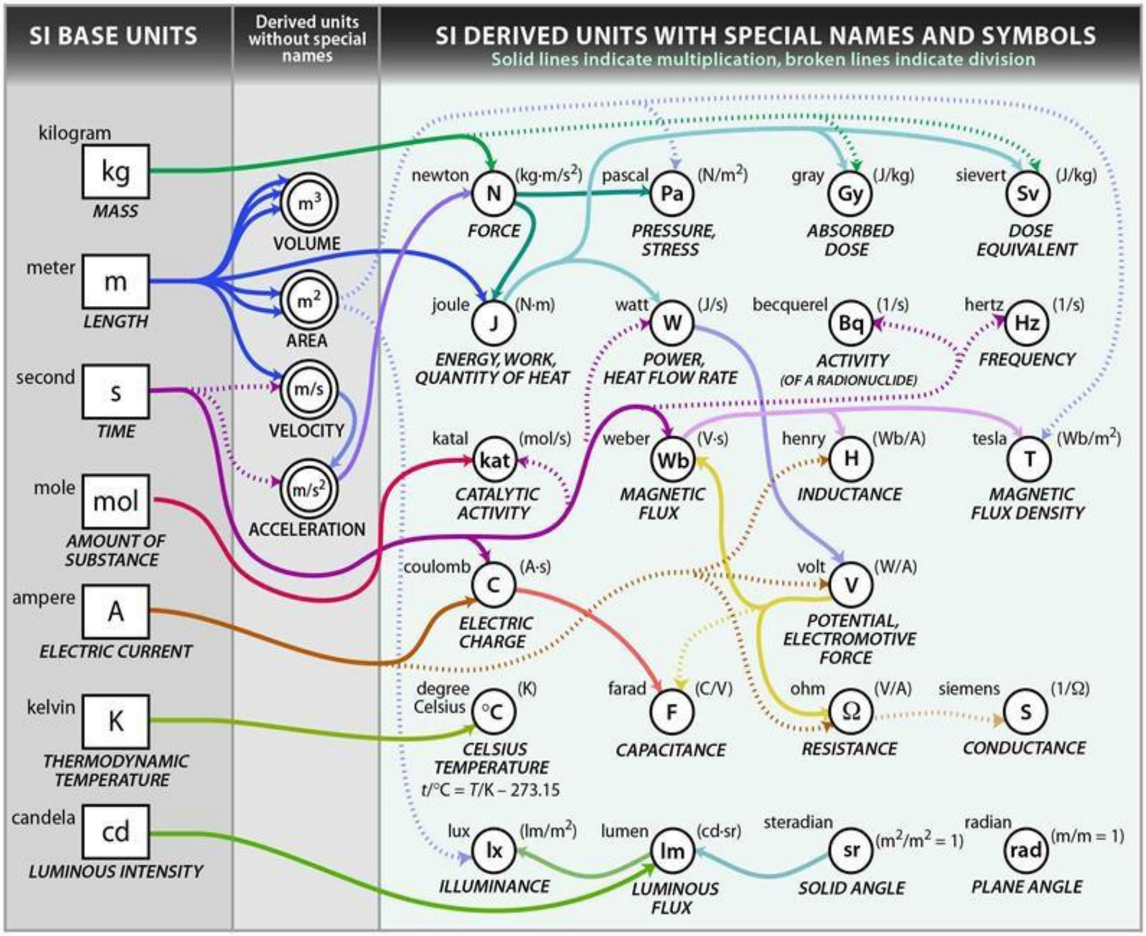
Establishing metrological traceability for force measurements



Two calibration laboratories can hand you the same number, 1000.0 lbf, and only one of them can prove it.

That is the whole problem in one sentence. A number on a certificate is worth exactly as much as the chain of evidence standing behind it, and not one bit more. One lab can walk that 1000.0 lbf back through a documented series of calibrations to the International System of Units. The other wrote down a confident reading. On paper, the two look identical. Only one is traceable.

The International Vocabulary of Metrology (VIM, item 2.41) gives the property its name and its definition: metrological traceability is the "property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty." Read that again with the working details in mind. Documented. Unbroken. And every link adds uncertainty, never subtracts it.

From there, almost every decision you make about how you calibrate force either protects that chain or quietly breaks it. The cleanest case is to calibrate the load cell and its indicator together as one system. The harder case is when you decide not to, and you have to establish traceability for the meter on its own, whether it reads millivolts per volt or milliamps. This article walks both.

