

Assessments are not about results, they are also about how well results do materialise for different types of students.

In some countries you can say that the impact of social background on learning outcomes is very, very strong. What this means is whether you come from a rich or a poor family matters a lot for the success of your children. But international assessments also tell you this is not true for every country.

Some countries do really well in moderating the impact of social inequalities. So you have again not only differences on the vertical axis in terms of quality, you do also have big differences in terms of the horizontal line. Everybody wants to be where the results are very good and the impact of social background is very small. This is the problematic area where results are not so good, and the impact of social background is very, very large. And that's not just a social question, that's also an economic question. When the impact of social background on learning outcomes is strong, it basically tells you that societies do not use their potential. There is a lot of human potential that isn't effectively utilised.

The green and red quadrants above are clear, nobody argues about this. It's harder to think about the quadrant *High average performance – Large socio-economic disparities*. Some people argue, "well, you know, in order to do well on average, you have to accept large disparities". And it is also hard when we come to the quadrant *Low average performance – High social equity*. Here people say, "well, you know, better focus on equity and accept mediocrity." Other people believe there is a trade between quality and equity: You can only maximize one and you will lose on the other. If that were true you would find all countries in either the left yellow quadrant or the right yellow quadrant. But the reality isn't like this.

If you look basically at how countries come out you see there are many countries in every corner of the world that are able to deliver strong results overall and an equitable distribution of learning opportunity.



This is how assessments can actually influence our thinking. Assessments can tell us that quality and equity are policies that can be tackled at the same time. A very important conclusion, you couldn't derive from a national debate. When you are stuck in a national debate people will always tell you "well, you know I don't want to put children from different social backgrounds together, because it goes to compromise the quality of my own child." Well, here you have results from other countries. You can use assessments. They allow you to use the world as a laboratory. You can think about what everybody else is doing, with what impact.

You can see that Portugal doesn't only struggle in terms of quality, Portugal also struggles in terms of equity, in terms of the fact that social background is a very strong determinant of success. Not as strong as Germany, my country. Results in my country are better than in Portugal. Germany does quite well by OECD standards, but social background is a very strong determinant of success. And we know why, because in my country, in Germany, the education system divides young people very early into different schools. At the age of 10, one group of young people is told, "well, you know, you did well at the age of 10, we will put you to a school leading towards university and you'll become a knowledge worker." And the other two thirds of the students are told, "well, you know, you didn't do so well, you'd better go to a vocational school and you're going to work for the knowledge workers." And that's basically how the system works. I exaggerated a bit, but more or less that's how the system reinforces inequality.

In France, very much the same story, even worse. On average they are doing ok but the impact of social background is very large. The United States... Portugal is not alone on that sort of dimension, there are many countries struggling. But you also can see there are many countries that are really very successful, the Nordic countries, Japan and Korea in Asia, Canada in North America, very, very successful.

So again, international assessments can set the ambition of what is possible, what we can strive to, what we can achieve.

And I just wanted to highlight a bit of what we have learned about.



That is my country. You can see the social background in the horizontal axis and the student performance on the vertical axis. Every dot is one school that has been assessed, the bigger dots are the bigger schools, the smaller dots are the smaller schools. And you can see how strong the relation is between social background and learning outputs. If we come from a privileged family we end up in a school that does really well.

The question is 'why? Why is it like this?' And again international assessments tell you something about it. Some people say it is all to do with families: Poor families don't spend money with their children, they don't have education and rich resources and so it's all to do with families.

You can actually estimate that. The yellow line shows you the impact of family background on learning outcomes inside a school. And as you can see, yes, there is some relation. Children of poor families don't do as well as children from richer families, but this line doesn't explain what we see. So you can't blame parents for all of it. The big part in the case of Germany comes from the school system, so basically international assessment tells you that's something public policy can do about. Most of the social and economic disparities observed in the case of Germany are attributable to the school system, not to parents.

Let's look at Portugal.



The first thing that strikes you for Portugal is that the dots are much more scattered. What does it tell us? It tells us that there are actually many poor schools from poor families that do really well and there other schools that are quite disadvantaged by Portuguese standards and the OECD standards that do not so well. And also have some schools from privileged backgrounds with so-so results. So, social background is not a guarantee for success, it is of course a strong determinant, but the cloud is quite scattered. You can also say for Portugal that the yellow line is flatter than the blue line, which tells you that the school system tends to reinforce social disparities.

