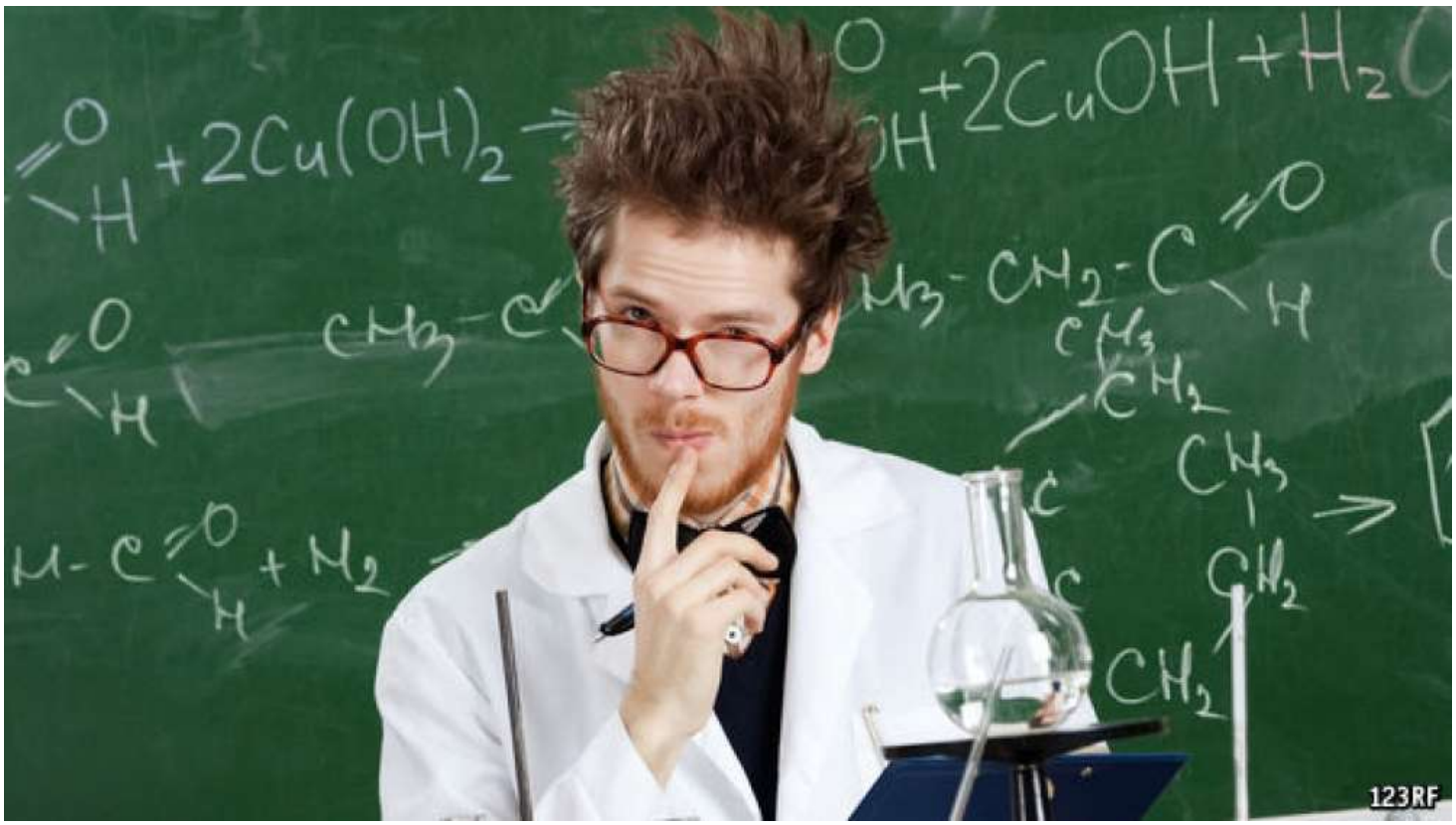
Detecting scientific sloppiness

Come again?

A surprisingly simple test to check research papers for errors



Print edition | Science and technology
Jun 16th 2016

"HOW extremely stupid not to have thought of that!" Many statisticians, confronted with the GRIM test might find themselves echoing Thomas Huxley's words when he read a the idea of natural selection. The GRIM test, short for

There is a wrinkle, though. Usually, the published value of the mean is rounded to two decimal places, for convenience. That rounding clearly affects whether the product of it and the sample size will be an integer. The GRIM test gets around this by rounding the product itself to the nearest integer (ie, 291), which is what the result would have to have been if the original numbers were accurate and the mean had not been rounded. That rounded product is then redivided by the sample size and the result of the calculation rounded to two decimal places. If this figure is not exactly the same as the original mean (and it is not, for it is 5.94) then either the original mean or the sample size is incorrect.

When Mr Brown and Dr Heathers test-drove their method on 71 suitable papers published in three leading psychology journals over the past five years, what they found justified the pessimistic sounding label they gave it. Just over half the papers they looked at failed the test. Of those, 16 contained more than one error. The two researchers got in touch with the authors of these, and also of five others where the lone errors looked particularly egregious, and asked them for their data—the availability of which was a precondition of publication in two of the journals. Only nine groups complied, but in these nine cases examination of the data showed that there were, indeed, errors.

The mistakes picked up looked accidental. Most were typos or the inclusion of the wrong spreadsheet cells in a calculation. Nevertheless, in three cases they were serious enough to change the main conclusion of the paper concerned.

That, plus the failure of 12 groups to make their data available at all, is alarming. But if knowledge that the GRIM test might be applied to their work makes future researchers less careless and more open, then Mr Brown's and Dr Heathers's maths will have paid dividends.