Some swans are grey

Scientific truth isn't as black and white as some claim. So why not change the definition of science to match the reality, suggests Robert Matthews

WHAT is the difference between astronomy and astrology? That's easy: astronomy is the scientific study of celestial objects, while astrology is a load of hokum. Anyone with the most basic understanding of science knows why. Astronomy passes the acid test of real science: its claims are always capable of being debunked – in other words, they are falsifiable.

Identified as the defining characteristic of real science by the philosopher Karl Popper more than 70 years ago, falsifiability has long been regarded by many scientists as a trusty weapon for seeing off the menace of pseudoscience.

The late Viennese thinker has been lauded as the greatest philosopher of science by the likes of Nobel prizewinning physicist Steven Weinberg, while Popper's celebrated book *The Logic of Scientific Discovery* was described by cosmologist Frank Tipler as "the most important book of its century".

Times change, though. Popper's definition of science is being sorely tested by the emergence of supposedly scientific ideas which seem to fail it. From attempts to understand the fundamental nature of space-time to theories purporting to describe events before the big bang, the frontiers of science are sprouting a host of ideas that are seemingly impossible to falsify.

So should the pursuit of such mind-boggling ideas be condemned as pseudoscience, or should scientists be more relaxed about falsifiability? It's a debate that's dividing the scientific community. Some are in no doubt about where they stand. "I never would have believed that serious scientists would consider making the kinds of pseudoscientific claims now being made," says theorist Peter Woit of Columbia University, New York, author of *Not Even Wrong*, a biting critique of current fashions in theoretical physics. For Woit, attempts to water down the falsifiability criterion are "an outrageous way of refusing to admit failure".

His bête noire is the recent explosion of interest in the multiverse, an infinite yet unobservable ensemble of universes of which our cosmos is supposedly just one part. "The basic problem with the multiverse is not only that it makes no falsifiable predictions, but that all proposals for extracting predictions from it involve massive amounts of wishful thinking," Woit says.

Others believe such criticism is based on a misunderstanding. "Some people say that the multiverse concept isn't falsifiable because it's unobservable – but that's a fallacy," says cosmologist Max Tegmark of the Massachusetts Institute of Technology. He argues that the multiverse is a natural consequence of such eminently falsifiable theories as quantum theory and general relativity. As such, the multiverse theory stands or fails according to how well these other theories stand up to observational tests.

In the meantime, says Tegmark, exploring the idea of the multiverse is no more pseudoscientific than pondering phenomena inside a black hole – another consequence of general relativity whose interior is just as unobservable as the multiverse.

In any case, dismissing a theory on the grounds that it fails Popper's acid test itself involves a huge leap of faith, says cosmologist Lawrence Krauss at Case Western Reserve University in Cleveland, Ohio. "You just can't tell if a theory really is unfalsifiable."

He cites the case of an esoteric consequence of general relativity known as the Einstein ring effect. In a paper published in 1936, Einstein showed that the light from a distant star can be distorted by the gravitational field of an intervening star, producing a bright ring of light around it. It was a spectacular prediction but also, Einstein said, one that astronomers stood "no hope of observing", as the ring would be too small to observe.

For all his genius, Einstein had reckoned without the ingenuity of astronomers, which in 1998 led to the discovery of the first example of a perfect Einstein ring – created not by a star, but by a vast galaxy billions of light years away.

Krauss admits he has fallen into the same trap, applying the falsifiability criterion to decide whether some or other idea is really "scientific" enough to be worth publishing. "I've decided not to write papers because I thought the claims would never be falsifiable, and yet [they] turned out to be so."

Still, for many scientists, Popper remains the only philosopher with any relevance to what they do. Much of his appeal rests on the clear-cut logic that seems to underpin the concept of falsifiability. Popper illustrated this through the now-celebrated parable of the black swan.