The chain that gives a number meaning

Force is a derived quantity. F = m a, built from the SI base units of mass, length, and time, so a force value is only ever as good as its connection back to those units. As Sir Isaac Newton would remind us, the same force is required to break the egg no matter where you stand on Earth, because the unit itself does not move. Realizing that unit from a deadweight machine, though, depends on local conditions: the force a given mass produces is m·g, so the chain actually runs through the calibrated mass, the local gravitational acceleration, and an air-buoyancy correction — not mass alone. National metrology institutes such as NIST maintain that connection through primary standards and keep it comparable across borders through the CIPM Mutual Recognition Arrangement and the Key Comparison Database (KCDB) maintained by the BIPM. When a laboratory publishes a calibration and measurement capability in the KCDB, that capability has been peer reviewed under the arrangement. That is what traceable to the SI means in practice. Not a logo on a certificate, but a published, reviewed link in a chain that ends at a base unit.

Now look at what the chain costs you (Figure 1). Typical relative uncertainties for force grow at every step down from the primary standard. At $k = 1$, they run from roughly 0.0004 % – 0.0005 % at NIST, to 0.0008 % at Morehouse as a primary reference laboratory, to 0.01 % – 0.05 % at accredited calibration suppliers (Morehouse near the low end, many labs 0.03 % – 0.05 %), to 0.1 % at working standards, and to 0.5 % at field measurement. Measurement uncertainty is cumulative. You inherit everyone above you and add your own.

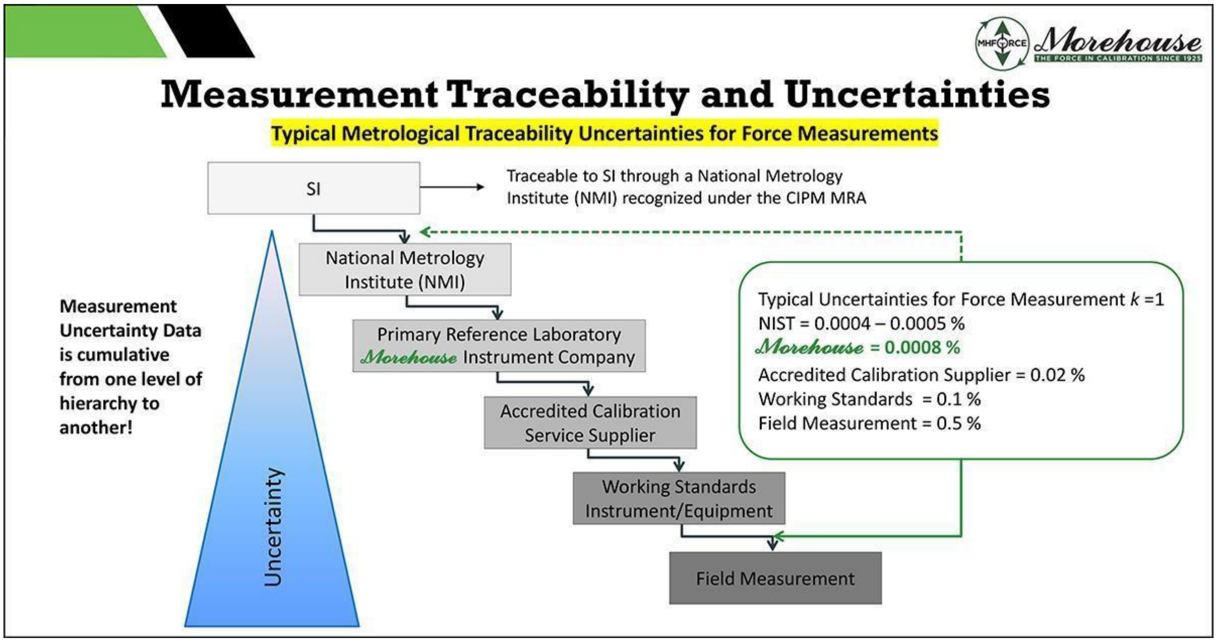


Figure 1. Measurement traceability for force. Uncertainty is cumulative and grows at each step below the primary standard. From Force Calibration for Technicians and Quality Managers.

So the choice of provider is not a formality. If you need a device held to within ± 0.02 %, you cannot get there through a lab whose own stated uncertainty is ± 0.04 %. The arithmetic forbids it. Morehouse keeps its force uncertainties low by sending its own standards, including several deadweight masses, to NIST for calibration. The closer your reference sits to the primary standard, the smaller your uncertainty and the more confidence you carry into every measurement below you.

