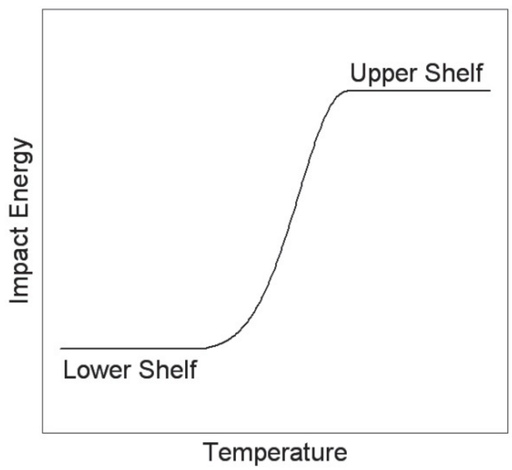


Figure 1. A temperature transition curve.

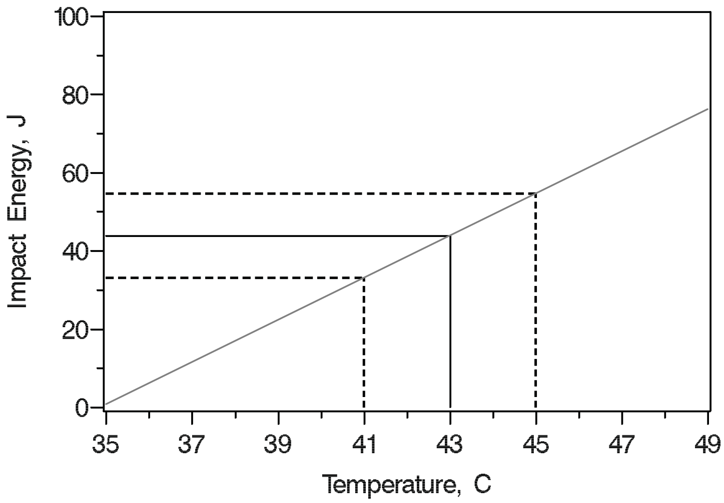


Figure 2. Mapping of temperature error into energy error in the temperature transition region.

Assuming we have data for a particular material that have been collected across a range of temperatures, we can fit a straight line to the data within the temperature transition region (ignoring the shelf data). The information from the regression fit can be used to quantify the effect of the temperature error on impact energy (Figure 2) for future samples of the same material.

For example, suppose we are interested in collecting some new data at 43 °C, but our temperature can be measured only to within ±2 °C. The true temperature could be anywhere in the range of 41 °C to 45 °C. Using the regression equation, we can compute the value of impact energy for both 41 °C and 45 °C, thus providing a range of potential impact energy values ΔE . Assuming the true impact energy has a rectangular distribution within ΔE , we can use the range of impact energy to compute the uncertainty as follows:

$$u(t) = \frac{|\Delta E|}{2\sqrt{3}} \ .$$

The degrees of freedom are $\mathrm{df}_t = n - 2$, where n is the number of observations used in the regression fit.

Optionally, multiple measurements could be made at each temperature (which is how the original curve is obtained) and define the uncertainty as the maximum uncertainty observed in the region. This procedure can also be applied in cases where upper or lower shelf regions have more gradual slopes.

7. Expanded Uncertainty

Sometimes we need to calculate an *expanded* uncertainty, U, which is just the combined standard uncertainty multiplied by a constant, or coverage factor, so that

$$U = k \cdot u(\bar{y}_{\text{corrected}}) .$$

The coverage factor k is determined by looking up the appropriate value in a *t*-table (Appendix C) based on the degrees of freedom associated with $u(\bar{y}_{\text{corrected}})$. The expanded uncertainty associated with a 95 % interval is

$$U_{95} = k_{95} \cdot u(\bar{y}_{\text{corrected}}) = t_{0.975;\text{df}_{\text{eff}}} \cdot u(\bar{y}_{\text{corrected}}) . \qquad (12)$$

The expanded uncertainty is interpreted as an uncertainty interval encompassing a large fraction of possible measurement results.

The degrees of freedom can be difficult to determine if there are many sources of uncertainty within $u(\bar{y}_{\text{corrected}})$. Fortunately, we can compute the *effective degrees of freedom* from the Welch-Satterthwaite approximation [2]

$$\text{df}_{\text{eff}} = \frac{u^4(\bar{y}_{\text{corrected}})}{\frac{1}{\text{df}}\left(\frac{s^2}{n}\right)^2 + \frac{u^4(\hat{b})}{\text{df}_{\text{b}}} + \frac{u^4(\hat{e}_{\text{systematic}})}{\text{df}_{\text{e}}}} , \qquad (13)$$

where $\text{df} = n-1$ and df_{e} are from the Type B uncertainty evaluation (see Section 3). We will also need to calculate df_{b} from

$$\text{df}_{\text{b}} = \frac{u^4(\hat{b})}{\frac{1}{\text{df}_{\text{V}}}\left(\frac{S_V^2}{n_V}\right)^2 + \frac{u^4(\hat{\delta}_{\text{systematic}})}{\text{df}_{\delta}} + \frac{u^4(R)}{\text{df}_{\text{R}}}} , \qquad (14)$$

where $\text{df}_{\text{V}} = n_V - 1$, df_{R} is provided by NIST with the indirect verification specimens, and $\text{df}_{\text{ä}}$ is from Type B uncertainty evaluation (see Section 3).

