How could this promote better learning?

So far we have reviewed the evidence about what great teaching looks like, and how it can be safely identified. This evidence is important for teachers to understand, but it is in some ways just a preamble to the crucial question of how that understanding can be used to improve students' learning. Before we can do that, we must first clarify some validity issues that arise out of any attempt to 'measure' teaching quality. Then we consider relevant evidence about how feedback about teaching quality can be used most effectively, and how this relates to the broader issue of teachers' professional development.

Validity Issues

Combining evidence from different evaluation approaches

One question we need to address early on is whether we are setting out to produce a single measure of teaching effectiveness. Today, many jurisdictions are using multiple sources of teacher evaluation, but with the intention of combining them into an overall measure (Burniske & Neibaum, 2012; Isoré 2009). A single measure will be required, for example, if we want to rank teachers in order of effectiveness, or to attach explicit consequences to different score ranges. On the other hand, if we want to focus on giving teachers feedback on a range of strengths and weaknesses, such a combined score may be unnecessary and unhelpful.

It may be that part of the reason researchers have not been more successful in achieving congruence across different methods and instruments for assessing effectiveness is that there is not just one kind of effectiveness. It may be, for example, that different teachers with very different sets of skills, knowledge and understanding can achieve similar ends in terms of students' learning. A measurement approach that starts from the assumption that the answer is a weighted sum of all the component parts may miss the subtlety of their interactions. If our investigative method is to feed potential explanatory factors into regression models we will be unlikely to find these kinds of relationships.

There may, for example, be threshold effects, so that once a particular teacher skill reaches an adequate level, further increases do not make much difference; below that level, however, and learning is likely to be diminished. Or there may be interactions, so that two (or more) particular skills can compensate for each other: as long as at least one of them is strong enough, the strength of the other is unimportant.

All of this is speculation, of course: any theory of teaching effectiveness would have to be developed fully and tested. But it may be important to keep an open mind about the kinds of relationships we may find.

Focus on student learning outcomes

We have already made clear that our definition of effective teaching is that which leads to enhanced student outcomes. An important corollary is that our criterion measure, against which we should validate all other sources of evidence about

effectiveness (such as from lesson observation, student ratings, etc.) must always be anchored in direct evidence of valued learning outcomes.

We need to stress that this does not mean that we have to privilege current testing regimes and value-added models. Existing measures and models may fall well short of what we need here. However, success needs to be defined not in terms of teacher mastery of new strategies or the demonstration of preferred behaviours, but in terms of the impact that changed practice has on valued outcomes. Because teachers work in such varied contexts, there can be no guarantee that any specific approach to teaching will have the desired outcomes for students.

Purposes: Fixing versus Firing

A key part of modern thinking about validity is that we need to know the purposes for which a measure is intended to be used before we can evaluate any evidence about whether it is fit for purpose.

James Popham (1988) has characterised two incompatible uses of measures of effectiveness as 'Fixing' (formative assessment, intended to improve practice) and 'Firing' (summative assessment, with consequences attached, e.g. merit pay or termination of employment). He pointed out that either may be fine alone, but together they make a counter-productive 'dysfunctional marriage'.

As Hinchey (2010, p6) explains

> "Assessment to improve practice requires that teachers be open to admitting weaknesses, which can happen only in a relatively non-threatening environment. … Teachers whose work can be improved but who are feeling at risk may understandably be inclined to hide, rather than confront, their problems—precluding valuable formative feedback."

The requirements for a measure to be used for 'fixing' may be very different from those for 'firing'. It will not be helpful to talk about 'validity' in a general sense without being clear about this.

Approaches to providing feedback

A range of studies suggests that the quality of feedback is a key component of any teacher assessment (Stiggens & Duke (1988), McLaughlin & Pfeifer (1988), Kimball (2002)).

Hattie & Timperley (2007) state that the main purpose of feedback 'is to reduce discrepancies between current understandings and performance and a goal' (ibid., p. 86). Although their review concerns teacher feedback to students, given that learning works in similar ways for adults and young people (Bransford, Brown, & Cocking, 2000) their findings can be adapted for our focus on feedback as a follow-up activity to an observation.

Hattie & Timperley argue that effective feedback answers three questions ('Where am I going?', 'How am I going?' and 'Where to next?') and operates at four levels:

the task ('How well tasks are understood/performed'); process ('the main process needed to understand/perform the task'); self-regulation; and self level ('Personal evaluations and affect […] about the learner').