Ideally we hope that education becomes a big equalizer. Ideally we sort of expect from education that it helps students from social backgrounds to move up society, but that's not what the reality shows. The reality is for Portugal and for Germany that the school system tends to reinforce social disparities.

Does it have to be like this? Again, this is the power of global assessments.



Look at Finland.

The yellow line, ignore the red line, the yellow line is like in Portugal. Finnish parents are not magic parents. If you have a poor parent, you end up with a child that does a little bit worse. If you have richer parents, children tend to do better. The yellow line is exactly like in Portugal.

What distinguishes Finland from Portugal is the blue line. In the case of Finland it's absolutely flat. If you have a child from an immigrant background, from a poor social background, you can put that child together with children of other similar situations and the school will address it, the school will help you with your child's development.

If you do that in Portugal or in Germany or in many other countries, France and so on, they don't do that, they will try to put your children in the most privileged school to be cared, in order to benefit from their neighbouring social context. Very important results!

Basically assessments can actually influence how we look at the world, how we look at the outcomes and what is possible to achieve. Finland is a very far remote country, and so on very special social circumstances, but in a way it tells us what's possible. And what's interesting is that most of the East Asian countries are very, very similar.



Of course there are some questions to be asked: Are quality and equity attainable goals? Do we measure what matters for people?

One important way to know how education assessment can influence the world is actually by looking at competences that matter for people. *How do we know that we found it*? How do you know that what you teach today is going to really help students in their future? How do you know that what you test today is shaping their future?

One way to do that is looking at what happens to these children later in their lives. And we have done that. We have done it in Canada. We've actually followed the students since the age of 15 in the year 2000 up to now and we'll continue to do that for another 15 years. So we can see what impact the PISA skills had at the age of 15 and what impact their attitudes had on their future lives. I will show you this, with some results.



This is the example of Canada. It is very expensive to do those kinds of long studies, we don't have them for many countries, but we have them for some.

This is at the age of 19. The people we tested when they were 15 years old in the year 2000 and who had achieved level 2 on the PISA test were by the age of 19 twice as likely to have moved on to higher education or a successful job. And this is after accounting for all social background factors, this is the real impact education has, it's the PISA skills; the people who had achieved level 3 had four times more chances to have moved on; the people who had achieved level 4, who were good readers in the year 2000 according to PISA, were eight times as likely; and the people with level 5 were sixteen times as likely to be successful in terms of education paths, of career paths.

Now what does this tell you? It tells you, on the one hand, that the kinds of skills that PISA measures, at least in Canada, are really important for the success of people: An odds ratio of 16 to 1. Doing well really matters for your future career. The other way of reading this is, "well, what you haven't achieved by the age of 15 is very, very hard to accomplish." You can also look at this in a way of telling yourself, "well, you know, what you haven't learned in school you can pickup later in life, there are lots of second chance opportunities." But that's not what these data tell us. What people didn't acquire by the age of 15 is very, very hard to achieve according to the probabilities.

You can do the same thing for economies. You can look at the long term impact of skills, in terms of economic output, and you will see the same picture.



You can see that basically the added level of economic output from improved performances on assessments on PISA is just very, very huge.

If the United States in the year 1989 would have actually done what they said... By that time they said, "we are going to become the best education system in the world..." If they had achieved that, by the year 2040 the added economic output would have paid for the entire education system. The education system would become an asset rather than a liability. By the year 2059 we would talk about 50% of added GDP with improved education performance. Again that's another way assessments can shape the world.

But I want to come to the last part of my presentation which is usually attributed great importance.



How do we understand what contributes to the success of education systems? What are international comparisons really about?



The first thing we need to think about is that money is important but not everything. The relationship between spending per student at the age of 15 and outcomes... you can say there's some relationship but you can see that some of the best performing systems are not spending very much, and some of the highest spending countries don't do particularly well. So money is no guarantee for success.

How do you spend your money is much more interesting. And I will show you this because the results for Portugal are very interesting.



