

# KNO-011 · Measuring Before Theorising

KNO · Ways of Knowing

Juan Segurola

v0.1 · EN

All Rights Reserved

## Article Abstract

Theory is often said to tell scientists what to measure, and often it does. This article follows the other direction, where numbers arrive first and stand until an explanation reaches them, and treats neither direction as the rule. It opens on the question KNO-010 handed forward, whether sorting into categories already counts as measuring, and names both answers without adopting either. It then takes three cases where the number came first: a thermometer that had to be trusted before anything could check it, long wavelength measurements a famous formula was written to fit, and an orbital residual no available theory accounted for. It closes on the reverse move of 2019, when a constant measured for a century became a definition instead. Experimental design, uncertainty analysis and the structure of data records belong to later articles.

## In 60 Seconds

Two orders both exist and neither is the order this field follows: sometimes theory names a quantity and an experiment goes looking for it, sometimes the numbers accumulate until a theory has to reproduce them. Whether classifying already counts as measuring depends on which definition is in force, and the permissive and the restrictive answer are both named here and neither adopted. An instrument has to be trusted before anything exists that could check it, and in the episodes discussed the circle was not avoided rather than broken. A formula can be written to hit a target that measurement had already set: a communication about long wavelength behaviour came first and the equation followed, which is an order of events and not an account of what caused what. A residual reports the size of a theory's failure and binds only successors covering the same observational domain, at the precision that matters there. In 2019 a measured constant became a defined one, and the defining value has no uncertainty at all; what carries uncertainty now is the realisation of the unit.

## Why This Question Matters

A number offered as a measurement does not explain itself. Read without asking what validated the instrument, whether the quantity was theorised first, and whether the figure is a value or a residual, the same number will be made to support conclusions it cannot carry. The distinctions here are what let a reader see that a formula was answerable to data before it meant anything, that a small stable correction can outlast every attempt to explain it away, and that a constant can stop being measured by decision rather than by discovery.

## 1. Two orders, and why neither is the rule

There is a tidy account in which theory comes first. A theory says that a quantity exists, says roughly how large it should be, and an experiment is then designed to find it. Much of modern physics works this way, and the account is not wrong.

There is an equally tidy account running the other way. People measure something because they can, the numbers accumulate, a pattern appears in them, and a theory is eventually built that has to reproduce the pattern or be discarded. Much of the history of physics and chemistry works this way too, and that account is not wrong either.

**What is wrong is treating either as the order of science.** The two directions coexist, alternate within a single research programme, and sometimes both operate on the same quantity at different moments. This article argues the second direction because it is the one usually left out, not because it is the true one, and section 6 shows the first direction winning decisively in a case where the second had held for a century.

### Where That Leaves Us

Two tidy accounts run in opposite directions and both describe real practice, which is why neither is the order this field follows. Theory naming a quantity and an experiment going after it happens; numbers accumulating until a theory has to reproduce them happens too, sometimes on one quantity at different moments. The second is developed here because it is the one usually left out.

## 2. What counts as measuring, and the question KNO-010 left open

KNO-010 stopped at the category and handed forward a question it did not argue: whether assigning a thing to a category is already a measurement. It has to be answered before anything else here, because the answer decides what counts as a case.

The permissive answer is the one most often quoted. Stevens defines measurement, in the broadest sense, as "the assignment of numerals to objects or events according to rules" [1]. On that definition his four scale types follow from what the assignment allows one to do. His table separates them by their basic empirical operation: the nominal scale by determination of equality, the ordinal by determination of greater or less, the interval by determination of equality of intervals or differences, and the ratio by determination of equality of ratios [2].

The restrictive answer objects that this makes the question unaskable. Michell observes that under Stevens's view "any rule for making numerical assignments always defines at least a nominal scale" [3]. If any rule at all produces a measurement, then whether the attribute being studied is quantitative in the first place is never put to test, and Michell's charge against a century of quantitative psychology is exactly that: the scientific question of the additivity of the attributes involved was hardly ever addressed [4].