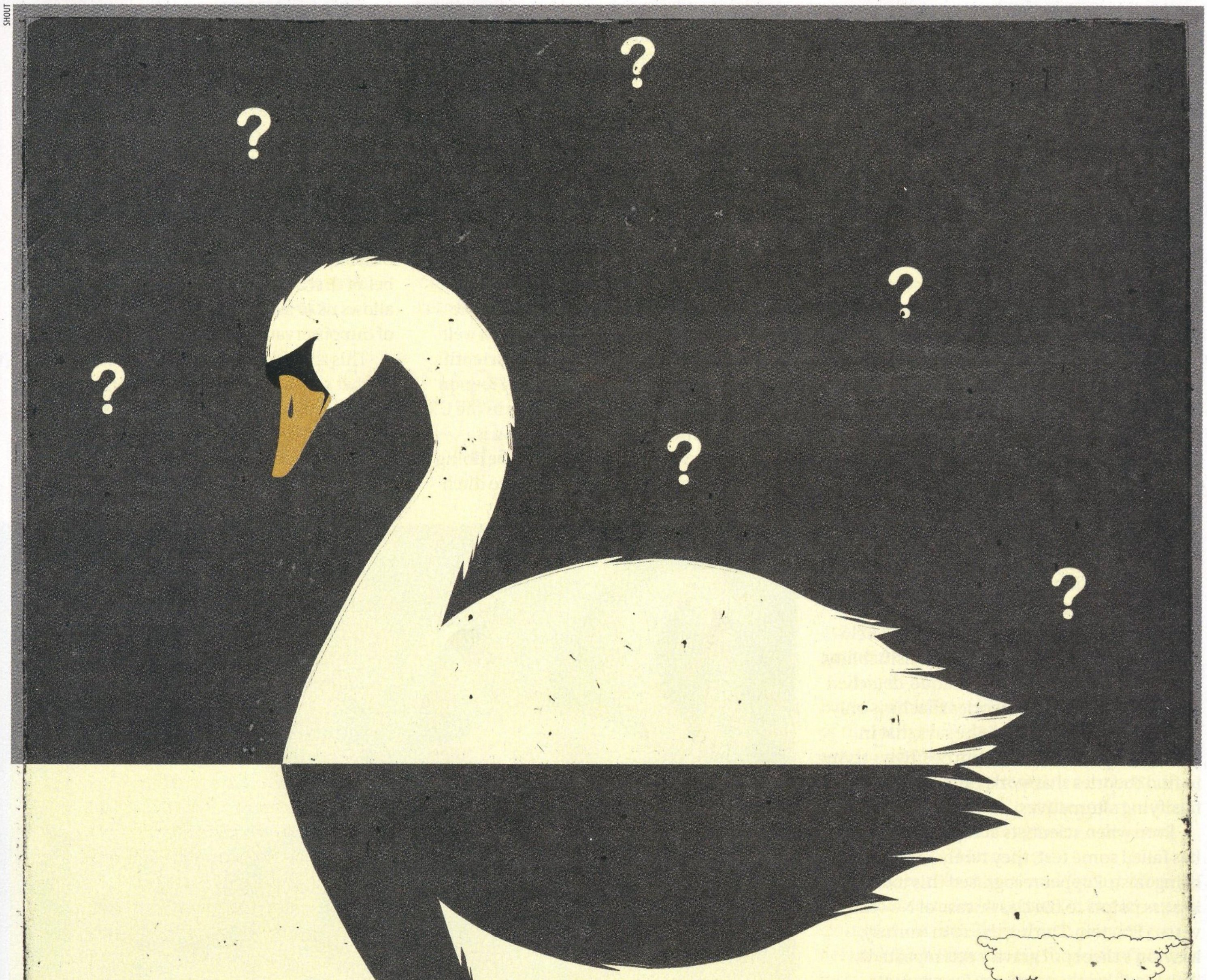
Suppose a theory proposes that all swans are white. The obvious way to prove the theory is to check that every swan really is white – but there's a problem. No matter how many white swans you find, you can never be sure there isn't a black swan lurking somewhere. So you can never prove the theory is true. In contrast, finding one solitary black swan guarantees that the theory is false. This is the unique power of falsification: the ability to disprove a universal statement with just a single example – an ability, Popper pointed out, that flows directly from the theorems of deductive logic.

