Assessment for learning: why, what and how?

Dylan Wiliam, University of London Institute of Education

Orbit **36**(2); 2006

Why do we need to raise achievement?

Education matters. It matters for individuals and it matters for societies. For the individual, greater levels of education are associated with greater control over one's life, greater economic wealth and a longer and healthier life (OECD, 2001 p. 33). For societies, increasing levels of educational achievement result in lower criminal justice costs, lower costs of state-provided health-care (Schweinhart et al., 2005) and increased economic growth. Quantifying the amount of extra economic growth inevitably involves some pretty heroic assumptions. Nevertheless, Hanushek (2004) has calculated that if we could increase student achievement by one standard deviation—equivalent to about 15 percentage points on a typical test where the scores range from around 20% to around 80%—then in 30 years, the increased growth in the economy would more than pay for the entire cost of K-12 education.

Where's the solution?

Around the world, many approaches to school improvement have been proposed, but most have been relatively ineffective, because they have failed to embrace a lasting and inevitable truth which is this: the single most important variable in the amount of progress that a student makes at school is the quality of the teacher.

This has not been obvious because for many years, we lacked the right kinds of data-sets. There is a huge variability in the levels of achievement of students in different schools, and simplistic analyses suggested that these differences must be attributable to the schools. However, it rapidly became clear that one of the major reasons for these differences in the "outputs" of schools was that the schools served very different communities. Analyses that took into account the socio-economic status (SES) of students in different schools painted a very different picture—it was SES, rather than school quality that was the major determinant of educational achievement. Still others attributed the differences to the students themselves; their prior achievement for example.

However, as more and more datasets began to consider the "value added" by the school —put simply the difference between what the student knew when they started at that school and what they knew when they left—it became clear that the quality of the teacher was the most significant variable. If you get one of the best teachers, you will learn in six months what an average teacher will teach you in a year, and if you get one of the worst teachers, the same amount of progress will take you two years—there is a fourfold difference between the productivity of the best and worst teachers (Hanushek, 2004; Wiliam & Thompson, in press).

Improving education is therefore, in effect, a labour-force issue with two solutions. One is to replace existing teachers with better ones. The problem with this is that there is no evidence that there are better teachers queuing up to get in, but can't get jobs (Darling-Hammond *et al.*, 2005). Nor is there that much evidence that better pay would have a substantial impact on teacher quality (there are many good reasons for paying teachers more, but the supply-and-demand argument is not one of them). The other solution is to invest in the teachers that we have—what my colleague Marnie Thompson calls the "love the one you're with" strategy.

Twenty years ago, this would have resulted in a very gloomy prognosis, since there was little evidence that it was possible to improve the quality of a teacher's practice, except very slowly, and a great cost (Fullan, 1991). Many people therefore concluded that the situation was hopeless, and that policy makers should, instead focus on solutions that did not require quality teaching (such as producing "teacher-proof" text-books, so that it didn't matter how good the teacher was).

However, the problem with this argument is that we can conclude very little from the failure of years of teacher professional development (PD) to improve teaching, because for the vast majority of teachers, the kind of PD they received flew in the face of what the research says about what makes for effective PD. That research says that PD needs to be continuing, rather than in the form of "one-shot" inputs (Cohen & Hill, 1998), it needs to be tailored to the local circumstances in which teachers work (Cobb et al., 2003), it needs to be directly related to the content of the subjects they teach (Supovitz, 2001), and in needs to involve teachers in active and collective participation (Garet, Birman, Porter, Desimone, & Herman, 1999).

So much for the "process" of teacher PD. What about the "content" of teacher PD? Does it matter what teachers focus on, as long as they are engaged in the processes listed above? It turns out that it does. It turns out that teachers can spend a lot of time developing their thinking, without changing their practice very much, and from the research it appears that one particular focus—the use of assessment for learning—has a bigger impact on student achievement than any other.