In general, an uncertainty interval for $\bar{y}_{\text{corrected}}$ is

$$\bar{y}_{\text{corrected}} \pm U_{1-\text{a}}$$

or

$$\bar{y}_{\text{corrected}} \pm t_{1-a/2,\,\text{df}_{\text{eff}}} \cdot u(\bar{y}_{\text{corrected}})\,. \qquad (15)$$

Typically a is 0.05, which corresponds to a 95 % interval. If a Charpy lab does not report results corrected for machine bias and systematic errors, they may want to indicate the magnitude of the estimated biases for informational purposes,

$$(\bar{y} - \hat{b} - \hat{e}_{\text{systematic}}) \pm U_{1-a} \;\text{ or }\; (\bar{y} \pm U_{1-a}) - (\hat{b} + \hat{e}_{\text{systematic}})\,. \qquad (16)$$

Thus, the interval would be shifted by $\hat{b} + \hat{e}_{\text{systematic}}$ if the laboratory wished to report the corrected mean absorbed energy; however, the expanded uncertainty would not be affected by the machine bias and systematic error corrections.

In practice, $k = 2$ is often used to compute the expanded uncertainty to approximate a 95 % interval, and the effective degrees of freedom are never calculated. However, if the effective degrees of freedom are small, then the level of confidence is thought to be less than 95 %.

8. Examples

All examples utilize the data displayed in Tables 1, 3, and 4.

8.1 Both $\hat{e}_{\text{systematic}}$ and $\hat{\delta}_{\text{systematic}}$ and Their Uncertainties Are Negligible

In the general case,

$$\bar{y}_{\text{corrected}} = \bar{y} - \hat{b} - \hat{e}_{\text{systematic}}\;,$$

but if the systematic error associated with the material variation $\hat{e}_{\text{systematic}}$ is assumed to be negligible, then

$$\bar{y}_{\text{corrected}} = \bar{y} - \hat{b}\,.$$

If the systematic error for the test machine variation $\hat{\delta}_{\text{systematic}}$ is assumed to be negligible,

$$\hat{b} = \bar{V} - \hat{\delta}_{\text{systematic}} - R = \bar{V} - R = 106.2\text{ J} - 109.9\text{ J} =$$

and

$$\bar{y}_{\text{corrected}} = \bar{y} - \hat{b} = 57.6 \text{ J} - (-3.7 \text{ J}) = 61.3 \text{ J} .$$

The combined standard uncertainty of $\hat{b}$ is

$$\begin{aligned} u(\hat{b}) &= \sqrt{\frac{S_V^2}{n_V} + u^2(\hat{\delta}_{\text{systematic}}) + u^2(R)} \\ &= \sqrt{\frac{S_V^2}{n_V} + u^2(R)} \\ &= \sqrt{\frac{(2.3 \text{ J})^2}{5} + (2.6 \text{ J})^2} \\ &= 2.8 \text{ J} , \end{aligned}$$

with effective degrees of freedom

$$\text{df}_\text{b} = \frac{u^4(\hat{b})}{\frac{1}{\text{df}_\text{V}}\left(\frac{S_V^2}{n_V}\right)^2 + \frac{u^4(R)}{\text{df}_\text{R}}} = \frac{(2.8)^4}{\frac{1}{4}\left(\frac{(2.3)^2}{5}\right)^2 + \frac{(2.6)^4}{102}} = 84.5 ,$$

which rounds down to 84. Thus, the uncertainty of the corrected mean value is

$$\begin{aligned} u(\bar{y}_{\text{corrected}}) &= \sqrt{\frac{s^2}{n} + u^2(\hat{b}) + u^2(\hat{e}_{\text{systematic}})} \\ &= \sqrt{\frac{s^2}{n} + u^2(\hat{b})} \\ &= \sqrt{\frac{(3.6 \text{ J})^2}{5} + (2.8 \text{ J})^2} \\ &= 3.2 \text{ J} , \end{aligned}$$

with effective degrees of freedom

$$\text{df}_{\text{eff}} = \frac{u^4(\bar{y}_{\text{corrected}})}{\frac{1}{\text{df}}\left(\frac{s^2}{n}\right)^2 + \frac{u^4(\hat{b})}{\text{df}_\text{b}}} = \frac{(3.2)^4}{\frac{1}{4}\left(\frac{(3.6)^2}{5}\right)^2 + \frac{(2.8)^4}{84}} = 43.5 ,$$

which rounds down to 43. A 95 % interval for $\overline{y}_{\text{corrected}}$ is

$$\begin{gathered}\overline{y}_{\text{corrected}} \pm t_{1-\alpha/2,\,\text{df}_{\text{eff}}} \cdot u(\overline{y}_{\text{corrected}}) \\ 61.3\text{ J} \pm t_{0.975,43} \cdot 3.2\text{ J} \\ 61.3\text{ J} \pm 2.017 \cdot 3.2\text{ J} \\ 61.3\text{ J} \pm 6.5\text{ J}.\end{gathered}$$

The expanded uncertainty, associated with a 95 % level of confidence is 6.5 J. The 95 % uncertainty interval is (54.8 J, 67.8 J).

If the value reported is not corrected for machine bias, we can express our interval as

$$\begin{gathered}(\overline{y} \pm U) - \hat{b} \\ (57.6\text{ J} \pm 6.5\text{ J}) - (-3.7\text{ J}) \\ (51.1\text{ J, } 64.1\text{ J}) + 3.7\text{ J}.\end{gathered}$$

The Charpy laboratory may or may not wish to disclose the estimated machine bias, however the information is available if needed. Notice that the interval for the uncorrected parameter is shifted just by the amount of the correction and the expanded uncertainty is the same regardless of whether or not the reported value is corrected.

