Fact-Checked Science

PAPER CHECK · NEWTON'S CONSTANT

Same instrument, two answers

NIST spent a decade rerunning BIPM's torsion-balance measurement of Newton's gravitational constant, G, on BIPM's own instrument [1 Abstract][1 p. 27]. Its result sits 250 parts per million below BIPM's and within the uncertainties of the consensus value. That says more about hidden laboratory errors than about gravity, though the cause hasn't been found.

By Fact-Checked Science · Published October 14, 2026

OUR ANALYSIS marks our own reasoning and calculations, and the verdicts are our assessment. Everything else comes from the numbered sources.

One instrument, two answers: the problem is most likely in the measuring

OUR ASSESSMENT

6.67387(38)
NIST's *G*, in 10^{-11} m^3 kg^{-1} s^{-2}; the (38) is the uncertainty in the last two digits, 57 parts per million [1 p. 27].

−250 ppm
Below BIPM's corrected result from the same balance [1 Abstract].

−64 ppm
Below the CODATA 2022 value, within the uncertainties [1 p. 27].

THE CLAIMS ABOUT THE RERUN

TRUE
NIST's value disagrees with BIPM's and is consistent with the CODATA value
Schlamminger et al. 2026 [1 Abstract][1 p. 27]

It sits 250 parts per million below BIPM's corrected 2014 value, about four standard uncertainties apart, and 64 below the CODATA 2022 value, about one, a difference the authors find not significant (p = 0.29). They add that their result "does not corroborate" the CODATA value either: it is consistent with it, not a confirmation. See section 1.

TRUE
The disagreement points to hidden errors in the measurements, not new physics
Schlamminger et al. 2026; the NIST release [1 p. 3][2]

One instrument gave answers 250 ppm apart in two laboratories. The authors call new physics "remote" and hidden instrumental effects "more plausible", but they found no single cause, and the test masses, the disk, the metrology and the analysis all changed between the two runs [1 p. 9][1 p. 28]. See section 5.

OVERSTATED
"The team found virtually identical results."
The NIST release, on copper and sapphire source masses [2]

By our arithmetic from the paper's tables, the sapphire runs sit 150 parts per million below the copper runs with the servo method, about four times their combined uncertainty, and 58 below with free deflection, well within it; and the paper finds "larger unexplained effects" in the sapphire measurements [1 p. 25][1 p. 26][1 p. 27]. See section 3.

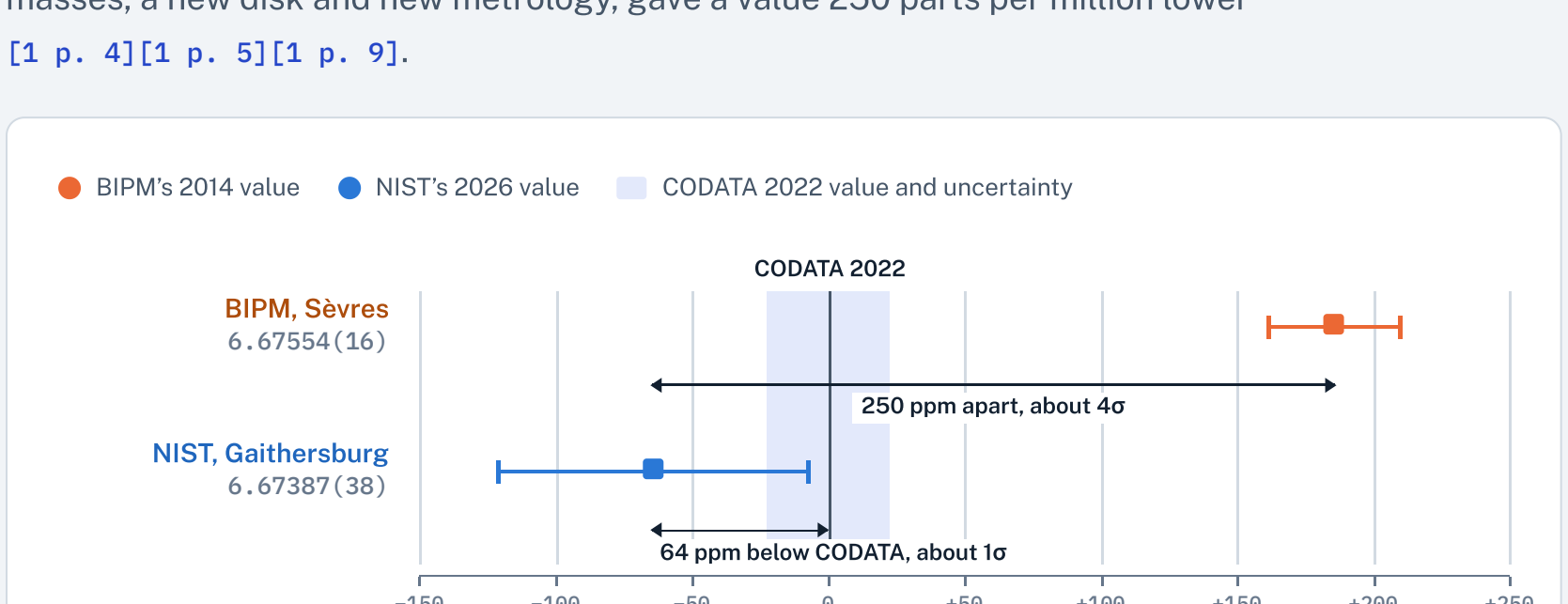
MISREAD
"the new result differs from the standard by 0.0235%"
Universe Magazine's summary of its April 21, 2026 article [3]

