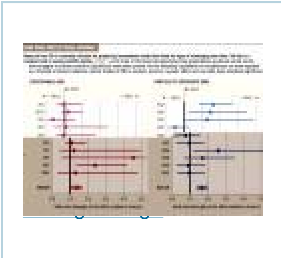
Opposites detract

Click to Print

13 March 2004
From New Scientist Print Edition. Subscribe and get 4 free issues.
Robert Matthews
Robert Matthews is a visiting reader in science at Aston University, Birmingham, UK



SO, YOU think you are rational, dispassionate and swayed only by hard evidence? Then try this little test. Last September two teams of respected scientists unveiled the outcome of research to prove the effectiveness of two very different agents. One team reported a powerful effect, much larger than expected by chance alone; the other could only muster an indifferent result with borderline significance. Which of these do you find the more convincing proof?

Most of us would view this as a no-brainer, and cite the first. But you probably sense a trap and would like to know more before deciding.

The weak result came from an international team of medical scientists studying a new drug aimed at reducing the chances of recurrent heart attacks. They found that the odds of another heart attack fell by just a few per cent, barely better than the reduction expected by chance alone. The far stronger finding came from a team at the Koestler Parapsychology Unit at the University of Edinburgh, UK, and seems to support the existence of extrasensory perception (ESP).

Still feel happy to put your trust in hard data? Or do you find yourself reaching for all kinds of reasons why, in this case, the experimental results alone just aren't enough to assess the merits of a scientific finding?

If so, you're in good company. For years, well-designed studies carried out by researchers at respected institutions have produced evidence for the reality of ESP. The results are often more impressive than the outcome of clinical drug trials because they show a more pronounced effect and have greater statistical significance. What's more, ESP experiments have been replicated and their results are as consistent as many medical trials - and even more so in some cases (see Chart). In short, by all the normal rules for assessing scientific evidence, the case for ESP has been made. And yet most scientists still refuse to believe the findings, maintaining that ESP simply does not exist.

Despite this relentless rejection of their work, parapsychologists such as those at the Koestler unit have ploughed on in search of clinching evidence they hope will convince the scientific community. Some believe it is a waste of time because the reality of ESP has now been put beyond reasonable doubt. Sceptics agree it is fruitless, but on the grounds that since ESP cannot exist, all positive results must be spurious. How has such a split arisen? After all, scientific evidence is supposed to drive everyone towards a single view of reality.

Over the years, sociologists and historians have often pointed out the glaring disparity between how science is supposed to work and what really happens. While scientists routinely dismiss these qualms as anecdotal, subjective or plain incomprehensible, the suspicion that there is something wrong with the scientific process itself is well founded. The proof comes from a rigorous mathematical analysis of how evidence alters our belief in a scientific theory. And it is not so easy to write off.

Its starting point is a profound result derived independently by the mathematicians Frank Ramsey and Bruno de Finetti in the 1930s. They showed that you can assign a number to the touchy-feely concept of belief using ideas drawn from probability theory. In particular, they proved that your faith in a theory can be quantified objectively on a scale ranging from near 0 for disbelief to near 1 for certainty. They also showed that scientific reasoning is logical provided your beliefs follow a rule known as Bayes's theorem.

Widely used in probability theory, Bayes's theorem shows how the chances of an event happening change in light of developments, such as the odds of a horse winning a race given that it won its last one. Ramsey and de Finetti showed that exactly the same rule applies to updating belief in a theory as new evidence comes in. The good news is that their rule is very simple: just take your original level of belief and multiply it by the strength of the new evidence, as captured by the so-called likelihood ratio. This gives the relative probabilities of getting such evidence if the theory is true, compared to if it were false. The likelihood ratio is large if the findings are consistent with theory, thereby boosting your level of belief in it.

But there is a nasty surprise lurking in the Ramsey-de Finetti analysis. How do you arrive at that original level of belief? In many scientific studies, there is a wealth of insight and evidence on which people can base their prior level of belief. But in novel or controversial areas of research, such as the paranormal, there isn't. And in those cases, it can only be based on gut feeling, instinct and educated guesses. In other words, it is entirely subjective.

This disturbing conclusion seems utterly at odds with the conventional view of science. Every week, research journals publish hard evidence supporting a host of theories, backed by statistical arguments for taking it seriously. Bayes's theorem implies that this whole process is nothing more than an elaborate attempt to dodge the subjectivity at the root of every scientific result.

While this prompts outrage among defenders of the scientific faith, many working scientists acknowledge that subjectivity plays a big role in their day-to-day thinking. Behind closed doors they routinely dismiss claims for, say, some new link between cancer and diet, simply because they find it implausible.

Nor is such prejudice the preserve of the life sciences. Even theoretical physicists routinely resort to subjective arguments to see off awkward results. Hearing that his new theory of special relativity had lost out to rival theories in its first experimental test, Albert Einstein simply brushed the evidence aside, arguing that the other theories were less probable.

Whether they realise it or not, scientists' thinking is influenced by Bayesian reasoning, and nowhere is this more apparent than in attitudes towards ESP research. By the standards of conventional science, the weight of evidence is now very impressive, but the scientific community refuses to accept the reality of ESP. Instead, they insist that extraordinary claims demand extraordinary evidence.

This is the perfect example of Bayesian reasoning. But who decides when an "extraordinary" level of evidence has been reached? It is something that can, and clearly does, mean different things to different people. Ultimately, it is not strength of evidence, or lack of it, that has been at the heart of the controversy over ESP. Yet the response of sceptics has been the same: whatever was responsible for the positive findings, it cannot be ESP. Something else must have happened: some flaw in the experiment, say, or a slip-up in the data analysis. Perhaps even fraud.