Popper went on to promote falsification as the essence of the scientific process, with the search for falsifiable predictions being the distinguishing feature between science and pseudoscience. Yet even at the time there were concerns his criterion wasn't up to the job.

The most obvious objection is that astrologers, soothsayers and quacks also make falsifiable statements – but that doesn't make them scientific. Yet could it be their cavalier attitude towards negative evidence that marks them out as pseudoscientific?

Worryingly, this doesn't work either, as was made clear over a century ago by the French philosopher and physicist Pierre Duhem. He pointed out that the predictions of a scientific theory often rest on a raft of other assumptions underpinning how the theory is tested. If an experiment seems to falsify the theory, it is often possible to pin the blame on one of these "auxiliary hypotheses" rather than the theory itself. ▶

SHOUT



This happens quite a lot in science. In fact, in the very year Duhem put forward his objections to falsification, experiments by a German physicist appeared to falsify Einstein's then-new special theory of relativity, lending support to rival theories. Yet Einstein blithely dismissed the results, saying the other theories were simply less plausible than his own.

He was hardly the last scientist to reject inconvenient results – as Popper was forced to admit. Even so, he remained convinced that at least looking for falsifiable consequences was the essence of doing science.

For Woit, it's precisely the absence of progress in finding such consequences of the multiverse theory that makes it pseudoscience. "If all you have to show is wishful thinking about the possibility of such progress, then you're not really doing science," he says.

Yet according to philosopher Rebecca Goldstein of Harvard University, this just highlights the idealistic view of scientists underpinning Popper's criterion: "Not only does Popper maintain that science as a field is unique, its borders fortified by falsifiability, but also that the scientist is unique, detached enough from his own theories that he is only out to shoot them down." She says that in reality the process is far more positive – trying to find theories that work, rather than falsifying alternatives.

Even when scientists accept that a theory has failed some test, they rarely junk it as being false. Popper recognised this too. Krauss points to the classic case of Newton versus Einstein. During the 20th century, Newton's theory of gravity was repeatedly "falsified" by observations: for example, by predicting only half the observed bending of light by the sun's gravitational field. Yet scientists are not about to ditch Newton any time soon, as his laws work perfectly well in everyday situations. "This is something we don't make clear enough," says Krauss. "We don't have true theories; we only have effective theories."

So after all these concessions, what remains of Popper's supposedly hard-and-fast criterion? It's hard to apply in practice, too vague to differentiate science from pseudoscience and bears little resemblance to what scientists really do. Why does it remain so popular? "Scientists like simple methodological theories which accord well with what they consider to be good scientific reasoning," says philosopher Colin Howson of the London School of Economics in the UK.

So if the simplicity of falsification is misleading, what should scientists be doing instead? Howson believes it is time to ditch Popper's notion of capturing the scientific process using deductive logic. Instead, the focus should be on reflecting what scientists actually do: gathering the weight of evidence for rival theories and assessing their relative plausibility.

Howson is a leading advocate for an alternative view of science based not on simplistic true/false logic, but on the far more subtle concept of degrees of belief. At its heart is a fundamental connection between the subjective concept of belief and the cold, hard mathematics of probability.

Talk of probabilities usually conjures up images of random events such as coin tosses, with the formulae of probability theory answering questions about the chances of getting, say, 20 heads from 30 tosses. That's not the only way to look at probability theory, though. It is also possible to turn it on its head and ask a far more interesting question: what are the chances that a coin really is dodgy, given we've seen 20 heads from 30 tosses? In other words, if we have a hypothesis – like the belief that a coin is dodgy – probability theory allows us to assess that hypothesis in the light of our observations.

This should sound familiar; after all, it is what scientists do for a living. And it is a view of scientific reasoning with a solid theoretical basis. At its core is a mathematical theorem, which states that any rational belief system obeys the laws of probability – in particular, the laws devised by Thomas Bayes, the 18th-century English mathematician who pioneered the idea of turning probability theory on its head.

