

MODERATOR: Dr. Dylan William is the Director of Learning and Teaching Research Center at the Educational Testing Service Center, ETS. Prior to joining ETS in September, the professor was the educational assessment and assistant principal of King's College of London, the fifth largest university in the United Kingdom.

He also was a teacher for eight years in inner city schools in London, a Nuffield Research Fellow at Chelsea College on the Graded Assessment and Mathematics Project and a lecturer in mathematics education at King's College where he ran the teacher's preparation program and developed a variety of statutory and non-statutory assessments for the national curriculum of England and Wales.

He has taught master's and doctoral-level courses on educational assessment, research methods, and the use of information technology and academic research. He has published over 150 articles, book chapters, and books in mathematics education, education law, and educational assessment. His current main interest and the reason you are here today is in exploring how assessments may be used to support learning. Please join me in welcoming Dr. Dylan William.

DR. WILLIAM: Thanks. There's only one slight problem. That CV is about four years old, but it doesn't matter. Okay. What I want to talk today about is why we need to raise achievement. I want to establish with you that this requires improving teacher quality. Economists like Rick Hanushek believe that we can't change teachers. You know the old light bulb jokes? How many psychoanalysts does it take to change a light bulb? One, but the bulb has got to want to change.

How many teachers does it take to change a light bulb? Change? So many economists believe that teachers can't change, so they actually think that we should improve the profession by raising the quality of entrants into the profession. The problem is that takes too long. So my argument is that we have to help the teachers we already have in post improve, what I call the love the one you're with strategy.

And it's not because, you know, these aren't good ideas. I mean, I think deselection of incompetent teachers is a good idea. I think raising the quality of people that come into the teaching profession is a great idea. But, by itself, those things will have a small effect, and it will take about 30 years to work through the system. So if we're serious about improving the quality of our students' achievement, we have to help the teachers we have improved.

Now the difficulty is the teachers can change what they do in a range of ways. A lot of teachers are interested in learning styles. That's very interesting work. Unfortunately, there's no evidence that it has any impact on student achievement. So you can spend a lot of time changing your teaching in ways that benefit nobody. Some changes in teachers' practice benefit students, others will not.

Those that do tend to involve changes in teachers' practice, and what we're discovering is that changing teachers' practice is hard. We need new kinds of models of teacher learning and new models of teacher professional development. And that's what I'm aiming to leave with you in two hours' time.

The first seven steps in that argument I think are matters of science. If you disagree with me, you are wrong. I can prove every single step in that argument. What I can't do is show that the ideas I'm going to share with you today are the best way to secure that goal. They may be. They may be not. So this is a matter of design.

We've played around with lots of ideas. We've worked with over 100 schools and districts in the U.S. and the U.K. finding different models that teach professional development. We've got some ideas now which we think are pretty good, but they'll change and improve, because this is not science, this is design.

Okay. Raising achievement matters but not for the reasons that most people assume. I know you're under pressure to raise test scores because of NCLB, but the real reason you need to raise achievement is because it matters for individuals, and it matters for society.

Increased lifetime summary. A male gets a 7% salary premium for having a degree. A female in the U.S. gets about a 10% salary premium. Improved health. People are healthier, and they have less disabilities toward the end of their lives, and they live longer. Now doing that research takes quite awhile, because you have to wait for people to die even before you can work out the impact that education had on their longevity.

But the latest estimates that we have are that one extra year of schooling adds 1.7 years to life. These are high stakes for the individuals, but they're also high stakes for society. Hank Levin at Columbia calculated for every student who would otherwise have dropped out of high school at 16 who we could get to stay on to 18, the net benefit to society would be right around \$210,000.

About \$140,000 of those dollars would be increased taxes those students would pay, because they were earning more money, \$35,000 reduced criminal justice costs, \$35,000 reduced medical costs. So these are high stakes. But recently, in January of this year, the OECD, the Organization for Economic Co-operation and Development, showed that the increased economic, sorry, increased educational achievement has really profound benefits.

What they showed that if we could actually increase students' scores on PISA, this big international test, the Program on International Student Assessment, we could increase every American kid's score up by 25 points. The net present value to the U.S. economy would be \$40 trillion, which is roughly the value of just about everything in the U.S. right now. That's a big amount of money.

And even more impressive, if we could get all students getting up to 400, which is two standard deviations below the mean, the net present value of that would be \$70 trillion. It would be like giving ourselves a great big \$70 trillion gift. So these are high stakes. Now these skills that we need to develop, we need to be very careful about what kinds of skills we're talking about, because the world of work is changing very rapidly in very surprising ways.

Now some jobs have been outsourced. Some jobs are being done by computers. But what's interesting is if you really want a safe job in the U.S., it's not clear what you should do anymore. Clearly, working with your hands on a planing machine in an engineering workshop is not a safe job, but people used to think it was high skill versus low skill.

But if you're a radiographer, you know, if you have a road traffic accident in New York, in Manhattan, say, at 3:00 a.m., your x-ray slides might well be looked at by a radiographer in Bangalore, which is actually pretty good news, I think, because he'll be awake. And what's interesting is that the jobs that could be outsourced are not just the low skills jobs. There are lots of high skills jobs being outsourced as well, programming, software.

There are certain jobs that can't be outsourced, hairdressing. No one's going to send their hair to Bangalore to be restyled no matter how inexpensive it is. Actually, I just did, but that's a different story. But what we're beginning to discover is that the skills that are really important are those that involve communicating with people, being with them, being physically present, and, in particular, in solving people's problems.

And this quote, I think, from Seymour Papert, is interesting. I mean, the important quote, a bit of it is right here at the end. We need to produce people who know how to act when they're faced with situations for which they were not specifically prepared. Seymour Papert actually studied with Piaget, and Piaget's definition of intelligence was intelligence is knowing what to do when you don't know what to do.

And so the interesting thing for us is how we help people develop that. Now we, I think, are often quite slow learners. This quote is very interesting. The test of successful education is not the amount of knowledge that a pupil takes away from school but his appetite to know and his capacity to learn.

If the school sends out children with a desire for knowledge and some idea how to acquire it, it would have done its work. Too many leave school with the appetite killed and the mind loaded with lumps of undigested information. The good schoolmaster is known by the number of valuable subjects which he declines to teach. Now the problem in states like Pennsylvania is that your standards writers never saw a standard they didn't like.

I'm on the validation committee for the core standards for this new national government's association, and I refused to sign up to the standards, because I think for English language arts, they have 24 standards. Now what's really interesting about every other high-performing country is they have like five or six, at a pinch, seven. It's about being really strategic about some really high-value goals.

And so, you know, we don't seem to have learned a lot. This guy, Sir Richard Livingstone, was actually president of an Oxford college in England in 1941. He could see the world coming, the world of getting people to think and knowing how to acquire information, and that's what you have to do at school.

So where's the solution to all of this? Well, of course, in the U.S., because of the influence of the Gates Foundation, it's smaller high schools. When I was living in New Jersey, our state capital, Trenton, got some money from the Gates Foundation and took Trenton Central High School, 3,000 kids, divided it up into six 500-student high schools in the same building. And they wondered why nothing had changed.

Well, something had changed. Their administrative costs had gone up, because they'd appointed six new principals to look after each of the high schools and had to increase the salary of the previous principal, because now she was looking after six baby principals. Many districts have tried K-to-eight reform, and the evidence is that the effects are very, very small, if not zero.

But much more damagingly, when you're doing a K-to-eight-type reorganization, you can't do anything else. Your entire attention is absorbed by that, so it actually deflects attention. In England, the idea is all through schools, the idea we have schools where, which cater to kids from the age of 3 to 19. I don't know who thought that putting all those hormonal adolescents in with kindergarteners was meant to be a good idea, but the evidence is that it makes no difference.

Now when politicians are in a hurry, one of the things they look for is a curriculum reform. We're going to have a curriculum for the 21st century. The Scots have just introduced a curriculum for excellence. That's a huge relief to me. I was so worried. Curriculum for excellence, curriculum for mediocrity, which shall we choose, you know? It's so hard, isn't it?

In the U.S., because most states have fixed their standards, the big move is on textbook replacement, and, you know, the average U.S. school district superintendent is in office less than three years. So what are you going to do to improve standards? We're going to get some new textbooks in. It looks like you're doing something. The evidence from the what-works clearinghouse is that textbook change doesn't make much difference at all.

Charter schools and vouchers, this research is actually difficult to pull out, because charter schools actually get, generally get less money per student than district schools, but the net effect appears to be about a half of charter schools are no better or no worse, about a third of them are slightly worse, and a sixth of them are slightly better than the schools that they're meant to replace. So the net effect of having charter schools is to lower student achievement slightly.

Vouchers appear also to have small effects. Computers. Computers have been about to revolutionize our classrooms for the last 30 years. I'm still waiting. Now computers have done some very, very impressive things.

So if you want to teach ninth grade algebra, there's a package available from Carnegie Mellon called the Cognitive Algebra Tutor which is better than all but about the best 5% of math teachers in teaching this stuff, but it's taken them 25 years to get to that point for two or three lessons of algebra just in ninth grade. So computers can help, but most of the time, they're an irrelevance.

Now the interactive white boards have also been very, very significant, you know, because, basically, computer manufacturers need to keep on selling you stuff, and once your closets were all full of computers, they needed to find something even bigger and more expensive to sell you, hence, the interactive white board.

