Statistical Significance of the New Physics Results

Prachurjya Pran Hazarika

IISER Pune, India

Statistical error is an integral part of any physical experiment, which arises due to some uncertainties or the limits of precision of the experiment. There is an inherent probability factor in particle physics experiments due to the quantum behaviour of the particles. It is therefore very important to understand whether one result is due to the actual physics, or due to random statistical fluctuations (which do happen in real life entirely by chance![1]). This is why the significance of physics results are calculated using statistical analysis. Statistical studies are also significant when we are dealing with a large amount of data. For example, you may ask how significant the effects of the new Covid-19 vaccine is, since it may not work on every individual. In this article I am going to discuss the statistical significance of two latest results from particle physics which claim to have found evidence beyond the Standard Model.

The Standard Model of particle physics comes with a "package" of elementary particles, and a set of laws that govern the interactions among them. Researchers at CERN are colliding two beams of protons that produces various particles which are detected in four giant particle detectors. The probability of how many of the various types of particles will be detected with what energy/momentum is given by the Standard Model, and we observe *almost* exactly the same. However, a recent study by LHCb detector at CERN, published in March 22, 2021, claims to have found that the fraction of electrons and muons taking part in some particular interaction processes to be different from what was predicted in the Standard Model. This problem was termed as the "*violation of lepton universality*", as it violates the rule that the class of particles called "leptons" (which includes electrons, muons and taus) must behave similarly; because the only practical difference among them is the value of their rest-masses. Another experiment, called the "*muon $g-2$ experiment*" in Fermilab, published a result on April 7, 2021, where the value of a parameter called the "Landé g-factor" for muons is found to be slightly different from what was

[1] For example sometimes you may agree with your horoscope, even though they are pretty absurd.

predicted in the Standard Model. These results are statistically significant enough to call them as “evidence” to something new, but not exactly “discoveries” to reject the Standard Model entirely. Let me first give a brief introduction to Gaussian errors, then I will discuss the standard deviation in the SM and beyond the Standard Model with the latest experimental results as example.

Gaussian Errors:

The unit of measurement when talking about statistical significance of such results is the Standard Deviation (σ), which refers to the amount of variability in a given set of data - whether the data points are clustered together, or very spread out. The mathematical definition of the Standard Deviation is well known. It is literally the root-mean-square of the set of data from the experiment, compared to the predicted set of data from the theory. We can intuitively put this as follows. For a set of N data points,

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\text{Deviation of } i\text{-th data from the mean prediction})^2}.$$

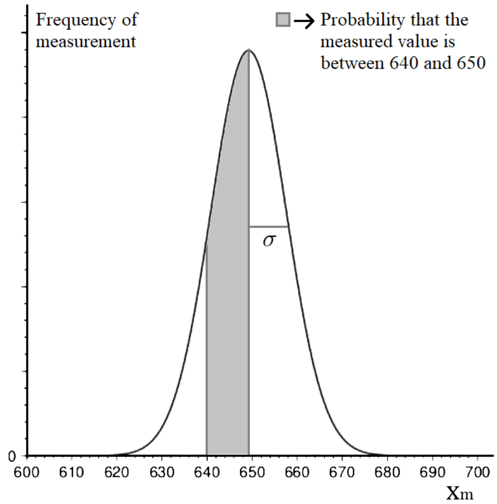
Now, you may have heard about “how much sigma significant” a particular result is. If we go through the “violation of lepton universality” paper, we see that they have found evidence of this violation with a significance of 3.1σ, and Fermilab stated that their measurement is 4.2σ away from the Standard Model. In order to understand what this means and why this is important in particle physics results, we must talk about the idea of **Gaussian Errors**. It is expected that the reader has some idea on basic calculus for this. A Gaussian distribution (also known as the Normal distribution) is described mathematically by the following expression.

