TRY IT AND SEE

In the social sciences, it is often supposed that there can be no such thing as a controlled experiment. Think again.

In the scientific pecking order, social scientists are usually looked down on by their peers in the natural sciences. Natural scientists do experiments to test their theories or, if they cannot, they try to look for natural phenomena that can act in lieu of experiments. Social scientists, it is widely thought, do not subject their own hypotheses to any such rigorous treatment. Worse, they peddle their untested hypotheses to governments and try to get them turned into policies.

Governments require sellers of new medicines to demonstrate their safety and effectiveness. The accepted gold standard of evidence is a randomized control trial, in which a new drug is compared with the best existing therapy (or with a placebo, if no treatment is available). Patients are assigned to one arm or the other of such a study at random, ensuring that the only difference between the two groups is the new treatment. The best studies also ensure that neither patient nor physician knows which patient is allocated to which therapy. Drug trials must also include enough patients to make it unlikely that chance alone may determine the result.

But few education programs or social initiatives are evaluated in carefully conducted studies prior to their introduction. A case in point is the 'whole-language' approach to reading, which swept much of the English-speaking world in the 1970s and 1980s. The whole-language theory holds that children learn to read best by absorbing contextual clues from texts, not by breaking individual words into their component parts and reassembling them (a method known as phonics). Unfortunately, the educational theorists who pushed the whole- language notion so successfully did not wait for evidence from controlled randomized trials before advancing their claims. Had they done so, they might have concluded, as did an analysis of 52 randomized studies carried out by the US National Reading Panel in 2000, that effective reading instruction requires phonics.

To avoid the widespread adoption of misguided ideas, the sensible thing is to experiment first and make policy later. This is the idea behind a trial of restorative justice which is taking place in the English courts. The experiment will include criminals who plead guilty to robbery. Those who agree to participate will be assigned randomly either to sentencing as normal or to participation in a conference in which the offender comes face-to-face with his victim and discusses how he may make emotional and material restitution. The purpose of the trial is to assess whether such restorative justice limits re offending. If it does, it might be adopted more widely.

The idea of experimental evidence is not quite as new to the social sciences as sneering natural scientists might believe. In fact, randomized trials and systematic reviews of evidence were introduced into the social sciences long before they became common in medicine. An apparent example of random allocation is a study carried out in 1927 of how to persuade people to vote in elections. And randomized trials in social work were begun in the 1930s and 1940s. But enthusiasm later waned. This loss of interest can be attributed, at least in part, to the fact that early experiments produced little evidence of positive outcomes. Others suggest that much of the opposition to

experimental evaluation stems from a common philosophical malaise among social scientists, who doubt the validity of the natural sciences, and therefore reject the potential of knowledge derived from controlled experiments. A more pragmatic factor limiting the growth of evidence-based education and social services may be limitations on the funds available for research.

Nevertheless, some 11,000 experimental studies are known in the social sciences (compared with over 250,000 in the medical literature). Randomized trials have been used to evaluate the effectiveness of driver-education programs, job training schemes, classroom size, psychological counselling for post-traumatic stress disorder and increased investment in public housing. And where they are carried out, they seem to have a healthy dampening effect on otherwise rosy interpretations of the observations.

The problem for policymakers is often not too few data, but what to make of multiple and conflicting studies. This is where a body called the Campbell Collaboration comes into its own. This independent non-profit organization is designed to evaluate existing studies, in a process known as a systematic review. This means attempting to identify every relevant trial of a given question (including studies that have never been published), choosing the best ones using clearly defined criteria for quality, and combining the results in a statistically valid way. An equivalent body, the Cochrane Collaboration, has produced more than 1,004 such reviews in medical fields. The hope is that rigorous review standards will allow Campbell, like Cochrane, to become a trusted and authoritative source of information.