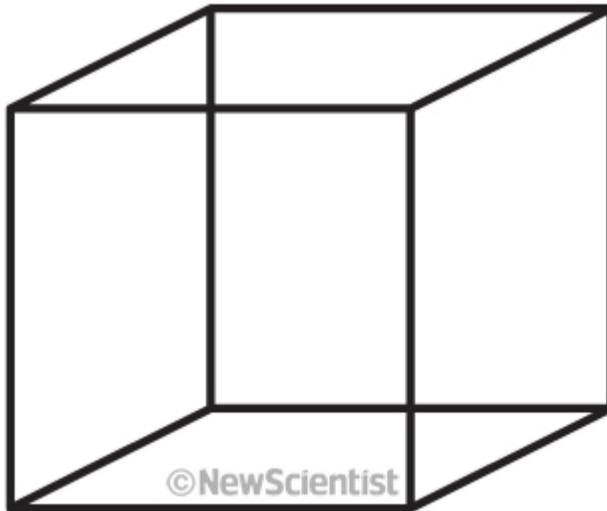


ESP evidence airs science's dirty laundry

- 11 January 2012 by [Bob Holmes](#)
- Magazine issue [2847](#). [Subscribe and save](#)



Take a look at this Necker cube, which can be seen in two different ways. How do you see it? From below or from above?

Now click either [here](#) or [here](#). If precognition is real, the image you are about to see should affect the way you have just perceived the illusion.

A barrage of experiments seems to show that we can predict the future – but they may tell us more about the scientific method

FEW sounds quicken the pulse like the clatter of the roulette ball as it drops. Fortunes can be won or lost as it settles into its numbered slot. Find a way to predict where it will come to rest, and you would soon become the envy of every other gambler and the worst nightmare of every casino.

Michael Franklin thinks he might be able to make that claim. Over thousands of trials, he seems to have found a way to predict, with an accuracy slightly better than chance, whether the ball will fall to red or black. You'll find no divining rods or crystal balls about Franklin's person. Nor does he operate from a murky tent swathed in lace and clouded with the fumes of burning incense; he works in a lab at the University of California, Santa Barbara.

Franklin is one of a small group of psychologists who are investigating precognition - the ability to foretell the future. Astonishingly, some of the groups, including Franklin's, are returning positive results. "I still want to see that I can actually win money with this," says Franklin, who rates his confidence in the data so far at about 7.5 out of 10. "As a scientist, I need to be agnostic."

If precognition does turn out to be real, it will shake the foundations of science and philosophy. Few researchers will be putting money on this conclusion, though; most expect that the puzzling results will begin to evaporate as others attempt to repeat the experiments. Even so, that could change science as we know it. Franklin and his colleagues are all using standard research methods that normally go unquestioned. If those methods can lead respected scientists to such startling errors, how many other studies might be similarly flawed?

A closer look

Despite widespread scepticism from mainstream scientists, studies of precognition and other forms of extrasensory perception crop up time and again. In the 1940s it was card-guessing, in the 1980s the ability to influence random-number generators, and in the last decade so-called "presentiment" - in which volunteers showed changes in skin conductance just before they saw disturbing images. In every case, however, independent researchers failed to repeat the initial results, eventually concluding that they were the result of procedural flaws or coincidence.

However, a year ago even the sceptics had to take a closer look when Daryl Bem, a well-respected psychologist at Cornell University, New York, reported some positive results in the *Journal of Personality and Social Psychology* ([vol 100, p 407](#)). "When Daryl Bem speaks, we listen," says Jeff Galak, of Carnegie Mellon University in Pittsburgh, Pennsylvania.

The appeal of Bem's work is that he used well-accepted psychological experiments, but with a twist. Countless studies, for example, have established that writing out a list of words makes it easier to recall those words later. Bem merely switched the order of the two events. His subjects viewed the list briefly and were tested on their initial recall. They were then given a smaller, random selection of words from the same list, which they were asked to type out and memorise. Surprisingly, they were more likely to have remembered these words during the initial memory test - *before* they had even seen the list. The difference between the two sets of words was slight - only 2.27 per cent - but statistically significant. Another of Bem's experiments focused on a psychological effect known as habituation; if volunteers are asked to choose between two similar images, they will tend to prefer an image they have seen before over one they have not. Again, Bem reversed the sequence - he showed subjects two new images, and asked them to choose which one they liked better, before showing them one of the images again immediately afterwards. Bem found that about 54 per cent of the time, people preferred the image they would end up seeing again later.