8.2 $\hat{e}_{\text{systematic}}$ Has One Component

Suppose $\hat{e}_{\text{systematic}}$ contains the error due to temperature so that $\hat{e}_{\text{systematic}} = \hat{t}$. The temperature error is systematic because it is likely to be in the same direction (always warmer or always cooler than the target temperature) for a single set of measurements. However, we do not typically estimate the temperature error, so we will we assume the value of $\hat{e}_{\text{systematic}}$ is zero. The uncertainty associated with $\hat{e}_{\text{systematic}}$ is

$$u(\hat{e}_{\text{systematic}}) = u(\hat{t}) .$$

The procedure outlined in Section 5 will be used to determine the uncertainty due to temperature $u(\hat{t})$. Figure 3 displays temperature data for the material of interest along with the regression line fit to the data in the transition region (ignoring the data on the “shelves”).

Suppose our test specimens from Table 1 were broken using a temperature of 80 °C, which is within the temperature transition region. A regression line was fit to the 21 data points in the transition region, resulting in the following equation:

$$\mathrm{E(J)} = -0.03973(\mathrm{J}) + 0.74084(\mathrm{J/°C}) \cdot \mathrm{T(°C)}\ .$$

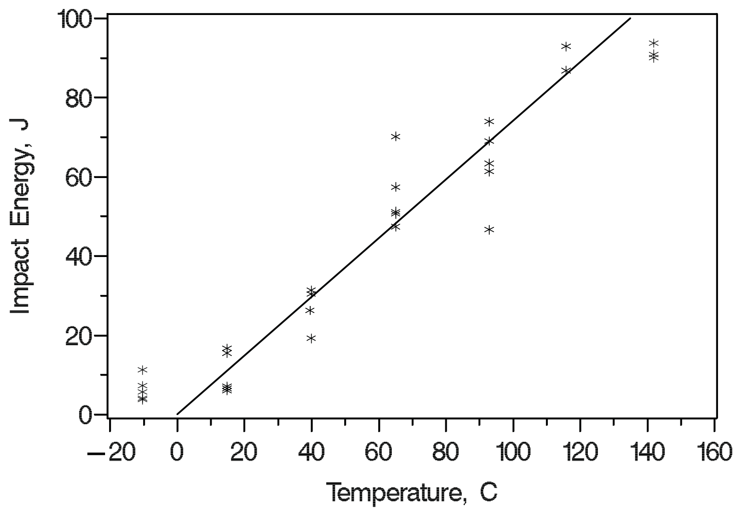


Figure 3. Temperature data for the test material. The straight line in the plot represents a regression fit to the data in the transition region only. We ignore data on the "shelves" at -10 °C and 141 °C.

If our temperature can be measured to within ±1 °C, then the true temperature is between 79 °C and 81 °C, and the energy range is defined by the following:

$$\mathrm{E(J)} = -0.03973(\mathrm{J}) + 0.74084(\mathrm{J/°C}) \cdot 79(\mathrm{°C}) = 58.5\ \mathrm{J}$$
$$\mathrm{E(J)} = -0.03973(\mathrm{J}) + 0.74084(\mathrm{J/°C}) \cdot 81(\mathrm{°C}) = 60.0\ \mathrm{J}$$
$$\Delta E = 60.0\ \mathrm{J} - 58.5\ \mathrm{J} = 1.5\ \mathrm{J}\ .$$

Next, the energy range is converted to a standard uncertainty based on a rectangular distribution,

$$u(\hat{t}) = \frac{1.5\ \mathrm{J}}{2\sqrt{3}} = 0.4\ \ \mathrm{J}\ ,$$

with degrees of freedom $\mathrm{df}_t = n - 2 = 21 - 2 = 19$. Thus the uncertainty of $\hat{e}_{\mathrm{systematic}}$ is

$$u(\hat{e}_{\mathrm{systematic}}) = 0.4\ \mathrm{J}\ ,$$

with 19 degrees of freedom.

The mean absorbed energy corrected for machine bias and other systematic effects is

$$\bar{y}_{\text{corrected}} = \bar{y} - \hat{b} - \hat{e}_{\text{systematic}} = 57.6\ \text{J} - (-3.7\ \text{J}) - 0\ \text{J} = 61.3\ \text{J}\ ,$$

where $\hat{b}$, its uncertainty, and degrees of freedom have not changed from example 7.1. The combined standard uncertainty is

$$u(\bar{y}_{\text{corrected}}) = \sqrt{\frac{s^2}{n} + u^2(\hat{b}) + u^2(\hat{e}_{\text{systematic}})} = \sqrt{\frac{(3.6\ \text{J})^2}{5} + (2.8\ \text{J})^2 + (0.4\ \text{J})^2} = 3.3\ \text{J}\ ,$$

with degrees of freedom

$$\text{df}_{\text{eff}} = \frac{u^4(\bar{y}_{\text{corrected}})}{\frac{1}{\text{df}}\left(\frac{s^2}{n}\right)^2 + \frac{u^4(\hat{b})}{\text{df}_{\text{b}}} + \frac{u^4(\hat{e}_{\text{systematic}})}{\text{df}_{\text{e}}}} = \frac{(3.3)^4}{\frac{1}{4}\left(\frac{(3.6)^2}{5}\right)^2 + \frac{(2.8)^4}{84} + \frac{(0.4)^4}{19}} = 49.2\ ,$$

which rounds down to 49. A 95 % uncertainty interval for $\bar{y}_{\text{corrected}}$ is

$$\begin{gathered} \bar{y}_{\text{corrected}} \pm t_{1-a/2,\,\text{df}_{\text{eff}}} \cdot u(\bar{y}_{\text{corrected}}) \\ 61.3\ \text{J} \pm t_{0.975,49} \cdot 3.3\ \text{J} \\ 61.3\ \text{J} \pm 2.010 \cdot 3.3\ \text{J} \\ 61.3\ \text{J} \pm 6.6\ \text{J}. \end{gathered}$$

The expanded uncertainty is 6.6 J, and the uncertainty interval encompassing 95 % of possible measurement results is (54.7 J, 67.9 J). If the value reported is not corrected for bias, we can express our interval as

$$\begin{gathered} (\bar{y} \pm U) - (\hat{b} + \hat{e}_{\text{systematic}}) \\ (57.6\ \text{J} \pm 6.6\ \text{J}) - (-3.7\ \text{J} + 0\ \text{J}) \\ (51.0\ \text{J},\ 64.2\ \text{J}) + 3.7\ \text{J}. \end{gathered}$$

9. Closing Remarks

We have developed a procedure for estimating the uncertainty associated with a reported mean absorbed energy from a Charpy test. The procedure is flexible enough to account for several systematic error sources, if necessary, and allows the user the choice of correcting the reported mean or not. The uncertainty procedure in this document applies to measurements completed in a Charpy laboratory.