One caveat is worth stating plainly. Accredited force capabilities vary widely by lab, range, and standard used. Morehouse appears in this chain at two tiers — near 0.0008 % (k = 1) with its deadweight primary standards, and around 0.01 % for its accredited secondary force work (achieved with HBM DMP40 meters and ultra-precision reference load cells that most labs do not use), so read the 0.01 % as the secondary-tier figure, not its best capability; many commercial labs land between 0.03 % and 0.05 %. Always read the specific scope of accreditation before you assume a lab can meet your needs. Even that might not be enough; if your needs include both force and electrical, you need to verify that both can be performed at the appropriate level.

Calibrate the load cell and the indicator as one system

ISO 376 is clear about what defines a force system. Section 3.1 defines a **force-proving instrument** as the whole assembly, from the force transducer through to and including the indicator. ASTM E74 treats the **force-measuring system** the same way, defining the force-measuring instrument as the elastic force transducer together with the instrumentation used to read and indicate its output. The subject of both standards is the pair, not the parts.

There is a good reason for that: it offers the cleanest traceability you can buy. When the load cell and indicator are calibrated together, force goes in from a deadweight or secondary machine, the system reads its output, and one documented chain ties the result to the SI. One certificate, one set of coefficients, one uncertainty budget. Nothing has to be reassembled later, and nothing has to be argued after the fact.

Pairing pays off again with every repeat calibration. When the reference laboratory reads and reports the actual mV/V values using a least-squares fit and hands you a fresh set of coefficients, your as-received calibration becomes your as-returned calibration. The new coefficients absorb whatever drift occurred and bring the bias back down, which lets you avoid the worst habit in the trade: tweaking.

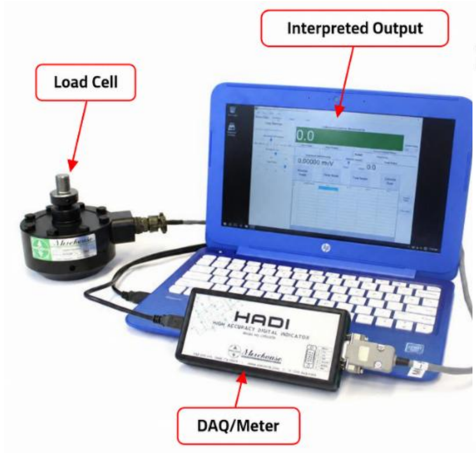
Coefficients also beat a two-point span by a wide margin, because almost no load cell is perfectly linear. In one worked case, fitting the curve instead of spanning two points cut the bias from 3.2 lbf (0.032 % of full scale) to about 0.1 lbf (0.001 % of full scale). That is roughly a 32-fold reduction in bias at the worst point, and it comes from arithmetic, not from turning a screw.

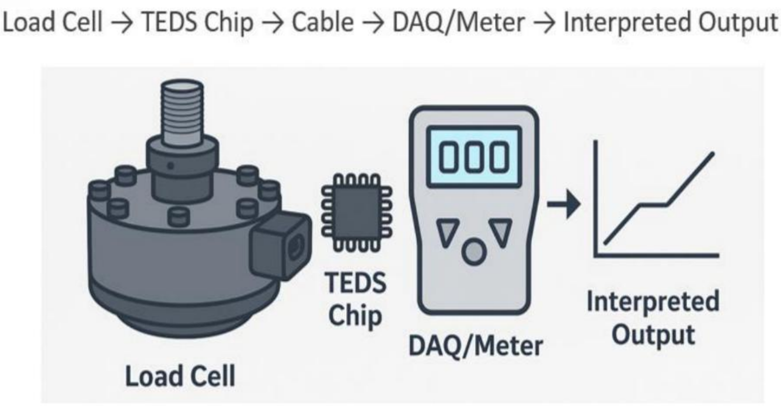
Why does that matter beyond the number? Because ISO/IEC 17025 expects you to monitor results so that trends are detectable, and a system you rarely adjust is a system whose drift you can actually see. Every time you tweak an indicator to chase nominal, you erase the trend and make the system less stable. W. Edwards Deming made the same point with his funnel experiment. Hold a funnel over a target, drop a marble through, and it lands a little off. Leave the funnel fixed, and the marbles fall in a tight, stable pattern around the target. Adjust the funnel after every drop to cancel the last miss, and the variance of the result doubles. Move it to wherever the last marble landed, and the pattern walks off the target entirely. The process was stable the whole time; the adjusting is what wrecked it. An indicator is no different. React to each reading, and you are correcting for noise, and every correction adds variation instead of removing it. Keep the cell and indicator married from one calibration to the next, read in mV/V, and let updated coefficients do the work. The bias stays low, and the history stays readable.

