

In Prof McGuinness's response to Hulme (Times Higher Education Supplement.24/3/06) she describes the problem with reading research studies that use the 'isolated-groups design'.

What Hulme objects to is my criticism that much of the reading research fails to adhere to proper research design which can then be tied to appropriate statistical tests. I refer here to studies in which there is no random selection of subjects into a data pool. Because statistics is based upon the mathematics of probability, which in turn is based on the world of random or chance events, statistical tests are only valid when certain criteria are upheld. One of these criteria is to take the world as you find it and don't divide your subject populations to suit your objective. This argument is spelled out in great detail in my book and is irrefutable, despite what Hulme thinks.

In reading research, it is common for children to be given a reading test battery, and then selected into the study solely on the basis of their reading test scores, as good versus poor readers. They are then given tests known to be correlated to, or likely to be correlated to reading test scores, and voila, a significant result appears. This practice voids the conditions for which the mathematics of probability can be applied, because it violates the principle of random selection (among other things). I christened this type of design *an isolated groups design* because of the way it looks when you graph the distributions of the data from the two groups. It will never be found in any textbook of statistics, nor in any account of how statistics works by the geniuses who invented these tests, such as Karl Pearson, William Gosset, and Ronald Fisher, because **no statistical tests can be applied to it.**

Hulme hasn't the faintest idea what I am talking about, and uses a spurious counter-example taken from medical research to refute my position. (Of course, one first has to believe that medical research is valid, when much of it is not. Interested readers should look at *The Limits of Biological Treatments for Psychological Distress*, Fisher and Greenberg, 1989, for a revealing assessment on how difficult it is to do medical research properly.)

Hulme argues that, "Her objections to these studies are based on a profound misunderstanding of basic statistics. Such designs (case-control studies) are widely used in medicine and epidemiology, and there is a whole raft of thorough statistical work on their design and interpretation. Based on McGuinness's view in chapter nine, we could throw out all studies that have ever compared a clinical group with a control group. He uses examples of people who have had heart attacks versus people who haven't, or people who are depressed versus those who are not."

Yes, I certainly would throw these studies out. What on earth would a design like this prove? One needs to know the connection between past behaviors and some medical outcome. The mind boggles as to which type of behaviors and conditions might become candidates here. Diet? Exercise? Family history? Job stress? Financial stress? – or more than one of these factors, or all 5, or perhaps something we haven't thought of yet. Furthermore, Hulme's example is nothing more than a correlational study in disguise, and correlations can never prove causality, proving a point I was making in my book. And if this is retrospective, how can any of this be verified?

Now, one might want to take a group of people all prone to heart attacks due to family history of heart attacks, and **randomly assign them to treatment conditions**, such as a particular diet, or exercise regime, and then see what happens over the long term. This is a bona fide research design, and, providing the patients stick to the regime, the data can be treated statistically by a number of valid tests. In the example for reading research which I outline below, they cannot.

In reading research, the researcher typically identifies a random population, roughly of the same age, and gives every child a reading test. The test is given IN ORDER TO split the children into two groups: One of these groups does extremely badly on the test, and the other scores in the normal range. Next, take these two extreme groups, give them some more tests (tests of memory, visual and auditory perception, etc.). Then compare the two groups statistically on each of these skills, using the statistical tests designed for independent groups or random groups. Run the math, and you will discover the poor readers scored significantly below the normal readers on most these tests. As the tests themselves are chosen because they are related to reading in some way, it is not surprising what the results show. The outcome is almost a forgone conclusion, because the research design dictates the outcome. As I wrote in my book, this is the only research design I am aware of in which a null result (no significant differences between groups) is far more unlikely (improbable) than a significant result.