Now in the hands of experts, these are stunning pieces of technology, but what we know from a wonderful natural experiment we did in London where the government decided these things are so cool, they were going to give every school one of these boards whether they wanted it or not, the net impact on student achievement was zero. There were as many schools where giving them these expensive bits of kit made things worse as they were where they made things better.

And here's another hot-off-the-press finding, classroom assistance, para-educators in classrooms. What we, we actually did some research for the government, because they spent something like \$3 billion on putting about 100,000 classroom assistants into English schools, so they wanted to find out the impact on student

achievement, and we came up with an answer, and the government told us it was wrong. Go back and do it again but properly, and get a different answer this time.

So we did the experiment again. We did the experiment eight times, and every time we did it, we found the same thing, putting classroom assistants, poorly paid adjuncts into classrooms with particular students lowers those students' achievement. So basically, if you have learning difficulties, and you're given an untrained volunteer to work with you, it'll lower your achievement. One hour of working with, one-to-one with an untrained person is less good than two minutes with a highly qualified professional.

Why? Because what we discovered was that these para-educators, or these volunteers, were taking their job as if it was to help the student do the work. So if there was an assignment given to those kids, they would help them do it by telling them what to write. And as long as the kid did the writing, they felt the kid was doing the work, but, of course, they were denied the cognitive activity that is what we're really interested in.

So what's interesting about all this stuff is that it seems to be really hard for educators to, for governments to improve. And one of the reasons that that is is because we've been looking in the wrong place for solutions. Now the big issue here is why do students get different results in school, why art comes different for different students, you know, because the model that's driving NCLB is basically a market model, isn't it?

Some schools get good results, some schools get bad results. I know, let's have more schools like the good ones. So you look at what's the effect, you know, what are the characteristics of these great schools? Well, actually, if you want to raise achievement in your school, you have to do three things. One, get rid of the boys. In every development in that country, now girls are beginning to overtake boys, and for the more girls you've got in the school, the better you're going to look.

Secondly, apparently become a parochial school. That's always a good move. But thirdly, and most importantly, move your school to Lower Merion, \$26,000 per student per year for that very needy community around Bryn Mawr College. We need to look much more carefully, you know, and that's the model behind NCLB, make all our schools like the schools of Lower Merion, and it doesn't actually get to the grips of the problem, because we're looking in the wrong place.

Of course, there are differences in students' abilities. There are differences in subject practice, there are differences between schools, but what the research shows is the biggest effect apart from the students' own differences, and, you know, those are substantial, but after the students' differences, it's the teacher differences. As long as you go to school, it doesn't matter very much which school you go to, but it matters very much which classrooms you're in.

And this is being found across the whole world. This charge from the OECD, these are the results for 15-year-olds in PISA, above the line is the U.S., I'm sorry, above the line is the within-school variation, below the line is the differences between schools. Now, for example, in Turkey, there's a lot different, you know, there's a lot of differences in schools and relatively little difference within schools. But a lot of these, the black here and the mid-gray here are social class effects.

They are the facts that it's not the same kids going to the different schools. And when you control those out, the red line here is the actual size of the school effect. So in the U.S., basically, the red size there is how much of a difference, the differences in

schools. I've published an article recently where I quantify this. The differences in kids' results, 8% is due to the school, 92% is due to other things. So the school effect is very small. It doesn't matter very much which school you go to as long as you go to school.

Eight percent is attributable to the quality of education brought by the school, so 92% of variability has nothing to do with the school. Now what this means in practice is that if you went to a so-called good school, two more students out of a class would reach proficiency, and if you went to a so-called bad school, two fewer students would reach proficiency. Now these are differences worth having, but they're actually quite small. I think they're much smaller than people think.

And when you think in the U.S. how much difference there is spent in terms of cash on children's education, the size of these differences is really quite remarkable. Why is it? Because, basically, teacher quality is pretty randomly distributed within the system. There are good teachers in every school. There are outstanding teachers in every school.

These differences between teachers are really very large. Take a group of 50 teachers, say, teaching fourth grade. If you are lucky enough to be taught by the best teacher in that group of 50, you will learn in six months what kids taught by the average teacher will take a whole year to learn, and if you're unfortunate enough to be taught by the least effective teacher in that group of 50, that same learning will take you two years.

The most effective teachers create learning in their students at four times the speed of the slowest. Now parents have kind of known this all the time. You know, they say they know that teachers make a difference, but only recently have we been able to quantify it to show that the most effective teachers are four times as effective.

And what's really interesting, some recent research for the University of Virginia by Robert Pianta and Bridget Hamre shows that in the classrooms of the most effective teachers, students from disadvantaged backgrounds learn at the same rate as those from advantaged backgrounds. I guess you're, in Pennsylvania, you're interested in achievement gaps? Yeah? Narrowing the gap?

Yeah, well, we had an interesting cycle about through in England. We were being, schools were being beaten up for not catering for gifted and talented students, so they had to have gifted and talented programs, and then they suddenly saw that that was making the achievement gap greater. Duh. So then we had a narrowing-the-gap team establishing our, getting our government ministry.

And I've got the solution for that, by the way. If you want to narrow the achievement gaps, it's very simple. Put lead in the drinking water of all the smart kids. Achievement gap gone. It's the wrong way of thinking about the problem. What we need is proficiency for all students, excellence for many, and all student groups represented in the excellent.

It's a focus on helping every student get up to a basic floor, and helping every student get as far as they can rather than actually if we were worrying about the gaps. Now as I said earlier, this is really a classic labor force issue. We can replace existing teachers with better ones, but that takes too long.

I've calculated that increasing the quality of entrants to exclude the lowest performing 30% of teachers would result in one extra student passing a class test every year, every three years, so we have to help the teachers we have improve. Now the

problem is that left to their own devices, teachers improve very slowly. These are data from Australia.

I mean, many of the American studies are quite floored, because they don't control for the fact that there's a lot of churn in the system, so a lot of people believe that teachers only get better for about six years and then stop improving. And often it's because they're now working in district office rather than in the classroom. So those kind of studies are not very well controlled.

This one from Andrew Lee in Australia is one of the best controlled studies. And what he's shown is that if you're taught by a 20-year veteran, you do learn more than if you're taught by a novice teacher. Let's take literacy or reading skills here. So if you are taught by a 20-year veteran, the blue line here, you will learn more than if you're taught by a novice teacher but the scary thing is without vertical accesses, extra months per year of learning.

So if you're taught by a 20-year veteran, you will learn half a month more per year than if you're taught by a novice. In other words, you will learn in 50 weeks taught by a 20-year veteran what a novice teacher will teach you in 52 weeks. So it is a difference, but it's not much of one. The average improvement of a student evaluated by a teacher over 20 years is one-tenth of the difference between a good teacher and a weak teacher on the first day of their teaching career.

So that's the answer to the question are teachers born or made. Well, it's both, but born is a bigger effect than made if we do the kinds of stuff we're doing right now. But I would say we've been doing the wrong kind of professional development. Now in England, these things are called baker days. You have, is it Act 48? And you have, what is it, 100 hours every four years? A hundred eight-five. Okay.

So what, I did a workshop, I think it was in this convention center about four years ago, and it was for science teachers, and it was really kind of spooky, because I started at 9:30, and at 9:35, the doors were locked. No one else could join. And then if they stayed until 12:00, everybody got a little bar code. They could prove that they had endured 150 minutes of professional development. It wasn't this, you know, so, but it's what other professions do.

I gave a, I was giving an address to the Association of Medical Educators in Europe last September. This was a conference which was, at which 2,300 people registered. No session was attended by more than 1,000 people. It was held in Malaga in Spain in September. But everyone who registered got three days of professional development credit.

So these kinds of days of endurance, I call it 100 days, because in England, we introduced this in 1988, and so every teacher's added 100 of these days of so-called training, and not one of them believes it's made them a better teacher. Professional updating. It's completely the wrong model. Now if you're a surgeon, there are new techniques, but in teaching, I haven't seen anything new in the last 2,000 years.

The reason you need professional development is not because there's new stuff, it's because this job is so difficult, so demanding, that one lifetime is not enough to master it. You show me a teacher who thinks they're doing a good job, and I will show you a teacher with low expectations of her students, because our daily experience is a failure. We teach these wonderful lessons, we take in the kids' notebooks, and we read what they wrote, and we wonder what planet they were on, don't we?

So we need an updating, we need to get better, because no teacher is good enough. And so, we're also trying recertification in England, and that's been abandoned. It's not working. So these easy kind of formulaic approaches don't work, because they don't address the issue. There's the famous joke about the drunk who's looking for his keys underneath the streetlight, and somebody says to him is this where you dropped your keys? He says no, but this is where the light is.

That's what we've been doing. We've been doing the PD, the professional development, that's easy to do, not the stuff that's going to make a difference to our kids, because it's incredibly hard to keep the focus on what you need to do. My wife is a high school principal, and she heard about, have you heard about Donald Schon's work on the reflective practitioner? She thinks it's a load of rubbish. She says, basically, the best principals are deflective practitioners.