**Both positions are recorded here and neither is adopted.** The consequence for this article is only procedural, and it is enough. **Under the permissive definition, classifying is measuring on the weakest scale; under the restrictive one, it is not measuring at all until the attribute is shown to be quantitative.** Everything that follows concerns quantities that satisfy the restrictive definition too, so the cases below stand under either reading. Where an article needs the distinction to do real work, it should say which definition it is using, because the two disagree about what has been achieved when a number is produced.

### Where That Leaves Us

Whether classifying already counts as measuring turns on the definition in force. The permissive one takes any rule-governed assignment of numerals and sorts the results by what the assignment supports; the restrictive one objects that this never puts to test whether the attribute is quantitative at all. Both are recorded and neither adopted, and everything below satisfies the stricter reading.

### 3. Validating an instrument with no standard behind it

The first difficulty of measuring before theorising is not the theory. It is that the instrument has to be trusted before anything exists that could check it.

A thermometer is the standing example. To know whether a thermometer reads correctly one needs a known temperature, and to know a temperature one needs a thermometer. The eighteenth and nineteenth century work on fixed points was an attempt to break that circle from the inside, and it is the subject of Chang's study, whose publisher describes it as examining "how they managed to assess the reliability and accuracy of these instruments without a circular reliance on the instruments themselves" [5].

**The circle is not a scandal, and in the cases this article discusses it was not avoided.** Several measurement standards were at some point established by instruments that the standard was later used to judge. Whether that holds of every standard is not established here and is not claimed. What can be asked of such a process is not that it start from certainty but that each round leave the next one better placed, and that the improvements be visible rather than assumed. This article does not develop the general epistemology of that process, which is the subject of Chang's book and beyond the scope of the series. **What is used from that book here is limited to what was checked directly on its official abstract and table of contents.** Its body was not read, so nothing in this section attributes an argument to it beyond the description its publisher gives.

**One consequence matters here.** A quantity can be measured well, consistently and usefully for a long time before there is any theory of what is being measured. Temperature was measured for a long period before thermodynamics gave an account of it, and this article does not fix the length of that period, which depends on what is counted as a measurement and on which account of thermodynamics is taken as the arrival. The measurements were not waiting for the theory; the theory had to accommodate them.

#### Where That Leaves Us

Nothing exists that could check the first instrument, which is the difficulty here, and the thermometer is the case always cited. In the episodes discussed the circle was not avoided; how far that goes is not settled here. What such a process owes is not certainty at the start but visible improvement each turn. A quantity can be measured well long before anything explains it.

### 4. A formula built to fit the measurements

The clearest case of a theory constructed to reproduce a number is also one of the most consequential, and its usual telling reverses the order.

By 1900 several competing radiation laws existed for the emission of a black body. González de Arrieta records that they "were all consistent with the expected properties of blackbody radiation, but, crucially, each predicted different behaviors at large values of the product  $\lambda T$ " [6]. The disagreement was therefore located precisely: the laws were indistinguishable in the accessible range and separated only at long wavelengths, which is a description of an experiment waiting to be done.

Rubens and Kurlbaum did it. They set out to improve an existing residual ray experiment, "using four blackbody cavities spanning a temperature range from  $-180$  to  $1600\text{ }^{\circ}\text{C}$ " [7]. The result was decisive in the negative direction first: "none of these expressions reproduced precisely the observed data for all temperatures" [8].

What happened next is the point of this section. On 7 October 1900 Rubens communicated to Planck that the limit of the radiation law at large  $\lambda T$  was a linear function of temperature, and Planck reworked the thermodynamic properties of his empirically based blackbody system and produced an equation that predicted the observed phenomena [9]. **The sequence is what the source records, and this article reads it as a sequence:** the communication came first and the equation followed, which is not by itself an account of what caused what. Rubens confirmed the fit, and the two results were presented together: "Rubens replied that the proposed equation fitted the results perfectly, and both Planck and Rubens and Kurlbaum presented their results at the same meeting of the Physical Society on October 19" [10].