Occasionally, there is some confusion about the NIST reference value, the reference value uncertainty, and Charpy verification limits with respect to results obtained in a

Charpy laboratory. The reference value is the measured mean absorbed energy of a batch of reference specimens. The reference value uncertainty describes the variability of the reference value and includes material, system, and machine variability. The reference value uncertainty does not describe the variability of a single verification specimen or the variability in the verification specimens (specimen variation cannot be estimated separately from machine variation). In the Charpy laboratory, the reference value and its uncertainty are used only to estimate the bias of a Charpy machine and the uncertainty of the bias; they provide no information regarding Charpy measurements for other materials. It is also important to remember that the reference value uncertainty is associated with a specific measurement result, while the verification limits describe the acceptable variation among means for a test method. These two items are not necessarily related.

10. References

[1] ASTM E 23-02a, "Standard Test Methods for Notched Bar Impact Testing of Metallic Materials," ASTM International, West Conshohocken, PA.

[2] International Organization for Standardization, *Guide to the Expression of Uncertainty in Measurement*, International Organization for Standardization, Geneva Switzerland, 1993 (corrected and reprinted 1995).

[3] ISO 148-1, "Metallic materials – Charpy pendulum impact, Part 1: Test method", ISO, Geneva, Switzerland.

[4] Yamaguchi, Y., Takagi, S., and Nakano, H., "Effects of Anvil Configurations on Absorbed Energy," Pendulum Impact Testing: A Century of Progress, STP 1380, T. A. Siewert and M. P. Manahan, Sr., Eds., American Society for Testing and Materials, West Conshohocken, PA, 2000.

[5] *NIST/SEMATECH e-Handbook of Statistical Methods,* http://www.itl.nist.gov/div898/handbook/, July 18, 2006.

Appendix A. Uncertainty Details

The following information is provided for completeness and to document the justification for the recommended uncertainty procedures.

A.1 Test Material

We define a single measurement for a test material measured in a Charpy laboratory as

$$y_i = \mu_Y + b_Y + e_{i,\text{inhomogeneity}} + e_{i,\text{repeatability}} + e_{\text{i, other random}} + e_{\text{systematic}}$$

for $i = 1, 2, \ldots, n$ measurements. Terms on the right side of the equation having the "*i*" subscript denote random errors that change from measurement to measurement.

represents the true mean breaking energy of the test material *if the material could have been tested on the three NIST reference machines.*

b_Y represents the true machine bias for the test material. This term includes all machine differences that are constant for the duration of the set of n measurements.

$e_{i,\text{inhomogeneity}}$ represents the material inhomogeneity.

$e_{i,\text{repeatability}}$ represents the machine repeatability.

$e_{i,\text{other random}}$ represents all other sources of error due to random effects.

$e_{\text{systematic}}$ represents errors due to all other systematic effects that are not already included in the machine bias (for example, operator error). Systematic errors remain constant for the duration of the set of n measurements.

The mean of n measurements of the test material is

$$\bar{y} = \mu_Y + b_Y + \bar{e}_{\text{inhomogeneity}} + \bar{e}_{\text{repeatability}} + \bar{e}_{\text{other random}} + e_{\text{systematic}} ,$$

and the true variance of $\bar{y}$ is

$$y_i = \mu_Y + b_Y + e_{i,\text{inhomogeneity}} + e_{i,\text{repeatability}} + e_{\text{i, other random}} + e_{\text{systematic}}$$

$$\text{var}(\overline{Y}) = \frac{\sigma^2_{\text{inhomogeneity}}}{n} + \frac{\sigma^2_{\text{repeatability}}}{n} + \frac{\sigma^2_{\text{other random}}}{n} \quad ,$$

which is estimated by $\frac{S^2}{n}$, with $\text{df} = n-1$ degrees of freedom. The three random errors *cannot* be estimated separately. The corrected value is

$$\overline{y}_{\text{corrected}} = \overline{y} - \hat{b}_Y - \hat{e}_{\text{systematic}} \; .$$

The indirect verification results will be used to estimate $\hat{b}_Y$ and its uncertainty, and we will assume that $\hat{e}_{\text{systematic}}$ is zero. There is uncertainty associated with each of the estimated systematic errors. The combined standard uncertainty of corrected value is

$$u(\overline{y}_{\text{corrected}}) = \sqrt{\frac{s^2}{n} + u^2(\hat{b}_Y) + u^2(\hat{e}_{\text{systematic}})} \; .$$

The effective degrees of freedom based on the Welch-Satterthwaite approximation are

$$\text{df}_{\text{eff}} = \frac{u^4(\overline{y}_{\text{corrected}})}{\frac{1}{\text{df}}\left(\frac{s^2}{n}\right)^2 + \frac{u^4(\hat{b}_Y)}{\text{df}_{\text{b}}} + \frac{u^4(\hat{e}_{\text{systematic}})}{\text{df}_{\text{e}}}} \quad .$$

