

Following is the speech delivered by Professor Dylan Wiliam at the ETS Europe breakfast salon which took place on the 11th July 2006.

DOES ASSESSMENT HINDER LEARNING?

Professor Dylan Wiliam, ETS

ETS Breakfast Salon, 11th July 2006

When teachers are asked how they assess their students, they typically talk about tests, examinations, quizzes and other formal methods. When they are asked how they know whether their students have learned what they have taught, the At one extreme we have a teacher questioning a student, trying to elicit evidence of (mis)conceptions that are likely to impede future learning. At the other extreme we have the use of A-level examinations by universities simply to decide which students to admit. The obvious conclusion is that the latter kind of assessment hinders learning while the former helps, but things are not that simple.

The teacher's questioning of the student can cause possibly irreparable damage to the student's sense of self if undertaken in a humiliating way. And at the other extreme, A-level examinations, by being made public as soon as they have been used, provide clear guidance to teachers and students about what the vague words in examination syllabuses mean. Furthermore, when used as "mock" examinations, past papers allow students to benchmark themselves against the standard established by the examination board, and to help each other rectify deficiencies. If we are to design assessment systems that help rather than hinder learning, we must go beyond looking at the assessments themselves and look at deeper issues about how the assessments help learners and their teachers know where the learners are in their learning, where they are going, and how to get there.

Through extensive reviews of the available research evidence, and through extensive field work with teachers, both in the UK and in the USA (see the suggestions for further reading at the end of this article), we have identified five 'key strategies', that when implemented appropriately, allow assessment to help, rather than hinder, learning.

1. Engineering effective classroom discussions, questions, activities and tasks that elicit evidence of student learning.

The first step in using assessment to help learning is to collect the right sort of evidence, and here it is clear that the tools that teachers use to find out where students are in their learning are given too little attention.

Few teachers plan the kinds of tasks, activities and questions that they use with their students specifically to elicit the right kind of evidence of student learning. As an example, consider the four diagrams shown below. In which of the following diagrams, is one quarter of the area shaded?



Diagram A is the obvious answer, but B is also correct. However, some students do not believe that one quarter of B is shaded because of a belief that the shaded parts have to be contiguous. Students who believe that one quarter of C is shaded have not understood that one region shaded out of four is not necessarily a quarter. Diagram D is perhaps the most interesting here. One quarter of this diagram is shaded, although the pieces are not all equal; students who rely too literally on the "equal areas" definition of fractions will say that D is not a correct response. By crafting questions that explicitly build in the under- and over-generalizations that we know students make, we can get far more useful information about what to do next. By equipping each student in the class with a set of cards with A, B, C and D on them, and by requiring all students to respond simultaneously with their answers, the teacher can generate a very solid evidence base for deciding whether the class is ready to move on. If every student responds with A, B and D, then the teacher can move on with confidence that the students have understood. If everyone simply responds with A, then the teacher may choose to re-teach some part of the topic. The most likely response, however, is for some students to respond correctly and for others to respond incorrectly, or incompletely. This provides the teacher with an opportunity to conduct a classroom discussion in which students with different views can be asked to justify their selections.

Of course, planning such questions takes time, but by investing the time before the lesson, the teacher is able to address students' confusion during the lesson, with the students still in front of them. Teachers who do not plan such questions are forced to put children's thinking back on track by marking, thus dealing with the students one at a time, after they have gone away...

2. Providing feedback that moves learning forward

The research on feedback shows that much of the feedback that students receive has, at best, no impact on learning, and can actually be counter-productive. One study (Kluger & DeNisi, 1996) reviewed over 3000 research reports on the effects of feedback in schools, colleges, and workplaces. They then rejected studies that failed to reach the highest standards of methodological rigour and were left with just 131 studies. Across these 131 studies,

they found that, on average, feedback did increase achievement, but that in 40% of the studies, feedback actually made people's performance worse than it would have been without feedback. The key feature of these studies was that feedback was, in the psychological jargon, 'ego-involving'. In other words, the feedback focused attention on the person rather than the quality of the work, for example by giving scores, grades or other forms of report that encouraged comparison with others. For the 60% of studies that found a positive impact on performance, Kluger & DeNisi found that the biggest impacts occurred when feedback told not just what to do to improve, but also how to go about it.