When you look at spending per student in relation to GDP per capita, we find Portugal number one. In relation to its spending capacity Portugal actually invests a lot in education. No country in the OECD invests more per student relative to spending capacity.

How do countries spend their money? It's another interesting question, not what they spend or how much, but how they actually spend it. One way you can spend a lot of money is by paying teachers well.

Look at the example of Korea. Korea, remember it is one of the best performing systems, puts a lot of money into teachers' salaries. They try to pay teachers really well, in relation to spending capacities. That drives the costs up, of course.

How else can you spend money? Instruction time. The longer you make school days, the more you end up spending. And you can see also

that in the case of Korea, that's more to spend. Students in Korea go for many hours to school per day; it costs a lot of money.

You can also spend money by giving teachers time to do other things than teaching: Professional collaboration, curriculum development, things like this. Korea does that as well. So you can see that's the bill they would have to pay if all those factors would matter.

It is a very, very expensive education system, but in reality they only spend that much. So, how do they save money? The answer is class size. They basically say, "We want to get the best people into teaching, we pay teachers a lot. We want that students have lots of opportunities to study, many hours. We want to make sure that our teachers continue to learn. We give them lots of opportunities for professional collaboration. All of this drives up costs, and how do we pay for this? With large classes!"

Luxembourg spends as much as Korea. But in Luxembourg parents and teachers like small classes, everybody favours that, that's what's driving costs up. Luxembourg pays a huge bill for having very small classes. And what does it mean? It means there is very little time for instruction, short school days, they can't pay their teachers well by relative standards, and they cannot give their teachers any time to do things other than teaching.

So what I am showing you here is that for all the things that you do, there is a cost associated.

Why is Portugal so expensive? Not because students have long school days, they have relatively shorter days than everybody else in other countries, but teachers are quite well paid, teachers have time for other things than teaching, class size is driving costs up as well. So in Portugal all of those things add up.

International comparisons can tell you a story about how you spend your money, in comparison to anyone else. What is the ideal class size? People argue about this for decades. What is the ideal teacher's salary? Who knows? But you can see actually how everybody else is spending their money. International comparisons cannot tell how Portugal should be doing, but they can tell Portugal what everybody else is doing.

Yet, money is not the only thing. International comparisons also show us some of the other factors that drive performance.



The first thing you notice when you look at outcomes from PISA and you look at the countries doing well, most of the countries doing well have very *clearly articulated ambitious standards*, universal standards for all students of what good performance is.

Finland doesn't have a very detailed curriculum. They don't tell teachers what to teach. But they are very clear about what good performances in Mathematics, in Science, and all of the other subjects.

They tell teachers what students should be able to do, as opposed to telling teachers what they should be teaching.

Combined with this and with the ambition system, there is also the capacity of a system to get the best people into teaching: how do you attract the brightest people into teaching? How do you make sure that teaching is an attractive profession? It has to do with money - we have just seen that - but not only. In some countries teachers are paid quite well, Portugal is also one of them, and still they find working conditions that are not as conducive as they are in other countries.

So it's the combination between the challenge and *support* that assessments tell us does matter.



When you have low ambitions and week support systems you don't expect very much. I think that's pretty obvious. Raising the challenge is a very favourite thing in some countries, you just make more ambitious standards, you make more demanding assessments, you assess more frequently and so on. These are things that are easier to do, but if they are not backed-up by that kind of support, you don't achieve that much. We have actually lots of good experiences across the world from those kinds of policies.

We have studied other countries like Denmark and Norway. They have very strong support systems, wonderful ways to support their people, but they are not very ambitious. They are not clearly articulated in what good performance is. And as a result some schools are good and other schools are not that good. So, according to our analysis, what really makes a difference is a combination of strong support and high levels of ambition. That's something that comparisons can tell you across countries: factors associated in a system.



There is another dimension that is very important.

We see, for example, that in most systems doing well there's a lot of responsibility on schools, schools have a lot of room to manoeuvre, but at the same time systems know how to intervene otherwise if necessary. I mean, systems do know when a school fails to have the capacity and the means to intervene. I'm going to show you this with some data.



