

The idea of pre and post testing of students is often accepted as a viable method to assess the extent to which an educational intervention has had an impact on student 'learning'. However, despite the fact that the results of such evaluative measures are extensively reported in the literature (generally as part of a more extensive discussion of particular CAL interventions) there is a dearth of detail on how the tests were conducted and of the criteria which have been used to measure their impact. Furthermore, there is rarely any discussion of the potential effects of other factors as contributing to what is often reported as 'improved student performance', least of all on the possibility that the testing process itself may have a significant effect on the outcome of the evaluation.

As with all evaluations, the starting point must be to define clearly why the evaluation is being performed, what is being evaluated, when the evaluation will be performed, and how it will be performed.

Why?

Because we know that students with different skills and backgrounds come to study a particular subject, we need to establish a base measure of their knowledge and understanding of a topic in order to be able to quantify the extent of any changes in this knowledge or understanding by the end of a particular period of learning. Ideally, we wish to know not only that the educational intervention has had an impact, on the student, hopefully a positive one, but we also want to be able to quantify that impact.

What?

Simplistically viewed, the process should entail students undertaking a test to determine some identifiable starting level of knowledge or understanding of a topic and a later point undertaking an exactly comparable test to determine the extent to which knowledge and understanding has been augmented by the educational intervention. However, unless we content ourselves with assessment instruments which simply seek to measure the ability to retain and recall 'known facts', it is difficult to provide an accurate scalar measure of the exact extent of the 'improvement' in performance. Furthermore, these measures do not directly inform us of the process by which the improvement was effected. There are a large number of variables introduced by factors which are intrinsic to the measurement of performance and also those which are related to the circumstances of the subjects being assessed. When combined, these create a level of tolerance which makes significance testing using statistical analysis of the empirical data virtually impossible.

Thus in the area of pre and post testing of students it is important to realise that one must be prepared to make only very limited claims for any findings which seek to provide an accurate measure of the change in learning which can be attributed to the intervention.

When?

Also it is important to bear in mind that the timing of the pre and post tests will have a critical impact on the results obtained. In some areas of study pre testing of students is not appropriate because at the outset of the course it is already known that students would be expected to have virtually no knowledge or experience of the subject. Pre testing would be pointless from the students point of view and so the pre test really has to be done at some point in time when we can expect the student to have acquired some relevant knowledge but before the student is exposed to the CAL materials. The results of post testing will vary considerably depending on how soon the test is administered after the student has used the CAL materials and ideally if the test is administered immediately after the learning session additional follow up tests at later dates should be used to provide some evidence of application and impact of what has been learned.

How?

The design of the pre and post questions is critical to success. The repetition of the same test questions is obviously not a sound solution to achieving comparability but it is a good idea to retain a proportion of the original test materials and to blend this with new questions which examine the same expected learning outcomes. It is also important to consider the type of questions used. Certainly we should not rely purely on objective questions but, extended questions which seek to test a whole range of issues are also inappropriate. However, the use of short definitions can make it easy to identify an accurate and unambiguous response. It is also possible to consider incorporating a requirement for students to indicate the level of confidence in their response. Design of pre and post tests is not easy and above all we have to be extremely careful in the test design to ensure that we are not simply 'testing the test'.

This all seems to paint a very negative picture of the value of pre and post testing of students. However, if one is willing to accept the fact that pre and post testing should be seen as assisting us to learn more about how students use CAL rather than as a means of demonstrating changes in knowledge or skills then we can set about using appropriately designed pre and post tests to achieve this objective. What we want to generate is not simply a performance measure of the number of 'correct' responses provided by a student but the manner in which the deliver mode of instruction has caused an alteration in their responses.

Robert Newton
The Robert Gordon University,
Aberdeen