An example from cricket may be helpful here. If a young bowler is taking one wicket every 200 runs he gives up, then we know that he is not doing well. This is the monitoring assessment. The monitoring assessment identifies that there is a problem, but doesn't identify what it is. By looking at his bowling, we might see that the reason that he is ineffective in taking wickets is that he is just not bowling fast enough. This is the diagnostic assessment. The diagnostic assessment identifies where the problem is, but by itself, doesn't give the athlete any clue about how to go about making improvements. However, if the bowling coach can see that the reason that the cricketer is struggling to bowl fast enough is because he has a 'mixed' action—the axis of his shoulders is not aligned with the axis of his hips at the moment he delivers the ball—then this gives the athlete something to work with. This is the formative assessment. Just as we use the term 'formative' to describe the experiences that shape us as we grow up, a formative assessment is one that shapes learning.

3. Clarifying and sharing learning intentions and success criteria with learners

In an article entitled "The view from the student's desk" over thirty-five years ago, Mary Alice White (1971) wrote: The analogy that might make the student's view more comprehensible to adults is to imagine oneself on a ship sailing across an unknown sea, to an unknown destination. An adult would be desperate to know where he [sic] is going. But a child only knows he is going to school...The chart is neither available nor understandable to him... Very quickly, the daily life on board ship becomes all important ... The daily chores, the demands, the inspections, become the reality, not the voyage, nor the destination. (p. 340)

In a similar vein, I have walked into many mathematics classrooms, and asked students what they were doing, only to be told something like "Page thirty-four" as if that were all I needed to know. For many students, school is just a series of tasks where the purpose is unclear, and even what counts as success is mysterious, especially for students from less advantaged backgrounds. For , many students from advantaged backgrounds know that control of narrative, effective characterization, and keeping the reader's attention are important aspects of good writing, but many less advantaged students think that what matters is that it's neat and it's long.

A crucial role that assessment can play in promoting learning, therefore, is to help students understand the learning intentions that the teacher has for them, and what would count as success.

4. Activating students as owners of their own learning

One of the great traps of teaching is the belief that teachers create learning. This is particularly important when teachers are under pressure to improve student results, because studies have shown that when teachers are told they are responsible for making sure that their students do well, the quality of their teaching deteriorates, as does their students' learning (Deci et al. 1982), hence the old joke about schools being places where children go to watch teachers work.

Only learners create learning, and so, when we look at the role that assessment plays in promoting learning, the crucial feature is not the validity of the assessment, or its reliability, but its impact on the student. No matter how reliable or valid the assessment, if it communicates to students that they cannot learn, it will hinder learning. Particularly important here is the work of Carol Dweck, who over a thirty-year period has examined the way that students make sense of their successes and failures in school (see Dweck, 2000, for a very readable summary of this huge volume of work). As a result of their experiences, some students come to believe that ability is fixed. The reason that this is so injurious to future learning is that every time a student with this belief is faced with a challenging task, her or his first reaction is to engage in a calculation about whether they are likely to succeed or not. If they feel confident that they will succeed, or if they feel confident that the task is so hard that many others will fail, they will attempt the task. However, if they feel that there is a danger that they will fail while others will succeed, they will disengage in order to protect their sense of self. Put simply, they are deciding that they would rather be thought lazy than stupid. Given the stark choice between these two, it is the same choice that most adults would make.

There are other students, who, for a variety of reasons, have come to regard ability is incremental rather than fixed. They believe that 'clever' is not something you are but something you get. For these students, challenging tasks are opportunities to increase their abilities, so whether their beliefs in their chances of success are high or low, they engage with task in order to grow. What is particularly interesting is that the same student can believe the ability in school subjects is fixed, while ability in sports or music is incremental. Most students believe that ability in, for example, triple jump, throwing the javelin, or guitar playing can be improved by practice. We need to inculcate the same beliefs about academic subjects.

In general, we need to activate students as owners of their own learning, so that they see challenge as a spur to personal growth, rather than as a threat to self-image. We need students who own their learning to the extent that they can self-manage both their emotional and their cognitive responses to challenges, so that all their energies are spent on developing capability rather than disguising its absence (see Wiliam, 2007 for a summary of research in this area).

5. Activating students as learning resources for one another

The research on collaborative learning is one of the success stories of educational research. Research in many areas of education produces ambiguous or contradictory findings. However, the research on collaborative learning, most notably the work of Robert Slavin (Slavin, et al., 2003), has shown that activating students as learning resources for one another produces some of the largest gains seen in any educational interventions, provided two conditions are met. The first is that the learning environment must provide for group goals, so that students are working as a group, rather than just working in a group. The second is individual accountability, so that each student is responsible for their contribution to the whole, so there can be no 'passengers'.