**The formula was answerable to the measurements from the moment it was written.** Its physical interpretation, the one that made it a founding result of quantum theory, came afterwards and had to be reconciled with a formula that already existed and already worked. That is measuring before theorising in its strongest form, and it is worth noticing that the article does not claim the measurement caused the theory. It claims something weaker and more defensible: **the theory had a target it had to hit, and the target was set by measurement.**

### Where That Leaves Us

Competing radiation laws agreed everywhere the instruments could reach and parted only at long wavelengths, which describes an experiment waiting to be done. It was done, none of the expressions survived it, and the equation that did come after a communication about how the limit behaved. That order of events is what the source records. The formula answered to the numbers from the moment it existed.

## 5. A residual that measurement established and theory did not

A measurement can also produce a number that no theory explains, and hold it in place for decades.

In 1859 Le Verrier addressed a letter to Faye on the theory of Mercury and the motion of its perihelion, published in the Comptes rendus of the Académie des sciences [11]. His finding was quantitative and narrow. As the passage is printed, it sufficed to add thirty eight arcseconds to the secular motion of the perihelion in order to represent all the observed transits to within one second [12].

Read carefully, that sentence has an unusual structure for a discovery. It does not report a new object or a new effect. **It reports the size of a correction that has to be applied to an existing theory for that theory to match the observations,** which is to say it reports the exact magnitude of the theory's failure. The correction was small, it was stable, and it was not going away.

**A residual of this kind is a peculiarly durable scientific object, and the rule for reading one is worth stating transferably.** A residual is an estimate that depends on the observations behind it, on the parameters used and on the model it is a residual of. It therefore constrains only those successor theories that claim to cover the same observational domain, and only to the precision that is relevant there. It is not an anomaly in the loose sense of something odd; it is a number, established by measurement, that a successor theory covering the same ground has to reproduce within that precision. It survived every attempt to explain it away by adjusting the inventory of masses in the solar system, and it was still there when a theory that predicted it arrived more than half a century later. **It was not there unchanged.** Laskar records Le Verrier's figure as roughly thirty eight arcseconds per century, obtained from twenty one transit observations between 1697 and 1848, against the roughly forty three arcseconds per century the later theory gives [12]. The two are close and they are not the same number, and the difference is the ordinary difference

between an estimate and a prediction rather than a defect in either. This article deliberately stops at the residual and does not treat the later explanation, which belongs to the history of gravitation rather than to the epistemology of measurement.

### Where That Leaves Us

Decades can pass with a figure in place that nothing accounts for. Neither a new object nor a new effect was reported, but the size of a correction an existing theory needed, which is the magnitude of its failure. Such a figure is an estimate, so it binds only successors covering the same ground at the precision that matters there, and the later value is close without being the same.

## 6. When a measured constant becomes a defined one

The three preceding sections all run from measurement towards theory. The most recent large change in metrology runs the other way, and it is worth setting beside them so that the article is not a one directional story.

Until 2019 the kilogram was an object. As the SI Brochure's own appendix on the kilogram states, "the previous definition of the kilogram was based on the mass of the international prototype of the kilogram (IPK) immediately after the prescribed cleaning procedure" [13]. Everything expressed in kilograms was, ultimately, a comparison with one artefact.

Resolution 1 of the 26th General Conference on Weights and Measures, adopted in 2018 and effective from 20 May 2019, replaced that with a constant. The kilogram "is defined by taking the fixed numerical value of the Planck constant  $h$  to be  $6.626\,070\,15 \times 10^{-34}$  when expressed in the unit  $J\,s$ , which is equal to  $kg\,m^2\,s^{-1}$ " [14].