Timperley et al. (2007) review the characteristics of the teacher 'knowledge-building cycle ' - a feedback loop for teachers - that are associated with improved student outcomes. Their synthesis 'assumes that what goes on in the black box of teacher learning is fundamentally similar to student learning'. Their findings suggest that teacher learning can have a sizeable impact on student outcomes.

They report that in effective interventions feedback was related to evidence and clear goals about developing teacher pedagogical content knowledge and student achievement or conceptual understanding, whilst providing the teacher with the skills to assess student outcomes. Moreover, professional instruction was followed by a range of opportunities to practice and learn.

The observation/feedback routine should be structured explicitly as a continuous professional learning opportunity that actively challenges teacher thinking and practice and enables them to work on improving, for it to be more likely to translate into student outcomes: teacher learning drives student learning. Principals can help by 'developing a vision of how teaching might impact on student outcomes, managing the professional learning environment, promoting a culture of learning within the school, and developing the leadership of others in relation to curriculum or pedagogy.'

Evidence of impact of feedback to teachers on student learning

This is some evidence, reviewed by Coe (2002), that the use of feedback information from school performance measures can have positive effects on subsequent school performance. However, as Coe points out, we are limited by the lack of both direct evidence and strong theory:

> Given the complexity of the kinds of feedback that can be given to schools about their performance, the varying contexts of school performance, and the range of ways feedback can be provided, it is extremely difficult to make any kind of generalised predictions about its likely effects.

One specific example of a positive impact of feedback from classroom observation is from Taylor and Tyler (2012). They used Danielson's *Framework for Teaching* to evaluate and feed back to teachers in Cincinnati over a period of seven years. They found a gain in students' performance in math test scores in the years following the intervention, equivalent to an effect size of 0.11. The cost of the observation intervention was estimated at $7,500 per teacher.

Enhancing teachers' professional learning

Timperley (2008) highlights a number of broad principles from an extensive research review on successful professional learning - and much of this advice can be translated to observation and feedback routines or programmes in general. To be effective, strategies:

- Must focus on and be measured against student outcomes;
- Encourage 'self-regulation' among teachers who need to embrace the experience as independent learners and sustain the techniques;
- Require some input from school leaders;
- Involve, ideally, collaboration with peers;
- Be a genuine challenge.

	Summary of advice from Timperley (2008)
1	'Focus on valued student outcomes', whether it is achievement or a deeper student understanding
2	'Professional knowledge and skills that do have a positive impact on student outcomes are consistent with evidence-based principles of teaching effectiveness', national associations' recommendations, or with rigorously-debated national policies.
3	'To establish a firm foundation for improved student outcomes, teachers must integrate their knowledge about the curriculum, and about how to teach it effectively and how to assess whether students have learned it'. We consider the last point to be especially relevant, as it is the basis for teacher monitoring of students but also self-regulation.
4	'To make significant changes to their practice, teachers need multiple opportunities to learn new information and understand its implications for practice. Furthermore, they need to encounter these opportunities in environments that offer both **trust** and **challenge**'
5	Whether the decision of engaging with professional development is voluntary or directed has no bearing on student outcomes.
6	'[I]f teachers are to change, they need to participate in a professional learning community *that is focused on becoming responsive to students* […]. As an intervention on its own, a collegial community will often end up merely entrenching existing practice and the assumptions on which it is based'.
7	'Expertise external to the group of participating teachers is necessary to challenge existing assumptions and develop the kinds of new knowledge and skills associated with positive outcomes for students', and this expertise can come from within or outside the school. When it is provided by the principal or other school leaders, these professionals should establish 'a vision of new possibilities […] through everyday activities', lead learning and organise learning opportunities.
8	'Sustained improvement in student outcomes requires that teachers have sound theoretical knowledge, evidence-informed inquiry skills, and supportive organizational conditions'.

One example of the importance of the school context in which professional learning takes place comes from a study by Kraft and Papay (2014). They provide a challenge to the now much quoted claim that teachers typically improve over their first 3-5 years and then plateau (e.g. Rockoff, 2004). Kraft and Papay found on average the same pattern: rapid improvement over the first three years, then much slower growth. However, they also found that teachers working in schools with ‘more supportive’ professional environments (assessed by teacher questionnaires) continued to improve significantly after three years, while teachers in the least supportive schools actually declined in their effectiveness.