With regard to assessment, then, a crucial feature is that the assessment encourages collaboration between students while they are learning. To achieve this, the learning intentions and success criteria must be accessible to the students (see above) and the teacher must support the students as they learn how to help each other improve their work. One particularly successful format for doing this has been the idea of 'two stars and a wish'. The idea is that when students are commenting on each others work, they do not give evaluative feedback, but instead have to identify two positive features of the work (two 'stars') and one feature that they feel merits further attention (the 'wish'). Teachers who have used this technique with students as young as five years old have been astonished to see how appropriate the comments are, and, because the feedback comes from a peer, rather than someone in authority over them, the recipient of the feedback appears to be more able to accept the feedback (in other words, they focus on growth rather than preserving well-being). In fact, teachers have told us that the feedback that students give each other, while accurate, is far more hard-hitting and direct than they would themselves feel able to provide. Furthermore, the research shows that the person providing the feedback benefits just as much as the recipient, because they are forced to internalize the learning intentions and success criteria in the context of someone else's work, which is less emotionally charged than doing in the context of one's own work.

The 'big idea': keeping learning on track

The 'big idea' that ties these strategies together is that assessment should be used to provide information to be used by students and teachers that is used to modify the teaching and learning activities in which they are engaged in order better to meet student needs. In other words, assessment is used to 'keep learning on track'.

That this is not common practice can be seen by imagining what would happen if an airline pilot navigated the way that most teachers teach. The pilot would set a course from the starting point (say London) to the destination (say New York). The pilot would then fly on this heading for the calculated time of travel, and then, when that time had elapsed, would land the 'plane at the nearest airport, and upon landing ask "Is this New York?" Worse, even if the 'plane had actually landed in Boston, the pilot would require all the passengers to leave, because he had to get on to his next job.

This would be absurd, and yet, this is how most teachers teach. They teach a topic for two or three weeks, and at the end of that teaching, they assess their students. And whatever the result of that assessment, the teacher is then on to the next topic, because she "has a syllabus to cover". If we are to 'keep learning on track' assessment cannot wait until the end of the topic. Instead, like the pilot, the teacher plans a course but then takes frequent readings along the way, adjusting the course as conditions dictate.

All this might seem obvious—perhaps little more than common sense—but there is substantial evidence, from Ofsted, from international surveys such as TIMSS, and from research studies, that such practices are rare. However, before we can argue that such changes in teachers' practices should be encouraged nation-wide, we need to be sure that the resulting changes in student achievement will be worth the investment that will be required to achieve them.

Value for money in educational reform

Many policymakers have focused on "what works" in education, but such a focus is misguided, since what is most important is the size of effect relative to cost. An intervention might "work" but the effects might be too small to be worth bothering with, or it might produce substantial effects, but be too expensive to implement. The focus in school improvement needs to be on cost-effectiveness.

When we adopt this perspective, we find that some effective interventions are too costly, and some interventions that produce only small impacts on student learning nevertheless turn out to be implementable at modest cost.

For example, the research shows clearly that reducing class size raises student achievement. So what? The important question should be how much improvement, and at what cost? And here the data are depressing.

For example, Jepsen and Rivkin (2002) found that reducing elementary school class size by ten students would increase the proportion of students passing typical mathematics and reading tests by 4% and 3% respectively. While there is evidence that the effects are larger for the early years, beyond this point, class-size reduction appears to be a very expensive way of increasing student achievement. To make this more precise, consider the cost of this intervention per classroom of 30 students. Reducing class-size from (say) 30 to 20 would increase the salary costs by approximately 50%, because we would need 3 teachers instead of 2 for a group of 60 students. Assuming an average teacher salary cost of £40,000 per year, this would cost £20,000 per year to implement.

Using tests such as those used in international comparisons such as TIMSS (Rodriguez, 2004) as a benchmark, we find that the effect of these class-size reductions would be equivalent to students in the smaller classes learning in 3 months what it would take the students in the larger classes 4 months to learn; a 33% increase in the speed of learning.