She sees her job as stopping most of the stuff that rains in from the district and saying we're not in the market to do that, and the district says you have to, and she says no, we're not going to do that, because it hasn't got anything to do with improving learning. And by focusing rigorously on the things that are going to help a teacher make a better job in the classroom, they've actually improved their results radically but only by actually ignoring everything that doesn't really make a difference to classroom learning.

Here's the cost-effect comparisons, class size reduction by 30%, maybe an extra four months learning per year, cost per classroom per year, \$30,000. Oh, and across the U.S., you're going to have to build an extra million classrooms. Nobody ever thinks of that. About four months extra learning per year, by the way? Great if you could get it. You won't, because what people forget about class size reduction is you need more teachers. Duh.

But they won't be as good as the ones you've already got, because, presumably, you hired the best ones when you were hiring. And what they found in California when they reduced class size there were some districts that lost all their good teachers, because there were loads of new jobs being created in much more attractive districts. So in some of the hard-to-staff districts, hard-to-staff schools, they were giving emergency permits to people in off the street to go and teach.

So in some schools, they lowered student class size and lowered student achievement at the same time, because these people are being trained by completely, being taught by completely untrained novices. Everybody says content knowledge is very strong, is very important. I'd like to believe that, but I don't actually see the evidence.

So we do know that the best work we've seen is Deborah Ball's work for the University of Michigan, and she has shown that if you're taught, say, fourth grade math by a teacher who is high in what Deborah Ball calls mathematical knowledge for teaching, so it's not just knowing a lot of advanced calculus but knowing how to explain why you invert and multiply fractions when you want to divide, say to kids, you know, for example, two shared by a half is like how many halves are there in two, those kinds of metaphors or analogies.

When you're teaching directed number, you can talk about heights above and below sea level, you can't talk about negative and positive bank balances, all those ideas about how to teach this stuff. But one standard deviation of teacher subject knowledge, in other words, being taught by a good person as opposed to an average

person, improves your speed of learning by about 4%. So it's again, it's about this idea of two or three weeks extra learning per year.

What we do know from the studies we've done so far is that when teachers pay attention to formative assessment or assessment for learning, we get an extra eight months learning per year at a cost of around about \$3,000 per classroom per year. So that's why I'm totally committed to formative assessment, which is actually, and I'll talk about this this afternoon, is really a special kind of response to intervention.

Response to intervention and formative assessment are really different ways of looking at exactly the same process. It's this idea that, hey, kids do not learn what we teach. See, if kids learned what we taught, then we'd never need to assess. We could just write down in our planners what we taught them, and we'd know that they knew it. But they don't learn what we teach, and so, therefore, we need to find out what they've learned before we try to teach them something else.

You can call that response to intervention. You can call that formative assessment. It's the same thing. It's based on this idea that teaching is unpredictable. But what we do have in formative assessment is a very strong research base. There are something like 17 studies there.

The Alala Lopez(?) reviews all the studies that have published only in French. The Curler(?) study there reviews the studies that have been done and published only in German. The El Sud More(?) reviews Dutch language literature.

So all over the world, we are seeing this consistent body of research coming out that doesn't matter whether you're teaching 5-year-olds or 25-year-olds, it doesn't matter whether you're teaching English language arts or music, it doesn't matter whether you're teaching in the U.S. or Portugal or the U.K. Attention to these processes, adjusting your teaching as a result of what you've learned about students learning pays big dividends.

Now here's my definition. And this is why it ties in with response to intervention. An assessment functions formatively. That's important. Let me give you an example. A teacher was preparing kids for a state test. So she gives them a practice test, a review test. The kids do the test on the test conditions. In most districts, what the teacher would do is then grade these tests, write comments on there to help the students, and give them back to them.

She actually hasn't got time to do that, so what she does is the next period, these kids get back their individual tests, and they are put in groups of four, and each group of four kids gets one blank test sheet, and they're to come up with their best composite exam test paper by putting together and fielding all their expertise. Who's got the best answer to question one, who's got the best answer to question two? And what's interesting about that is that's a formative process.

That's causing learning, but it's using a test that was designed entirely as a summative test. So it's not, there's no such thing as a formative assessment. If you hear people talk about common formative assessments, tune your radar. If you hear people talk about formative assessments, they're not talking about what I'm talking about, because the one I'm talking about is a process. It's a process of constantly adjusting your teaching to better meet your students' learning needs.

So an assessment functions formatively to the extent that evidence about student achievement elicited by the assessment is interpreted and used to make decisions, and that's the crucial thing here is decisions.

When I started to work at ETS, I started researching formative assessment across the U.S., and there were lots of companies like the Grow Network, there was a company called Pulliam in California, and what they do is they will come in and test your kids, and they'll take away the test papers, and they'll score them for you, and you get full-color printout of where each kid is, only the full-color printout arrives about three weeks after the kids last did any work on this topic.

And what these people are saying to you is here's some really interesting colorful information about your students, make some use of it. I call that data push. It's pushing data at teachers saying do something smart with this. And it doesn't work, because by and large, teachers don't have any use for that data, because they've already moved on. The best formative assessment is decision pull formative assessment.

You decide on the decision you're going to make, and you pull towards you the information you need to make that decision. And it could be at two extremes. It could be at does this kid need to repeat the grade? So that's a formative decision. Let's collect the right data for that. The other decision is do I need to go over this topic one more time, or can I move on to the next chapter? Collect the right information for that decision, and you'll be okay.

The next steps in instruction that will likely do better or better founded than the decisions that we've been taking in the absence of our evidence. That qualifies importance. Sometimes a teacher is teaching something, and she'll just check if the whole class is still with her.

So, you know, for example, a teacher teaching AP calculus just asks the students quickly to sketch a graph of Y equals one over one plus X squared on these little dry erase boards, these show-me boards. So every student holds up their answer. The teacher eyeballs the whole class, sees they've got it and moves on, because she knows they get it. But the important point that that formative assessment didn't change what she was going to do.

She wanted to check that it was okay to move on, and that's just what she did. She didn't change what she did, but now she knows that the decision she made was better founded than the decision she would have made just by checking whether the kids looked bored or glazed or confused. See, teachers do this all the time. All teachers rely on cues by looking at the class to see whether they get it or not.

But the quality of data that most teachers get is very poor. And kids are very good at deceiving teachers, because kids know that if they frown or look puzzled, the teacher may well pick on them and say what don't you understand? So kids learn to nod and smile, and the teacher just carries on. So it's about improving the quality of the data that you're getting to make those instructional decisions. That's simply all it is.

Now when I first arrived in the U.S. back in 2003, I went to a conference in D.C., and I was talking to a woman, and she said, oh, yeah, we do formative assessments in our district, yeah, once every ten weeks. And I, so I talked about it, and it was one of these testing systems where they come in, and they test your kids every ten weeks, and they use it to predict which kids are going to fail the state tests following March.

It doesn't tell you what to do about it, but at least you're impending failure comes with plenty of advanced warning, and you begin to start polishing the resume. Having said that, if you've got very poor alignment between what you're teaching and what's being tested, then that can be, that can have a big effect.

I was working in Trenton, and I don't, in New Jersey, they have this high school proficiency assessment, and you can't get the high school diploma without passing HSPA in 11th grade. And in Trenton Central High School, the pass rate was about 27%. So we worked with the teachers, and we did some of these common formative assessments, and we discovered that these teachers were not teaching the things that were in the HSPA. And so we said wouldn't it be a good idea if you did?

And the following year, the pass rate went up from 27% to 40%, really quite significant, and that was just a quick win by actually aligning what the teachers are doing with what the test was going to test. So it's not a bad thing to do, but the problem is it's a one-shot deal. You'll get that quick win once, but that's it. There's no further gains to be made that way.

Medium cycle, this is the sort of thing that Rick Stiggins mostly talks about. It's this idea of student-involved classroom-based assessment, and there's no doubt that it's effective, but it's not as effective as this minute-by-minute and day-by-day stuff. And the problem in the U.S. is that because the long-cycle stuff and the medium-cycle stuff is easier to do, that's where the attention has become dominated, so, hence, this idea of interim assessments, benchmark assessments, common formative assessments.

Those are all very easy things to do. Get the kids to do a task, bring the teachers together after classroom time to talk about their kids' results. The problem is the evidence that it makes a big difference is pretty thin. What we do know is this stuff works, because it changes classroom practice and increases student engagement. But it's also very hard to do. So what is formative assessment? Well, I think there are three key processes here, where learners are in their learning, where they're going, how to get there.

And when we think of roles of teachers, peers, and students, we get this complicated matrix of nine cells, which I think can be grouped into five what we call key strategies of formative assessment, clarifying, understanding, and sharing learning intentions, what is your curriculum philosophy, engineering effective discussions that elicit evidence of learning, providing feedback that moves learners forward.

Now most American teachers give kids grades, and so those comments, we'll talk a bit about, more about that in a minute. But in 1998, when Paul Black and I said we should stop giving grades and start giving comments, a lot of teachers followed our advice, and they gave comments. Unfortunately, they weren't great comments.

There were things about what was wrong with the last piece of work rather than what they were going to do next. And because the piece of work they've just turned in is never going to get done again, telling them what was wrong with that is not particularly helpful. As Douglas Reeves points out, it's the difference between having a medical and a post-mortem. And most of our kids get post-mortems.

