

Table of Contents

About the Author	ix
Introduction	1
CHAPTER 1	
Why Educational Achievement Matters	3
<i>The Increasing Importance of Educational Achievement</i>	3
<i>Why Is Raising Student Achievement So Hard?</i>	10
<i>Three Generations of School Effectiveness Research</i>	15
<i>The Impact of Teacher Quality</i>	17
<i>How Do We Increase Teacher Quality?</i>	22
<i>Conclusion</i>	26
CHAPTER 2	
The Case for Formative Assessment	27
<i>Professional Development</i>	27
<i>The Origins of Formative Assessment</i>	33
<i>What, Exactly, Is Formative Assessment?</i>	37
<i>Strategies of Formative Assessment</i>	45
<i>Assessment: The Bridge Between Teaching and Learning</i>	46
<i>Conclusion</i>	50
CHAPTER 3	
Clarifying, Sharing and Understanding Learning Intentions and Success Criteria	51
<i>Why Learning Intentions Are Important</i>	51
<i>When Are Learning Intentions Useful?</i>	56
<i>Issues in Constructing Learning Intentions</i>	61
<i>Practical Techniques</i>	65
<i>Conclusion</i>	69
CHAPTER 4	
Eliciting Evidence of Learners' Achievement	71
<i>Finding Out What Students Know</i>	71



CHAPTER 2

The Case for Formative Assessment

We've discussed how increasing the educational achievement of students is a national economic priority, and the only way to do that is to improve teacher quality. We also saw that deselecting existing teachers and improving the quality of entrants into the profession will have, at best, marginal effects, and so securing our economic future boils down to helping teachers who are already in post become more effective.

This chapter reviews the research on teacher professional development and shows that while there are many possible ways in which we could seek to develop the practice of serving teachers, attention to minute-by-minute and day-to-day formative assessment is likely to have the biggest impact on student outcomes. The chapter concludes by defining what, exactly, formative assessment is.

Professional Development

Andrew Leigh (2010) analysed a data set that included test scores on 90,000 Australian primary school students and found that, as in the American research, whether the teacher had a master's degree made no difference. He did, however, find a statistically significant relationship between how much a student learned and the experience of the teacher, as can be seen in figure 2.1 (page 28).

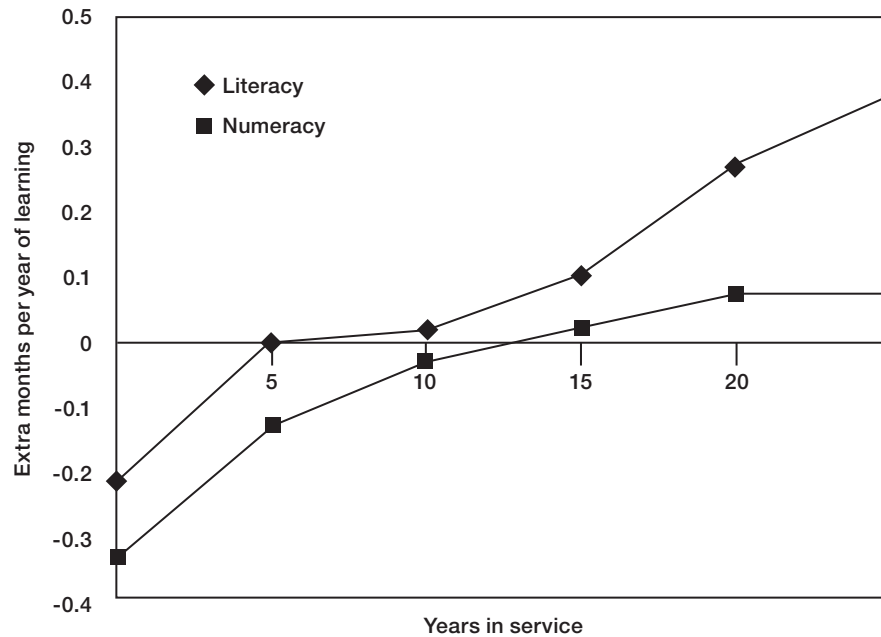


Figure 2.1: Increases in teacher productivity with experience (Leigh, 2010).