A recent study by Jenkins et al (2006) found that additional resources, if used to reduce class size, might have a larger increase on student achievement in secondary school science (up to four times the effect size found by Jepsen and Rivkin), but this model assumed that additional science teachers of equivalent quality to those already in post could be found, which is at the very least questionable. Nevertheless, it may be that there are special cases, such as early years, and secondary school science, where class size reduction may have a substantial effect on student achievement. However, in general, class-size reduction would appear to offer only modest increases in student achievement at very high cost.

How else could we obtain the same increase in student achievement? One obvious candidate is to increase teacher subject knowledge. Hill, Rowan & Ball (2005) found that an increase of one standard deviation in what they termed "mathematical knowledge for teaching" was associated with a 0.1 increase in student achievement. This was not a direct experiment, and so we cannot infer a causal relationship. However, it suggests that the same improvement in achievement that would be gained by reducing class size from 30 to 20 might be secured by increasing teacher subject-matter knowledge by one standard deviation. Unfortunately, there is currently little evidence about how to do this.

In contrast, supporting teachers in developing the use of assessment for learning has been shown to roughly double the speed of learning (William, Lee, Harrison & Black, 2004; Hayes, 2003). In other words, students learned in six months what would have taken a year to learn in other classrooms.

The cost of this intervention was around £5,000 per teacher, but of course, unlike class-size reduction, the cost of which is an annual cost, investing in teachers' professional development is a one-off cost, which can be depreciated over the teacher's remaining career. Against this, some cost of annual renewal needs to be allowed. Assuming that teachers in post will continue to teach for at least 5 years, and allowing time for 4 hours of meetings with colleagues per month for renewal, the cost of teacher professional development focused on assessment for learning would appear to be around £2,000 per teacher (and therefore per classroom) per year.

Compared to class-size reduction, therefore, improving teachers' use of assessment for learning would appear to promise two or three times the increase in student learning, for around one-tenth the cost. Even in the special case of secondary school science discussed above, assessment for learning produces the roughly the same size of benefit as reducing class-size by 30%, at less than one-fifth the cost.

Putting it into practice: the case for teacher learning communities

The fact that we know what needs to be done does not, of course, mean that we know how to do it. While the work Black et al. (2003) has shown what we can achieve, the track record of professional development in producing significant effects on a large scale is rather unimpressive. However, this should not worry us unduly because very little of the professional development that teachers have received in the past is consistent with what we know makes for effective teacher change (William & Thompson, 2007).

Why this is the case is complex, and beyond the scope of this article (see William, 2003, for an extended discussion). What is clear is that in general, researchers have underestimated the complexity of what it is that teachers do, and in particular, have failed to understand how great an impact context has on teachers' practice. That is why "what works?" is not the right question, because everything works somewhere, and nothing works everywhere. The right question is "Under what conditions will this work?" And even if we might be able to answer this question scientifically at some point in the future, we are so far away from an answer now that we have to rely on the professional judgment of teachers.

This is why we cannot tell teachers what to do. This conclusion does not stem from a desire to be nice to teachers. Indeed, if I could identify a way of telling teachers what to do that would raise student achievement, I would have no hesitation in mandating it if I had the power to do so. The reason that we cannot tell teachers what to do is that we cannot provide them with reliable guides to action. The situations that they face in their classrooms are just too varied for us to predict. That is why professional judgment is important; we have to develop the ability of teachers to react appropriately to situations for which they have not been specifically prepared.

The specific changes that I am arguing for here appear to be quite difficult for teachers to implement, because they involving changing habits. A teacher with 20 years experience will have asked something like half a million questions in her career. And when you've done something the same way, half a million times, it's quite difficult to start doing it in any other way. But there is a deeper reason why change is difficult, even for inexperienced teachers. Teachers learn most of what they know about teaching before they are eighteen years old. In the same way that most of us learn what we know about parenting through being parented, teachers internalize the 'scripts' of school as students. Even the best four-year teacher education programmes will find it hard to overcome the models of practice that students teachers will have learned in the thirteen or fourteen years that they spent in school.

The process that I believe provides the best mechanism for supporting teachers in making these changes is through the use of teacher learning communities (TLCs). This is not out of some ideological commitment to the benefits of teachers talking to each, but because of the nature of the changes we are seeking to produce. If we were trying to increase teacher subject knowledge, then TLCs would not be a very sensible approach—it would be far better to arrange for high-quality direct teaching. But when we are trying to change deeply ingrained, routinized practices, then it seems that TLCs offer the best hope, and indeed, the results we have achieved in the United States have been very encouraging.