The expanded uncertainty associated with $\overline{y}_{\text{corrected}}$ is

$$U = t_{1-a/2,\,\text{df}_{\text{eff}}} \cdot u(\overline{y}_{\text{corrected}}) \quad .$$

The corrected value reported by the Charpy laboratory has the form $\overline{y}_{\text{corrected}} \pm U$. If a Charpy laboratory does not report results corrected for machine bias, they may want to indicate the magnitude of the estimated bias for informational purposes as

$$(\overline{y} - \hat{b}_Y - \hat{e}_{\text{systematic}}) \pm U \;\text{ or }\; (\overline{y} \pm U) - (\hat{b}_Y + \hat{e}_{\text{systematic}}) \; .$$

A.2 Indirect Verification Test

The Charpy laboratory's indirect verification test will be used to estimate machine bias in conjunction with the associated NIST reference value. A single measurement in the indirect verification test is defined as

$$V_i = \mu_Z + b_V + \delta_{i,\text{inhomogeneity}} + \delta_{i,\text{repeatability}} + \delta_{\text{i, other random}} + \delta_{\text{systematic}} \quad ,$$

where $i = 1, 2, \ldots, n_V$ measurements (n_V is usually five). The "*i*" subscript denotes errors that change from measurement to measurement.

μ_Y represents the true mean breaking energy of the reference material if the material could have been tested on the NIST reference machines.

b_V represents the machine bias for the reference material. This term includes all machine differences that are constant for the duration of the set of n_V measurements.

$e_{i,\text{inhomogeneity}}$ represents the reference material inhomogeneity.

$e_{i,\text{repeatability}}$ represents the machine repeatability.

$e_{i,\text{other random}}$ represents all other sources of error due to random effects.

$e_{\text{systematic}}$ represents errors due to all other systematic effects that are not already included in the machine bias. Systematic errors remain constant for the duration of the set of n_V measurements.

The mean of n_V measurements is

$$\overline{V} = \mu_Z + b_V + \overline{\delta}_{\text{inhomogeneity}} + \overline{\delta}_{\text{repeatability}} + \overline{\delta}_{\text{other random}} + \delta_{\text{systematic}} \quad ,$$

and the variance of $\overline{V}$,

$$\text{var}(\overline{V}) = \frac{\sigma^2_{\text{inhomogeneity}}}{n_V} + \frac{\sigma^2_{\text{repeatability}}}{n_V} + \frac{\sigma^2_{\text{other random}}}{n_V} \quad ,$$

is estimated by $\frac{S_V^2}{n_V}$, with $\text{df}_\text{V} = n_V - 1$ degrees of freedom. The three random errors *cannot* be estimated separately.

A.3 NIST Reference Value

The NIST reference value will be used to estimate machine bias in conjunction with the customer's associated verification test.

According to ASTM E 23-06, the reference value of Charpy indirect verification specimens is established using three master machines maintained by NIST.

In the NIST Charpy verification program, the reference value and its associated uncertainty are based on two sets of measurements. The first set of measurements involves breaking 75 verification specimens (25 on each master machine) from a "pilot" lot to determine if the material meets the rigid specifications of the verification program. If the material is acceptable, the remaining verification specimens in the lot are machined and a second set of measurements are performed from the full "production" (25 on each master machine). Assuming the production lot has not changed significantly from the original pilot lot, the material is sold to the public in sets of five specimens as a Standard Reference Material. The reference value R is established using the 75 verification lot and 75 production lot specimens.

We make the following assumptions when determining the reference value and its uncertainty.

1. The reference value is defined to be the "truth," so there is no bias associated with the reference value.
2. There is no difference between pilot lot specimens and production lot specimens. (Differences are evaluated using a t-test for means and an F-test for variances.) In the event that the verification lot and production lot have significantly different means and/or variances, the reference value will be based solely on the production lot data.

A.3.1 Reference Machine

We define a single measurement taken on a NIST reference machine as

$$Z_{1k} = \mu_1 + \gamma_{k,\ \text{inhomogeneity}} + \gamma_{k,\ \text{repeatability}} + \gamma_{k,\ \text{other random}} + \gamma_{\text{systematic}} \quad ,$$

where $k = 1, 2, \ldots, n_1$ measurements (n_1 is usually 50). The "k" subscripts on the right hand side of the equation denote errors that change from measurement to measurement.

μ_1 represents the true mean breaking energy of the reference material as measured by the NIST reference machine.

$\gamma_{k,\text{inhomogeneity}}$ represents the reference material inhomogeneity.

$\gamma_{k,\text{repeatability}}$ represents the machine repeatability.

$\gamma_{k,\text{other random}}$ represents all other sources of errors due to random effects.

$\gamma_{\text{systematic}}$ represents the errors due to all systematic effects. Systematic errors remain constant for the duration of the set of n_1 measurements. Although we assume $\gamma_{\text{systematic}}$ is zero, it does have some uncertainty.

The mean of n_1 measurements taken on a NIST reference machine is

$$\bar{Z}_1 = \mu_1 + \bar{\gamma}_{\text{inhomogeneity}} + \bar{\gamma}_{\text{repeatability}} + \bar{\gamma}_{\text{other random}} + \gamma_{\text{systematic}} \quad ,$$

and the variance associated with the mean,

$$\text{var}(\bar{Z}_1) = \frac{\sigma^2_{\text{inhomogeneity}}}{n_1} + \frac{\sigma^2_{\text{repeatability}}}{n_1} + \frac{\sigma^2_{\text{other random}}}{n_1} \quad ,$$

is estimated by $\frac{S_1^2}{n_1}$, with $\text{df}_1 = n_1 - 1$ degrees of freedom. The three random errors *cannot* be estimated separately.

