

Supporting the development of classroom formative assessment: the role of leaders.

Expert article series n.4

Dylan Wiliam, *UCL Institute of Education*

Over the last fifty years, policymakers around the world have realized that improving educational achievement is important not just for economic success, but also for the maintenance of healthy democracies, and, most importantly, to allow individuals to take greater control over their lives.

The search for solutions has included increasing, or reducing, government control of schools, increasing, or reducing the use of the private sector, specifying the curriculum in more, or less, detail, and dozens, if not hundreds, of other ideas. However, over the last twenty years, there has been growing acceptance that the quality of individual teachers is one of, if not the, most important factor in determining the quality of an education system.

Predictably, in some countries, this has led to calls for raising the bar for entry into the teaching profession, retaining more effective teachers by paying them more, or removing less effective teachers, but unfortunately, none of these things is likely to have much impact.

First, it does not seem to be possible to predict who will be good teachers from their qualifications or the training courses they attend. Second, identifying more effective teachers is almost impossible. Most people in education believe that they “know good teaching when they see it” but the available evidence suggests that they don’t. Evaluations of teachers are affected by the prior achievement of the students—every teacher looks better when they are teaching higher-achieving students—and learning that looks effective may be quickly forgotten. Furthermore,

even using measures of student achievement rather than observations doesn’t work, because teachers who appear to be effective in the short-term often do not prepare students well for future learning—every teacher builds on the foundations laid by her predecessors. Third, because we aren’t able to identify less effective teachers with any accuracy, if we do get rid of the teachers our measures indicate are less effective, we are likely to be getting rid of quite a few teachers who are above average.

Perhaps more importantly, even if we were able to remove less-effective teachers and replace them with better ones, the number of such teachers is small, so that the impact on average teacher quality, and therefore the impact on student achievement,



will be small. The best way to improve an education system is to invest in the teachers we already have—what we might call the “love the one you’re with strategy”.

Many school leaders realize this and focus their efforts on supporting less effective teachers—the idea is that we support the teachers who are “struggling” and leave those identified as more effective alone to get on with their job. But in

most education systems, there are simply not enough “struggling” teachers for this to produce much improvement across the whole system. The only way to provide substantial, system-wide improvements to student achievement is to expect that every single teacher needs to get better, not because they are not good enough, but because they can be even better.

The main task of the school leader, therefore, is to create a culture where every single teacher accepts that they need to improve, even if they are already highly effective. Some teachers may resist this idea, saying, “I’m already the best teacher in the school. I get great results. Why should I improve?” There are several responses to this, but in my experience, two are particularly effective. The first is to re-engage teachers with the moral purpose of teaching. I don’t know of any teachers who joined the profession to improve test scores. They became teachers in order to make a difference in the lives of students, and reminding teachers that when they do their job better, their students will be healthier, live longer, and contribute more to society can often re-energize teachers. The second is to ask the teacher whether they think there are any ways in which their teaching could be improved. I do not know of any teacher who would reply that their teaching is perfect in every single respect. As soon as a teacher identifies one small thing that they might improve, the leader can then say, “How can I help you do that?” By keeping the “top-down” component of improvement to a minimum—every teacher at the school will be getting better at something—the “bottom-up” component, is maximized. By treating each teacher as an expert in their own practice, and not dictating what they should get better at, the amount of resistance from more experienced teachers is likely to be reduced.

Once teachers have accepted the need to improve continually, the leaders also need to ensure that teachers get better at the things that matter to their students. It is frankly self-indulgent for teachers to get better at things that do not benefit their students, and this is where research comes in.

Research will never tell teachers or leaders what to do—schools and classrooms are just too complex for this ever to be possible—but what research can do is to identify things that research shows are unlikely to be of much help, like paying attention to students’ learning styles, and also to point to the “best bets” for improvement, by asking four

questions of research:

1. *Does it solve a problem we have?*
2. *If we do this, how much extra achievement will we get?*
3. *How much will it cost, in money, and, more importantly, in teacher time?*
4. *Can we implement it here?*

Does it solve a problem we have?

Many research studies show that certain factors are associated with higher student achievement, and it might be tempting to work on these factors. For example, research shows that when students are taught by more knowledgeable teachers, they make more progress. However, if the teachers in a particular school already have good subject knowledge, then increasing it further is unlikely to have much impact. In addition, to be useful, the research must identify factors that can be changed, and not just which factors are influential. We know that taller people have an advantage in basketball, but as one coach famously remarked, “You can’t teach height”.

If we do this, how much extra achievement will we get?

Often, policymakers ask about “What works” but the fact that a study shows that a particular intervention has a statistically significant impact on student achievement just tells us that the result is unlikely to be due to chance. It doesn’t tell us how big the impact is.

Many studies report the size of the impact on student achievement using standardized effect sizes. However, because the measures used in different research studies differ in how sensitive they are to the effects of teaching, results of different studies cannot really be compared unless they use the same assessments to measure achievement, in which case, you can simply use the scores on the assessment. A stark illustration of this was a review by Maria Ruiz-Primo and Min Li that looked at the effects of feedback in mathematics, science and technology education. In those studies that measured student achievement with assessments that matched what students were learning, the average effect of feedback was more than five times greater than those studies that measured student performance with standardized tests, which are generally rather insensitive to the effects of teaching. The important point here is that different measures of achievement differ in

how sensitive they are to the things that teachers are changing, and so analyses that combine the effects of studies that use different achievement measures are not really comparing like with like. Strictly speaking, effect sizes are justifiable only when they are unnecessary.