You can measure school autonomy in many ways; it doesn't really matter for the purpose of this analysis. Here I've just measured whether schools have an influence on whom they hire as teachers and to what extent the system has some kind of external reference framework for performance. And on the vertical axis I show you the PISA difference.

What does this show you? Schools that have very little autonomy and no external reference for performance tend to be the worst performing schools. Schools that have no autonomy but they do operate on a clear framework of external standards come out 46 points higher and that's more than one school year advantage. Schools that have a lot of autonomy but no external frameworks are often private schools and come out 41 points. But the biggest price comes with what you've seen before, with the combination of local responsibility and at the same time an external framework of standards. I just want to make sure that this is not misunderstood as a plea for private schooling. I mean, local responsibility has nothing to do with private schooling; in fact every public school in Finland has more room for discretion than any private school in Portugal.



So basically this is not an issue of who runs the school, but how the schools perform. When you look at the results, you can see that in most countries private schools do a bit better than public schools. It's the bottom line performance. But when you account for the social composition of the schools you can actually see there's no performance difference left.

So again what I am showing you here in terms of local responsibilities is not to do with who runs the school but with how the school operationalizes its degrees of freedom. In fact, there are some countries where you can see quite a big performing advantage of public schools. If you look at Japan or at Switzerland, for example, public schools do quite a lot better than private schools after accounting for social background in particular. Just to look at the slide below, at what we have learned from international assessments, when you take all the school factors together.

Measured effect is what we see and *Effect after accounting for the socio-economic background of students, schools and countries* is what we would have seen, if all students, all schools, all countries would be equal.



You can see that *schools that group their students by ability* usually don't do so well, that's something that drags on the performance.

But when you have *schools on average*, across countries, *being academically selective* , they do better. That's very clear: if you can select your students with an entrance test you have an advantage, and that is true before and after accounting for social background. You may wonder, "Well, you know, that seems a good trick, to make every school selective. What if you make a school policy a system policy?" And then you can test it. We will actually see there's no system with effect left, so that's not so important.

This is a very controversial finding; *schools that post their results publicly* **____** typically come out better.

This is also important, *one hour of Science learning per week at school* has quite a big performance advantage. Some people say, "You know, it's not quantity that matters, it's quality." It's true, but at the end of the day the outcomes are the product of quality and quantity. And you can see quantity alone has quite a big influence on outcomes: one hour of Science learning at school in terms of Science performance.

Out of school learning has negative impact, but I think that's just a selection effect.

One hour of self-study or homework also has a positive effect, but it's much smaller than what schools can achieve. Why is that important? Because of course the most precious resource in education is the student's learning time. It's the only thing you can increase. When you spend that hour in school you get a lot, when you spend it out of school you get little. So, both are positive but in very different ways.

School activities to promote Science learning _____. Here you can see schools that did more than just teaching: in Science competitions, Science exhibitions, they have brought science to life. You can actually compare to tell you that this actually is already a positive outcome.

We already talked about *autonomy* , it's a big factor influencing results.

The rest of the factors are relatively small and they usually don't usually come out in the model, *Effect after accounting for the socio-economic background of students, schools and countries*, so I'm not going to talk about them.

What you see on the above chart is almost worth two school years. You get all those factors right, the most successful schools are two years ahead with those kinds of combinations of factors. Finally, and with this I want to conclude, we do see that success is associated with systems that are highly inclusive, highly integrated, and that at the same time combine this with access to high level of personalised education.



It's not just putting all students together, doing the same thing with all students, but it's about working in an integrated system in a highly individualised way. That's what we see, and we see that happening in some countries even with large classes, where we ask students for example, "do you perceive that your teacher understands what you could do?" "Does your teacher support you to help you achieve your maximum goals?"

In some countries, even with large classes student's respond very positively. In other countries even despite small classes, students don't perceive them to be an individual, they perceive them to be one of many in the class. And those things seem to be mattering, at least in the perception of students.



And I want to show you this; it's my last slide.

On the vertical axis you see again the quality of outcomes, we've seen that before. On the horizontal axis, equity. And now you can simplify this and you can divide systems into two types: High degree of stratification and low degree of stratification.