The release's 0.0235% is the gap between NIST's and BIPM's values [2]. NIST's result is 0.0064% below the CODATA value, a difference the authors find not significant [1 p. 27]. The article's own text comes closer, giving the figure as the difference from "the previous reference measurement" [3]. See section 5.

1 · TWO LABORATORIES

One torsion balance, measured in two laboratories

BIPM's torsion balance measured the gravitational constant, *G*, in Sèvres, France, in 2007. Shipped to NIST in Gaithersburg, Maryland, in 2016, the same apparatus, with new test masses, a new disk and new metrology, gave a value 250 parts per million lower [1 p. 4][1 p. 5][1 p. 9].



BIPM: 6.67554(16), its corrected 2014 value; NIST: 6.67387(38); CODATA 2022: 6.67430(15), all in 10^{-11} m^3 kg^{-1} s^{-2} [4 Table XXX][4 Table XXXI][1 p. 27]. The σ figures are our arithmetic: they combine the published standard uncertainties and ignore any correlation between two runs of the same apparatus.

What the authors conclude

Their value is "inconsistent with the earlier determination of *G* obtained using the same apparatus", while its difference from the CODATA value "does not differ significantly from zero (p-value = 0.29)". Yet "Our result does not corroborate either the earlier BIPM value or the current CODATA recommendation" [1 p. 27].