He reported nine such time-reversed experiments, in eight of which later events appeared to influence people's earlier behaviour. The effects were subtle, typically swaying the results by 2 to 4 per cent, but Bem's statistical tests suggested that they were unlikely to have occurred by chance. Taken together, the experiments strongly suggested that people could indeed "feel the future", Bem concluded.

Since Bem's study appeared, two other researchers presented similar results at the Towards a Science of Consciousness conference in Stockholm, Sweden, in May 2011. Franklin was one of them. He asked volunteers to identify certain complex geometric shapes, some of

which they would use for practice later on. In line with Bem's results, their performance on the first task seemed to be affected by the shapes they saw in the later practice run.

Franklin wondered whether he could adapt his method to make more useful predictions. Suppose the choice of practice shapes was determined by the spin of a roulette wheel, for example. He could then use the subject's test results to predict whether the ball would fall on red or black ahead of time.

"If you know which practice condition they're going to get, then you can say OK, I'm going to predict that the roulette spin is going to be red," Franklin says. Sure enough, when he actually tried this he correctly called the fall of a roulette ball in another room about 57 per cent of the time - enough to make a tidy profit, had he bet real money on a game.

Dick Bierman, a physicist and cognitive psychologist at the University of Amsterdam in the Netherlands, meanwhile, showed volunteers a standard optical illusion known as a Necker cube - a two-dimensional drawing of a cube that appears to alternate between top and bottom views (see diagram). The subjects had to indicate which view they saw and when their perception switched.

Then Bierman showed each volunteer a filled-in drawing of a cube that was unambiguously either top or bottom view. Those who later saw the top view spent more time perceiving the Necker cube in top view, while those who later saw the bottom view tended to see the Necker cube in bottom view, he found. It was as if their initial view predicted what they would be shown the second time. As with Bem's study, the difference was slight - about 3 per cent - but unlikely to have occurred by chance.

At a first glance, these experiments appear to blow the conventional notion of cause and effect out of the water. It would suggest your hunches may really tell you something about the future. Perhaps you could even "remember" next week's lottery results.

Surprisingly, none of this would break the laws of physics as we know them, since the equations of physics are mostly "time-symmetric", notes Daniel Sheehan, a physicist at the University of California, San Diego. "If you were to say the past influences the present, everyone would say OK. But you can equally say that the future boundary conditions affect the present," says Sheehan, who organised a symposium on reverse causation in 2011 for a regional meeting of the American Association for the Advancement of Science in San Diego. "The future has an equal say in the present as does the past."

Cracks appear

Indeed, even some physics experiments can be interpreted as causation moving backward in time. A photon passing through one or two narrow slits will behave as either a wave or a particle, depending on how many slits are open - even if the slits are adjusted after the photon has passed through ([Science, vol 315, p 966](#)). The experiment can also be explained without invoking reverse causation, but the arguments tend to be tortuous, says Sheehan. The vast majority of scientists won't be embracing these interpretations just yet. "Before you jump to attempts to accommodate these phenomena in physics, show us that the

phenomena are there," says James Alcock, a psychologist at York University in Toronto, Canada. And if you take a closer look at the latest studies, cracks begin to appear in their seemingly convincing facade.

To begin with, the statistical analyses, which are meant to weed out results that could have occurred by chance, may in fact be hiding false positives. Using the standard technique, researchers try to estimate the probability of their results occurring under a so-called "null hypothesis" that nothing unusual is going on - in this case, the idea that precognition does not occur. If this probability is suitably small, they conclude that an alternative hypothesis must be true instead - in these experiments, the idea that precognition really exists.

But nowhere in this process do the experimenters calculate how likely the observations would be under the alternative hypothesis. When you do perform this calculation, using a different method known as Bayesian statistics, the results can be disconcerting. Suppose you toss a coin a 1000 times, and end up getting 527 tails. The odds of getting this far away from a 50:50 divide using a fair coin are a little less than 1 in 20. It's tempting to think this means the odds are 20 to 1 that the coin is biased - but that's not the case, says Jeff Rouder, a psychologist at the University of Missouri in Columbia. Even the best alternative - a coin that gives tails 52.7 per cent of the time - yields a probability of just 0.025 of getting exactly 527 tails, which is only a four-fold improvement over the odds of producing exactly 527 tails with a fair coin. "In other words, this new alternative is not much more probable than the one I rejected," says Rouder. Other alternatives - that the coin comes up tails 55 per cent of the time, say - offer even less of an edge over the null.