The corrected value for the NIST reference machine is

$$\bar{Z}_{1,\text{ corrected}} = \bar{Z}_1 - \hat{\gamma}_{\text{systematic}} \quad .$$

The combined standard uncertainty of the corrected value is

$$u(\bar{Z}_{1,\text{ corrected}}) = \sqrt{\frac{S_1^2}{n_1} + u^2(\hat{\gamma}_{\text{systematic}})} \quad ,$$

which has effective degrees of freedom

$$\text{df}_{Z_1} = \frac{u^4(\bar{Z}_{1,\text{ corrected}})}{\frac{1}{\text{df}_1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{u^4(\hat{\gamma}_{\text{systematic}})}{\text{df}_\gamma}} \quad ,$$

based on the Welch-Satterthwaite approximation.

The procedure for computing $\bar{Z}_{1,\text{ corrected}}$, $u(\bar{Z}_{1,\text{ corrected}})$, and df_{Z_1} for one reference machine also applies to the remaining two NIST reference machines so that we obtain

$\bar{Z}_{2,\ \text{corrected}}$, $u(\bar{Z}_{2,\ \text{corrected}})$, and df_{Z_2} for the second reference machine, and $\bar{Z}_{3,\ \text{corrected}}$, $u(\bar{Z}_{3,\ \text{corrected}})$, and df_{Z_3} for the third reference machine. The results from all three reference machines are needed to compute the NIST reference value, as we discuss below.

A.3.2 NIST Reference Value

The NIST reference value based on data observed for the three reference machines is defined as

$$\mu_Z = \frac{\mu_1 + \mu_2 + \mu_3}{3} \qquad ,$$

where μ_1, μ_2, and μ_3 denote the respective true mean breaking energies for each of the three reference machines. The NIST reference value defines the true breaking energy of the material.

We estimate the NIST reference value using

$$R = \frac{\bar{Z}_{1,\text{corrected}} + \bar{Z}_{2,\ \text{corrected}} + \bar{Z}_{3,\ \text{corrected}}}{3} ,$$

which has combined standard uncertainty

$$u(R) = \sqrt{\frac{1}{9}\left[u^2(\bar{Z}_{1,\ \text{corrected}}) + u^2(\bar{Z}_{2,\ \text{corrected}}) + u^2(\bar{Z}_{3,\text{corrected}})\right]} ,$$

and effective degrees of freedom

$$\text{df}_R = \frac{u^4(R)}{\dfrac{(\frac{1}{3})^4 u^4(\bar{Z}_{1,\text{corrected}})}{\text{df}_{Z_1}} + \dfrac{(\frac{1}{3})^4 u^4(\bar{Z}_{2,\ \text{corrected}})}{\text{df}_{Z_2}} + \dfrac{(\frac{1}{3})^4 u^4(\bar{Z}_{3,\text{corrected}})}{\text{df}_{Z_3}}} \qquad ,$$

based on the Welch-Satterthwaite approximation. The reference value expanded uncertainty is

$$U_R = t_{1-\alpha/2,\ \text{df}_R} \cdot u(R) \qquad .$$

A.4 Estimating Machine Bias

Assume the machine bias is the same for the new material (b_Y) and the reference material

(b_V), so that $b_Y = b_V = b$. The best estimate of the machine bias b is

$$\hat{b} = \overline{V} - \hat{\delta}_{\text{systematic}} - R \quad ,$$

which has combined standard uncertainty

$$u(\hat{b}) = \sqrt{\frac{S_V^2}{n_V} + u^2(\hat{\delta}_{\text{systematic}}) + u^2(R)} \quad .$$

The effective degrees of freedom based on the Welch-Satterthwaite approximation are

$$\mathrm{df}_\mathrm{b} = \frac{u^4(\hat{b})}{\frac{1}{\mathrm{df}_\mathrm{V}}\left(\frac{S_V^2}{n_V}\right)^2 + \frac{u^4(\hat{\delta}_{\text{systematic}})}{\mathrm{df}_\delta} + \frac{u^4(R)}{\mathrm{df}_\mathrm{R}}} \quad .$$

Appendix B. Direct Verification Components of Uncertainty

B.1 Anvils and Supports, *A*

A paper by Yamaguchi, Takagi, and Nakano [4] provides some information regarding the uncertainty associated with anvil configurations. Assuming that other Charpy machines behave similarly to the machine tested in the paper, we can use the uncertainties listed in the paper (Table 9) as ballpark estimates. Table B.1 lists the uncertainty estimates for low, high, and super-high energies.

Table B.1. Estimated uncertainties due to the anvil and support bias.

Standard uncertainty	Low energy	High energy	Super-high energy
$u(A)$	0.05 J	0.29 J	0.77 J

Since degrees of freedom are not provided in the paper, we will also assume that $\mathrm{df}_\mathrm{A} = \infty$, which implies that we know $u(A)$ exactly.