Unlike Popper's concept of science, the Bayesian view doesn't collapse the instant it comes into contact with real life. It relies on the notion of accumulating positive evidence for a theory which, according to Tegmark, is what scientists really spend their time doing. "What we do in science isn't falsifying, but 'truthifying' – building up the weight of evidence," he says.

"The multiverse is no more pseudoscientific than the inside of a black hole"

The Bayesian approach quantifies this practice. Scientists begin with a range of rival explanations about some phenomenon, the observations come in, and then the mathematics of Bayesian inference is used to calculate the weight of evidence gained or lost by each rival theory (*New Scientist*, 22 November 1997, p36). Put simply, it does this by comparing the probability of getting the observed results on the basis of each of the rival theories. The theory giving the highest probability is then deemed to have gained most weight of evidence from the data.

Chief among them is that, while Bayesian methods show how observations add weight of evidence to initial beliefs or theories, they say nothing about what those initial beliefs should be. And if a theory is completely new, the beliefs behind it may be based on nothing but subjective intuition.

Advocates of the Bayesian approach point out that such prior beliefs typically become less important as the results accumulate. In other words, Bayesianism confirms another maxim of scientists: that as the observations come in, the truth will out. Wrong-headed initial beliefs are never totally falsified, but they do end up buried by the sheer weight of evidence against them.

"The frontiers of science are sprouting a host of ideas seemingly impossible to falsify"

It captures many other features of real-life science too. For example, it shows that seemingly implausible theories require a hefty weight of evidence before they can be taken seriously – reflecting that familiar maxim that "extraordinary claims require extraordinary evidence". The Bayesian view also gives vague or contrived theories that fit pretty much any data set a tough time in the quest for credibility.

With its mathematical rigour and natural fit with real-life science, it's an approach that now commands the attention of many philosophers of science. "The most interesting views these days are to be found in Bayesianism. It's where much of the current research impetus is directed," says philosopher Robert Nola of the University of Auckland in New Zealand. He adds, though, that the approach is not without its problems.

It is not just philosophers of science who see Bayesianism as the way forward: so do working scientists in fields from archaeology to zoology. Among the proponents of this view are cosmologists, who are now using Bayesian methods to extract the most plausible model of the universe from signals flooding in from observatories. One of their prime roles is constraining speculation and deciding whether current theories are compatible with observations, or if some extra ingredient is needed.

Take the mysterious force said to be driving the ever-faster expansion of the universe. Theorists are exploring the idea that this "dark energy" may have varied over the course of cosmic history, rather than stayed constant. Such ideas might keep theorists in work but they also make for a more complex model of the universe, says Andrew Liddle at the UK's University of Sussex in Brighton. "The question is whether the observational data support a simple or a complex model."

He and his colleagues have applied Bayesian methods to assess the plausibility of the intriguing idea of varying dark energy and found that the standard model with constant dark energy remains a far better bet. That could change, but the smart money is on variable dark energy being a dead end (*New Scientist*, 8 March, p 32).

Talk about "best bets" and "smart money" might not sound very scientific, but it's much closer to how real-life research priorities are decided. With Bayesian methods, that process is captured in rigorous, quantitative detail – the black and white of falsification being replaced with the shades of grey of the real world. "I think it's absolutely the way to go," says Liddle.

So where does all this leave the debate about whether concepts like the multiverse are really scientific? According to Howson, the multiverse is entirely scientific in Bayesian terms, as it is based on theories carrying huge weights of evidence. "If Popper condemns it as pseudoscience because it is 'unfalsifiable' – and it may not always be – then so much the worse for Popper."

But whatever one regards as the essence of science – black-and-white falsification or subtle shades of grey – in the end it is still empirical observations that decide if a theory gets taken seriously. "At some level, you cannot give up the idea of falsification," says Krauss. "Rumours of the death of science have been greatly exaggerated." ●

Robert Matthews is visiting reader in science at Aston University, Birmingham, UK