When you do not calibrate as a system

Sometimes you have a reason to split them. Maybe you want to check out two load cells and one meter for a job and free the rest of the inventory for other work. Maybe you keep an identical indicator as a backup so a single failure does not strand a system. The flexibility is real, and the standards allow it. They simply attach conditions, and the conditions are where laboratories get into trouble.

Here is the principle the textbook states twice, because it is that important. A meter used independently of its load cell is not made traceable by simulation alone. Simulating a bridge output proves the meter is consistent. It does not connect that meter to the SI. To establish traceability when the meter stands alone, the meter must be calibrated and its measurement uncertainty carried, and you now need two chains where you had one: traceability for the electrical quantity, and traceability for force. That is more work, more documentation, and more uncertainty.





This shows up in ordinary ways. Program a TEDS chip on a load cell without the specific indicator it will be read on, and that indicator still has to be calibrated, and its uncertainty included. The convenience of swapping cells across a meter does not buy you out of the chain; it just moves the calibration somewhere else.

Both force standards spell out the rules. ISO 376, in section C.2.11, says the deviation between the original and replacement indicators should be determined by calibrating both or by using a common bridge simulator, and the uncertainty of that deviation should be estimated and either corrected or carried. Section 5 adds that the indicator calibration shall be traceable to national standards, over a range equal to or greater than the range of use, with resolution at least equal to the original, and in the same excitation quantity (e.g., 5 V or 10 V) and type (AC or DC carrier), and recommends the replacement indicator's uncertainty be no greater than one third of the system uncertainty.

ASTM E74 is more prescriptive. Section 12 is titled, plainly, substitution of electronic indicating instruments used with force-measuring systems. Section 12.1.1 requires that the original and the substitute indicator have each been calibrated with their uncertainties determined, that the substitute be calibrated with traceability to the SI over the full range of intended use, including positive and negative values if the system is used in tension and compression, with a minimum of five points, and that each indicator uncertainty be less than or equal to one third of the system uncertainty. Section 12.1.2 sets the bar for the artifact you calibrate against: the **transducer simulator** uncertainty shall be less than or equal to one-tenth of the force-measuring instrument uncertainty. Section 12.1.3 adds that the excitation voltage amplitude, frequency, and waveform shall be maintained.

That last clause is not boilerplate. It is the hinge of the entire electrical side, and it splits the discussion in two because the quantity you have to make traceable depends on what the device actually outputs.

The mV/V path: a DC voltage ratio traceable to the volt

Most reference load cells are unamplified strain-gauge bridges. Excite the bridge with a regulated voltage, apply force, and the bridge returns a small signal that scales with the excitation. The natural measurand is the ratio of output to excitation, expressed in millivolts per volt (mV/V). So when the meter stands alone, the electrical quantity you must anchor is a DC voltage ratio, and the reference it is anchored to is the SI volt.

In practice, that means a precision digital multimeter operating in DC voltage-ratio mode, reading output against excitation (DCV: DCV), together with a calibrated load cell simulator that supplies stable, known mV/V steps. Both have to be traceable to the volt. As a sense of what is achievable, Morehouse runs a Fluke 8588A in DC mV/V ratio mode, calibrated by the NIST voltage group. The 8588A itself has a measurement uncertainty of better than 0.001 % ($k = 2$, approximately 95 % confidence). Calibrating a Morehouse load cell simulator with it, Morehouse achieved an expanded uncertainty of under 3.0×10^{-5} mV/V ($k = 2$, approximately 95 % confidence), with stability the dominant contributor in both cases. That is the order of capability a working laboratory can hold for the electrical link. The simulator then becomes the reference for calibrating other meters, and because uncertainty only accumulates downward, its capability sets the ceiling for everything below it.

ASTM E74 Section 12.1.3 requires that the excitation voltage amplitude, frequency, and waveform be maintained. The signal a bridge returns depends on how it is excited, so an indicator that does not match the original excitation is not a drop-in. Morehouse compared two high-end indicators, an HBM DMP40 reading an AC carrier and a Fluke 8508A reading DC, against a NIST-tested simulator (Figure 2). The two disagreed, and the disagreement was not linear: about 0.003 % on one test, with the DC side consistently higher. The lesson is direct. If a laboratory uses a DC reference, its traceability is to DC, and an AC indicator cannot stand in for it without recalibrating the system.