B.2 Height of Pendulum Fall, *h*

The height of the pendulum fall is

$$h = S \cdot (1 - \cos\beta) \quad ,$$

where S is the measured length of the pendulum, and β is the measured fall angle. Letting ΔS and $\Delta\beta$ denote the manufacturer's stated error bounds, respectively, and assuming a rectangular distribution bounded by $\pm\Delta S$ and $\pm\Delta\beta$, the uncertainties for S and β are

$$u(S) = \frac{\Delta S}{\sqrt{3}} \quad \text{and} \quad u(\beta) = \frac{\Delta\beta}{\sqrt{3}} \quad .$$

Then the uncertainty of h is

$$\begin{aligned} u^2(h) &= \left(\frac{\partial h}{\partial S}\right)^2 u^2(S) + \left(\frac{\partial h}{\partial \beta}\right)^2 u^2(\beta) + 2\left(\frac{\partial h}{\partial S}\right)\left(\frac{\partial h}{\partial \beta}\right) u(S,\beta), \\ &= c_S^2\, u^2(S) + c_\beta^2\, u^2(\beta) + 2 c_S c_\beta\, u(S,\beta), \end{aligned}$$

where

$$c_S = \frac{\partial h}{\partial S} = 1 - \cos\beta \quad \text{and} \quad c_\beta = \frac{\partial h}{\partial \beta} = S \cdot (\sin\beta).$$

If (S, β) are independent, then only the first two terms are needed to determine the uncertainty. We can assume that $\mathrm{df}_S = \infty$ and $\mathrm{df}_\beta = \infty$, which implies that we know $u(S)$ and $u(\beta)$ exactly. (See the ISO-GUM, G.4.3 [2] for details.) The effective degrees of freedom associated with $u(h)$ are

$$\mathrm{df}_h = \frac{u^4(h)}{\dfrac{c_S^4\, u^4(S)}{\mathrm{df}_S} + \dfrac{c_\beta^4\, u^4(\beta)}{\mathrm{df}_\beta}},$$

based on the Welch-Satterthwaite approximation.

B.3 Potential Energy, *E*

The potential energy is

$$E = h \cdot F \ ,$$

where F is the measured supporting force exerted by the pendulum in horizontal position, and h is the height of the pendulum fall defined in Section B.2. Letting ΔF denote the manufacturer's stated error bound of the measurement instrument, and assuming a rectangular distribution bounded by $\pm\Delta F$, the uncertainty of F is

$$u(F) = \frac{\Delta F}{\sqrt{3}} .$$

Then the uncertainty of E is

$$\begin{aligned} u^2(E) &= \left(\frac{\partial E}{\partial h}\right)^2 u^2(h) + \left(\frac{\partial E}{\partial F}\right)^2 u^2(F) + 2\left(\frac{\partial E}{\partial h}\right)\left(\frac{\partial E}{\partial F}\right) u(h,F) \\ &= c_h^2\, u^2(h) + c_F^2\, u^2(F) + 2 c_h c_F u(h,F), \end{aligned}$$

where

$$c_h = \frac{\partial E}{\partial h} = F \quad \text{and} \quad c_F = \frac{\partial E}{\partial F} = h .$$

The uncertainty associated with h is defined in Section B.2. If (h, F) are independent, then only the first two terms are needed for the uncertainty. We can assume $\mathrm{df_F} = \infty$, which implies that we know $u(F)$ exactly. The effective degrees of freedom associated with $u(E)$ are

$$\mathrm{df_E} = \frac{u^4(E)}{\dfrac{c_h^4\, u^4(h)}{\mathrm{df_h}} + \dfrac{c_F^4\, u^4(F)}{\mathrm{df_F}}} ,$$

based on the Welch-Satterthwaite approximation.

B.4 Impact Velocity, v

The impact velocity is v

$$v = \sqrt{2 \cdot g \cdot h} \qquad ,$$

where g is the local acceleration of gravity, and h is the height of the pendulum fall defined in Section B.2. Letting Δg denote the manufacturer's stated error bound of the measurement instrument (0.001 m/s^2 according to ASTM E 23), and assuming a rectangular distribution bounded by $\pm\Delta g$, the uncertainty for g is

$$u(g) = \frac{\Delta g}{\sqrt{3}}.$$

The uncertainty of v is

$$\begin{aligned} u^2(v) &= \left(\frac{\partial v}{\partial g}\right)^2 u^2(g) + \left(\frac{\partial v}{\partial h}\right)^2 u^2(h) \\ &= c_g^2\, u^2(g) + c_h^2\, u^2(h), \end{aligned}$$

where

$$c_g = \frac{\partial v}{\partial g} = \frac{h\sqrt{2}}{2\sqrt{g \cdot h}} \quad \text{and} \quad c_h = \frac{\partial v}{\partial h} = \frac{g\sqrt{2}}{2\sqrt{g \cdot h}}.$$

The uncertainty associated with h is defined in Section B.2. We can assume $\mathrm{df_g} = \infty$ which implies that we know $u(g)$ exactly. The effective degrees of freedom associated with $u(v)$ are

$$\mathrm{df_v} = \frac{u^4(v)}{\dfrac{c_g^4\, u^4(g)}{\mathrm{df_g}} + \dfrac{c_h^4\, u^4(h)}{\mathrm{df_h}}},$$

based on the Welch-Satterthwaite approximation.

B.5 Center of Percussion, *L*

The center of percussion is

$$L = \frac{g \cdot p^2}{4\pi^2},$$

where g is the local acceleration of gravity defined in Section B.4, and p is the mean period of the swing of the pendulum from three measurements for 100 swings. (There may be some systematic error associated with p that should taken into account.) The standard deviation of three p measurements is s_p, so the uncertainty of the mean period is

$$u(p) = \frac{s_p}{\sqrt{3}},$$

with $\mathrm{df_p} = 3 - 1 = 2$ degrees of freedom.

