

put things in context, but to write a good story. Fortunately this is not the view of most responsible journalists!

Education is fundamental to so many aspects of life that one would think this is an area which might be given priority, in the sense that educational methods should be tried and tested before becoming the norm. However, Henderson points out that this is far from the case. There is little attempt to perform rigorous trials. Even when these are attempted, as with the “Every Child a Reader” programme, there was no randomisation and the trial was terminated after the first year (it was planned to last for three years) and rolled out nationally. It was argued that a rigorous trial would have been too expensive and it would have been difficult to recruit schools. It is generally accepted that we carefully test drugs before they are introduced, yet that is not so for how we educate our children.

Henderson urges scientists (‘geeks’) to stand up and be counted, both to influence politicians and to counter irresponsible media stories. If more had done when the MMR scare was brewing, we might not have been in the current situation where, at the time of writing, Wales is in the grip of the largest outbreak of measles in over 20 years, with over 1,000 reported cases and healthcare professionals in England are embarking on a campaign to immunise about a million under-immunised school age children against measles with the MMR vaccine.

There is hope. In the recent US budget, the Obama administration stated its intention to look at ways of implementing evidence-based decision making. Nearer home, the doctor and journalist, Ben Goldacre was commissioned to write a report for the Secretary of State for education on how to introduce more evidence-based practice into education, an area renowned for being evidence free. Perhaps the 3 Rs could be expanded to reading, ‘riting, ‘rithmetic and reasoning, with the latter being taught as part of science?

I strongly suggest that you read this interesting and enjoyable book.

*David Elliman*  
*Consultant Community Paediatrician*

## COMMENT

### **An Inherent Complication in Assessing Tests of Scientific Hypotheses**

**Larry Brownstein**

The testing of scientific hypotheses is not as straightforward as it often looks and perhaps contributes to the fact that many politicians do not pay as much attention to the character and context of scientific evidence as they should.

Most of us are familiar with the probabilistic character of many scientific hypotheses and the ways in which these can affect the testing scenarios. But what is not as well-known is a strategy for “saving the hypothesis” that is independent of probabilistic considerations. This strategy makes the testing of scientific hypotheses more complicated than Karl Popper and others thought it was.

The methodology described briefly by the reviewer is that developed by Karl Popper, which is that scientists should concentrate on falsifying their theories rather than corroborating them. Many ordinary scientists use Popper's falsification model as their guide. Effectively, what you do is list the initial conditions, and the lawful regularities, and in the case of the more formalized sciences, you deduce consequences from these, for instance a particular event – a procedure that has been discussed in physics and in other more formalized disciplines. If the event predicted by the theory under test does *not* occur, that is, is *not* observed, then Popper concludes that the theory has been decisively falsified. If the event does occur and has been observed, the theory has only been corroborated, that is, the likelihood that it is true has increased. For strictly logical reasons, no theory can ever be conclusively verified, that is, proved to be true, though Popper never explicitly points this out.

There is, however, a kind of get out of jail free card that can be used to “save the hypothesis”. This is known as the Duhem-Quine gambit. When you set up your theory testing scenario, in addition to listing all the initial conditions that apply and the regularities that are involved, there are a number of auxiliary hypotheses that usually go unstated with respect to the testing scenario. In physics, this is because experimental physicists often know what these possible contaminants of the testing process are and attempt to control for them. In less formalized sciences, such influences may be unknown or only informally considered. The Duhem-Quine gambit, when utilized appropriately, is a legitimate procedure, not an attempt to cheat the testing process.

Basically, if the result of the test of a given scientific hypothesis is not observed, for whatever reason, instead of directly falsifying the hypothesis or theory under test, the testers can select one or more of the auxiliary hypotheses as the culprit, such as the nature and operation of the equipment, any presumed biases on the part of the testers, sampling problems, the general environment in which the test takes place, and so forth. Selecting one or more of the auxiliary hypotheses that inevitably accompany any scientific testing situation will enable the scientist to “save the hypothesis”. Of course, one cannot continue to blame failure to obtain a result the theory says should have been observed but wasn't on the auxiliary hypotheses. Continued failure to observe a result predicted by a theory must eventually lead to that theory (or a portion of it) being considered to be falsified.

Elliman's point that more scientists should be involved in independent ways in policy decisions not subject to direct ministerial control becomes even more salient in light of the inevitable availability of the Duhem-Quine gambit, a central feature of any scientific testing scenario, which renders assessment of the relevance of scientific evidence for policy purposes even more problematic than it might otherwise be considered to be.

*Larry Brownstein*

*Email: zen143717@ZEN.CO.UK*