

Redesigning Schooling

Principled assessment design



REVISED AUSTRALIAN
EDITION

Dylan William



Contents



Introduction	2
Chapter 1	Why assess?	6
Chapter 2	What should we assess?	10
Chapter 3	Quality in assessment	18
Chapter 4	Reliability	29
Chapter 5	Designing assessments	49
Chapter 6	Interpreting evidence	62
Chapter 7	Targets, recording and reporting	69
Chapter 8	Putting it all together	78
Appendix	Estimating reliability	84
References	86

Introduction

The premise of this booklet is that any assessment system should be designed to support the curriculum in place in a school, rather than having the curriculum designed to fit the assessment system. Or to put it another way, assessment should be the servant, not the master, of the learning.

The aim of this booklet, then, is to provide schools with some ‘tools for thinking’ as they begin the task of designing their own assessment system. A school’s assessment system has many jobs to do. It must provide useful information for students, parents, teachers and leaders about the progress being made by students. At all times, the system should provide information about important checkpoints or milestones in learning. But, as students approach high-stakes assessments such as the National Assessment Program – Literacy and Numeracy (NAPLAN), it should also provide information about the likely outcomes of those assessments. Some of the information generated by the system will be carefully recorded while other assessment information will be of a more ephemeral nature, informing decisions about next steps in teaching, for example, where formal recording could be unduly

burdensome or even counterproductive. The information also needs to be generated in a timely manner. For some kinds of assessment information, the 'sell-by' date is very close to the collection of the evidence, while other kinds of assessment information has a longer 'shelf-life'. Inevitably, these different demands on the assessment will create conflicts and trade-offs, many of which cannot be resolved easily, if at all. The aim of this booklet is not, therefore to tell schools how to design their assessment system, but rather to help them understand better what trade-offs they are making in designing their assessment system and, in particular, what the costs and benefits of various decisions will be.



In some cases, formal recording could be unduly burdensome, or even counterproductive.



In presenting these issues, I have tried to steer clear of jargon and technical issues as much as possible, but the field of assessment is one where it is impossible to understand what is going on without some engagement with the technical details. I have kept this to what I think is the minimum necessary to engage properly with the issues. Inevitably, however, some parts of the booklet may appear to be rather like a textbook on assessment.

It is also important to realise that there is a great deal of science here. By that I mean that there are some issues in which personal opinions play very little part, and this is rather rare in education. As many people have pointed out (see, for example, Lagemann 2000) education lacks strong disciplinary foundations, and there is relatively little that could be regarded as 'reliable knowledge' (Ziman 1978). In the field of assessment, however, there are some very well-understood technical aspects where failure to appreciate the technical details means that one

can do things that are just wrong. This is particularly true for the issue of reliability, where schools routinely say things to students or parents that are not just unprofessional (in that they contradict professional standards or codes), but actually incorrect. For example a teacher may state that a student who gets the same level of achievement at the end of the term that he did at the beginning has made no progress. As we will see there is no certainty that this is true.

Chapter 1 begins the discussion by outlining the main reasons we assess student performance in schools, while chapter 2 introduces some features of the design of school assessment systems, specifically what should be assessed. Chapter 3 explores in some depth the issues of the quality of assessment, focusing on the central concept of validity. Reliability, which is a key aspect of validity, is discussed in more detail in chapter 4. Chapter 5 deals with assessment design, focusing explicitly on the idea that assessments should be designed backwards from the inferences they are designed to support. The interpretation of evidence is discussed in chapter 6 and in chapter 7, I discuss how the results of assessments can be recorded and reported to various stakeholders. Chapter 8 draws the various threads of the other chapters together.

Assessment is a vast field, and I have inevitably been highly selective in my coverage of aspects of assessment, both in what aspects to include and what to say about the aspects that have been included. No doubt others would make a different selection, and indeed, had I written this booklet a few years ago, I may well have made a different selection – this booklet is very much a product of its time.

© In making these selections, I have steered clear of trying to propose a model for a school assessment system – such a model might be suitable for some schools but not for others. Instead, I have tried to focus on some key principles for assessment that will help school

leaders and others to think about assessment in a more disciplined way. These principles will often, if not always, be in tension, so there definitely won't be any assessment system that satisfies all the principles. But by making these principles explicit, my hope is that the compromises being made are clear and explicit and above all planned, rather than just being the unintended consequences of decisions taken by different people, in different parts of the school, at different times.



I have tried to focus on some key principles for assessment that will help school leaders and others to think about assessment in a more disciplined way.