The last two, I think, are the most important, activating students as learning resources of one another and activating students as learning owners of their own learning. So this brings in collaborative learning, reciprocal teaching, peer assessment, and mental cognition, motivation, interest, attribution, and, of course, self-assessment.

The big idea is just using evidence about learning to adapt instruction to meet students' needs. It's as simple as that.

Now I flew in from London to Dulles last night, and I am very glad that our pilot didn't navigate the way that most teachers assess, because if he had, he'd have set a course from London to Washington, D.C., on a bearing of about 250 degrees with a planned flight time of eight hours, and after seven and a half hours, he'd begin looking for a landing strip and set down at the nearest landing strip and say is this Washington Dulles? And the ground crew say, nope, it's Atlanta. There was a strong wind.

And the captain says, everybody off, I've got another job to go on to, because that's what we do, isn't it? We teach kids stuff, and right at the very end, we give them a quiz, and some kids get A's, and some kids get D's, and we say, but don't worry, because we're on to a different topic next week. So the simple idea is just what our pilot did.

He started out with a very clear idea of where he was starting from, London Heathrow, he was very clear about where he was going, Washington Dulles, he carefully planned the journey, he began the learning journey, but he made regular checks on progress along the way, and he made adjustments to the course as conditions dictate.

And that's all Rtl is about. That's all formative assessment is about. It's about undertaking this learning journey with your students but understanding that because what they learn is not predictable, you have to find out where they are before you plan your next steps.

Okay. So let's look at some practical examples of how you can begin to find out where kids are in their learning. Here are two questions from the Third International Math and Science Study, TIMS, and these are the data on Israeli middle school students. Question one, which fraction is the smallest? Success rate, 88%. Question two, which fraction is the largest? Success rate, 46%.

Okay. In your tables, I'd like you to spend a minute or so just discussing why that first item was to much more difficult, so much easier than the second item. Why could 88% of the kids get the first item right, but only 46% get the second item right? Now rather than shouting over you, when I want to bring the whole group back together again, I'll raise my hand like this.

And when you see me raise my hand, I'd like you to do two things, one easy, one hard. The easy one is to raise your own hand, and the hard one is to stop talking. Okay? So a minute on this. Why is that second question so much more difficult than the first? Okay. Thank you. Could you give me a number between one and six please?

WOMAN: Two.

DR. WILLIAM: Two, okay. One, two, one, two, one, two . . . it's you, I'm afraid.

WOMAN: What? I didn't hear what you said.

DR. WILLIAM: What did you discuss on your table?

WOMAN: We talked about the numerators were smaller, and item A, and that it was just easier to visualize, those fractions.

DR. WILLIAM: Okay. Thanks. Number between one and six, please . . . a number between one and six.

WOMAN: Five.

DR. WILLIAM: Five. Okay. One, two, three, four, one, two, three, four, five, one, two, three, four, five. Okay. Anything to add?

WOMAN: I couldn't hear what the original person said, but . . .

DR. WILLIAM: Okay. She just said because the numerators were easier, I'm sorry, the denominators were easier on the first one.

WOMAN: Correct. We said that also, and we also said that the second one requires multiple steps, which some people wouldn't be able to process through.

DR. WILLIAM: Okay. Fine. Any other ideas? Any other contributions about why that second item is so much more difficult?

WOMAN: Sometimes, I don't know if the questions were side by side, but sometimes if you do one thing in one question, they think it's the same question the second time. They don't look at, read carefully.

DR. WILLIAM: Right. So they might have tried to do the smallest twice. Although, what is the smallest one in the first, in the second question?

WOMAN: . . .

DR. WILLIAM: No, it's five-eighths, isn't it? And, in fact, the interesting thing is that the kids, they didn't choose that by and large. You had a contribution here?

WOMAN: I just thought . . .

DR. WILLIAM: Hang on.

WOMAN: I just thought that the kids know what a half is, they know what a third is. If they know what a third is, then they know what two-thirds are.

DR. WILLIAM: So the explanations you've come up with are pretty typical. There's the gestalt idea that they just see in the first one, they know that two-thirds is bigger than one-third, so they can reject that, and the others are just familiar fractions. There's the idea that the lowest common multiple is given to you in the first question but not in the

second, so it makes it a two-step problem. It makes it more complicated, because you might misread it and just do the same thing twice.

And these are pretty much the kinds of things that the Israeli researchers came up with, but they weren't particularly convinced by any of these, so they decided to ask the students, which is always a good idea, and what they discovered was that a lot of kids were using a rather naïve strategy of biggest bottom makes the smallest fraction, smallest bottom makes the biggest fraction.

So when they were using the first question, which fraction's the smallest, look for the largest denominator, find six, choose A, correct. Which fraction's the largest, look for the smallest denominator, find four, choose B, incorrect. And what's interesting here is that 39% chose B. Three-fourths of the kids who got it wrong chose the same incorrect answer. And that shows you there's something going on here.

And in particular, if you look, if you add 46% to 39%, you get something very close to 88%, which is quite strong evidence that a lot of the kids who got the first question right got it right for the wrong reason. Why does this matter? Because when teachers ask kids questions, and they get the answers they were hoping for, what do we conclude? That they get it.

And if the questions we're using are more like the first than the second, we will conclude that the kids' learning is on track when, in fact, they're heading off in a completely different direction. So the starting point for effective Rtl or formative assessment is asking the right question in the first place, and that's much more difficult. People think it's about the feedback, but it's not. It's about looking for the right information in the first place.

Let's say you've got a young fast-pitch softballer whose ERA is 10. Now for those who know nothing about fast-pitch softball, that is not good. So you say to her you need to get your ERA down. And she says, gee, thanks. So we try to be a bit more helpful.

So we look at her pitching, and we see that where she's going wrong is that her rising fastball isn't rising, so the idea of a rising fastball is a pitch thrown with sufficient backspin so that as it reaches the plate, it rises quite sharply, and it's quite unhittable. Of course, if it doesn't rise, it's a fastball over the middle of the plate, and that's why this poor athlete is getting the ball hit out of the park. So you say to her it's your rising fastball, it's not rising. And she said, yeah, gee, thanks, I knew that.

But if you, as a coach, can see that what she's doing is not dropping her pitching shoulder sufficiently to deliver the pitch from below the knee, then that gives the athlete something to work with. But most, a lot of coaches don't know enough about the mechanics to know that that's the thing to look for. So good feedback starts from looking for the right information in the first place. And that's why questioning is the starting point.

Okay. Now can you draw an upside down triangle? Can you vote for me, please? Thumbs up if you think you can draw an upside down triangle, thumbs down if you say it's impossible. Can I just point out that I can tell whether you voted or not? Okay. Thanks. Now the purists will say that it can't be done. That's just a triangle. It's not the way that we normally see it.

But the point is when a kid says that's an upside down triangle, it might be that they don't know the rules of what a triangle is, or they might be trying to, just being extra

helpful by saying it's a triangle, and it's not the way it normally is. So, for example, if you see that emergency exit sign over the doorway there, that red exit sign, if I ask you is that square, is it square on the wall? Different question, isn't it? One's a math question, one's a real-world question.

So what we often find is that these kids' misconceptions, these ideas that kids have, are not really misconceptions at all. They're just perfectly good conceptions in the wrong place, rules that work most of the time but have been taken beyond where they work.

I spented all my money. That's so smart, because I don't know how many verbs there are in the English language, but there's probably around about 5, about 100,000, and only about 50 of them are irregular, so just playing the odds, you know, I spented my money is a pretty good bet, isn't it? So, typically, what these questions do is they, when you get that right, it's a reveal things that kids are doing that you wouldn't have picked up had you not had a smart question.

You've got, I give this to kids in middle school. If $3A$ equals 24, A plus B equals 16, and a lot of kids will say to you I can't do it. You say why can't you do it? Well, I keep on getting B is 8, but it can't be, because A is. Somewhere, this kid acquired the belief that A and B had to be different. Where'd they get that from? From me, because every previous activity I'd given them always had A and B as different numbers. And this kid thought that was important.

Now what's significant here is had that number 16 been any other number at all, the kid would have gotten it right provided his arithmetic was okay. And I would have assumed his learning was on track. So good questions that find out where kids are are really hard to come up with.

The science example. A science teacher asked the kids to draw a sketch of the molecular structure of water, not a particularly valuable thing to do, because water doesn't look anything like this, but the kids did as they were told, pairs of hydrogen atoms connected to an oxygen atom. The teacher salvaged something by asking the kids, okay, then, so what's in between the molecules? And every single kid in the class said water.

The kids thought the molecules were in the water not that the molecules were the water. But they've all produced the right sketch. Now some teachers say to me I don't ask questions like this, because we've got a pacing guide in our district. So I think we have to, you know, we shouldn't be like ostriches and bury our heads in the sand. We should know. Whether we, you'll take decisions where you go on knowing that some kids don't get it and saying to them, you know, I'll see you later.

That's your professional decision, but you should not fail to collect the evidence in the first place. So I would crystallize all the research as coming out of these big ideas. Basically, there are two good reasons to ask a question, to cause thinking or to give you data that informs your teaching, your instruction. Most teacher questions don't satisfy either of those.

Okay, class, now yesterday, we talked about triangles, we had two sides the same length. Any, had a big name for it, anybody remember what it is? And now the teacher is in a trap. You've got to get the kids to say the word isosceles. You can provide hints, but if you have to provide the missing word yourself, it's a fundamental

mission of failure, isn't it? And so the kids will play a game of guess what's in teacher's head. These are the good questions, cause thinking.

