

# Hacer de la evaluación formativa una realidad: siete cambios necesarios

Dylan Wiliam  
Instituto de Educación de la UCL

Durante los últimos 25 años, hemos reunido numerosas pruebas que indican que la evaluación puede mejorar la educación, así como medir sus resultados. Sin embargo, también hemos aprendido que utilizar dicha evaluación para mejorar el aprendizaje va mucho más allá de añadir algunas «soluciones rápidas» a la labor del profesorado. Aprovechar el poder que tiene la evaluación para mejorar el aprendizaje implica introducir una serie de grandes cambios en la forma en la que concebimos la educación y la política. Si no se introducen esos cambios, es probable que no se logre aprovechar el potencial que tiene la evaluación para mejorar el aprendizaje. A continuación, se resumen los cambios que parecen más pertinentes para el contexto político actual en Europa.

## 1. Curriculum: from coverage to outcomes

One common reason given by teachers for not using formative assessment is that they have a broad curriculum to cover, and if they take time to make their teaching responsive to their students' needs, they won't get through all the material they have to cover. This is a valid point. There is no doubt that using formative assessment slows the teaching down, because teachers find out that they need to spend more on a topic than they had planned. One response to this is never to do formative assessment. After all, if you have no intention of doing anything with the evidence you collect, why collect it in the first place? That belief is logically consistent, but in my view, immoral. The teacher's job is not to "get through the curriculum" but rather to ensure that as many students learn as much as possible. Teachers must focus less on what they are putting in to the process, and more on what the students are getting out of it.

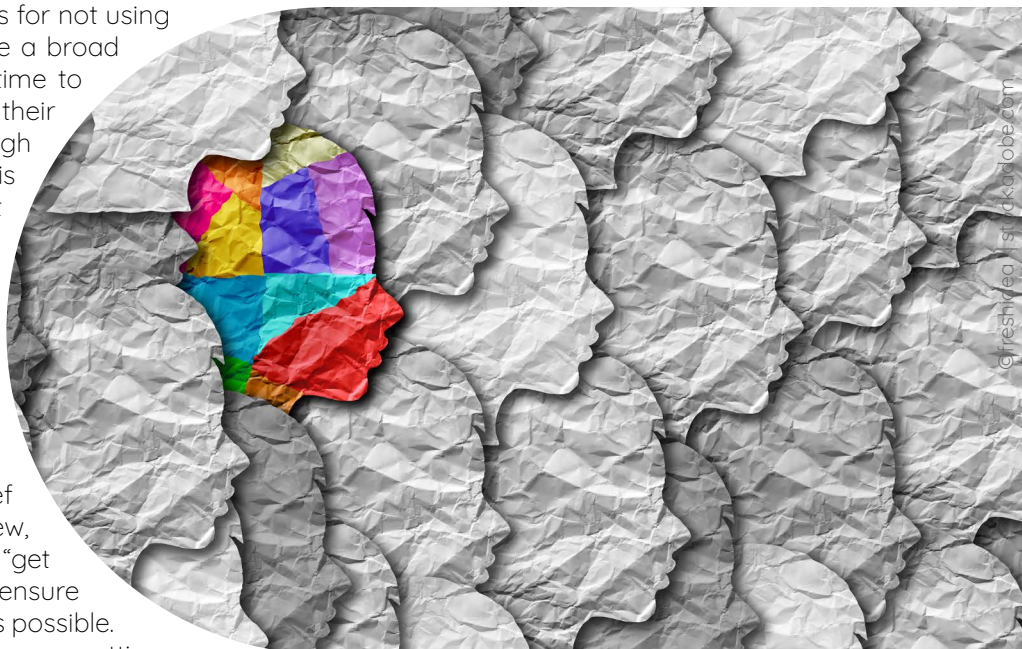
## 2. Aptitude: from 'proportion of material retained' to 'time needed to reach mastery'

Many people think of 'aptitude' as how much students can learn. Exposed to the same teaching, some students learn a great deal, and others learn less. The traditional

view of aptitude is how much can students learn? Building on the work of John Carroll, Benjamin Bloom suggested that we should turn this idea on its head and instead think of aptitude as the time taken to reach mastery. The key insight here is that while some students find learning easier than others, there is no limit to what students can achieve, if they are given enough time and support. This, in turn, suggests a further shift:

## 3. Variability: from inevitable to an indicator of ineffective teaching.

With the traditional view of aptitude, the fact that students' scores on tests are spread out—and often resemble a normal distribution or a "bell curve"—is treated as something natural—a direct consequence of differences



in aptitude (as traditionally defined).

