

# Why aren't our test scores reflected in our sales quality?

*How multiple choice test success might not be an accurate predictor of a salesperson's ability to apply knowledge and understanding in the field.*

---

Here at GR Business Process Solutions the question that we are asked more often than any other forms the title of this article. Financial Services, Pharmaceuticals and many other business areas are now either formally or informally regulated in respect of the technical ability of both their front and back office staff.

One of the common complaints that we receive from compliance departments is that, while their salesforce regularly achieves average percentage scores in their in-house testing programmes in the high 70s or low 80s, these scores are not reflected in their sales quality review figures.

“Why is it,” they ask us, “that our salesforce have demonstrated that they can answer technical questions in relation to our products and the advice we give and yet they cannot complete sales documentation that substantiates the sale to the same standard?”

All too often our reply will be along the lines of, “The scores you have recorded in your in-house testing programme are NOT indicative of their overall ability.” and in this article I would like to explain (briefly) why that is.

Let's begin by examining a typical training department and look for the areas that will impact on overall scores.

Typically a training department will be responsible for identifying the training need, performing the training needs analysis, designing developing and delivering the training and then designing and delivering the validation for that training. Where this involves technical knowledge this almost invariably involves the use of Multiple Choice Question Tests (MCQs) and this is where the process tends to break down.

The Qualifications and Curriculum Authority (the regulatory body for most examination awarding bodies in the UK) says quite clearly that there should be clear water between those responsible for formulating tests and those responsible for training the syllabus that is to be tested. However, in our experience, the same personnel tend to be responsible for both. Almost inevitably this leads to them training the content of the test and not the syllabus. Let me give you a very simple example.

Part of my syllabus is that all students will know the twelve times table up to 12 x 12, understand how to multiply any number by twelve and be able to apply that knowledge to any given number. This should mean that the training should enable the students to confidently recite the twelve times time up to 12 x 12, show that they understand how to perform multiplication generally and be confident that given any number they could multiply it by 12 and achieve the correct result.

One of the ways in which this is to be validated is by giving them a question asking them to multiply a number greater than 12 (e.g. 17 x 12).

All well and good. But supposing the person responsible for delivering the training knows full well that this question will appear in the validation test and, like most people involved in teaching, wants to help their students achieve the best possible result. There will be a strong temptation to include this particular example within the training course and get them to learn it by rote, so that when it appears in the test the answer will be recalled immediately.

Understandable, but in the process the question does not now test their ability to apply the principles of multiplication, simply to recall previously learned information which requires much less brainpower. If this is endemic throughout the training process you will end up with a group of students who can remember specific facts but who have no ability to apply their knowledge in novel or more complex situations.

This, unfortunately, is very typical of our experiences in Financial Services. If the question bank has not been 'ring fenced' from the people responsible for delivering the training, they will, in our experience, always ensure that the specific examples used within the test environment are covered on the training course. The consequence of this is that as soon as the advisers find themselves in the real world with real life examples that do not fit comfortably within the examples they have learned in their training programme they do not have the necessary ability to apply the learning.

The second reason why test scores are not indicative of overall ability is where the construction of the test itself inevitably leads to the correct answer rather than any innate understanding by the candidate. Let me give you an example. Suppose you came upon the following question: -

1. Which was the 16<sup>th</sup> state to join the Union in the USA?

I doubt very much that one in a thousand people would know the correct answer. However (as is more usual) it would probably be presented to you as part of a multiple choice question which might look like this: -

1. Which was the 16<sup>th</sup> state to join the Union in the USA?
  - a. Tennessee
  - b. Hawaii
  - c. Vancouver
  - d. Alaska

The chances are that a large number of people would now get the question correct providing they know that Vancouver is in Canada and Alaska and Hawaii were the 49<sup>th</sup> and 50<sup>th</sup> states to join the Union. In fact we have conducted research using this very question and 75% of respondents chose this option. So in the process of changing this from an open question to a MCQ the overall score has gone from less than 0.1% to about 75%!

Clearly 75% of the population did not KNOW that Tennessee was the 16<sup>th</sup> state to join the Union and in a multiple choice test with four options and no inherent knowledge by the candidates of the correct answer we would expect the result to vary little from chance (25%). This question is demonstrably performing three times better than chance and is therefore producing a misleading result.

So how to construct a MCQ that gives a truer result? How about this: -

2. Which was the 17<sup>th</sup> state to join the Union in the USA?
- a. Louisiana
  - b. Indiana
  - c. Mississippi
  - d. Ohio

Now there are no clues as to which might be the correct answer (the list contains the 17<sup>th</sup>, 18<sup>th</sup>, 19<sup>th</sup> and 20<sup>th</sup> states to join the Union) and again we have used this question in our research. The scores for each answer are: -

- a. 23.2%
- b. 23.2%
- c. 27.7%
- d. 25.9%

These are so close to chance that they indicate that across the population of people who answered this question there was no real knowledge of the correct answer. So we can clearly see that the construction of test items can (and does) have an enormous impact on the quality of the result. (N.B. We will be publishing a more comprehensive report of our research at a later date).

If one or more of your distractors are never or only rarely being chosen the ability of the candidate to guess correctly increases from a one in four chance where all options are equally attractive to one in three or one in two if it is possible to discount one or two of the options. If this situation extends to more than a few of the questions in your question bank then the chances are that the results will be significantly skewed upwards by the increased ability of candidates to guess correctly.

So, if there is a mismatch between your test scores and sales quality here are some places to begin your search for the reasons why.

In our experience it can be very difficult for training departments to find the time or have access to the necessary levels of expertise internally to be able to effectively remedy the situation. The Qualifications and Curriculum Authority provide some helpful guides and there are a few consultancies, like our own, which specialise in working alongside firms in order to achieve test results which accurately measure the ability of the candidates concerned.

It can seem quite daunting (and potentially expensive) to apply these standards but in closing let me tell you about a situation which we recently encountered which has had a very positive outcome.

Our client was determined to make its testing environment more robust and stringently applied the rules outlined above. Initially, this resulted in a considerable drop in the firms overall test results and, as a consequence, the training department (without recourse to the test questions) set about improving the quality of their design and delivery in order to ramp up the overall scores. By strongly tying their learning objectives to the knowledge syllabus (on which the test items were based) they brought about a significant increase in scores, the analysis of which demonstrated that they were a much truer reflection of the candidates' overall ability.

As a consequence their sales quality has improved and the training department now has a much greater degree of confidence in themselves, their training materials and their test results.