$$f(x;\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

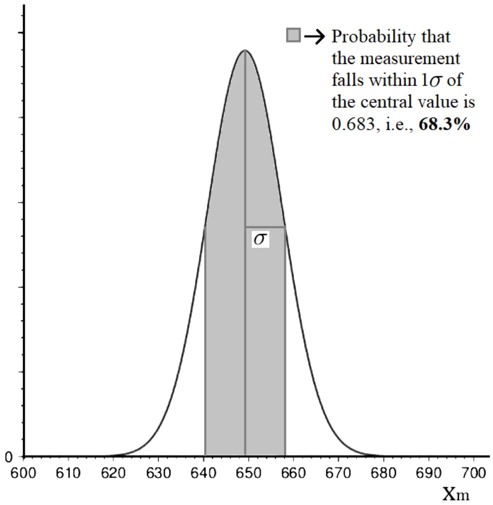
It is a function of x, and contains two parameters - the mean, μ and the Standard Deviation, σ - which determines the bell-shaped peak. The peak of the exponential function is centred at μ and σ controls the width. The numerical term in front to the exponential function is a normalisation term, which sets the area under the entire distribution (from $-\infty$ to $+\infty$) to unity.

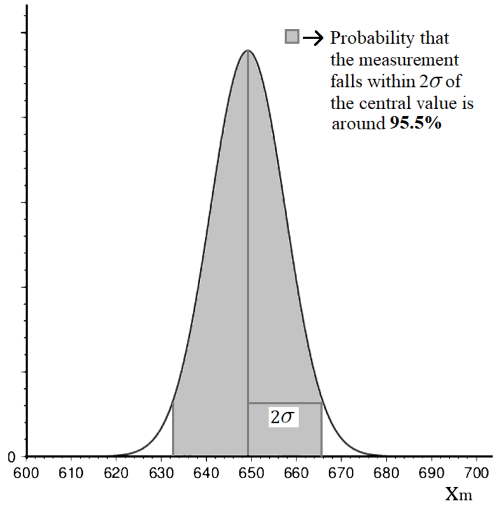
Now, let’s measure a quantity called X, whose theoretical value is X_t and measured values are X_m. **We assume that the measurement errors are Gaussian distributed.** That means, all the measured values X_m will fall around the true value X_t with the statistical frequency determined by a Gaussian distribution. Let’s discuss a very specific case to understand this. The following example is taken from the YouTube channel Think Like a Physicist.

Suppose the quantity, X has a true value of 649.2 with a measurement uncertainty 8.33. This is typically written as 649.2 ± 8.33. The plot shows the distribution of X_m around the "true value" $X_t = 649.2$ with the standard deviation $\sigma = 8.33$. Here is the key idea. **The probability that a particular measurement takes the value in a particular range is given by the area of the distribution in that range.** For example, the probability that the measured value is between 640 and 650 is the area under these two points. This is why, the area under the whole Gaussian distribution is unity by design, as the probability of measuring any value from $-\infty$ to $+\infty$ must be 1. We can see that the probability of getting very ridiculous results away from the centre at 649.2 (such as measuring a value between 690 and 700) is highly unlikely. You are most likely to measure something close to the true value.



The idea that the measured values are Gaussian distributed around the true value is very powerful and significant. Any measurement which is away from the true value is less likely to occur, and therefore, on repeating the experiment multiple times, we get a set of measured values which are distributed around the true value. The important thing to note here is that, in a repeated experiment, the unlikely results may still appear. But there is nothing to worry here. Such measurements do not necessarily "break" your theory; they are just **less significant**! Before going to the discussions of disagreements with the Standard Model, let's see how the data in our example agrees with the true value. From the definition of the Gaussian distribution, we can now calculate the probability of the measured value falling within 1σ of the central value. For this, we simply integrate the Gaussian from $\mu - \sigma$ to $\mu + \sigma$, and arrive at a result close to 0.683, i.e, 68.3%. In our example, $\mu = X_t = 649.2$ and $\sigma = 8.33$, but this result is independent of these particular values. Similarly, the probability of the measurement falling within 2σ, 3σ, 4σ and 5σ of the central value respectively are 95.5%, 99.7%, 99.994% and 99.999943%. That's why, a measurement which is 5σ away from the central value is extremely unlikely!