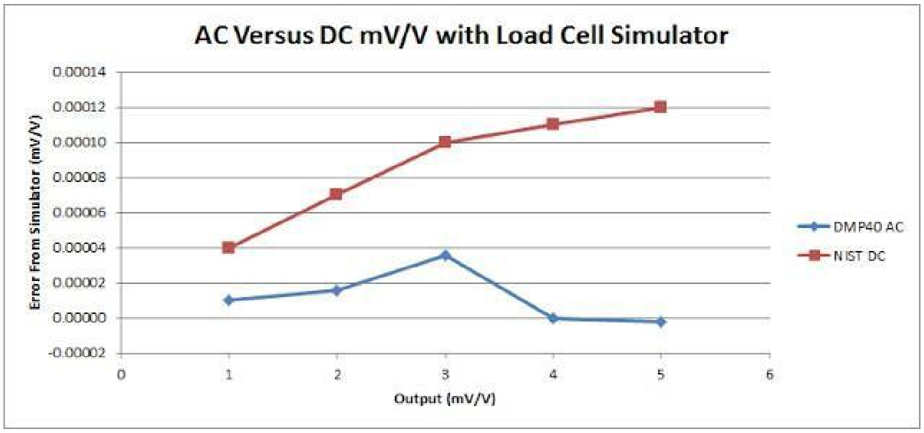
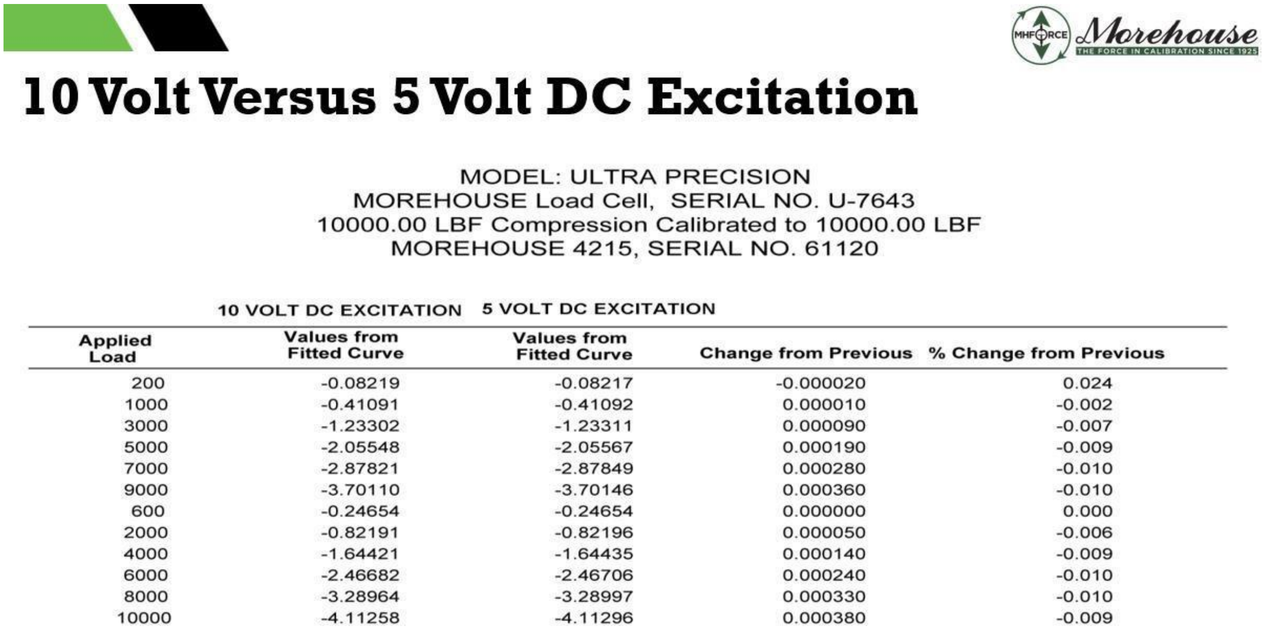


Figure 2. The same bridge output read on an AC carrier indicator and a DC indicator, against a NIST-tested simulator. The two are not interchangeable.

The excitation level matters too. Holding everything else fixed on one shear-web cell in a 12 000 lbf deadweight machine, itself within 0.0016 % (k = 2, i.e., 0.0008 % at k = 1), Morehouse measured a consistent difference of about 0.01 % between 5 V and 10 V DC excitation (Figure 3). For a reference standard, 0.01 % is too much to give away to a setting. Match the excitation, frequency, and waveform the cell was calibrated with, or measure what the mismatch costs and carry it.