OUR ANALYSIS

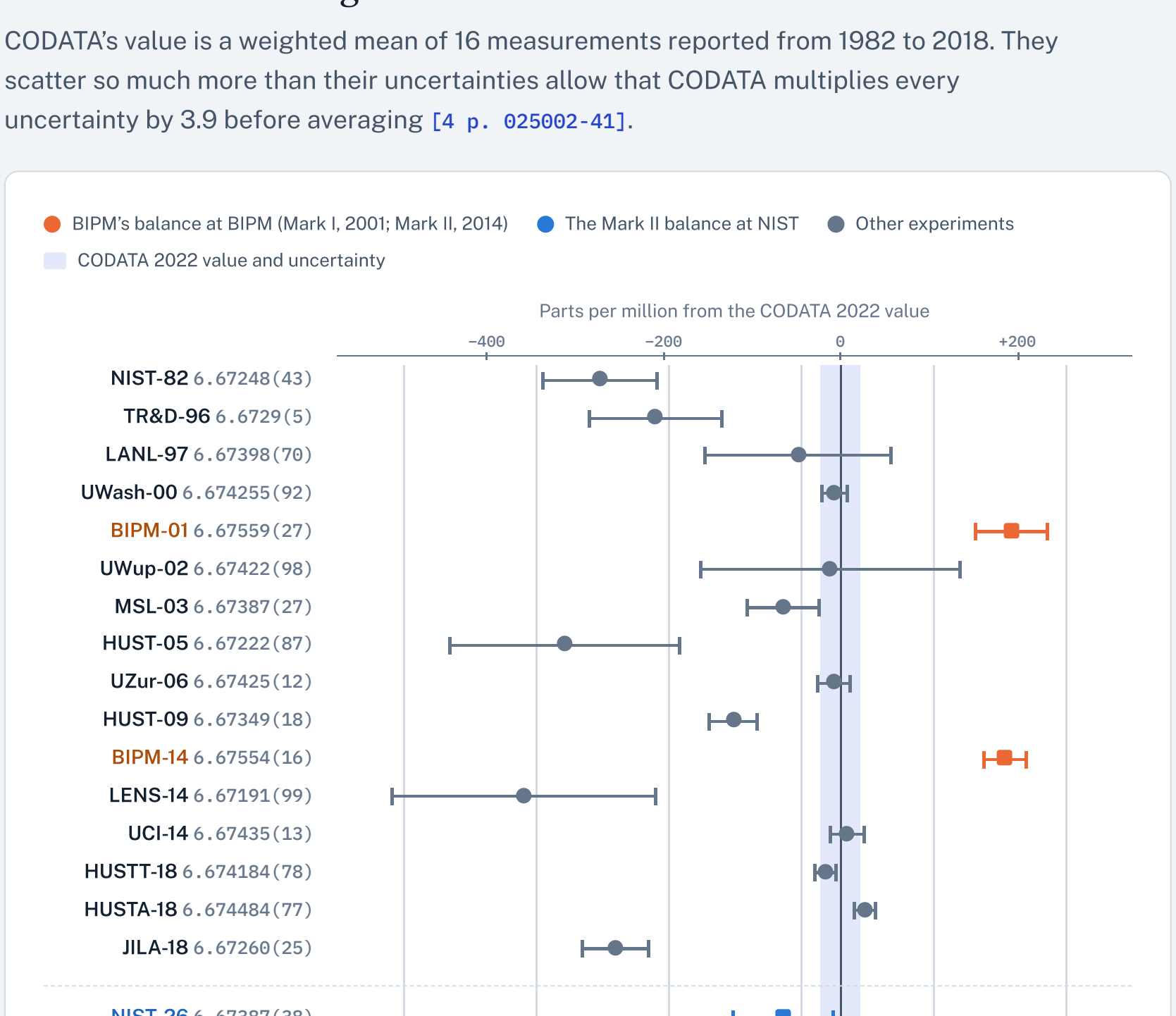
What "agrees" means here

A 64 ppm gap with a 57 ppm uncertainty is consistent with the CODATA value, but it is not a precise confirmation of it. NIST's uncertainty is more than twice CODATA's, and the final value combines four sub-results that disagree with one another, which the authors handle by adding "dark uncertainty" and by using a prior centred on the CODATA value (sections 3 and 5).

2 · FOUR DECADES

Where it lands among four decades of measurements

CODATA's value is a weighted mean of 16 measurements reported from 1982 to 2018. They scatter so much more than their uncertainties allow that CODATA multiplies every uncertainty by 3.9 before averaging [4 p. 025002-41].



The 16 results CODATA 2022 combines, labelled by laboratory and the year the data were reported, plus NIST's 2026 result [4 Table XXX][1 p. 27]. BIPM-01 came from the first version of BIPM's balance, Mark I; BIPM-14 and NIST-26 used the rebuilt Mark II [1 p. 4][1 p. 9]. LENS-14 is the only one that used cold atoms, as its test masses; its source masses were about 516 kg of tungsten alloy [5]. By coincidence, NIST-26 matches New Zealand's MSL-03 value to all six digits.

SPREAD · OUR ANALYSIS
≈ 550 ppm
From LENS-14, the lowest of the 16, to BIPM-01, the highest [4 Table XXX].

