

## **David J. Hand**

Imperial College, London

### **What's the question?**

Science progresses through an iterative process of formulating theories and comparing them with empirical real-world data. Different camps of scientists will favour different theories, until accumulating evidence renders one or more untenable. Not unnaturally, people become attached to theories. Perhaps they invented a theory, and kudos arises from being the originator of a generally accepted theory. A theory might represent a life's work, so that being found wanting might be interpreted as failure. Perhaps researchers were trained in a particular school, and acknowledging its shortcomings is difficult. Because of this, tensions can arise between proponents of different theories.

The discipline of statistics is susceptible to precisely the same tensions. Here, however, the tensions are not between different theories of "what is", but between different strategies for shedding light on the real world from limited empirical data. This can be in the form of how one measures discrepancy between the theory's predictions and observations. It can be in the form of different ways of looking at empirical results. It can be, at a higher level, because of differences between what is regarded as important in a particular context. Or it can be for other reasons.

Perhaps the most familiar example of this tension within statistics is between different approaches to inference. However, there are many other examples of such tensions. This paper illustrates with several examples. We argue that the tension generally arises as a consequence of inadequate care being taken in question formulation. That is, insufficient thought is given to deciding exactly what one wants to know - to determining "What is the question?".

The ideas and disagreements are illustrated with several examples.