10 Volt Versus 5 Volt DC Excitation

MODEL: ULTRA PRECISION
MOREHOUSE Load Cell, SERIAL NO. U-7643
10000.00 LBF Compression Calibrated to 10000.00 LBF
MOREHOUSE 4215, SERIAL NO. 61120

	10 VOLT DC EXCITATION	5 VOLT DC EXCITATION		
Applied Load	Values from Fitted Curve	Values from Fitted Curve	Change from Previous	% Change from Previous
200	-0.08219	-0.08217	-0.000020	0.024
1000	-0.41091	-0.41092	0.000010	-0.002
3000	-1.23302	-1.23311	0.000090	-0.007
5000	-2.05548	-2.05567	0.000190	-0.009
7000	-2.87821	-2.87849	0.000280	-0.010
9000	-3.70110	-3.70146	0.000360	-0.010
600	-0.24654	-0.24654	0.000000	0.000
2000	-0.82191	-0.82196	0.000050	-0.006
4000	-1.64421	-1.64435	0.000140	-0.009
6000	-2.46682	-2.46706	0.000240	-0.010
8000	-3.28964	-3.28997	0.000330	-0.010
10000	-4.11258	-4.11296	0.000380	-0.009

Figure 3. The same load cell read at 5 V and 10 V DC excitation. The difference held near 0.01 % across the range.

The simulator is the other half of the reference, and not every simulator qualifies. ASTM E74 §12.1.1 requires the calibrated range to bracket the transducer output at the lower force limit and at the maximum applied force, with a minimum of five points across that range; the non-mandatory Appendix

X2.2 adds that the simulator provide at least one point for every 20 % interval throughout the range. A budget simulator that starts at 0.5 mV/V and tops out at 4 mV/V (Figure 4) usually cannot reach low enough: on a 2 mV/V cell, its lowest 0.5 mV/V step is already 25 % of full output. To anchor the low end, you would need points near 0.04 mV/V — a 2 % lower limit factor (LLF) — and 0.1 mV/V, which is 5 %, and the budget unit cannot produce them.

Figure 4. A budget mV/V load cell simulator, 0.5 mV/V to 4 mV/V. Excellent for checks, not adequate for meter substitution.

The high-end units start far lower, near 0.02 mV/V (Figure 5), which is what allows them to support substitution down to better than 2 % of the verified range for many cells. Morehouse holds about ± 0.00003 mV/V on its high-end simulator. None of this makes the budget simulator useless. It is a strong tool for cross-checks, for confirming that coefficients were entered correctly, for catching a linearity problem, and for ruling the meter in or out as the source of a fault. It simply cannot, by itself, calibrate a meter for substitution.

Figure 5. A high-end load cell simulator starting near 0.02 mV/V, low enough to support compliant substitution for many cells.

What does separating actually cost in the budget? In our experience, meter substitution adds roughly 0.125 % or more to the low end of the range, below 10 % of capacity, and about 0.002 % or more at the high end. Set those numbers against the classes. ASTM Class AA expects better than 0.05 %, ISO 376 Class

00 and Class 0.5 are tighter still, and even Class 1 and Class 2 are likely out of reach at the low end once you add that much. ASTM Class A, at better than 0.25 %, has room. And the contributions you are adding are not exotic, only numerous: gain error, zero offset, temperature effect on sensitivity, and possibly on zero, non-linearity, gain and zero stability, etc., if the system is not a true six-wire connection, the difference in cabling between the simulator and the load cell. Every one of them has to be evaluated. None of them existed when the cell and indicator were one calibrated system.

The mA path: a DC current traceable to the ampere

Not every load cell hands the meter a millivolt ratio. Amplified load cells build the signal conditioning inside and put out a standardized 4 mA – 20 mA current loop or a 0 V – 10 V DC voltage, the signals a PLC, data acquisition system, or process controller expects. The 4 mA – 20 mA loop is popular in industry for a good reason: a current loop shrugs off voltage drop and electrical noise over long cable runs in a way a low-level millivolt signal never could.

The traceability logic does not change; the quantity does. When the output is current, the electrical quantity you must anchor is DC current, in milliamps, and the reference, the same class of precision multimeter, must be traceable to the SI ampere rather than the volt. The Fluke 8588A that Morehouse validated includes a DC current budget alongside its DC voltage and voltage-ratio budgets, specifically for this case. Everything else from the mV/V discussion still applies. The amplified system is calibrated by applying real force, the reference reads the output, and the chain ends at a base unit.