Here's a math question. Can you have a triangle with two right angles? Now that's, for most kids, that's an interesting question, because they don't know. So I was working with a group of kids, and one kid was trying to build a really long, thin triangle, because his teacher told him that parallel lines meet at infinity. Smart kid, stupid teacher, but smart kid.

Another kid knew that the angles had to add up to 180, knew they used up two 90's, so the group, then this kid was asking the group can you have an angle of zero degrees? And a third kid, a girl in the group, she said I can actually do it. I can have a triangle with two right angles provided I'm allowed one vertex and, you can tell what sort of kid it was, can't you, vertex, provided I'm allowed one vertex at the North Pole and the other two in the equator.

And the two boys in the group said, nah, that's not a proper triangle. But the interesting thing was that although this is a really boring, closed question which you thought had a yes or no answer, it caused thinking, so it was valuable. Provide data that informs teaching.

Ask kids which way does light go? Does it go from my eye to what I'm seeing or from what I'm seeing to my eye? It's a pretty boring question, isn't it? There's two answers, and one of them is incorrect. But it's a good question to ask, because a lot of kids think that light goes this way. So those are the two good reasons to ask questions.

So if we're going to do that, first of all, generating questions with colleagues, closed versus open, which is what most teachers worry about, is actually not so important. What's important is low order versus high order. I asked the math teachers what's three threes? And they gave me answers like nine. And I said can you come up with any other different answers? And pushed to it, they'll say 27. It's three times three times three. They may even go really, you know, really, really experimental and go 333.

What's interesting is how much worse adults are at that kind of question than fourth graders. Let me give you some fourth grade answers to the question what's three threes. Eighty-three. Three. They're on top of each other. A flower. A small flock of birds. What's important about your questioning is that it's higher, it's causing thinking, not testing something that kids have recalled previously, it's actually generating new knowledge and then giving kids appropriate wait time.

The other thing that we've seen a lot of work done on is getting away from IRE. IRE is what the researchers call what they see in most classrooms, initiation, response, evaluation. So a teacher asks a question, kid responds, and the teacher evaluates that kid's response. So everything comes back to the teacher. So it's like spokes on a wheel. And I call that serial table tennis. Duh-duh, duh-duh, duh-duh, duh-duh, duh-duh, duh-duh.

And what we see teachers experimenting is more what we call basketball, so I ask you a question, I'll ask you if you think her answer's correct, and I'll pick on you for an explanation, so that the classroom questioning goes around the classroom rather than always coming back through the teacher. No hands up except to ask a question. That's probably the most important rule I'm going to say to you today, and it really is incredibly important.

I talked earlier about the achievement gap. When a teacher asks a question in most classrooms, there are four, five, or six kids who are dislocating their shoulders in their eagerness to show you they have an answer. And there are other kids from a stable of the radar, aren't there? Now the point is that work by people like Neil Mercer of Cambridge University has shown that those kids who were engaging in everything the teacher's doing are actually getting smarter. We can show their IQs are going up.

And there are other kids in the same classroom who are foregoing that opportunity to get smarter. So if you're serious about achievement gaps in your classroom, if you're serious about creating a maximal opportunity for all students, then you have to get away from allowing kids to volunteer to participate, which is why we say no hands up except to ask a question.

So any kid could raise their hand to ask a question, but if the teacher's asked a question, the initial response must always be to pick a kid at random. And what we find also is that teachers can't do random.

You know, teachers say to us I'll pick it at random, but when it gets towards the end of the period and you need a kid to give you a good answer so you can get on and finish the topic off, it's, you're just drawn to the kids who you know are safe pairs of hands, you know, just, so you have to use some sort of randomization device.

And what we found works best is kids' names written on Popsicle sticks. I think, and it works wonderfully, because it's, I mean, you've actually got an app for the iPhone now which will choose kids names at random. You can download things off the Internet. But I like the Popsicle sticks, because it provides ourselves opportunities for teacher play. So, for example, if a kid's really irritating you, you can write their name on ten additional Popsicle sticks and add it to the mix.

You also put a kid's stick back in a special place for easy subsequent retrieval. I saw one kid's name come up 4 times in a 42-minute lesson, and he just thought he was having a really unlucky day. Now the other thing about these things is how you handle it. So in some classes, when the teacher picks a kid's name at random, asks the kid, and the kid says, don't know. And the teacher says, okay, and asks somebody else.

Now let's see what happened here. The teacher said, I want a classroom where everyone's thinking. The kid says, I don't want to think. Teacher says, okay. The kid's won. So a smarter response is to say, okay, if you don't know, I'll come back to you, and go and get three or four different responses from other kids, and then come back to that kid and say now which of those answers do you like best? So it's always, so it's never letting kids get away without thinking.

And the other thing you can do, you have to use it carefully, because you need to be sure that the only reason the kid is saying don't know is just to blow you off, but if you're sure that's what's happening, you can say, yes, but if you did know, what would you say? So these techniques, you know, can make classroom questioning a bit more fun.

A lot of teachers now actually let the kids, one kid have the responsibility of picking the names on the Popsicle stick. So they're given this cup with the sticks. And so whenever the teacher asks a question, she asks the cup holder to pick a name at random. And you've got to make it a bit more fun, so you can actually allow the kids to go, if it's a multiple choice question, go 50/50 or ask the audience or allow them to phone a friend.

You know, so you make the classroom questioning fun without getting away from this idea that thinking is required here. Everybody has to be on their toes. And hot seat questioning is where you just pick on one kid and then ask him a second, and a third, and a fourth follow-up question, and people say they'll be off task over here. Not if they know that the next move is going to be, okay, summarize what she just said. There's other ways of keeping kids awake without sacrificing the quality of the dialogue.

Okay. We'll just take a, we'll take a five-minute break in a couple of minutes. Before I do that, I just want to talk about all student response systems, because this is very important. When I was a teacher teaching in urban schools in London, the decision I made most often every day was do I need to go over this one more time, or shall I move on? How did I make that decision?

I make up a question on the spur of the moment, I ask the class, six kids raise their hands, I pick one of them, they give me the correct answer, and I say good and move on. How dumb is that? I'm assuming that the other 29 kids either already knew it or now get it, because they've heard one kid give the correct answer.

So if we're serious about high quality, engaging teaching, then we have to make, I would say, once every lesson, you have to get information from every single kid about achievement. We call these all-student response systems. So A, B, C, D cards, it's a very simple idea, just cards with just letters of the alphabet on a ring.

You know, you can spend \$2,000 on a clicker system, or you can get, batteries not included or required, and, in fact, what people don't realize is that these do one thing that you can't do with the clickers, because you can have two correct answers or three. You can have one correct answer that's only going to be spotted by your gifted students, so you can really push them.

And mini white boards, I think mini white boards are tremendous. I think mini white boards are the most important development in educational technology since the slate. There's nothing new here. The best teachers were doing this 200 years ago. And, you know, the great thing about mini white boards is you can ask a, like I said earlier, you know, sketch a graph of Y equals one over one plus X squared, eyeball the whole class, if everybody gets it right, you move on.

If nobody gets it right, you teach it again but slower and louder or preferably in a different way, but the most likely outcomes is that some kids get it right, and some kids get it wrong, and you've got this wonderful teachable moment, because then you can actually draw in the kids. These things, these A, B, C, D cards can be used for right/wrong questions, but they can also be used for opinion questions.

So in English language arts, you might have five different responses to a Shakespeare question, and you just say which do you think is the best? And then you hear from the A's first of all. Now you thought the answer was A, why? You thought also that the answer was A, why? You thought the answer was B. Do you think A is wrong?

So do you see how by actually getting information from kids you can make yourselves more skilled as a conductor or an orchestrator of classroom questioning, because you're collecting information, and you're adjusting your instruction on the basis of the information you collect to better meet your students' learning needs.

And then, finally, exit passes, the idea here, you know, if you're a history teacher, and you're doing some work on historical sources and bias, you might ask kids why are

historians concerned with bias when analyzing historical sources? And you give every kid a three-by-five index card. And then at the end of the period, the kids hand these things in as they walk out the door. So that's why it's called the exit pass. The kids can't get out without handing one of these in.

And what you do is, as soon as the kids have left, you read through these things, and then you put them in the trash. You don't grade it, because you know now where to start the next period of instruction. And that's a response to intervention, isn't it? Now, unfortunately, you haven't got a grade for every single kid in the grade book on how good their exit pass question was, but what you have got is much more responsive pedagogy.

You're constantly attuning your teaching to better meet your students' learning needs, and it does two things. It makes your teaching better, and it makes the students more engaged, and that's why that has the biggest impact on student achievement. Okay. One TV program that I think should be required watching for all teachers is *The Dog Whisperer*.

I mean, I'm amazed how few teachers understand what that guy is talking about in terms of, because when teachers are shouting at classes, it's not commiserative energy, is it? It's just a contest of egos, and so when I saw a drama teacher use this technique to get a group quiet, I thought yup, I'm going to do that.