SCATTER
4.2×
The standard deviation of the values is 4.2 times their median reported uncertainty [6 p. 4].

EXPANSION FACTOR
3.9×
CODATA multiplies every input uncertainty by 3.9, which brings each value within two of its expanded uncertainties of the mean [4 p. 025002-41].

3 · TESTS

How the rerun tested itself

Four source masses pull on four test masses on a disk that hangs from a thin strip; the twist, or the voltage needed to stop it, gives *G* [1 p. 4][1 p. 5].

- Source masses on a carousel: copper, or sapphire
- Test masses on the disk
- Electrodes, two beside each test mass
- The copper–beryllium strip the disk hangs from, 30 µm thick

Schematic top view, not to scale [1 p. 5][1 p. 7][1 p. 21].

1. **Two ways to read the same pull**
In free deflection the disk twists and the angle is measured, as in Cavendish's method; in the servo method, voltages on the electrodes hold it still. Each gives its own estimate of *G* [1 p. 4][1 p. 7].

2. **Copper and sapphire**
Bryan Kibble had suggested that magnetic coupling between the masses inflated BIPM's value; a 2023 study put an upper limit of about 20 ppm on that effect, and NIST added source masses of single-crystal sapphire "to test this conclusively" [1 p. 9][7]. The sapphire results came out 150 ppm (servo, about 4σ) and 58 ppm (free deflection, within the uncertainties) below copper, by our arithmetic from the paper's tables, and the paper flags "larger unexplained effects" in them [1 p. 25][1 p. 26][1 p. 27]. It doesn't say the test settled Kibble's idea.

3. **Tighter geometry**
Diamond-turned test masses pin down where each mass sits; a new disk replaced the original, which had slightly non-parallel faces [1 p. 9].

4. **A blind analysis**
NIST's mass group multiplied every source-mass value by a secret factor within ±0.1% and sealed it in an envelope, to prevent experimenter bias [1 p. 27]. The release says the worry was drifting toward BIPM's value, and describes Patrick Abbott subtracting a number from the weights of some of the masses [2].

5. **A torque from the residual gas**
The team found a spurious torque that a temperature difference of just five millionths of a degree over 10 cm inside the vacuum can could produce through the little gas left inside, one of several effects "not previously treated for this apparatus" [1 p. 19][1 p. 28]. At BIPM's likely pressure the authors put it at no more than about 10 ppm, too small to explain the gap, but it may matter in balances with smaller signals [1 p. 28].

4 · HISTORY

A measurement kept sealed for years

2007
BIPM measures *G* with the rebuilt Mark II balance in Sèvres, France [1 p. 4].

2013–14
BIPM publishes 6.67545(18). A 2014 erratum, after a source-mass density correction had been applied twice, gives 6.67554(16) [8][6 p. 4].

2016
After BIPM's Terry Quinn proposed it at a 2014 NIST workshop, the balance is shipped to NIST for an independent team to rerun [1 p. 3][1 p. 5].

2022
A planned unblinding is called off when the team spots an overlooked effect "related to air pressure" [2].

July 11, 2024
The envelope is opened during a talk at the CPEM conference in Aurora, Colorado. The secret offset is large and negative, but too large for the result to match BIPM's [2][1 p. 27].

April 16, 2026
After about two more years of analysis, the result appears in *Metrologia*: 6.67387(38). The value shown at the unblinding was 44 ppm lower; later measurements of where the test masses sit moved it [1 p. 27].

5 · READ WITH CARE

A careful rerun that deepens the puzzle rather than solving it

This section is our assessment, checked against the paper, the release, the CODATA tables and the coverage.

OUR ANALYSIS

Not "different from the standard"

One outlet's summary says the result differs from the standard by 0.0235% [3]; its article text attributes the figure to "the previous reference measurement". That figure is the release's gap between NIST and BIPM [2]. Against the CODATA value the gap is 64 ppm, 0.0064%, and not significant [1 p. 27].

OUR ANALYSIS

0.0235% or 2.5×10^{-4}?