What schools need is evidence about the likely impact on student achievement in months of extra learning per year. Researchers often argue that this is difficult, and argue for more easily calculated measures like changes in percentile ranks, but this



is a bit like the drunk man looking for his keys underneath the streetlight, not because that's where he dropped the keys, but that's where the light is. Any improvement in educational processes will result in students learning more in a given time period, and leaders need to know how much more before they can make smart decisions about what to prioritize.

How much will it cost, in money, and, more importantly, in teacher time?

While it might be tempting to ignore interventions that have a small impact on student achievement, such interventions can be an important part of any school's improvement efforts if it doesn't take too much time, and doesn't cost too much money. For example, while interventions that help students see ability as something that they can change (a "growth mindset") have a relatively small impact on student achievement, such changes can often be produced by an intervention that takes less than an hour, and lasts at least a year. Other examples, discussed in more detail below, are providing teachers with improved resources such as "off-the-shelf" lesson plans or improved textbooks, and supporting teachers in making greater use

of classroom formative assessment. In education, we are always looking for the "next big thing" but as far as we can tell right now, improvements are more likely to come from old, mostly small, things.

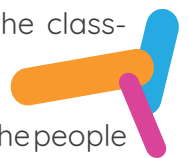
Can we implement it here?

Perhaps the most neglected aspect of applying research to the challenge of improving education is that "What works?" is usually the wrong question, because everything works somewhere and nothing works everywhere. The important question is "Under what circumstances does this work?"

For example, when the same teacher teaches a smaller class, students usually make faster progress, especially when the teacher is provided with professional development that helps them use teaching approaches that take advantage of the smaller classes. However, as well as being expensive, reducing class size means

that more teachers are needed—reducing class-size from 30 down to 20 requires 50% more teachers and the big question is whether the additional teachers are as good as the ones already employed. If there is a plentiful supply of highly effective teachers waiting for jobs, then class-size reduction is likely to improve student achievement (albeit at high cost), but where teacher recruitment is challenging, reducing class size might actually make things worse, because the additional teachers being employed might be so much less effective than existing teachers that the teacher-quality effect overpowers the class-size effect.

The important point here is that generally the people "on the ground"—teachers and leaders—are likely to be better able to judge whether a particular reform can be implemented in a particular setting than those in ministries of education and district offices. School leaders therefore need to become "critical consumers" of educational research. The four questions discussed above are probably the best starting point for engaging with research evidence.



Because different things will work differently in different systems and schools, there can't be a single guaranteed solution. However, right now, for most schools, there are two things that the available research evidence suggests are two "best bets" for educational improvement—the things that are most likely to have the biggest impact on student achievement for the lowest cost. These are curriculum development, and helping teachers develop their practice of formative assessment.

With the same teachers, a more effective curriculum can result in 25% more progress, for very little additional cost—after all good textbooks usually cost the same as bad ones. The problem with focusing on curriculum is that while we know that curriculum makes a difference, we don't know what makes the difference in curriculum. We can, retrospectively, determine whether some curricula are more effective than others (which is where the 25% figure quoted above comes from) but right now, we are not able to predict in advance which curricula are going to be more effective. Plus, given the differences in the way that different jurisdictions measure student achievement, what works well in one system might be less effective in another.

Which leaves classroom formative assessment. From the pioneering work of Benjamin Bloom over fifty years ago, we now know that supporting teachers in making greater use of classroom formative assessment—not just week to week but also day to day and minute to minute—can have a profound impact on student learning. More importantly, over the last twenty years, a great deal has been learned about how to help teachers develop their practice of formative assessment. Some of these approaches have involved intense face-to-face work with teachers. This approach can provide valuable lessons about how teachers can incorporate formative assessment into their work, even in systems that are driven by standardized tests and examinations, but obviously, such an approach is not easy to scale up across tens of thousands of classrooms. However, by studying in detail the way that teachers can incorporate formative assessment into their regular teaching, we have been able to make substantial progress on effective, scalable, teacher professional development in formative assessment. In particular, we have learned that teacher professional development is more effective when it gives teachers choice about which aspects of formative assessment are most relevant to them, encourages

them to adapt techniques to their own classrooms, acknowledges that changes to classroom habits take time, makes teachers accountable for making changes in their classroom practice, but also provides support for them to do this. These principles have been implemented in a two-year teacher professional development programme titled Embedding formative assessment (EFA).

EFA is designed to be implemented by schools without any external support, and provides teachers with practical ideas for how they might develop their practice of formative assessment, together with videos of practice, and detailed agendas and hand-outs for monthly teacher-led meetings. A randomized control trial involving 140 high schools in England found that students taught by teachers given access to the EFA programme made 25% more progress in 9th and 10th grade, even though the monthly meetings took up only 1% of teachers' time, and the programme's cost was around €1.50 per student per year. Right now, we know of nothing else that has such a large impact on student achievement, for such a small cost (in money and in teachers' time)

As noted above, research will never tell leaders what to do. But by creating a culture where every teacher accepts the need to improve, not because they are not good enough, but because they can be even better; by focusing teachers' efforts on the things that have the greatest benefits for their students; by giving teachers time, support, and permission to innovate, and by supporting teachers in taking risks, leaders can produce substantial, lasting improvements to their schools, and prepare our young people to flourish in the complex, unpredictable world they will face when they leave school.