Assessment for *learning*

For over 20 years, Paul Black and I have been researching the role that assessment can play in raising levels of student achievement. In particular we wanted to find out if using assessment to *support* learning, rather than just to measure its results, could improve students' achievement, even when such achievement is measured in the form of state-mandated tests. In reviewing 250 studies from around the world, published between 1987 and 1998, we found that a focus by teachers on assessment *for* learning, as opposed to assessment *of* learning, produced substantial increases in students' achievement—typically doubling the rate of learning (Black & Wiliam, 1998). Since the studies also revealed that day-to-day classroom assessment was relatively rare, we felt that

considerable improvements would result from supporting teachers in developing this aspect of their practice.

One of the things that we have struggled with is what, exactly, is assessment for learning. Many people have proposed definitions, all of which differ slightly, but it seems that the "big idea" is that evidence about student achievement is used to adapt instruction to better meet learning needs. The key strategies involved in this are (Wiliam & Thompson, in press):

- Clarifying and sharing learning intentions and success criteria with learners
- Engineering effective classroom discussions, activities and tasks that elicit evidence of student achievement
- Providing feedback that moves learners on
- Activating students as owners of their own learning; and
- Activating students as instructional resources for one another.

There is ample evidence that focusing on any one of these five strategies is highly effective in raising student achievement (Wiliam, in press). Together they provide the highest leverage focus that we know, for improving schools, classrooms, and the achievement of students. The problem is that knowing what needs to be done is one thing. Doing it is quite another.

Putting it into practice

We know that "one size fits all" does not work in teacher professional development. What might be exactly the right thing to do in one situation might be exactly the wrong thing to do in a similar, but different situation. Expert teachers do not use general, all-purpose approaches to solve problems, but rather generate solutions that take advantage of specific details in the challenges they face (Berliner, 1994). That is why "what works" is not the right question in education; everything works somewhere, and nothing works everywhere. The right question is "under what conditions does this work?" That is why if we are serious about developing assessment for learning, we must help each teacher find her or his own way of doing this.

The opposite of "one size fits all" is to allow each teacher to choose what they want to do. This may be attractive, but it is not rigorous. Teachers may choose to change things that have no impact on student achievement. The trick is, therefore, to generate teacher professional development that is open enough to allow each teacher to adapt the new ideas into their own practice, but structured enough so that the adaptation does not produce a "lethal mutation" (Pellegrino, 2006) that renders the innovation impotent.

Over the last three years, I and my colleagues at the Learning and Teaching Research Center at the Educational Testing Service in Princeton, New Jersey, USA have been working with many groups of teachers on the best ways to implement assessment for learning. Some approaches have been extremely successful, and others less so, but even from our failures, we have usually been able to learn something (and often more than

from our successes). As a result of reflection on both our successes and failures, we have been able to identify five elements that increase the successful implementation of assessment for learning: accountability, support, choice, flexibility, and gradualism.

Accountability

Most professionals involved in teacher development will have had the experience of generating considerable enthusiasm for, and commitment to, change during a summer workshop, only to find that all the good intentions seem to be erased by the demands of the new school year. That is why we suggest that teachers should make a commitment in writing about what they are going to change about their practice (the action plan) and then to be held accountable by colleagues at monthly meetings for making those changes in their practice. Each month, every teacher describes what he or she tried and how it went. Teachers have told us that having to face their colleagues and "deliver" on the promises they made the previous month helps them move their "change" task to the top of their in-tray.

Support

The other side of the coin of accountability is support. We have found that the at the monthly meetings, teachers are able to offer each other advice when the planned changes are not going well, and the fact that these meetings are groups of peers, rather than an expert teaching a novice, appears to be particularly beneficial. It must be borne in mind however, that ultimately, implementing assessment for learning involves changes in day-to-day, and even minute-to-minute classroom practice. That is why we also recommend that teachers engage in peer observation. To clearly distinguish these observations from those routinely carried out to manage performance, these observations should be done by genuine peers rather than those in a hierarchical relationship. Another important requirement is that the teacher being observed must set the agenda for the observation and spell out for the observer what should count as evidence, by reference to her or his action plan. 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.