Okay. Feedback. Response to intervention is all very well, but how do we actually get feedback about the intervention? Well, this study is interesting. This is a study of 12 classrooms. In 4 of the classrooms, the kids were given scores as they worked somewhere between 40 and 99. In four of the classrooms, they were given comments, and in four of the classrooms, they were given both the scores and the comments.

In the classrooms given scores, they made no gain. When they asked the kids do you want to carry on with this unit, the kids who got high scores said, yes, please, the kids that go low scores said, no, thank you. The kids given comments, their scores were up by 30%. They didn't know that, because they hadn't been given their scores, but their work got better, and when the teachers asked the kids do you want to carry on with this work, the high scorers were positive, and the low scorers were positive.

So what do you think happened for the kids given both scores and comments? Was it, well, here's a reminder, scores produce no gain, high scores positive, low scores negative, comments, 30% gain, high scores positive, low scores positive. So what do you think happened for students given both scores and comments? Take two minutes on your tables to discuss before I ask you to vote.

Okay. Thanks. So now can you vote, please? One finger for A, two for B, three for C, four for D, five for E, and vote now, please. What do you think happened? Okay. Thanks. Quite a wide group. Some people thought A. The typical view for A is that it's the best of both worlds, that the gain of 30% you get from the comments, but you also get useful information from the numbers.

B is a kind of averaging effect. Yes, you get a 30% gain, but you also get a polarization of attitudes, so the kids get despondent, because they get low scores. It turns out that what actually happened was D. No gain, high scores positive, low scores negative. Giving kids both scores and comments produces exactly the same outcome as just giving them the scores.

So what this study and others like it show that if you're going to write careful, detailed diagnostic comments and then go and put a grade or a score on it, you're wasting your time, because the kids who get the high scores don't need to read the comments, the kids who get the low scores don't want to read the comments. You'd be better off just giving them the grade. They wouldn't learn anything, but you will save yourselves a heck of a lot of time grading.

A second study by the same team, here, there were two classrooms given comments as they worked, two classrooms given grades, two classrooms given written praise, and two classrooms given no feedback at all. Experimental group two, the two classrooms given grades and the two classrooms given praise made the same amount of progress as the control group. Giving grades and giving praise is no more effective than giving absolutely no feedback whatsoever.

So much for this kids need to know where they are. Well, actually, your kids know where they are. Your kids can probably put themselves in rank order for math, language arts, science, physical education just as well as any teacher can. Kids know where they are. What they don't need is to be reminded of it all the time. But they want that feedback, don't they? They say what did I get? How did I do?

And you know what? We're the same. I'm trying to lose weight at the moment, and I know you're only meant to weigh yourself once a week, but after a day on lettuce, I want some feedback, don't I? So I get on the scales, and if I lost some weight, I feel good, and I don't want to eat, but if I haven't lost weight, or, even worse, I've put weight on, I get depressed and go looking for the nearest Dunkin Donuts.

See, I understand the psychology perfectly well. I want that feedback, but I only want it if it's going to be good, and I don't know if it's going to be good until I get it. And that's a trap we've created for our kids. So the kids say, you know, what grade did I get? And you can either give in, or you can be strong. And I think we need to be strong. I think we need to have less grades, particularly in elementary and middle school.

Arthur Cohen(?) says never grade students while they're learning. By all means, give them a grade at the end of their learning, but grading during learning actually impacts learning. And those are just two studies. This is the important study here, Kluger and DeNisi, 1996. They reviewed every single research study on the effects of feedback going back over 90 years.

They found 131 which were done with sufficient methodological rigor to be reliable, and out of those 131, they found in 50 out of the 131 cases giving people feedback made their performance worse than it would have been had you not given them feedback at all. In almost 40% of the cases, you would have better off shutting up than giving the feedback.

And what they found was the common characteristic of that ineffective feedback or that negative feedback was ego involvement, you're a good employee, you're a bad employee, you're a good student, you're a bad student. And the grades that we give students are ego involving. In our earlier study where kids were given both the comment and the grade, what was the first thing they looked at?

ALL: The grade.

DR. WILLIAM: What was the second thing they looked at?

ALL: . . .

DR. WILLIAM: Somebody else's grade. What did you get? And that comparison immediately puts you in a fight-or-flight reaction. One school I've been working with, there the principal insists there's a grade on each piece of work, so the English language arts teachers, who are usually the most rebellious in most schools that I work with, they give kids a minus, an equals, or a plus according to the work, whether the work is not as good as, about the same as, or better than the last work you did on this unit.

The kids who used to get the A's hate it, because they have to keep on getting better to get a plus. I mean, they will get the A at the end of the marking period, but during the marking period, all you're getting is feedback about whether you're making progress or not. I thought that was quite interesting, because the way we give feedback currently is quite profound in the way the kids make sense of their success and failures.

So Carol Dweck has spent a lifetime asking kids questions like when you get an A, why is that? When you get an F, why is that? And what she found out is that a lot of kids think that I got an A, because I did a good piece of work. That's an internal attribution. I got an F, because Mr. Smith hates me, external.

I got an A, because I'm smart she calls stable, because if you think the reason you got an A is because you're smart, you'll be guaranteed an A next time, because you'll still be smart. I got an A, because I worked hard on this assignment is not stable, because it's not going to guarantee future successes unless you also work hard.

And she found that there was a tendency for boys to attribute their successes to internal, stable causes like ability and their failures to external, unstable causes like teachers not liking them. Girls, on the other hand, are more likely to attribute their successes to internal, unstable causes like effort and their failures to internal, stable causes like not being smart enough.

And I think this is beginning to explain why girls are beginning to outperform boys all over the planet. Basically, girls are half right, and boys are all wrong. Because what the research shows is the best learners consistently attribute both their successes and failures to internal, unstable causes. It's down to you, and you can do something about it.

When I'm talking to kids about things like this, I use examples like Michael Jordan, who was dropped from his high school varsity basketball team for not being good enough. Mike Piazza, when he was drafted by the Los Angeles Dodgers, he was taken in the 63rd round of the draft. That means the L.A. Dodgers took 62 players ahead of him, and he was actually the 1,390th player taken in the draft that year.

Tom Brady, when he was drafted for the New England Patriots, was drafted at 199th. There are 30 teams in the NFL, each one needs about six new players, so if you drafted past 180, that's the league's way of saying to you, we probably don't need you, but we're going to give you a tryout. He goes on to win two Super Bowl MVPs. So this idea that nothing that you do changes where you are right now, but everything you do feeds into where you can be, I think is absolutely crucial.

And it also plays out in terms of changes in kids' views of ability. Some kids see ability as fixed. They say there are smart kids and not so smart kids. If kids think that, then when they are faced with a challenging piece of work, they immediately do a calculation, will I succeed? Because if they think they'll succeed, that's fine, they'll go for it in order to have their views of their abilities confirmed.

If they think they might fail, they do a second calculation, what about everybody else? Because it's okay to succeed, and it's okay to fail if everybody else fails. But it's disastrous to fail when everybody else succeeds. And what you see kids are not calculation deciding that I would rather be thought lazy than dumb, which from that kid's point of view is a smart move. But there are other kids who see ability as incremental. For them, a challenging task is a chance to get smarter.

So they don't care whether they'll succeed or not, they know they'll get smarter in the attempt. And what I find most intriguing is there are lots of kids who see ability in athletics, like triple jump, as being incremental, and ability in math as fixed. Those kids believe that by training, they improve their ability in track and field. And we have to get the same idea over in math.

So there are some practical techniques. Well, my summarization of all the research on feedback is this. If your feedback causes kids to think, then you're on the right track. If your feedback is causing an emotional reaction, you're sunk. Who, here, believes that your kids spend as long taking on board your feedback as it takes you to provide it? That's got to change. Feedback should be more work for the recipient than the donor.

So comment-only grading. A lot of teachers have tried this, but it's quite hard to get the kid to pay attention. So one English language arts teacher I've been working with, she did an essay on Shakespeare recently, and she got the comments on the kids' essays, but she writes them on strips of paper.

And each group of four kids gets back their four essays and the four comments on four separate strips of paper, and the whole group has to work out which comment goes with which essay. Isn't that cool? What she's managed to do is to slow down that immediate reaction, and now the kids are actually reading the feedback.

Focus grading. Don't grade everything, explicit reference to rubrics, obviously. Not giving complete solutions. Math teachers say this comment-only grading is all very well in language arts. You can't do it in math, because even if you checked 50 of these correct and put a cross next to five others, the kids' are going to look out for themselves, they got 15 out of 20, whether you tell them or not. So why don't you just tell them five of these are wrong, you find them, you fix them.

Feedback should be more work for the recipient than the donor. And here's why that's so important. We did some research in looking at kids' attitudes to the, the expressions on their faces when they got feedback. And we talked to these kids afterwards, and they said feedback means that you didn't do very well. And I said what do you mean? Well, the good kids, they just get A, good job. And if there's lots of writing or comments on your work, it means it wasn't very good said this one girl.

So we were quite surprised, but we realized that these girls, these kids, thought that the more comments the teacher wrote on your work, the worse it was. So they saw feedback as a punishment. So what we started trying to do is to say don't give the kids

feedback unless you can make the first ten minutes of the next period that you have with them a time when everybody works on improving taking on board the feedback.

So it doesn't matter whether you did a perfect piece of work or a very poor piece of work, there's some feedback there that says what's next for you. And that means that everybody is looking forward rather than backward.