The uncertainty of L is

$$u^2(L) = \left(\frac{\partial L}{\partial g}\right)^2 u^2(g) + \left(\frac{\partial L}{\partial p}\right)^2 u^2(p),$$
$$= c_g^2\, u^2(g) + c_p^2\, u^2(p),$$

where

$$c_g = \frac{\partial L}{\partial g} = \frac{p^2}{4\pi^2} \quad \text{and} \quad c_p = \frac{\partial L}{\partial p} = \frac{g \cdot p}{2\pi^2}.$$

The uncertainty associated with g is defined in Section B.4. From the Welch-Satterthwaite approximation, the effective degrees of freedom associated with $u(L)$ are

$$\mathrm{df_L} = \frac{u^4(L)}{\dfrac{c_g^4\, u^4(g)}{\mathrm{df_g}} + \dfrac{c_p^4\, u^4(p)}{\mathrm{df_p}}} .$$

B.6 Friction Loss, *D*

The friction loss is

$$D = E_0 - E_1,$$

where E_0 is the potential energy due to the combined indicator and pendulum, and E_1 is the potential energy due to the pendulum. The uncertainty of D is

$$u^2(D) = u^2(E_0) + u^2(E_1) + 2u(E_0, E_1) .$$

Assuming perfect correlation between E_0 and E_1, a conservative estimate of the covariance $u(E_0, E_1)$ is

$$u(E_0, E_1) = \sqrt{u^2(E_0) + u^2(E_1)} .$$

The effective degrees of freedom associated with $u(D)$ are

$$\mathrm{df_D} = \frac{u^4(D)}{\dfrac{u^4(E_0)}{\mathrm{df}_{E_0}} + \dfrac{u^4(E_1)}{\mathrm{df}_{E_1}}},$$

based on the Welch-Satterthwaite approximation.

B.7 Scale Accuracy, *r*

Let r represent the bias in the scale mechanism and $\pm\Delta r$ be the specified error bounds of the measurement instrument. Assuming a rectangular distribution, the uncertainty of r is

$$u(r) = \frac{\Delta r}{\sqrt{3}}.$$

We will assume $\mathrm{df_r} = \infty$, which implies that we know $u(r)$ exactly.

Appendix C. t-Table

The following *t*-table values were taken from *NIST/SEMATECH e-Handbook of Statistical Methods* [5].

Table C.1 Upper critical values of Student's t distribution with degrees of freedom, df.

df	0.900	0.950	0.975	0.990	0.995	0.999
1	3.078	6.314	12.706	31.821	63.657	318.313
2	1.886	2.920	4.303	6.965	9.925	22.327
3	1.638	2.353	3.182	4.541	5.841	10.215
4	1.533	2.132	2.776	3.747	4.604	7.173
5	1.476	2.015	2.571	3.365	4.032	5.893
6	1.440	1.943	2.447	3.143	3.707	5.208
7	1.415	1.895	2.365	2.998	3.499	4.782
8	1.397	1.860	2.306	2.896	3.355	4.499
9	1.383	1.833	2.262	2.821	3.250	4.296
10	1.372	1.812	2.228	2.764	3.169	4.143
11	1.363	1.796	2.201	2.718	3.106	4.024
12	1.356	1.782	2.179	2.681	3.055	3.929
13	1.350	1.771	2.160	2.650	3.012	3.852
14	1.345	1.761	2.145	2.624	2.977	3.787
15	1.341	1.753	2.131	2.602	2.947	3.733
16	1.337	1.746	2.120	2.583	2.921	3.686
17	1.333	1.740	2.110	2.567	2.898	3.646
18	1.330	1.734	2.101	2.552	2.878	3.610
19	1.328	1.729	2.093	2.539	2.861	3.579
20	1.325	1.725	2.086	2.528	2.845	3.552
21	1.323	1.721	2.080	2.518	2.831	3.527
22	1.321	1.717	2.074	2.508	2.819	3.505
23	1.319	1.714	2.069	2.500	2.807	3.485
24	1.318	1.711	2.064	2.492	2.797	3.467
25	1.316	1.708	2.060	2.485	2.787	3.450
26	1.315	1.706	2.056	2.479	2.779	3.435
27	1.314	1.703	2.052	2.473	2.771	3.421
28	1.313	1.701	2.048	2.467	2.763	3.408
29	1.311	1.699	2.045	2.462	2.756	3.396
30	1.310	1.697	2.042	2.457	2.750	3.385
40	1.303	1.684	2.021	2.423	2.704	3.307
50	1.299	1.676	2.009	2.403	2.678	3.261
60	1.296	1.671	2.000	2.390	2.660	3.232
70	1.294	1.667	1.994	2.381	2.648	3.211
80	1.292	1.664	1.990	2.374	2.639	3.195
90	1.291	1.662	1.987	2.368	2.632	3.183
100	1.290	1.660	1.984	2.364	2.626	3.174
∞	1.282	1.645	1.960	2.326	2.576	3.090

Appendix D. Glossary of Terms

n	Number of test material samples measured
$\bar{y}$	Mean absorbed energy of test material samples
s	Standard deviation of test material samples
df	Degrees of freedom for test material standard deviation
n_V	Number of indirect verification samples measured
$\bar{V}$	Mean absorbed energy of verification samples
S_V	Standard deviation of verification samples
df_V	Degrees of freedom for verification material standard deviation
R	NIST reference value
$u(R)$	Standard uncertainty of NIST reference value
df_R	Degrees of freedom for reference value standard uncertainty
$\hat{e}_{\mathrm{systematic}}$	Systematic error estimate associated with test material
$u(\hat{e}_{\mathrm{systematic}})$	Standard uncertainty of test material systematic error
df_e	Degrees of freedom for standard uncertainty of test material systematic error
$\hat{\delta}_{\mathrm{systematic}}$	Systematic error estimate associated with verification material
$u(\hat{\delta}_{\mathrm{systematic}})$	Standard uncertainty of verification material systematic error
df_δ	Degrees of freedom for standard uncertainty of verification material systematic error
$\bar{y}_{\mathrm{corrected}}$	Corrected test result
$u(\bar{y}_{\mathrm{corrected}})$	Combined standard uncertainty of corrected test result
$\mathrm{df}_{\mathrm{eff}}$	Degrees of freedom for combined standard uncertainty of corrected test result
U	Expanded uncertainty of corrected test result