Looking at Bem's data in this way, Bayesian reanalyses by Eric-Jan Wagenmakers, a mathematical psychologist at the University of Amsterdam, and Rouder found the case for precognition to be unconvincing. Such analyses have their own issues, though, since Bayesian statisticians must initially estimate the size of the possible effect - in this case, how much precognition should sway the results of the tests. Bem argues that Wagenmakers's estimate was too high. His own Bayesian analysis concluded that his experiments, taken as a group, provided odds of about 13,000 to 1 in favour of precognition.

Whatever the outcome of this debate, most sceptics agree that another factor - prevalent throughout science - may have boosted the chances of producing these puzzling results. The issue arises from the difficulties of setting up a study, when you choose which variables to measure, how many samples to take, and which confounding factors to consider.

Ideally, all these decisions are made ahead of time, before any measurements have been collected. But in practice, experimenters often wing it, adjusting their techniques based on the results they see. For example, a researcher might measure two related variables, then report the one that gives the clearer (that is, more significant) result - providing two places for a false positive to pop up. The common practice of "optional stopping", which involves analysing the data, then collecting more if the result is not quite significant, can also sway the chances of a false result.

Torturing the data

When researchers pick over their results in this way, they are in essence "torturing the data until they confess", says Wagenmakers. The effect can be huge: a study with four common errors of this sort could end up with false positives more than 60 per cent of the time, according to simulations by Leif Nelson, a psychologist at the University of California at Berkeley, and his colleagues.

Bem says he has been careful to avoid these pitfalls. However, the sceptics - and even some who are inclined to think the results are real - wonder whether some unknown factors may have remained unchecked. "There's a genuine concern that these effects - including the ones that I've observed - could be a product of all the different ways that you could analyse the data to make the research fit the hypothesis," says Jonathan Schooler, a psychologist at the University of California at Santa Barbara who is a co-author of Franklin's study.

Fortunately, there is a straightforward way to settle the question - repeating the studies, with all procedures specified ahead of time in a public forum where everyone can see them. This eliminates the flexibility that can inflate the risk of false positives. If these studies still find evidence of precognition, then it deserves a closer look. If they do not, then the initial results were probably just false positives.

Bem, to his credit, has made this easier by making his set-up available to other researchers, and several replications have already been done. One, by Eugene Subbotsky of the University of Lancaster, UK, found results almost identical to Bem's, while six others, including [tests by Galak and Nelson](#), Schooler, [Wagenmakers](#) and Richard Wiseman of the University of Hertfordshire, UK, have failed to find any supporting evidence. Franklin, too, has finished a replication of his results that gave just barely significant support for precognition.

In the end, most scientists expect that most replications will fail and the current controversy over precognition will fade away, as all previous theories about ESP have before. "People are going to fail to replicate it, and that's what's going to make his statistics unimpressive in the end," says Rouder.

But even if precognition does amount to nothing, Bem and Franklin may have made an important contribution to science by drawing our attention to how easily good researchers can be misled. After all, ESP researchers are not the only ones who torture their data. Thanks to increased computing power, researchers in every field have started to test one variable after another in search of interesting results, laying the ground for false positives to pop up like mushrooms after rain. "I think an awful lot of what's published out there is wrong," says Jim Berger, a statistician at Duke University in Durham, North Carolina.

To remedy this, Schooler, Nelson and others are calling for a more rigorous approach, in which researchers routinely do what some are now doing for Bem's work: declare the experimental set-up beforehand, and report the results no matter what they look like.

The proposal may be more radical than it seems, because most researchers are reluctant to do "mere replications" that aren't breaking new ground.

"It's systemic," says Wiseman. "Funders don't like funding replications, scientists don't like carrying them out, and journals don't like publishing them." Changing that mindset may be the biggest challenge of all.

Bob Holmes is a consultant for New Scientist based in Edmonton, Canada



- From issue [2847](#) of New Scientist magazine, page 38-41.
- As a subscriber, you have unlimited access to our online archive.
- Why not [browse past issues](#) of New Scientist magazine?