Over the last three years, we have tried a number of different approaches to establishing and sustaining TLCs, and as a result of this experimentation, it appears to us that five principles appear to be particularly important: gradualism, flexibility, choice, accountability and support.

Gradualism

Asking teachers to make wholesale changes in their practice is a little like asking a golfer to change her swing during a tournament. Teachers have to maintain the fluency of their classroom routines, while at the same time disrupting them. Teachers should develop an action plan that specifies a small number of changes — ideally two or three — that they will make in their teaching. As teachers establish new techniques in their practices, they can take on additional ones. For those in the school leadership team whose job it is to improve the quality of teaching, there will be a temptation to push teachers to change faster than they might otherwise do, but the result will only be a shallow adoption of the new practices while the teachers are being monitored. As soon as the supervision is relaxed, the teachers will revert to their earlier practices, and nothing will have been achieved.

Flexibility

Teachers should modify techniques to make them work in their classrooms; in the process of adapting techniques, teachers often refine and improve them. One high school mathematics teacher heard about the “traffic light” technique in which at the end of a piece of work, students indicate their confidence in their understanding of a piece of work with a green, yellow or red circle, representing complete, partial or little understanding. She decided that she would not wait until the end of the lesson to engage students in this self-assessment, and gave each student a disc, green on one side and red on the other. At the start of each lesson students place the disk on their desk with the green side of uppermost. A student can indicate confusion at any time by turning the disc over to show red. The teacher found that students who had never asked a question all year in class were prepared to signal their confusion in this way. Another teacher tried this approach, but found it difficult to see the discs from the front of the class so she provided each student with three paper cups, one green, one yellow and one red, nested inside each other on the students’ desks. Students used these cups to indicate whether they were following the teacher’s explanation (green), wanted the teacher to slow down (yellow), or wanted her to stop to ask a question (red). The teacher made students accountable for signaling correctly by establishing that whenever one student showed red, a student who was showing green or yellow would be chosen at random to come to the front of the class to answer the question posed by the student showing red. In this classroom there is nowhere to hide!

Choice

As noted above, teachers often describe the process of changing their practice as “scary”, but when they are responsible for choosing what they will change about their practice they feel empowered, especially when they can choose among a range of techniques those that appeal to them. This choice lies, however, within a framework of accountability. While teachers are free to choose what they change, they are accountable for changing something.

Accountability

Most professionals involved in teacher development will have had the experience of generating considerable enthusiasm for, and commitment to, change during a workshop, only to find that all the good intentions seem to be erased once the teachers return to the classroom. Teachers should be held accountable for making changes by colleagues at monthly meetings of their teacher learning community. Each teacher describes what he or she tried and how it went. Teachers repeatedly tell us that having to face their colleagues helps them move their “change” task to the top of their in-tray.

Support

Along with ideas for what to change and the support of a teacher learning community, two elements are highly desirable if not essential for teacher learning. The first is training for those who will lead the learning communities. The person leading the learning community must be clear about his or her role. The leader is not to create teacher change, but to engineer situations in which the teacher change can take place. Those in supervisory roles often find this more difficult than do teachers, because those teaching every day understand how difficult it is to change practice. The second element is peer-observation. Collaborative planning in the monthly TLC meetings can help teachers focus on what they want to develop in their practice, but teachers need support in enacting these practices in their classrooms. To clearly distinguish these observations from those routinely carried out to manage performance, these observations should be done by peers rather than supervisors. The teacher being observed must set the agenda for the observation and spell out for the observer what should count as evidence. By defining the observer’s role, both in terms of what is to be looked for and what counts as evidence, the observer’s own prejudices are minimized, and the difference between this and supervisory observation is emphasized.

Balanced assessment

The available research evidence, as well as our experience of working with LEAs, suggests that the use of teacher learning communities, focused on the use of minute-to-minute and day-to-day assessment to adjust teaching to meet student needs, represents the most powerful single approach to improving student achievement. However, if we are to maximize the impact on student learning, other parts of the system need to be ‘in sync’. As well as

the minute-by-minute and day-by-day assessments that allow teachers to “keep learning on track”, teachers also need a range of more formal assessment tasks and activities that support valid and reliable conclusions about the extent of student learning. Our experience, both in the US and in the UK, is that teacher-made assessments often focus on shallow aspects of learning, rather than the ‘big ideas’. Developing high-quality assessments that involve students, and motivate them to improve takes time, but there are good examples of how to go about this (Stiggins et al., 2004). On a longer time-scale, termly assessments that are paced to the curriculum being taught can provide school leadership teams with valuable information about the progress—or lack of it—that is being made by students, and annual diagnostic analyses of high-stakes tests can provide important insights into the alignment between the teaching and the national curriculum. None of these different kinds of assessment is in conflict with any of the others. Each represents an important part of complex machine providing information at the right level of specificity for the decision that needs to be made. Together they form a balanced assessment system that can produce unprecedented increases in student achievement, benefiting both the individual and society as whole.