It is clear that the value added by a teacher increases particularly quickly in the first five years of teaching, but what is most sobering about figure 2.1 is the vertical axis. If a student is taught literacy by a twenty-year veteran, the student will learn more than he will if taught by a novice, but not much more. In a year, with a twenty-year veteran, a student will make an extra half-month's progress—in other words, a twenty-year veteran teacher achieves in fifty weeks what a novice teacher will take fifty-two weeks to achieve. Because of the size of the study, this result is statistically significant, and the improvement is worth having, but it is not a large difference. Therefore, it's not surprising that many have argued that the answer is more, and better, professional development for teachers.

Indeed, it would be hard to find anyone who would say that teacher professional development is unnecessary. Professional development for serving teachers is a statutory requirement in many jurisdictions. However, most of these requirements are so loosely worded as to be almost meaningless. Pennsylvania's Act 48 requires teachers to undergo 180 hours of professional development that is related to an educator's certificate type or area of assignment. Note that there is no requirement for teachers to improve their practice or even to learn anything. The only requirement is to endure 180 hours of professional development.

some of the ideas they had discussed in the workshops and could discuss how their ideas could be put into practice more effectively.

Because each teacher had made his own decisions about what aspect of formative assessment to emphasise and which classes to try it with, it was impossible to use a traditional experimental design to evaluate the effects of our intervention. Therefore, we designed a “poly-experiment”. For each class with which a teacher was trying out formative assessment techniques, we looked for the most similar comparison class and set up a mini-experiment in which the test scores of the class that was using formative assessment were compared with the test scores of the comparison class. This experimental design is not as good as a random-allocation trial, because the teachers participating in the experiment might have been better teachers to begin with, and so the results need to be interpreted with some caution. Nevertheless, in this study, using scores on externally scored standardised tests, the students with which the teachers used formative assessment techniques made almost twice as much progress over the year (Wiliam, Lee, Harrison & Black, 2004).

What, Exactly, Is Formative Assessment?

As the evidence that formative assessment can have a significant impact on student learning has accumulated, a variety of definitions of *formative assessment* has been proposed. In our original review, Paul Black and I defined *formative assessment* “as encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (Black & Wiliam, 1998a, p. 7). Writing around the same time, Bronwen Cowie and Beverley Bell qualified this slightly by requiring that the information from the assessment be acted upon while learning was taking place. They defined *formative assessment* as “the process used by teachers and students to recognise and respond to student learning in order to enhance that learning, *during the learning*” (Cowie & Bell, 1999, p. 32, emphasis added). Others have also emphasised the need for action during instruction and defined *formative assessment* as “assessment carried out during the instructional process for the purpose of improving teaching or learning” (Shepard et al., 2005, p. 275). A review of practice by the OECD across eight countries defined *formative assessment* as “frequent, interactive assessments of students’

progress and understanding to identify learning needs and adjust teaching appropriately” (Looney, 2005, p. 21).

What is notable about these definitions is that, however implicitly, formative assessment is regarded as a process. Others have tended to regard formative assessment as a tool. For example, Stuart Kahl (2005) defined formative assessment as “a tool that teachers use to measure student grasp of specific topics and skills they are teaching. It’s a ‘mid-stream’ tool to identify specific student misconceptions and mistakes while the material is being taught” (p. 11). Indeed, it appears that the term *formative assessment* is now more often used to refer to a particular kind of assessment instrument than a process by which instruction might be improved.

The difficulty with trying to make the term *formative assessment* apply to a thing (the assessment itself) is that it just does not work. Consider an AP calculus teacher who is getting her students ready to take their examination. Like many teachers, she has her students take a practice examination under formal test conditions. Most teachers would then collect the papers, score them, write comments for the students, and return the papers to the students so that they could see where they went wrong. However, this calculus teacher does something slightly different. She collects the papers at the end of the examination, but she does not score them. Instead, during her next period with the class, each group of four students receives their unscored papers and one blank examination paper, and has to compile the best composite examination paper response that they can. Within each group, the students review their responses, comparing their answers to each question and discussing what the best answer would be. Toward the end of the period, the teacher reviews the activity with the whole class, asking each group to share with the rest of the class their agreed answers.

The assessment that the teacher used—an AP calculus examination—was designed entirely for summative purposes. AP exams are designed by the College Board to confer college-level credit so that students passing the exam at a suitable level are exempt from introductory courses in university. However, this teacher used the assessment instrument formatively—what Black and I have called “formative use of summative tests”. Describing an assessment as formative is, in fact, what Gilbert Ryle (1949) called a “category error”: the error of ascribing to something a property that it cannot