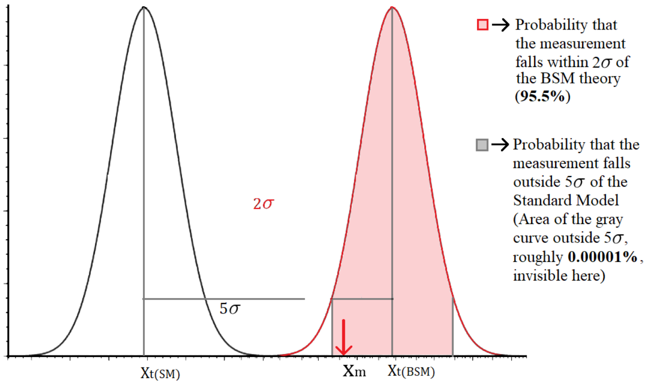




5-Sigma Deviation from the Standard Model:

Since we have established the idea of Gaussian errors, we are ready to talk about what we meant by a 5σ discovery or observation. For this, we take two models of physics explaining the particle physics phenomena - the Standard Model (SM), and a theory Beyond the Standard Model (BSM). In today's particle physics, the Standard Model of elementary particles is the current best model for particle interactions. Its predictions agree with the results of literally thousands of experiments with very high precision. However, it leaves a lot of unanswered questions (such as neutrino mass, dark matter, matter-antimatter asymmetry etc.). That's why we are motivated to go for "New Physics", also termed as Beyond the Standard Model theories, to answer some of these questions, which have to be consistent with *all* the experiments to date. They should make testable predictions which differ from those of the SM. Sometimes experimental results disagree with the predictions of the SM. This may mean that either there is some problem with the experiment, or the data has large statistical fluctuation, or we might have discovered New Physics!

Now, let's consider an example. Suppose we are constructing an experiment to test the BSM theory. Since SM has already passed a wide array of experimental tests and has proven to be a good working hypothesis for explaining a wide range of observations, and we consider it our default hypothesis while making predictions about the experiment. After the experiment, if we concluded that BSM is correct, that would be a significant scientific advance. The BSM theories predict the existence of some new elementary particles, in addition to the SM particles. Therefore, in our experiment, we must determine whether what we see in the data results are from the production of SM particles only, or also from the particles predicted from the BSM theories in addition to the regular SM particles. Suppose we are focusing on measuring a particular parameter X (which could be an interaction cross-section or the Landé g factor or anything else). Both of these theories predict some value for X_t. We label them as $X_{t(\mathrm{SM})}$ and $X_{t(\mathrm{BSM})}$ respectively. We perform the experiment and obtain a value X_m. We will assume that our measurement errors are Gaussian distributed, with an uncertainty σ.

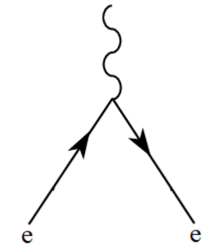


If the measured value X_m is 5σ away from $X_{t(\mathrm{SM})}$, and within 2σ of $X_{t(\mathrm{BSM})}$, then we can say that our measurement is more compatible with BSM. Because, the probability of the measurement being 5σ away from any prediction is of the order of 0.00001%, while staying within 2σ of any prediction is 95.5%. Therefore, in this scenario, there is 95.5% chance that the BSM theory is correct, and there is 0.00001% chance that the SM still holds. This is a situation in which many physicists would consider the term "discovery" appropriate. In fact, this 5σ disagreement with the Standard Model has become a norm for using the term "discovery" in experimental particle physics papers. This case can also be alternatively stated as "a discrepancy compatible with the BSM has been observed at 5σ". Sometimes the result may disagree with SM up to 5σ, but may not be compatible with the new physics. In that case, physicists refuse to call it a discovery of new physics.

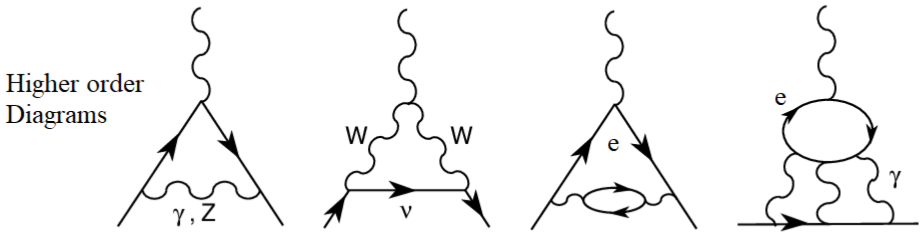
Measured Value of Muon g-2 in Fermilab:

