

## Background

Why do we still know so little about the efficacy of the technologies into which we have invested much energy, time and money in our language teaching and learning endeavours? A question that sounds quite simple on the face of it is, of course, immensely complex when we look at it more closely. How can we hope to arrive at a sensible answer to the often posed question *How effective are technologies in promoting learning?* when the scope of the investigations required to come even close to a valid generic conclusion is beyond most researchers' capacity. We might as well capitulate and say that it is '42' which was writer Douglas Adams' answer to the question of life, the universe and everything.

It appears that we have two choices in tackling the question which remains interesting and worth investigating, if only to confirm our instincts that what we are doing with a great deal of effort is worthwhile. On the one hand, we can reduce the scope by focusing on one particular piece of technology being used having one or several effect(s) on a measurable learning process or outcome. On the other hand, we can widen the scope by analyzing results of a large body of research and synthesizing findings related to one or several variable(s) under investigation. In both cases, the question needs to be qualified and matched closely to the project's context, methods and analyses. Curiously, all of this has already been done in one way or another over the past two decades, yet general conclusions and claims remain largely equivocal.

In order to provide a clearer picture, we spent the last four years compiling and analyzing a vast amount of data related to CALL effectiveness. Because of the overwhelming volume of papers located in our searches, including articles written in other languages than English was at this stage beyond our capacity. We began by examining the entire meta-research conducted in the field (Felix 2005a). While this provided us with a little more certainty regarding outcomes that might be generalised, it did not dispel frustrating worries about the validity of some research designs. Clearly more close attention needed to be given to this.

Our next study (Felix 2005b), therefore, was dedicated to examine the sorts of designs researchers had used to ascertain effectiveness. The study highlighted strengths and weaknesses in this body of work and singled out models of good design practice. In this context, the usefulness of meta-analyses to determine clear cause and effect results, relying solely on effect sizes, hence solely measuring *outcomes*, appeared questionable. It became clear that a series of systematic qualitative syntheses of findings related to one particular variable such as *learning strategy* or *writing quality* might produce more valuable insights into the potential impact of technologies on *learning processes* as well as *outcomes*.

The next step was to test this assertion, and our final project (Felix 2006) was designed to look closely at all studies which had dealt with L2 writing. It was hoped that this would add important qualitative information, complementing the quantitative findings of the meta-research, especially since the latter had largely concentrated on L1.