If, instead, we think of aptitude as the time taken, and the support needed, to learn something, then a normal distribution is a sign of educational failure. After all, if students vary in the speed at which they can learn, then if we treat all students the same, we will get the bell curve. The job of teachers is to destroy the bell curve, and make

sure that students who need more time and support to succeed, get it.

#### 4. The role of data: from data-driven decision-making to decision-driven data collection

A common finding in studies of teachers implementing formative assessment is that teachers frequently do not use the assessment evidence they are given. There are many reasons why this might be the case. Sometimes, the information contradicts what the teachers has learned about her students. Sometimes, the information often arrives too late to be of practical use; the teaching (and the students!) have moved on so the information is out of date. The major problem, however, in my experiences, is there isn't any 'slack' in the system. The teacher is under pressure to move on to another unit or another chapter of the textbook because the curriculum is too full. Perhaps most importantly, even when the information is potentially helpful, the rhythm of the school day, week, or semester provides no opportunity to use the data to make adjustments to the teaching. Rather than pushing data at teachers, hoping that the data will be helpful, we need to reverse the process, and start from the decisions that teachers need to make, and only then figure out what data will inform those decisions.

essential material must be taught in the first 12 class periods, and then the students do a test in period 13. The teacher then uses the students' performance on the test to decide whether lessons 14 and 15 should be spent on consolidating the essential material, or going on to the desirable material. On a longer timescale, 20% of the material allocated to a five week block is designated as desirable rather than essential, so that the essential material is covered in the first four weeks, culminating in a test, which the teacher uses to plan the focus for week 5. The important feature of all these scenarios is that the teacher does not know what will be taught next until she sees the effects of the teaching so far—teaching becomes a contingent or responsive process rather than a linear one.

#### 6. Teacher expertise: from explicit to implicit knowledge

When professional development is provided for teachers, the underlying assumption is often that teachers lack certain knowledge, and if the professional development sessions provide that knowledge, then the teachers will as a result become more effective. The problem here is that while experienced teachers are indeed more knowledgeable than novices, the knowledge of expert teachers is more like the knowledge of how to ride a bicycle than the knowledge of how to solve equations. It is relatively straightforward to explain to someone how to solve equations; it is impossible to explain to someone how to ride a bicycle. It is something that each teacher has to work out for themselves.

#### 7. Professional development: from knowledge acquisition to habit change

Even when we can make the things that expert teachers do explicit, actually implementing these elements is far from straightforward. For example, the research on "wait time"—the amount of time teachers allow students to answer a question before providing a hint or moving on to another student—shows that teachers frequently allow less than one second for a student to begin a response. The research also shows that when teachers allow more wait time, their students learn more. However, reminding teachers about the importance of wait time seems to have little impact; it has about as much impact on teachers' practice as reminding smokers of the harmful effects of smoking, because it is not a knowledge problem, it's a habit change problem.

I have lost track of the number of times I have described a formative assessment technique to a group of teachers and someone says to me, "Oh, yes. I used to do that. It was good." The teacher used to employ that technique, but something new came along, and he doesn't do it any more. The problem in education is that each new wave of reform displaces the previous one, rather than building on it, so professional expertise rarely accumulates systematically.

#### 5. Teaching: from a linear to a contingent process.

One concrete way to build 'slack' into teaching is within any block of available time—be it a lesson, a week, a term or a semester—is to designate at least 20% of the material to be taught in that block of time as 'desirable' rather than essential. This could be a 50-minute lesson, where the content has to be taught in the first 40 minutes, giving the teacher time to assess the students, and then decide whether the last few minutes should be spent on review or extensions (the desirable material), or making sure that the essential material is understood. If a unit on 'pulleys and levers' is allocated 15 class periods, the



Fortunately, there is quite a lot of research on habit change, especially in the field of health education, although this is rarely used in designing professional development for teachers. By recognizing increasing teacher expertise as a process of habit change, we can produce significant, lasting improvements in teachers' practice, and this having a profound and sustainable impact on student achievement.

There are no doubt many other shifts in attention or focus that could support teachers in developing their practice of formative assessment, but in my work with teachers around the world, the seven shifts I have outlined above seem to me to be the most important. Nothing can guarantee that any attempt to improve education will be successful—schools and classrooms are much too complex for that to be the case.

If we focus on what our students are learning, rather than what we are teaching, if we accept that with time and support all our students can succeed, and that our teaching needs to be responsive to their needs, if we start with the decisions that we need to make rather than with data that might inform those decisions, and if we understand just how complex teacher expertise is, then we can make effective formative assessment a reality in every classroom. After all, if each teacher reflects on two questions—"What did I just do?" and "What did my students learn?"—there is no more powerful focus for teacher learning.