When Paul Dirac was developing Quantum Electrodynamics (QED) - a theory of electromagnetic interactions in the mathematical framework of Quantum Field Theory (QFT), he considered the very simple, trivial interaction of electrons with the electromagnetic field to calculate a parameter called the Landé g factor, also known as *anomalous magnetic dipole moment*. This parameter becomes relevant when a charge particle interacts in a magnetic field. Subatomic particles have inherent magnetic spins, and can be thought of as small rotating magnets. In presence of a magnetic filed, they orient themselves in specific ways, and gains or loses energy. The Landé g factor is a multiplicative term appearing in the energy levels of particles in such a magnetic field. It depends on how particles interact with the electromagnetic field, and therefore is a very handy parameter to test the known interactions of matter. Paul Dirac found it to be exactly 2 in his trivial calculations.

However, later it was established that there are "higher order" processes with much complicated looking Feynman diagrams which also contributed to this g factor. There are literally infinite such processes in the Standard Model, but the contributions to g from the higher order terms reduces drastically. Therefore, the infinite series for the g factor becomes a convergent quantity, which is slightly greater than 2. This deviation from 2 corresponding to the higher order diagrams in represented as $g-2$. The muon $g-2$ experiment at Fermilab is designed to calculate this $g-2$ factor for muons.

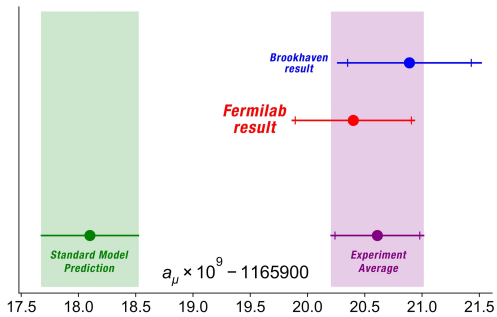


Trivial Diagram for g measurement



Therefore, the $g-2$ term represents a very small value corresponding to those higher order diagrams. In case of electrons, the experimental value of $g-2$ agrees remarkably, up to 1 part in a billion, with the Standard Model predictions. The SM prediction for half of $g-2$ for electrons (labelled as a_e) is 0.00115965218**1643(±764)**, and the experimental average is 0.00115965218**073(±28)**. These two results agree well within the statistical error, and there is no issue. This is by far the most accurate predictions in all of physics. However, the same measurements for the muons are apparently not agreeing that much.

During 1990s, the Brookhaven National Laboratory conducted an experiment to measure the half of Landé g of muons. They tried to calculate the half of $g-2$ (labelled as a_μ) which was predicted to be 0.00116591804(±51) in the SM, but it was found to be an average value of 0.0011659209(±6) instead, which is significant enough to worry about. For the last 4-5 years, Fermilab independently repeated the experiment, and by combining the data with the previous experiments, they published a more accurate value of 0.00116592040(±54) on April 7, 2021. This new and improved result is 4.2σ away from the Standard Model prediction. **But does it suggest the end of the Standard Model?**



A disagreement of 4.2σ means, there is a chance that the results are a statistical fluctuation of about 1 in 40,000. That means, there is about one in 40,000 chance that this result is purely by chance! However, new physics theories are constantly being developed to account for this disagreement. Since the muon is about 200 times heavier than the electron, therefore it is more likely to undergo the "higher order" interactions which may even involve any undiscovered *exotic* particles. Such exotic interactions are thought to be more sensitive in case of muons, and this is why this experiment is more exciting than the one involving electrons. There is a possibility that we might have ignored some of the Feynman diagrams involving these undiscovered BSM particles which also contributes to the value of g for muons. For example, one such potential new theory involves a new kind of particle called a "lepto-quark" which contains properties of both leptons and quarks. **Because of the 5σ convention in particle physics, researchers are calling this result an *evidence* for a potential new physics theory, but not a *discovery* of new physics.** At this point one may conclude as follows.

