

The Economist March 2 2002, p97-98

Science and technology

Social Science

Try it and see

PHILADELPHIA

In the social sciences, it is often supposed, 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. Real scientists do experiments to test their theories – or, if they cannot, 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.

The Campbell Collaboration, whose second annual conference has just taken place in Philadelphia, exists to change both this perception and the reality behind it by advancing the cause of "evidence-based" social policy. The collaboration is an international, independent, non-profit organisation that brings together social scientists, statisticians and policymakers. Its aim is to assemble and evaluate the best available evidence for the effectiveness of various social interventions. In particular, that means evidence from experiments.

Get real

Governments require sellers of new medicines to demonstrate the safety and effectiveness of their products. The accepted "gold standard" of evidence is a randomised controlled 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. This "double-blinding" reduces the risk that wishful thinking or other potential biases may influence the outcome. Drug trials must also include enough patients to make it unlikely that chance alone may determine the result.

Yet the medical industry is held to a higher standard of evidence than that to which governments hold themselves. This is bad, because, as Carol Fitz-Gibbon, a Campbell Collaboration participant from Durham University, in England, points out, school education amounts to about 15,000 hours of compulsory treatment. Social welfare and criminal-justice interventions can be similarly invasive. But few education programmes 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. Whole-language 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 randomised trials before advancing their claims. Had they done so, they might have concluded, as did an analysis of 52 randomised 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

about to begin in the English courts. The experiment, initiated by Lawrence Sherman, a criminologist from the University of Pennsylvania (with the support of England's Lord Chief Justice), will include criminals who plead guilty to robbery or serious assault. Those who agree to participate will be assigned randomly either to sentencing as normal or to participation in a conference in which the offender meets his victim and discusses how he may make emotional and material restitution. The purpose of the trial is to assess whether such restorative justice reduces reoffending. If it does, it might be adopted more widely.

Other randomised trials going on in Britain include an evaluation of the educational, nutritional, social and psychological effects of free breakfasts in English schools, a trial of smoke-alarm installation, and a trial of sex education in Scottish secondary schools.

We have control

The idea of experimental evidence is not quite as new to the social sciences as sneering natural scientists might believe. In fact, randomised trials and systematic reviews of evidence were introduced into the social sciences long before they became common in medicine. Iain Chalmers, a founder of both the Campbell Collaboration and its older and better established medical sibling, the Cochrane Collaboration, identifies an apparent example of random allocation in a study carried out in 1927 of how to persuade people to turn out to vote in elections. And randomised trials in social work were begun in the 1930s and 1940s. But enthusiasm later waned. Brian Sheldon, a social worker from the University of Exeter, in England, suggests this loss of interest can be attributed, at least in part, to the fact that early experiments produced little evidence of positive outcomes.

Thomas Cook, a pioneer of controlled experiments in education at Northwestern University, in Chicago, suggests 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). Randomised trials have been used to evaluate the effectiveness of driver-education programmes, 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. An examination of 308 studies from the criminal-justice literature, by David Weisburd of the University of Maryland and his colleagues, found that randomised trials were significantly less likely to report positive outcomes than non-randomised studies. Analysis of work in other areas gives similar results.

The problem for policymakers is often not too few data, but what to make of multiple and conflicting studies. This is where the Campbell Collaboration comes into its own. Rather than initiating research, it is designed to evaluate existing studies, in a process known as 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. The Cochrane Collaboration has produced more than 1,000 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.

The evidence-based policy movement has its detractors, however. Notably, many object on ethical grounds to the idea of randomly denying half of the population the potential benefit of a new service or initiative. But there is a rejoinder: to ask for evidence that the intervention in question is not itself harmful.

For example, who could object to driver education programmes in schools? But three different studies of a total of about 15,000 students have shown that such programmes are likely to increase road deaths. That is because the training programmes, while not producing significantly safer drivers, cause young people to obtain driving licences at an earlier age. And more young drivers means more accidents.

Or take the approach to criminal deterrence commonly known as "scared straight". The first completed Campbell Collaboration review is an analysis of such programmes, which introduce juvenile delinquents to prison inmates who portray conditions inside in harsh, or at least, realistic, terms. The theory is that exposure to the grim realities of life in prison will deter at-risk youths from future crime. A nice idea – but wrong.

Using a variety of statistical techniques, Anthony Petrosino, a criminologist at the US Academy of Arts and Sciences, combined the results of the seven available randomised studies of scared-straight programmes. This meta-analysis, as it is known, strongly suggests that participation in a scared-straight programme substantially increases the likelihood of subsequent arrest among participants.

Because they are cheap to run and politically popular, scared-straight programmes are widespread in the United States. Yet the obvious conclusion from Dr Petrosino's analysis is that such programmes are not only harmful to participants, but also place everybody in society at increased risk of crime. A few more counter-intuitive results such as that, and experiment-based social science might at last be given the respect that it deserves.