The catch with amplified cells is usually not traceability but the amplifier. In Morehouse's search for a good one, most amplifiers were stable to three decimal places at best, and an unstable amplifier quietly throws away the performance of a good load cell. To show what a stable one can do, Morehouse coupled a 10 000 lbf shear-web cell to its **Load Cell Amplifier for Controllers** (LAC), set it to 4 mA – 20 mA, and calibrated the assembly in the 12 000 lbf deadweight machine by the ASTM E74 procedure, reading the current to the last stable digit (Figure 6). The lower limit factor came out near 0.28 lbf, about 0.003 % of full scale, against the 0.02 % of full scale that most amplifiers manage. A second test on a 1000 lbf ultra-precision cell held maximum non-linearity to about 0.007 % of full scale, comfortably inside the cell rated 0.02 %. Read to five decimal places, the amplifier did not degrade the cell.

Factor	Output mA	LBF	%FS
LLF	*4.0004401*	*0.28*	0.003
Class A	*4.1760276*	*110.24*	1.102
Class AA	*4.8801378*	*551.20*	5.512

Figure 6. ASTM E74 results for a 10 000 lbf load cell paired with the Morehouse LAC set to 4 mA – 20 mA. The lower limit factor is about 0.003 % of full scale.

The point for traceability is the same one the mV/V path makes. If you separate an amplifier or current-loop indicator from its cell, you still owe two chains, and the electrical one now lives in milliamps. The reference has to be traceable to the ampere, the substitute indicator and the standard you calibrate it against still have to satisfy the one-third and one-tenth rules, and the output stability and resolution become first-order contributors to the budget rather than afterthoughts.

Figure 7. The Morehouse Load Cell Amplifier for Controllers (LAC 65.1), with bipolar voltage and 0 mA to 20 mA or 4 mA to 20 mA current output.

So which do you choose?

Put the two paths side by side, and the trade is clear. Marrying one indicator to each load cell and calibrating them as a system gives you the cleanest chain, the lowest uncertainty, the simplest compliance, and a useful side effect for scheduling: if a cell or an indicator goes down, only one item is out of service, and the rest of the inventory keeps working. Morehouse can calibrate a load cell with a fresh indicator as a system and turn it around quickly. Separating the meter buys flexibility and a backup, but it asks for a compliant simulator, a separately calibrated and traceable meter, more documentation, tighter control of excitation and cabling, and a larger uncertainty budget. You are paying for two calibrations instead of one.