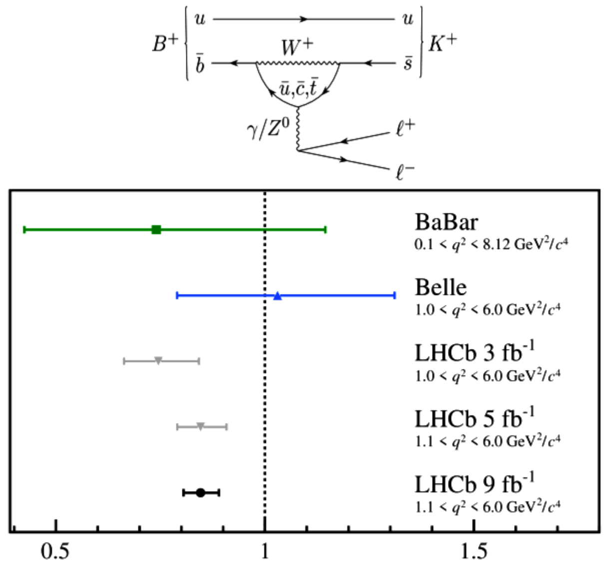
- *Either* the Standard Model is correct but this is a highly unusual result (1 in every 40,000), which just happened purely by chance because of the statistical uncertainties,
- *Or*, this is a significant deviation from the SM, so we must be looking for new theories such that this result is one of the common predictions of those theories.

Once physicists gather more data so that the uncertainties are low enough, this result may have higher significance than before. If the combined data from the future experiments deviates from the SM by more than 5σ, then, by convention, we may simply ignore the 1 lucky result in every 1,744,278+ fluctuations, and call it a discovery of new physics. Physical significance of this discovery may open up to new portal for new forces or new interactions, which dictates the deviation of anomalous magnetic moment from its SM behaviour. That would be really groundbreaking!

Violation of Lepton Universality in LHCb:

In the Standard Model, the leptons, i.e, electrons, muons and taus are very similar except for their masses (approximately 0.5 MeV, 100 MeV and 1777 MeV respectively). This is why they are expected to behave in a similar manner. For example, an 80 GeV Z^0 boson is equally likely to decay to a pair of muons ($Z^0 \to \mu^+\mu^-$), as well as a pair of electrons ($Z^0 \to e^+e^-$), since the rest-mass of the electrons and muons are negligible compared to the gigantic rest-mass of the Z^0 boson. (However, tau being comparatively much heavier, its mass is not negligible in this case.) In general, there is a sort of "universality" in the kind of interactions these leptons take part.

In a recent study of b-quark decays in the LHCb detector at CERN ($b \to sl^+l^-$, where $l = e, \mu$), physicists compared the fraction of muon pairs to the fraction of electron pairs produced in the interactions, and found a significant mismatch (even after taking their difference in mass into consideration). This result is a 3.1σ deviation from the Standard Model. Now we understand that this result is even less significant than the Fermilab result, therefore nothing to worry here. Interestingly, similar BSM theories like "lepto-quarks" along with other theories have been hypothesised in order to account for this disagreement. Tackling with b-quarks is difficult, and we will have to wait for future experiments in order to fully understand the importance of this result.



Standard Model predicts that the ratio parameter should be exactly 1, but the final result (at the bottom) is slightly less than 1 with a 3.1σ deviation.

Additional Information:

The following is an interesting table from Wikipedia which illustrates how unlikely it is to be away from a mean value by certain σ level. Pritam Das (research scholar in Theoretical High Energy Physics, Tezpur University) helped me with some of the corrections in this article.

Range	Expected fraction of population inside range	Approximate expected frequency outside range	Approximate frequency for daily event
μ ± 0.5σ	0.382 924 922 548 026	3 in 5	Four or five times a week
μ ± σ	0.682 689 492 137 086	1 in 3	Twice a week
μ ± 1.5σ	0.866 385 597 462 284	1 in 7	Weekly
μ ± 2σ	0.954 499 736 103 642	1 in 22	Every three weeks
μ ± 2.5σ	0.987 580 669 348 448	1 in 81	Quarterly
μ ± 3σ	0.997 300 203 936 740	1 in 370	Yearly
μ ± 3.5σ	0.999 534 741 841 929	1 in 2149	Every 6 years
μ ± 4σ	0.999 936 657 516 334	1 in 15 787	Every 43 years (twice in a lifetime)
μ ± 4.5σ	0.999 993 204 653 751	1 in 147 160	Every 403 years (once in the modern era)
μ ± 5σ	0.999 999 426 696 856	1 in 1 744 278	Every 4776 years (once in recorded history)
μ ± 5.5σ	0.999 999 962 020 875	1 in 26 330 254	Every 72 090 years (thrice in history of modern humankind)
μ ± 6σ	0.999 999 998 026 825	1 in 506 797 346	Every 1.38 million years (twice in history of humankind)
μ ± 6.5σ	0.999 999 999 919 680	1 in 12 450 197 393	Every 34 million years (twice since the extinction of dinosaurs)
μ ± 7σ	0.999 999 999 997 440	1 in 390 682 215 445	Every 1.07 billion years (four occurrences in history of Earth)
μ ± xσ	$\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)$	1 in $\frac{1}{1-\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)}$	Every $\frac{1}{1-\mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)}$ days