And then Japanese teachers often have what they call the three-fourths of the way through a unit test. So in Japan, in middle school science, a unit on pulleys and levers, for example, might be allocated 14 class periods, but the content is only allowed to take up the first 11 periods. In period 12, the kids will do a quiz. The teacher doesn't grade the quiz. The teacher reads through the answers that the kids give to the quiz questions, and on the basis of what she sees, she plans lessons 13 and 14.

Response to intervention, using information about the students' learning to adjust her instruction to better meet her students' learning needs, and so, therefore, she ends up at the end of period 14 having used the information to adjust her instruction.

Sharing on the intentions. These are all very obvious, I think. Perhaps the last two are probably not so familiar. Before in science middle school, before you get kids to do lab reports, give them some lab reports from last year's class at different levels of quality. And get them to critique.

We know that kids are much better at spotting errors in other people's work than they are in their own, and so if you are to give them a chance to look at somebody else's work, they'll actually begin to think, oh, yes, and we've heard kids say, oh, yes, I must do that in my lab report when they look at good features. So getting them to critique other people's work is a way of teaching them the rubric. I actually think that peer assessment is a stepping-stone to self-assessment.

Teachers give kids rubrics, and they complain that the kids can't use the rubrics, because it's too painful. I think the best way to learn the rubric is in the context of applying it to somebody else's work. Then you internalize the rubric, and then you can apply it to your own work. But asking kids to apply a rubric to their own work is actually far too difficult for beginners in this process, so, for me, peer assessment is a stepping-stone to self-assessment.

And then, finally, when you're doing test preparation, it's been shown that more effective than either practice tests or revision guides or giving kids the chance to study on their own is to get them to write test questions with correct answers about what they've been learning. If you have two classes, say a period-one class and a period-two class, you can get the period-one class to write the tests, the practice tests for the period-two class, and vice versa. And that's been shown to increase kids' test scores by more than anything else.

Then, finally, students assessing their own or peers' work. This starts very young, kindergarten. A kindergarten teacher I know, she gets the kids to sign in every day, so the kids' names down the side, Monday, Tuesday, Wednesday, Thursday, and the kids have to sign in. For most of the kids, of course, it's just a squiggle. But on Friday, they have to go along with a buddy, and they have to look at their five sign-ins, and they have to circle the best one. So she's beginning to do peer assessment or self-assessment right away.

Shoes swap shoes. So kids learning to write the letters of the alphabet, write the letter D ten times, circle your best D, swap with a neighbor, circle your neighbor's best

D. If you disagree about what's the best D, have a discussion about it. Plus, minus, interesting. Kids coming in having done homework, this is fourth, fifth grade. You know, they actually have to post up what they thought was easy, plus, what they thought was hard, minus, and something they thought was interesting.

A learning portfolio. This, we got this idea from a reception teacher. She gets the kids to do a self-portrait every month, September, October, November, December. In January, the kids have to get out their self-portraits, and they have to talk to the teacher about what's got better in their self-portraits. And this girl said my arms don't go into my head anymore.

But what this teacher is trying to do is to encourage the kids to keep all the old stuff in the portfolio to see that there's a learning journey here, that there's a progression, building into this idea that ability is incremental rather than fixed. There aren't good painters and bad painters, there are painters who are getting better, and there are painters who are not.

And then, finally, two stars and a wish. When you get kids to start with peer assessment, quite a nice protocol is two stars and a wish where if anybody is comfortable with anybody else's work, they say I like that, I like that, I wish you'd done that. And that idea of twice as much positive as negative is a very nice protocol for getting started.

Training students to pose questions. So at the end of a teacher topic, you know, the teacher's typical move is to say, right, anybody got any questions? No, right, good, onto the next thing. So rather than doing that, you just say, okay, in your groups, decide if you've got any questions. And then, if they have, they have to write them on strips of paper or post-it notes for you. And what you find is that kids become smarter at asking questions.

And then, finally, self-assessment of understanding, traffic lights. This we got from a history teacher, so the idea was, you know, is it objective for the lesson, and at the end, the kids have to do a self-assessment. Green means I got it, yellow means not so sure, red, I have no idea what's going on. The trouble with that is it's a self-report, and it's risky. So we have to unpack that a little bit.

So what one teacher does is she says reds over here with me, greens help the yellows, yellows make sure the greens understand this quite as well as they think they do. So green, in her classroom, means I'm now ready to teach this to somebody else, make it a bit more grounded. And one teacher decided to, this is AP calculus, every kid was given a disc, red on one side, green on the other, beginning of the lesson, green disc showing. If you think the teacher's going a bit fast, you flip the disc over to red.

And what's interesting is how willing kids are to say I don't get it. In one class we heard about, this was a class who was taking Algebra I for the third time, okay, and one girl turned the disc over to red, and the teacher had forgotten to check. So the teacher was, you know, at the overhead projector going through this complicated bit of algebra, and this girl was getting more and more frustrated.

She looked at her two friends, realized that they didn't get it either, so she grabbed their discs as well and started flagging the teacher to slow down. This is a kid taking algebra for the third time, and yet she felt so empowered that she felt she had a right to understand. We thought what an amazing kid. But, actually, the more we've

done this, we've had more and more stories of that kind where if the teacher opens up a channel of communication with the kids, the kids will use it to respond.

Another teacher, have you ever had one of those lessons where you know it's not working, but you can't stop yourself? At the bottom of a deep hole and still digging. But we heard about this lesson, this was in England, and it was July last year, July, it's hot, seventh grade science, and this teacher is just, you know, it's, everybody's gone to sleep, I say, it's not working.

And he's trying to get more and more in the kids' faces trying to wake them up, and all of the sudden, a lad at the back of the class puts his hand up and says sir, this isn't working, is it? And I was talking to the teacher afterwards. He said a year ago, I'd be so angry with that kid for embarrassing me, but this teacher had been trying to get the kids to take more responsibility for their own learning, so the teacher just said you're right, what should we do? And not one kid suggested having a recess.

They had a sensible conversation about their learning and how to improve it. So it's about, and what's interesting about that story is this good, formative assessment is about getting the kids' ego out of the classroom, so they're okay about showing that they don't get it, but it's also about getting the teacher's ego out of the classroom. That teacher had the grace to realize that kid was throwing him a lifeline. And it was painful, it was embarrassing, but that was an important moment for that teacher.

Another teacher tried these. She couldn't see them in her classroom because of the fluorescent lights, so she went down to the party store and got every kid three cups. This is A level, this is AP calculus. So at the beginning of the lesson, green cup showing. If you want to signal the teacher's going a bit fast, you show yellow. And if you want to stop and ask a question, show red.

Why would anyone show red? Because in this classroom, the rule is as soon as one kid shows red, the teacher uses the Popsicle sticks to choose at random from all the other kids, and the chosen kid has to come to the front of the classroom to answer the question being posed by the kid who showed red. This teacher describes what she's doing as converting her classroom into one gigantic game of chicken.

And her current problem is that kids are stopping the lesson all the time. Isn't that great? Isn't that what we want? Isn't that we want, kids saying we don't get it when they really don't get it? Actually, to make it a bit more manageable, she now says you can only stop the lesson once, and if you want to stop it a second time, you have to get your friend to do it. But I think that kids telling us that I don't get it, I'm wanting to understand what causes problems, but I think they are problems we should want to have.

Okay. So so far, so good. So these techniques, I mean, you know, these have been around for 100 years, at least, most of them, so why are we doing it already? Well, it turns out is because most teachers' knowledge is not of the kind that we recognize. This is by some guys who work in business called Noncontingency(?), but they pointed out that most of teachers' knowledge is like other people's knowledge, it's not explicit. It's implicit.

And that's very revealing, because, in fact, the whole model of PD we talked about earlier, it's based on the idea that you don't know stuff, and we get you into a room like this, and people like me give you knowledge that you lack, and you go away and be better teachers. And it hasn't worked, has it, really? And it hasn't worked,

because the diagnosis is not correct. What you, you don't lack explicit knowledge. You lack knowledge of how to do this in your own classrooms, which is often tacit.

Most teachers can't describe what they do. So if one person's explicit knowledge is given to somebody else as explicit knowledge, we call that combination people telling people what to do. When you just pick up how things work in an organization where you pick up other people's tacit knowledge as tacit, that's called socialization.

But there are also interesting processes, whereby, people are forced to make their tacit knowledge explicit through a process of externalization, and it's also when people's explicit knowledge gets internalized as tacit knowledge. I've yet to have the experience of hearing advice from somebody, and you think you get what they're saying, but then months later you realize, oh, so that's what she meant.

That's a process of making that explicit knowledge your own personal, tacit knowledge. And what these guys, Nonaka and Takeuchi, suggested, if we're serious about recognizing the role of implicit knowledge, not just in schools but everywhere else, then we need these four processes of dialogue, networking, learning by doing, and sharing experience. And that's the way we need to drive professional development.

So my model for learning is content then process. So the evidence earlier on, all the ideas of formative assessment I've shared with you, are only part of the story. There's also how to go about change. And I think there are five important components, choice, flexibility, small steps, accountability, and support.

Why give teachers choice? Well, it's simple. Teachers teach differently. A guy called Meredith Belbin about 30 years ago now wrote a book about management teams, why they succeed or fail. And what he pointed out was that people tend to adopt one of a small number of roles in any group situation they're in. And they can't do the other things very well.