It is a response that provokes understandable resentment among parapsychologists. They complain that exactly the same approach could be used to reject unwelcome findings in any other field of science. It is too easy, they argue, for critics to dream up endless ways to explain positive ESP findings. Sceptics, meanwhile, insist it is only right to eliminate every alternative explanation before reaching a final conclusion.

Bayes's theorem shows that both camps are right. But it also reveals another disturbing fact: wrangling over alternative explanations can never be ended objectively. The reason is that every attempt to test a scientific theory involves a slew of "auxiliary hypotheses" - assumptions made about the design of experiment, the data analysis, and even the mindset of the researchers. For instance, medics confronted with the results of a clinical trial they find implausible routinely check the researchers' affiliations to see if they have a reasons to show the results in a particular light. And perhaps this is justified, given that academic studies funded by industry are more prone to producing positive findings (New Scientist, 1 February 2003, p 8). If the medics do suspect that the research findings are skewed, they will water down their faith in the results no matter how statistically significant they may be.

Never the twain shall meet

In the case of ESP research, the mindset of researchers has become a key issue (see "The power of belief"). Some studies suggest believers are more likely to get positive results than sceptics. For sceptics, this is proof of the auxiliary hypothesis that they insist explains all evidence for ESP: slipshod or dishonest practice by true believers hell-bent on proving their case. Others, however, see it as just another case of an intriguing effect that has been observed in many other areas of research.

Even so, it is only after all these alternative explanations have been dismissed that researchers can claim their results have been vindicated. Once again, the Ramsey-de Finetti analysis provides a mathematical rule for deciding when it is safe to say that evidence best matches the theory under test, rather than some auxiliary hypothesis. The bad news is that the rule demands estimates for the plausibility of competing explanations, which is again subjective.

The worst suspicions of parapsychologists are thus entirely justified. It is impossible to find evidence for ESP that will win round the sceptics. But those who see this as final proof of the futility of parapsychology should ponder this: exactly the same holds true for all scientific research. There are always auxiliary hypotheses, and deciding whether the evidence backs them or the theory being tested is just a matter of judgement.

The famous criterion of "falsifiability", the notion that scientific theories can never be proved, only disproved, is therefore a comforting myth. In reality, scientists can (and do) dream up ways to explain away awkward findings. The only difference with parapsychology is that scientists have no qualms about invoking everything from incompetence to grand conspiracy to explain the results.

It therefore seems that all that parapsychologists can do is collect ever more evidence, in the hope of gradually persuading more scientists of the reality of ESP. In this, they are appealing to one of the central tenets of the scientific process: that as more evidence builds up, the case for a theory becomes ever stronger. Yet the mathematics of scientific inference reveals even this to be a myth.

Bayes's theorem shows that belief in a theory increases with the strength of evidence. Mathematically, this is captured by the likelihood ratio (LR) - the likelihood of getting such evidence if the theory is true, compared to if it were false. So, for example, if the evidence is 10 times as likely to emerge if the theory is true rather than false, the LR is 10, and belief in the theory increases tenfold. If, however, the evidence is twice as likely to emerge if the theory is false, then the LR is 0.5, and the level of belief is halved.

All of this is perfectly reasonable - except how do you convert raw data into the all-important LR? The answer is, there is no hard and fast rule. It is yet another occasion for judgement, opinion and educated guesswork. Subjectivity has once more reared its head, and this time it undermines the most cherished principle of the scientific process: that, in the end, the accumulation of evidence ensures the truth will come out.

Once again, parapsychology provides an important lesson. Over and over again, reputable researchers have found strong evidence for the existence of ESP in tightly controlled experiments. Many would conclude the evidence is more consistent with the existence of ESP than any other explanation, such as sloppy methodology or fraud. As such, the LR is a large number, and should greatly increase belief in the existence of ESP - or so the parapsychologists would argue.

Sceptics, on the other hand, claim that pretty much any explanation for the evidence is more plausible than ESP, so the LR is far less than 1. So any fresh evidence actually reduces their belief in ESP.

The upshot could hardly be more different from the standard view of the scientific process. Both camps can look at precisely the same raw data and legitimately reach utterly different conclusions, because they have radically different models for the cause of the data. One camp insists that the results are more plausibly caused by of ESP than anything else; the other camp simply does not agree.

It gets worse. As the evidence accumulates, the two camps will not only fail to reach consensus but actually be driven further apart - propelled by their different views about the LR. And worst of all, there is no prospect of such a consensus unless the two sides can agree about the cause of the data.

Does the evidence accumulated over all these years prove the existence of ESP? By now, it should be clear that there is no objective answer. It should also be clear this is not the fault of parapsychologists. It simply reflects the fact that science alone cannot give us what we seek: an objective view of reality.

More than any other scientific discipline, parapsychology pushes the scientific process to its limits, and reveals where its faults lie. In particular, it has highlighted that, contrary to the insistence of many scientists, data alone can never settle this or any other issue.

This does not bode well for parapsychologists hoping to amass enough evidence to convince even hardened sceptics of the reality of ESP. It shows instead that there is only one way forward: for both sides of the debate to agree on their models for the results that emerge from ESP experiments. That, in turn, means working together in good faith to devise tests whose outcome can be agreed by all. For the key lesson of the mathematics of scientific inference is ultimately very simple: the credibility of all evidence is a matter of trust.