How do you deal with the variation of student's performance? There are two philosophies. One philosophy is, and I mark this in green, that you basically have a very open education system, highly inclusive and it's the teacher's responsibility, the school's responsibility to constructively engage with this diversity, to personalise learning. Marked red are those education systems with a lot of what we call stratification in the system, those which say, "well, we deal with variation through the system. If students haven't succeeded, we let them do the same thing next time with another teacher. If student don't do well, we put them into a vocational school". I mark the inclusive systems in green and the stratified systems in red. And you can see that on the vertical axis, maybe the distinction is not so clear, but on the horizontal axis the distinction is very clear.

The more stratified an education system is, the greater the impact of social background on outcomes. Systems are not so good in selecting students by their academic potential, but most school systems are very good in sorting students by a social background. It's different. And that's true for Portugal as well, and you'd tell me, "well, how can it be? Portugal has a very sort of standardised education system."

Well, in Portugal that kind of stratification very much comes with great repetition. A lot of students repeat year after year and many students for more than once. We can see that in the PISA results. And usually that is also a social divider. We can see that very clearly associated with the impact of social background on learning outputs. Not just in Portugal.

And, by the way, it's a very expensive way of addressing problems. And we can say for example that the educational cost for a great repetition is not very large: we just put that student into another classroom and you have one student more on the class. But if you think about the social cost, we've estimated that cost of one student repeating one year being about 15 to 20.000 euros in OECD countries. Assume Portugal at the lower end of that spectrum; 15.000 euros being wasted by a student repeating one year more.

What does Finland do? Finland doesn't spend that money the same way, it gives that money to the school. Schools can devote 30% of their instruction time to learning outside formal classes. Schools have a lot of resources where they can basically identify problems, get students into a special class, put them back into their classroom, devote a lot of resources: social resources, educational resources for the students. And the money they spend for this is less than what Portugal loses in terms of high repetition rates.

Finally, what we see here is that tracking stratification always works against equity in systems. It's quite consistent across countries with very, very few exceptions.



With that I want to conclude the presentation.

I just want to give you one more picture of our future in the context of PISA. Where do we see the future of assessments in terms of the value for policy, in terms of what we believe we can do, there are also things we want to do but not everything that we do.

Of course we want to do the things that are important, that we can do. We don't want to do the things that are difficult to do and maybe not so important. Assessment must have the ambition to move the agenda forward to do some of the things that are hard to do but very, very important for public policy like assessing the kind of competences we are not able to capture today. And of course we always pick up some things that are maybe not so difficult to do but not so important.

This is what I've shown you that we can actually do quite well: examine the individual institution and systemic factors associated with high *performance*. Today we can account for about 70% of the performance variation of schools across countries. We know what factors are associated with success. We know nothing about the cause or relationship, that's not for PISA to solve, but we have a quite good idea.

Monitoring educational progress is very, very important and hard to do. People tell us "why is it so difficult? You just repeat the same test." Yes, but who wants to know in the year 2020 what was important in the year 2000? The kind of skills that matter change and so keeping that balance between innovation and continuity is a very tricky challenge for us and for everyone. *Establishing the relative standing of countries in terms of quality and equity in basic subjects* is not such a problem.

Measuring growth in learning – one of the things that are most important for teachers – is where we are knowing very little as much on how well students do, but how they are progressing. When you get a student at a primary school, what do you know about that student from early childhood education? When you get a student at secondary education, how do you know what the learning difficulties of that student were earlier?

Extending the range of competences through which we evaluate quality is very, very important. Assessment still is a very narrow perspective. We focus on things that are easy to test, not always on things that are important, extending, that is important. And we have a lot of things in PISA on the way to improve that.

And then finally, the biggest prizes will come to building *real-time* assessment environment that bridges the gap between what we call summative assessment and formative assessment. How do we actually connect what you do with PISA with what is happening in classrooms and build those kinds of bridges? That's the sort of picture of where we see the challenges for assessment in itself, but the real message I really wanted to convey here is that we know that some systems are very successful in quality, equity and efficiency, and we can understand many of the factors that relate to the success of these systems.

So the assessment can influence our thinking in terms of understanding and learning from other systems. It can influence our thinking in terms of the kind of competences we want to value in our society and compare it to others, and things like this. Thank you very much.