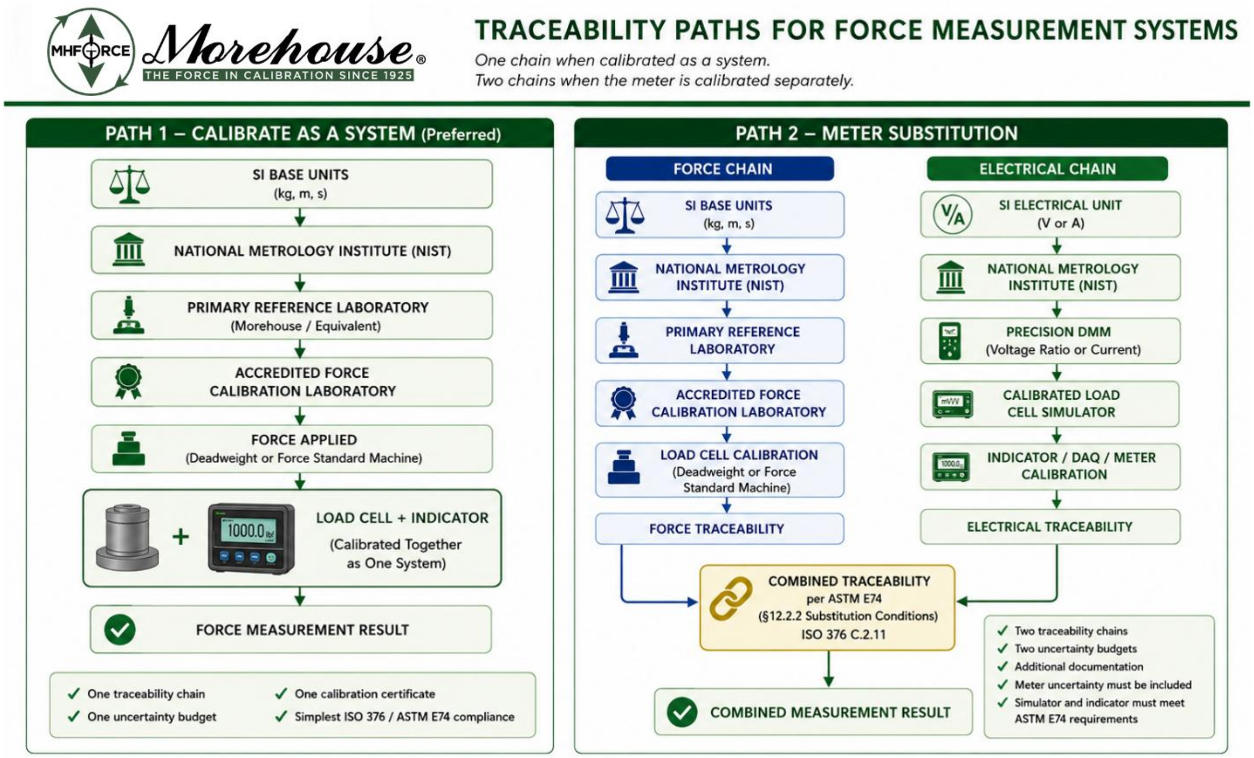


Figure 8. Traceability Paths for Force Measurement Systems.
Calibrating the load cell and indicator as a system ties the result to the SI through one chain.
Separating the meter requires two chains, one electrical and one for force, combined under the ASTM E74 substitution conditions.

Figure 8. The two traceability paths. Calibrating the load cell and indicator as a system ties the result to the SI through one chain. Separating the meter requires two chains, one electrical and one for force, combined under the ASTM E74 substitution conditions.

Table 1. Choosing between a system calibration and meter substitution.

Consideration	*Calibrate as a system*	*Separate the meter*
Chains to the SI	*One*	*Two: electrical and force*
Uncertainty budget	*One, lowest*	*Two, larger*
Certificate and coefficients	*One set*	*Separate, must be combined*
Compliance	*Simplest*	*ASTM E74 §12.1.1, §12.1.2, and §12.1.3; ISO 376 §5 and §C.2.11*
Documentation	*Minimal*	*More*
Excitation and cabling control	*Fixed at calibration*	*Must be matched and carried*
Flexibility and backup	*Limited; one item out of service*	*Swap cells or meters; keep a backup*
Best when	*Accuracy and simplicity matter most*	*Flexibility and a backup matter most*

So the rule of thumb is short. If accuracy and simplicity matter most, calibrate as a system. If flexibility and a backup matter more, and you can carry the added uncertainty and do the compliance work correctly, substitute the meter, with a properly calibrated simulator and a meter traceable to the right quantity. **What you cannot do is split the system, skip the second chain, and still call the result traceable.**

Traceability is not paperwork for its own sake. It is the reason a bridge cable, an aircraft fastener, or a hardness block can be trusted at the number on its certificate, and the reason a force measurement made in one lab means the same thing in another. Decide deliberately whether your meter is part of a calibrated system or a standalone instrument with its own chain, build the budget that choice requires, and document it. If you want the full treatment, with the worked uncertainty budgets, the simulator requirements, and the data behind these figures, it is all in Force Calibration for Technicians and Quality Managers, and the Morehouse team is glad to help you build the chain that fits your work.

Henry Zumbrun, Morehouse Instrument Company

For more information on force measurement, consider downloading our e-book for free here or purchasing the Kindle edition or hard copy on Amazon.

References

JCGM 200:2012. International Vocabulary of Metrology — Basic and General Concepts and Associated Terms (VIM), 3rd edition. Bureau International des Poids et Mesures (BIPM).

International Committee for Weights and Measures. CIPM Mutual Recognition Arrangement (CIPM MRA) and the Key Comparison Database (KCDB). Bureau International des Poids et Mesures (BIPM).

Zumbrun, H. Force Calibration for Technicians and Quality Managers. Morehouse Instrument Company.

ISO 376. Metallic materials — Calibration of force-proving instruments used for the verification of uniaxial testing machines. International Organization for Standardization, Geneva.

ISO/IEC 17025. General requirements for the competence of testing and calibration laboratories. International Organization for Standardization, Geneva.

ASTM E74. Standard Practices for Calibration and Verification for Force-Measuring Instruments. ASTM International, West Conshohocken, PA.

Deming, W. E. Out of the Crisis. Massachusetts Institute of Technology, Center for Advanced Engineering Study, Cambridge, MA, 1986.