References :

[1] MIT news : David L. Chandler, MIT News Office. Explained: Sigma

[2] Lepton Universality Violation paper : LHCb Collaboration, Aaij, Roel and Beteta et al. Test of lepton universality in beauty-quark decays. *arXiv preprint arXiv:2103.11769*

[3] Fermilab Muon g-2 Results : B. Abi et al. (Muon g-2 Collaboration). Measurement of the Positive Muon Anomalous Magnetic Moment to 0.46 ppm. *Phys. Rev. Lett. 126, 141801*, 2021

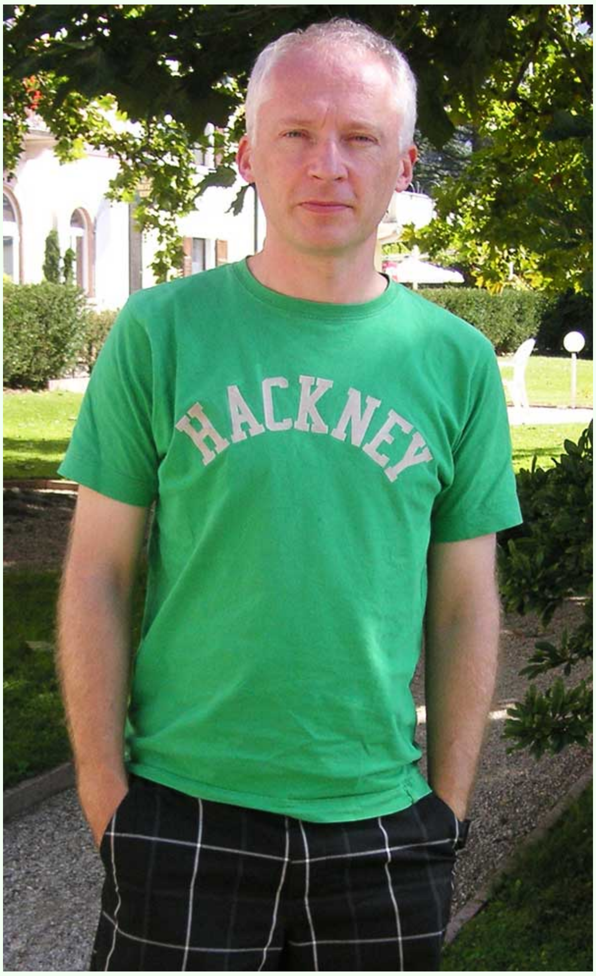
Related YouTube Videos :

[1] Fermilab : Scientific Seminar: First results from the Muon g-2 experiment at Fermilab
[2] Fermilab : Muon g-2 experiment finds strong evidence for new physics (outline)
[3] Fermilab : The physics of g-2
[4] PBS Space Time : Why the Muon g-2 Results Are So Exciting!

"For me, science is about discovery but it is also about communication. A scientific discovery barely exists until it is communicated and brought to life in the minds of others.

I am deeply committed to continuing my own research which seeks to uncover some of the deep eternal mysteries of number theory and symmetry. But at the same time, I am passionately dedicated to giving as many people as possible access to the exciting and beautiful world of mathematics and science that I inhabit and revealing to them why it is such a powerful way to understand the world.

A mathematically and scientifically literate society is essential given the huge role science now plays in our world. But my belief is that message can A mathematically and scientifically literate society is essential given the huge role science now plays in our world. But my belief is that message can best come from someone actively involved at the cutting edge of their science."



– Marcus du Sautoy FRS