Choice

Teachers often describe the process of changing their practice as "scary". Implementing assessment for learning in particular makes many teachers feel as if they are being asked to "give up control" of their classrooms. However, teachers responsible for choosing what they will change about their practice feel empowered, especially when they can choose among 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. They are also accountable for showing to their peers how their innovations are consistent with the principles of assessment for learning. One especially powerful way to support this is for peers to ask each teacher, "What's formative about that?" in order to emphasize the need for the changes in practice to be focused on using assessment evidence to adapt instruction to meet student needs.

Flexibility

A technique that works for one teacher may not work for another, but may do so after some modification, Teachers need to be encouraged to modify techniques to make them work in their classrooms, but they also need to understand enough of the research evidence so that the changes they make do not render the innovation ineffective.

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. The action plans that teachers develop should specify a small number of changes—ideally two or three—that they will make in their teaching. In our experience, teachers who try to change more than three things in their practice at the same time are never successful. They try to do too much, and their classroom descends into chaos, as a result of which they revert to what they know how to do. In other words, going slower produces faster real change.

Conclusion

Our experience is that, equipped with some basic ideas about assessment for learning, teachers can support each other in making radical improvements in their students' learning, provided each teacher is responsible for her or his own development targets. It is perfectly OK to hold teachers accountable for making changes in their practice, provided the teachers have choice in how they put the basic principles of assessment for learning into effect, that they are not pushed to make changes faster than they can incorporate them into their normal practice, and that they have the support and accountability that a group of like-minded group of peers can provide. Teachers doing it for themselves…

References

Berliner, D. C. (1994). Expertise: the wonder of exemplary performances. In J. N. Mangieri & C. C. Block (Eds.), *Creating powerful thinking in teachers and students: diverse perspectives* (pp. 161-186). Fort Worth, TX: Harcourt Brace College.

Black, P. J., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles Policy and Practice,* **5**(1), 7-73.

Cobb, P., McClain, K., Lamberg, T. d. S., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher,* **32**(6), 13-24.

Cohen, D. K., & Hill, H. C. (1998). *State policy and classroom performance: mathematics reform in California*. Philadelphia, PA: University of Pennsylvania Consortium for Policy Research in Education.

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Vasquez Heilig, J. (2005). Does teacher preparation matter? Evidence about teacher certification, teach for America, and teacher effectiveness. *Education Policy Analysis Archives,* **13**(42).

Fullan, M. (1991). *The new meaning of educational change*. London, UK: Cassell.

Garet, M. S., Birman, B. F., Porter, A. C., Desimone, L., & Herman, R. (1999). *Designing effective professional development: lessons from the Eisenhower Program.* Washington, DC: US Department of Education.

Hanushek, E. A. (2004). *Some simple analytics of school quality* (NBER working paper no. W10229). Washington, DC: National Bureau of Economic Research.

Pellegrino, J. W. (2006, Mar 17). Personal communication.

Schweinhart, L. J., Montie, J., Xiang, Z., Barnett, W. S., Belfield, C. R., & Nores, M. (2005). *Lifetime effects: the High/Scope Perry Preschool Study through age 40.* Ypsilanti, MI: High/Scope Educational Research Foundation.

Supovitz, J. A. (2001). Translating teaching practice into improved student achievement. In S. H. Fuhrman (Ed.), *From the capitol to the classroom: standards-based reform in the States* (Vol. Part 2, pp. 81-98). Chcago, IL: University of Chicago Press.

Wiliam, D. (to appear, 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.

Wiliam, 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.

Dylan Wiliam is deputy director of the University of London Institute of Education. He can be reached at dylanwiliam@mac.com. For further details of the professional development materials developed by the Learning and Teaching Research Center at ETS, please contact Susan Baer at sbaer@ets.org.