Following are some of the links which were referred to during the speech:

Classroom assessment: minute by minute and day by day can be downloaded free of charge by clicking [here](#).

The “Inside the black box” series is available from NFER-Nelson:

<http://www.nfer-nelson.co.uk/>

American versions of two of the publications can be downloaded from the Phi Delta Kappan web-site for \$3 apiece:

<http://www.pdkintl.org/>

References

- Deci, E. L., Speigel, N. H., Ryan, R. M., Koestner, R., & Kauffman, M. (1982). The effects of performance standards on teaching styles: the behavior of controlling teachers. *Journal of Educational Psychology*, 74, 852-859.
- Dweck, C. S. (2000). *Self-theories: their role in motivation, personality and development*. Philadelphia, PA: Psychology Press.
- Hayes, V. P. (2003). Using pupil self-evaluation within the formative assessment paradigm as a pedagogical tool. Unpublished EdD thesis, University of London.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers’ mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 317-406.
- Jenkins, A., Levacic, R., & Vignoles, A. (2006). Estimating the relationship between school resources and pupil attainment at GCSE (Vol. RR727). London, UK: Department for Education and Skills.
- Jepsen, C., & Rivkin, S. G. (2002). What is the tradeoff between smaller classes and teacher quality? Cambridge, MA: National Bureau of Economic Research.
- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284.
- Rodriguez, M. C. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education*, 17(1), 1-24.
- Slavin, R. E., Hurley, E. A., & Chamberlain, A. M. (2003). Cooperative learning and achievement. In W. M. Reynolds & G. J. Miller (Eds.), *Handbook of psychology volume 7: educational psychology* (pp. 177-198). Hoboken, NJ: Wiley.
- Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment for student learning: doing it right—using it well*. Portland, OR: Assessment Training Institute.
- White, M. A. (1971). The view from the student’s desk. In M. L. Silberman (Ed.), *The experience of schooling* (pp. 337-345). New York, NY: Rinehart and Winston.
- William, D. (2003). The impact of educational research on mathematics education. In A. Bishop, M. A. Clements, C. Keitel, J. Kilpatrick & F. K. S. Leung (Eds.), *Second International Handbook of Mathematics Education* (pp. 469-488). Dordrecht, Netherlands: Kluwer Academic Publishers.
- William, D. (2007). Keeping learning on track: formative assessment and the regulation of learning. In F. K. Lester Jr (Ed.), *Second handbook of mathematics teaching and learning*. Greenwich, CT: Information Age Publishing.
- William, D., & Thompson, M. (to appear, 2007). Integrating assessment with instruction: what will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: shaping teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- William, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education: Principles Policy and Practice*, 11(1), 49-65.
- Further reading

A) The ‘black box’ series

- Black, P. J., & William, D. (1998). *Inside the black box: raising standards through classroom assessment*. London, UK: NFER-Nelson.
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2002). *Working inside the black box: assessment for learning in the classroom*. London, UK: NFER-Nelson.
- Black, P., & Harrison, C. (2002). *Science inside the black box: assessment for learning in the science classroom*. London, UK: NFER-Nelson.
- Hodgen, J., & William, D. (2006). *Mathematics inside the black box: assessment for learning in the mathematics*

classroom. London, UK: NFER-Nelson.

Marshall, B., & Wiliam, D. (2006). *English inside the black box: assessment for learning in the English classroom*. London, UK: NFER-Nelson.

B) Other publications

Leahy, S., Lyon, C., Thompson, M., & Wiliam, D. (2005). Classroom assessment: minute-by-minute and day-by-day. *Educational Leadership*, 63(3), 18-24.

Wiliam, D. (2006). Assessment: learning communities can use it to engineer a bridge connecting teaching and learning. *Journal of Staff Development*, 27(1), 16-20.

Wiliam, D. (2006). Assessment for learning: why, what and how. *Orbit: OISE/UT's magazine for schools*, 36(3).