**The inversion is complete and it is worth stating plainly.** The Planck constant was measured, with improving precision, over a long period before its value was fixed; this article does not put a figure on that period. It is not measured any longer, because it no longer has a value to be found: its numerical value is now fixed by definition, and the official statement is unambiguous that "the numerical values of the seven defining constants have no uncertainty" [15]. The uncertainty did not disappear. It moved. Where there was once an uncertain constant and an exact unit, there is now an exact constant and a unit whose realisation carries the uncertainty.

**The resolution itself records what happened to the prototype, and it is worth reading exactly.** Resolution 1's own Appendix 2, under *Status of constants previously used in the former definitions*, states that from 20 May 2019 the mass of the international prototype of the kilogram is 1 kg within a relative standard uncertainty equal to that of the recommended value of the Planck constant when the resolution was adopted,  $1.0 \times 10^{-8}$ , and that its value is determined experimentally thereafter [16]. **This is not a conserved quantity changing hands.** The defining constant has no uncertainty at all, which is the separate statement above with its own source; what this appendix adds is that the artefact now has an experimentally determined mass carrying an inherited uncertainty, so *moved*, in the paragraph above, describes where the work is done and not a quantity passed along.

**This is a metrological convention, adopted deliberately, and not a case of theory telling measurement what to do.** A committee decided to fix a number that had until then been sought, and it was only possible because measurement had made the constant known well enough that fixing it changed nothing practical. The two directions are not rivals here. The second was earned by the first.

## Where That Leaves Us

The most recent large change ran the other way. A constant measured for a century was fixed by decision, and the defining value has no uncertainty because there is no longer anything to find. The prototype keeps a mass determined experimentally and carrying an inherited relative uncertainty. Nothing was passed along: what changed is where the work of determining is done.

## 7. Reading a measurement well

What the preceding sections support is a short set of questions to ask of any number offered as a measurement.

**Which definition of measurement is in play?** On the permissive one, a classification with numerals attached qualifies. On the restrictive one, nothing qualifies until the attribute is shown to be quantitative. The two disagree about what has been established.

**What validated the instrument, and what validated that?** Every chain ends somewhere, and the interesting question is whether the last link was examined or assumed.

**Was the quantity theorised before it was measured?** If not, the measurements are not confirmations of anything and should not be read as such.

**Is this number a value or a residual?** A residual reports the size of a theory's failure, and it constrains successors in a way that an ordinary measurement does not.

**Is the constant still being measured?** Some are not, by decision. A defined constant has no uncertainty because uncertainty has been moved elsewhere, not because it has been eliminated.

**What would it take for this number to change?** A better instrument, a redefinition, or a reinterpretation are different answers with different consequences, and a number that cannot change for any of the three is probably not a measurement.

Measuring before theorising is neither a lesser activity nor a preliminary one. It produces the constraints that theories are answerable to, and it occasionally produces a number that stands unexplained for fifty years and is right the whole time.

## Where That Leaves Us

The article ends in questions to put to any number offered as a measurement. Which definition is in force, what validated the instrument and what validated that, whether the quantity was theorised first, whether the figure is a value or a residual, whether it is still being measured, and what would have to happen for it to change. What the sections establish, and how, comes next.

## What We Know, and How

The permissive definition takes measurement to be the assignment of numerals to objects or events according to rules, and separates four scale types by the empirical operation each supports. The restrictive objection is that under such a view any rule at all already defines a nominal scale, so whether the attribute is quantitative is never put to test. Competing radiation laws differed only at long wavelengths; a residual ray experiment across four cavities found none of them reproducing the data at all temperatures; a communication about the limit at long wavelengths preceded the equation that worked, and the two results were presented at the same meeting. Le Verrier's correction to the secular motion of Mercury's perihelion is recorded by Laskar as roughly

thirty eight arcseconds per century from twenty one transit observations, against roughly forty three from the later theory. Resolution 1 defines the kilogram by fixing the Planck constant, the seven defining constants have no uncertainty, and that resolution's own Appendix 2 puts the prototype at 1 kg within the inherited relative standard uncertainty, determined experimentally from then on. Chang's study is used only as its publisher describes it; its body was not read here.