I did a, do you know what a 360-degree appraisal is? It's where you're appraised by people at your level and people above you and people below you. And I got my feedback on this 360-degree appraisal. And on one side was strengths. What do you think was written on the other side? It wasn't, it was actually areas for development. What a lot of rubbish. I'm 55. I'm not going to change. Deal with it.

What's really interesting about our approaches to teaching is that they appear to be predicated on this assumption that we want to make every teacher into a clone of every other teacher so that if you have weaknesses, we have to work on those weaknesses. Well, no, I'm afraid that's not right. Most students will benefit most by helping the teacher become outstanding at what they're already good at.

Now there are some teachers whose weaknesses are so egregious they require immediate attention, but for most teachers, their kids will benefit more by helping those teachers work on the things they're already good at rather than actually try to make every teacher into a clone of every other teacher. That's why I think teachers need choice.

Flexibility. These five strategies I've talked about are important for everybody to do. How you do it in your classroom is up to you. Those techniques allow customization, creating ownership, and sharing responsibility. And what, the interesting thing about this is that every teacher can then customize it to make it work in their own classroom.

Small steps. Why do we allow teachers to take small steps? Because if you push teachers faster than they want to go, they will humor you while you are around, but as soon as your back is turned, they'll go back to doing what they feel safe with. Now I said earlier, I started off the day with a joke about teachers not wanting to change. Actually, that's not right. Teachers do want to change, but they find it hard because of their existing expertise.

Teachers are very good at what they do. Watch novices doing things like trying to get the kids into the classroom. They're useless, aren't they? I mean, it's like herding cats. I mean, novices are just so useless at so many things. I mean, you know, an expert teacher takes about five minutes to plan an hour's instruction. A novice teacher takes about four hours. So, basically, a novice teacher does this thing 50 times slower than an expert.

Teachers are hugely expert, and what David Berlin has shown is that they are very similar to experts in other domains. What that means is that because they're expert, it's hard to change. Now expertise is also quite hard to spot. Kline and Kline did some video clips of people doing CPR. Five of the video clips were students who were learning how to do CPR, one was an experienced paramedic.

These clips were shown to different people, and they were asked who would you want doing CPR on you if you needed it. Okay? The experts chose the expert 90% of the time. The students chose the expert 50% of the time. The instructors, the teacher who teach other people how to do CPR, picked the expert 30% of the time. Why? Well, Kline and Kline concluded that the reason for this was that these experts, I'm sorry, the instructors, were looking for people to enacted the rules that they taught.

They taught a set procedure, and the people who followed those rules they thought were good paramedics. In fact, they often weren't, they were just novices learning about being paramedics. But what's interesting is the experts could spot the expert even though experts basically differed from the standard practice. So expertise in teaching is more than we can say.

The most powerful teaching knowledge is not explicit. That's why telling teachers what to do doesn't work. Involving, improving practice requires changing habits, not adding knowledge. We have a lot to learn from Weight Watchers, seriously. Weight Watchers, I mean, Weight Watchers ought to be the least successful organization on the planet, because everyone who wants to lose weight knows what to do, eat less, exercise more, that's it.

There is no magic third secret that you get told by joining Weight Watchers, that if you stir your cornflakes clockwise, it'll turn the dextrose and the sugar into a left-rotated form of sugar that is no longer absorbed by your body's physiology. No, I'm sorry, it's just eat less and exercise more. But Weight Watchers have got one fundamental insight. They are not in the knowledge-giving business, they are in the habit-changing business.

And that's what we have to get in with teachers. The hardest bit is not getting new ideas into people's heads, it's getting the old ones out. And if it happened naturally, the most experienced teachers would be the best, and they're not.

Now the other reason this is very hard is because, actually, most of the time, you have very little control over things that you do in teaching, because teaching is a largely

unconscious process. In fact, most teachers are rarely conscious while teaching. It's just too demanding, because there isn't time to think.

You're just in the zone in the same way that Joe Montana said he was never aware of being in a football game. He was just so much in the game that that's all there was. So he wasn't aware of doing what he was doing, because he was doing what he was doing. Your eyes take in 10 million bits of information, but you only process consciously 40 of them.

At the end of the day, if I have an extra bag to take home, I'd like to try and put it by my office door, and at the end of the day, I just walk straight past it. I saw it, but I didn't see it. Teaching is a largely unconscious process. Merleau-Ponty, the French philosopher, once said when I speak, I discover what it is I wanted to say. And the reason that changing teaching is so hard is because it's changing an unconscious process, not a conscious process.

Now nobody's got this cracked, so, you know, often education research is beaten up by comparison with medical research. But medical research has made huge advances in terms of things like surgery, but there are certain places in which medical research has made very little progress. Do you think that doctors agree that it's a good idea for doctors to wash their hands, for example, in a neonatal intensive care unit? In a study by Larsen and Natal in 1992, average compliance rate was 29%.

And you won't find one person who says, oh, I don't agree with hand washing, that's one of these modern ideas, I know . . . everybody knows that they should be doing it, but they don't do it, because habit change is hard. Nobody's got that cracked. So what, if you're serious about improving education for your districts, you need to create time and space for teachers to reflect on their practice and to learn from each other's mistakes.

Esther Dyson points out that there's no excuse for making the same mistake over and over again, but there's also not any excuse for not making any mistakes at all, as Mario Andretti, the formula one racing driver, used to say, if everything seems under control, you're not going fast enough. So what she, Esther Dyson says is always make new mistakes.

I come back to Samuel Beckett, ever tried, ever failed, no matter, try again, fail again, fail better. And I think that should be out motto as teachers, because we will fail, because we have such high expectations of our students. What we can do, and this is crucial, we can do a better job tomorrow than we did today.

If you don't believe that, you should be fired immediately. Because what happens when a teacher doesn't believe they can be better is they start blaming all their future failures on the kids. They say what can you expect from these kids? The fact is that this job is impossibly difficult, you know. It's not rocket science. It's much harder than that. The great thing about it is you never get any good at it, and that gives you a lifelong career.

Andre Previn was at one time the highest paid film score composer in Hollywood, and one day he just walked into his office and he quit. And people said to him why did you quit this amazing job, and he said I wasn't scared anymore. His job held no more challenges. He was going to his office every single day knowing he could do the job. You don't need to worry about that problem. This job that you've got is so hard that we're never going to actually crack it. What we can do is do a better job in the future.

So what we need to recognize is that in the same way that teachers can't do the learning for the students, you, as learning leaders, can't do the learning for your teachers. But what you can do is secure from your teachers a commitment to the continuous improvement of practice.

So here's a question to ask every teacher. Do you need to get better? If they say no, do everything you can to get them fired. If they say yes, work with them. Most teachers asked that bold question do you need to get better would actually say yes. And once you've got that commitment, then you can start working with them.

What's needed with leaders, then, is creating expectations for the continuous improvement of practice, keeping a focus on the things that make a difference, and providing time, space, dispensation for innovation, and supporting risk taking. Get teachers to promise what they're going to try out. Try something out. These are lists of why we think action planning is important.

But here's an interesting quote from a teacher who's reflecting back a year later about the process of being asked to commit to making a change. He said I think, specifically, what was helpful was those ridiculous no carbon required forms. I thought that was the dumbest thing. When I'm sitting with my friends, and on the NCR form, I write down what I'm going to do next month.

When it turns out to be a sort of I'm telling my friends I'm going to do this, and I really actually did it, and it was because of that, it was because I wrote it down. I was surprised at how strong an incentive that was to actually do something different, that idea of writing down what you're going to do, and then because when they come by next month, you better take out that piece of paper and say did I do that?

Just the idea of sitting in a group working out something and making a commitment, I was impressed about how that actually made me do stuff. So it's that notion that Weight Watchers have of making a commitment knowing that that's going to be what's going to keep you going until the next meeting. So that's the big idea here is that the research on a formative assessment suggests that this is the biggest change you'd be making in terms of classroom practice.

The research shows that, the ideas or the techniques I've shared with you, and then the support model. Choice, giving every teacher a choice about how to take this forward within that agenda of formative assessment techniques, because that's what matters. Flexibility. The teacher who changed this into this made it much better, and that's why I love working with teachers, because the creativity of teachers to take an idea of somebody else's and make it even better is just unlimited.

Small steps, allowing teachers to go as slowly as they like, because that's actually how they can actually can slow this. Somebody described changing your teaching as a bit like engine repair in flight. It's that kind of messing about with something that's kind of working and getting through the day. And then, finally, those two things, accountability and support. Actually, two sides of the same coin.

Support of accountability we sometimes call it, the idea that creating an accountability structure also provides support. And the evidence is when teachers engage in this process, we get improvements of something like one standard deviation of teacher effects. So, you know, we're seeing significant improvement on student outcomes in as short a time as a year when teachers change their classroom, and we get kids who are more positive in their learning.

In the second workshop after lunch, I will dig a little bit deeper into the exact details of how to do this in practice, but we just have a couple of minutes left before we finish those. Does anybody have any burning questions? Perhaps I should have made you write them on strips of paper or an exit pass question, yeah. Two hundred exit passes, yeah. Okay. In that case, I suggest we steal and march on the others at the coffee cube.