The paper's abstract and Figure 27 compare NIST with BIPM's corrected 2014 value: 2.5×10^{-4} [1 Abstract][1 p. 27]. Its closing section instead places BIPM 1.71×10^{-4} above CODATA and NIST 0.64×10^{-4} below, which adds up to the release's 0.0235% [1 p. 27][2]. By our arithmetic the corrected BIPM value is 1.86×10^{-4} above CODATA; 1.71 is closer to BIPM's superseded 2013 value, 1.72 above.

OUR ANALYSIS

Quote the published number

NIST's publications database still lists 6.67366 ± 0.00020, a relative uncertainty of "3.0×10^{-6}" (that ratio is 3.0×10^{-5}) and a 2.81×10^{-4} gap to BIPM, apparently from an earlier version [9]. The paper gives 6.67387(38); the release, 6.67387 [1 p. 27][2].

OUR ANALYSIS

What the blind covers

The blind protected the analysis until July 2024. The value shown then was 44 ppm lower than the published one, which the authors attribute to later measurements of the test masses [1 p. 27]. The final uncertainty also includes "dark uncertainty" for the scatter among the four sub-results, and the combination uses a prior centred on the CODATA value, with a standard deviation of about 135 ppm [1 p. 26].

OUR ANALYSIS

Systematics, not new gravity

One instrument giving two answers in two laboratories points to errors in how the measurement was made or analysed, at BIPM, at NIST or both. The authors quantified several effects not previously treated for this apparatus but found no single dominant cause, "leaving the larger discrepancy unresolved" [1 p. 28].

OUR ANALYSIS

Not fully independent

Clive Speake contributed to both the BIPM and NIST efforts, so the authors call their determination "not entirely independent" [1 p. 3].

OUR ANALYSIS

Bottom line: a meticulous rerun that is consistent with the consensus value and disagrees with the instrument's own earlier answer.

The original infographic's reading holds: the 250 ppm gap says more about hidden laboratory errors than about gravity. But copper and sapphire didn't give "virtually identical" results, the sapphire test didn't settle Kibble's idea, and the cause of the BIPM discrepancy is still unknown [1 p. 27][1 p. 28].

Sources

Bracketed numbers refer to this list; p. gives a page (for the journal articles, the journal's page numbers), Table a table, and Abstract the part of a work we used. Links checked October 14, 2026.

1. Schlamminger, S., Chao, L., Lee, V., Shakarji, C., Possolo, A., Newell, D., Stirling, J., Cochrane, R., and Speake, C. (2026). Redetermination of the gravitational constant with the BIPM torsion balance at NIST. *Metrologia* 63(2), 025012. Published April 16, 2026. CC BY 4.0.
2. National Institute of Standards and Technology (2026). NIST Weighs In on the Mystery of the Gravitational Constant. News release, April 16, 2026 (page updated May 5, 2026).
3. Nehriienko, R. (2026). Problem G: New precise measurements of the gravitational constant disagree with previous ones again. *Universe Magazine*, April 21, 2026; the page's metadata names Liliia Alaieva. The quotation is from the article's search and link-preview summary (its meta description); its text says the new value "differs from the previous reference measurement by 0.0235%".
4. Mohr, P. J., Newell, D. B., Taylor, B. N., and Tiesinga, E. (2025). CODATA recommended values of the fundamental physical constants: 2022. *Reviews of Modern Physics* 97(2), 025002. Table XXX lists the 16 input values for G; we read the copy NIST hosts.
5. Rosi, G., Sorrentino, F., Cacciapuoti, L., Prevedelli, M., and Tino, G. M. (2014). Precision measurement of the Newtonian gravitational constant using cold atoms. *Nature* 510, 518–521. We read the accepted manuscript, arXiv:1412.7954.
6. Merkatas, C., Toman, B., Possolo, A., and Schlamminger, S. (2019). Shades of dark uncertainty and consensus value for the Newtonian constant of gravitation. *Metrologia* 56(5), 054001. CC BY 3.0.
7. Speake, C. C., Bryant, J. L., Davis, R. S., and Quinn, T. J. (2023). The influence of time varying magnetic fields on the determinations of Newton's constant at the Bureau International des Poids et Mesures. *Metrologia* 60(2), 024001. CC BY 4.0.
8. Quinn, T., Parks, H., Speake, C., and Davis, R. (2013). Improved determination of *G* using two methods. *Physical Review Letters* 111, 101102; erratum, *Physical Review Letters* 113, 039901 (2014). We read the 2013 abstract; the erratum's content is as Merkatas and colleagues and CODATA report it.
9. National Institute of Standards and Technology. Redetermination of the gravitational constant with the BIPM torsion balance at NIST. NIST publications database entry (updated May 2, 2026), read October 14, 2026.