## What We Still Don't Know

Whether the circularity described holds of measurement standards in general, which section 3 declines to establish. How long the Planck constant was measured before its value was fixed, and how long temperature was measured before thermodynamics accounted for it, neither of which this article puts a figure on. What Chang's book argues, as opposed to what its publisher says it examines. And what caused what in October 1900: the order of events is recorded and the causal question is left open.

## What Could Change Our Minds

A reading of Chang's argument rather than its publisher's description would settle what that study establishes about breaking the circle. A demonstration that some measurement standard was validated without ever relying on instruments it later judged would narrow the claim section 3 declines to generalise. On the residual, a successor theory covering the same observational domain and failing to reproduce the figure at the relevant precision would matter; one covering different ground would not. And a metrological decision to unfix a defining constant would return it to being measured.

## If You Remember Only One Thing

Numbers often arrive before the theories that explain them, so ask what a measurement was answerable to before asking whether it is true.

## Related Articles

- KNO-010 · *Classification as a Knowledge Tool*

## References

The bracketed numbers in the body are claim numbers, not reference numbers. Each resolves to a row of KNO-011\_CLAIMS\_AND\_SOURCES\_RECONSTRUCTION.md, which gives one academic or official source for each and names the entry below that supports it. Uncertain claims, declared scope limits and excluded sources are listed in KNO-011\_UNCERTAINTIES\_RECONSTRUCTION.md.

BIPM. 2018. *Resolution 1 of the 26th CGPM (2018): On the revision of the International System of Units (SI)*. Conférence Générale des Poids et Mesures. URL estable: <https://www.bipm.org/en/committees/cg/cgpm/26-2018/resolution-1> DOI 10.59161/CGPM2018RES1E

BIPM. 2021. *SI Brochure – 9th edition (2019) – Appendix 2*, 7 de julio de 2021. Bureau International des Poids et Mesures. URL estable: <https://www.bipm.org/documents/20126/41489673/SI-App2-kilogram.pdf>

BIPM. Página «SI defining constants». Bureau International des Poids et Mesures. **Sin fecha mostrada en la página.** URL estable: <https://www.bipm.org/en/measurement-units/si-defining-constants>

Chang, Hasok. 2004. *Inventing Temperature: Measurement and Scientific Progress*. Oxford Studies in Philosophy of Science. Oxford University Press. ISBN impreso 9780195171273. URL estable: <https://global.oup.com/academic/product/inventing-temperature-9780195171273> **Manifestación en línea, Oxford Academic:** ISBN en línea 9780199835003, DOI 10.1093/0195171276.001.0001, <https://academic.oup.com/book/5530>

González de Arrieta, Iñigo. 2022. «Beyond the infrared: a centenary of Heinrich Rubens's death». *The European Physical Journal H* 47: artículo 11. DOI 10.1140/epjh/s13129-022-00044-x

Laskar, Jacques. 2017. «Des premiers travaux de Le Verrier à la découverte de Neptune». *Comptes Rendus Physique* 18: 504–519. DOI 10.1016/j.crhy.2017.10.011

Michell, Joel. 1997. «Quantitative science and the definition of measurement in psychology». *British Journal of Psychology* 88: 355–383. DOI 10.1111/j.2044-8295.1997.tb02641.x URL estable: <https://errorstatistics.com/wp-content/uploads/2015/02/michell-1997-1.pdf>

Stevens, S. S. 1946. «On the Theory of Scales of Measurement». *Science* 103 (2684): 677–680. **Sin DOI localizado sobre página consultada.** URL estable: <https://dl.icdst.org/pdfs/files3/38396a3ee53de9478a57af4605f0e1c1.pdf>