A question of maths, science …

Would you describe yourself as anti-CLIL?

No, of course I’m not ANTI-CLIL. To use a personal example, fact, when I was invited by Lorenzo and colleagues to attend a research meeting for Andalusi an teachers in the programme bilingüe I had high positive expectations. I was curious listening to the teachers informally, generally more vocal dissidents and to the exaggerated claims being made, I started to wonder. I then started looking more closely at the research design and interpretations.

How much is the reported weakness of students in EFL streams compared to those in CLIL due to the weakness of EFL provision?

After an informal study done with my research group of over 1,800 students in over sixty classes, and triangulated concluded that the communicative approach had not arrived in most Andalusian classes. The reason for this delay was the grammar instruction, always linear, making listening to the teachers informally, generally more vocal dissidents and to the exaggerated claims being made, I started to wonder. I then started looking more closely at the research design and interpretations.

How much is the reported weakness of students in EFL streams compared to those in CLIL due to the weakness of EFL provision?

After an informal study done with my research group of over 1,800 students in over sixty classes, and triangulated concluded that the communicative approach had not arrived in most Andalusian classes. The reason for this delay was the grammar instruction, always linear, making listening to the teachers informally, generally more vocal dissidents and to the exaggerated claims being made, I started to wonder. I then started looking more closely at the research design and interpretations.

In a beautifully crafted if conceptually response to the argument on pre-test, Lorenzo quoted Diane Larsen-Freeman in his defence: “Although many extant research methods can be used with complexity theory, an analysis may vary, which is the case with the classical pre-test-post-test experiment design. Pre-test-post-test experiments are problematic from a complexity theory perspective. Most importantly, linear causal modelling is ecologically valid. Further, they can only at best lead to claims about direct, linear cause and effect while not allowing for multiple or reciprocally interacting factors that change over time. In addition, they ignore nonlinearity.”

In his analysis he stated the main argument in terms of the development of a theory, not of testing it – as used in the defense of Lorenzo.

To examine the question of scientific method, we sent the paper to a climatologist Dr Tim Osborn (box) and Dr Lance Willey, a postdoctoral researcher at the Oxford University Department of Geography who has taught applied linguistics and has a professional interest in experimental design.

‘The problem with complex, multi-variate, non-linear structures is that it’s hard to ‘get at’ what’s going on (cause and effect) from a single experiment. ’Take Darwin and the theory of evolution. ’It wasn’t experimentation alone which led him to this but mainly systematic observation. ’It simply wouldn’t have been possible for him to “get at” all of the underlying structures and processes he theorised were in place through experimentation alone. However, once this theory was put forward it could be experimentally tested and over again in different contexts. Thus the complex structure of evolution was (strongly) theorised from systematic observation, but then backed up by rigorous repeated scientific experimentation. ’If you want to develop a theory of second language acquisition (SLA), experiments aren’t going to help that much. If you want to back up a theory of SLA, experiments are your best friend – designed properly, repeated, and with appropriate conclusions drawn at any one time.’

Hall also pointed out, that in the context of research, randomisation is not possible, testing the learners’ language at the end of the period, a post-test, by itself did not give us very much information: ‘By simple application of logic, applying a pre-test-post-test design will return more information than just a post-test design. The pre-test-post-test design tells us whether we can expect positive, negative or no change over the time period studied. Failing to pre-test means failing to establish baseline levels of ability. Thus we would have to limit our conclusions to “absolute levels of ability” rather than “change over time”.’

Another of Bruton’s concerns is that the children were not randomly assigned to the particular programme – it was their parents who chose. Given that the groups weren’t randomised and the design was post-test only, what can the data tell you about cause and effect? Dr Hall is blunt: “Virtually diddly-squat, I’m afraid. If kids weren’t randomly assigned to each experimental condition it opens the door for erroneous (false conclusions) because of extraneous factors. The real power of the experimental method as used in the social sciences comes from randomising, which balances all extraneous factors (measured or not) between the two (or more) groups and thus removes their influence.” In a non-randomised post-test only study you have virtually no information to draw upon – you don’t know base levels, and the lack of randomisation means all extraneous factors are at play.’

Q&A with Dr Tim Osborn

Do you agree that pre-test-post-test design should not be used to test a theory about a complex non-linear process?

I disagree, especially if the alternative is the post-test. I don’t think there is any fundamental disadvantage with using the pre-test approach even with complex systems. However, systems that behave non-linearly tend to be harder to ‘diagnose’ than simple, linear systems. Larger samples and multiple experiments are required to make progress. Does what does post-test tell us about change over time?

In climate science you can say very little about change over time with only post-test measurements. I think the same applies to most fields.

What can a non-randomised post-test tell us about cause and effect?

Post-test measurements with no control for confounding factors and with non-random allocation to groups can tell you little about cause and effect. Reducing confounding factors via random allocation can help. But with complex non-linear systems, it can be difficult to achieve genuine random allocation unless the sample sizes are very large. I don’t think that complex systems negate the use of pre-test-post-test methods, and particularly they don’t favour post-test instead. But the power of both methods is weakened compared with their utility with simpler systems.

Dr Tim Osborn is a reader at the Climatic Research Unit, School of Environment Sciences, University of East Anglia.