METHOD

How this was checked

On October 14, 2026, we checked every statement, number and source on the original infographic, 60 items, against the published paper, the NIST release, the CODATA 2022 tables, the 2019 analysis of the CODATA inputs and the papers they cite. We recomputed every difference, ratio and σ value and measured the original charts against their axes. The check corrected several statements; the change log lists them.

After the page was built, a second reviewer who hadn't seen the drafting recomputed every number and chart mark from the sources, which matched, and checked every quotation and reference. The changes that followed: the Universe Magazine article is credited to its visible author, and the quoted phrase is identified as its search and link-preview summary, while its text gives the 0.0235% as the difference from "the previous reference measurement"; the copper–sapphire comparison gives the uncertainties (the servo results differ by about four standard uncertainties, the free-deflection results agree within them); the temperature difference behind the gas torque is over 10 cm inside the vacuum can, and the authors' 10 ppm is an upper bound; the blind's purpose is the paper's, to prevent experimenter bias; the dark uncertainty and the prior appear where the page discusses agreement with CODATA; and several locators were added.

The verdicts and the items tagged "Our analysis" are our own assessment, built on the cited sources: the σ comparisons, the spread of the 16 values, the copper–sapphire differences and our reading of the release's 0.0235%. The paper is open access under CC BY 4.0 [1]; the charts are our own, drawn from the CODATA input values and the paper's result.

This page was researched and written with Claude, an AI assistant made by Anthropic, and then re-checked against its sources in a separate verification pass.

The verdicts on this page

- OVERSTATED — Has some basis but claims much more than the evidence supports.
- MISREAD — Cites a real source but gets wrong what it says.
- TRUE — The evidence supports the claim.

All ten verdicts, and how we check

CORRECTIONS

Corrections and changes

If we get something wrong, we fix it and record the change below with the date. If a correction changes a verdict, the log says what the old verdict was.

Change log

Oct 14, 2026 — Published on factchecked.science after a full check against the sources. Changes from the infographic: added our verdicts on four claims. NIST's result is "consistent with" the CODATA value rather than in agreement with it, as the authors say it "does not corroborate" that value either. Copper and sapphire source masses did not give "virtually identical results", the release's phrase: by our arithmetic from the paper's tables, the sapphire runs sit 58–150 ppm below copper, the servo pair by about four standard uncertainties. Nor does the paper say the sapphire test ruled out Bryan Kibble's magnetic-coupling idea, on which a 2023 study had put an upper limit of about 20 ppm. In the 2024 data sapphire came first, so "Copper, then sapphire" became "Copper and sapphire". The blind is described as the paper does, a secret factor applied to every source-mass value to prevent experimenter bias, with the release's version attributed to it. The gas torque is "not previously treated for this apparatus", the paper's words, rather than "never described before", and the authors put it too small to explain the gap. The new disk replaced one with non-parallel faces. HUSTA-18's uncertainty is (77), not (78). We found no source for Mark II's rebuild in "2001–03", so the years are dropped. LENS-14 used cold atoms as its test masses, with tungsten source masses. The legend's "Mk II 2013" became "Mark II, 2014", the year of the corrected value plotted. The release's 0.0235% matches the paper's own closing section, rather than suggesting the release used BIPM's old value. "Lab-specific errors" became errors at BIPM, NIST or both, since the authors found no single cause, and the claim that matching copper and sapphire results rule out a material effect is gone. We added the 44 ppm change after unblinding, the dark uncertainty and prior in the final combination, two electrodes per test mass in the schematic, and the sources' full references.