Conferência

International assessments, quality and equity of education systems

Andreas Schleider¹

First of all, please accept my apologies for doing this presentation in English. It's not my native language either, but I think it is the only language we can use to communicate this morning.

I would like to thank the CNE for this invitation. It's a pleasure to present our work to you here, since Portugal has actually been a very active participant in our work on international comparisons, which I will be speaking about.

My presentation is about the impact of international assessments and one of my aims is really to look at the impact beyond the rankings.

One of the things we always look at first when we see the international results is where countries come out in terms of quality and in terms of equity, but there are many other dimensions that can influence the debate in participating countries.

¹ Head of Indicators and Analysis Division – OCDE



First of all, *why are you looking at international assessments?* Why do we compare internationally? Because we live in a world in which improving by national standards is very important but to look at how everybody else is changing is also important. If the context around us is changing fast, then improving in national standards alone is not enough. One of the things international assessments can tell us is the pace of change, the pace of improvement, catching up with the rest of the world.

Second, *what are we looking for*? How do we define quality in education? How do we define equity in education?

There is no straightforward answer, but we can say that global comparisons, global collaboration, can provide important answers to questions like What are those kinds of skills that make people successful?, How do we measure them? How do we compare them?

This is a debate which no country can answer by its own. However, through working together countries can collaborate in finding answers to those kinds of questions. Third, *Search systematically*. Even here there are big changes to be considered and understanding those changes is a very important topic where assessments can influence debate: The kinds of competences that made our parents successful are no longer the kinds of competences that will help our children to succeed.

How do we benchmark quality, equity, efficiency in education? These are also very important dimensions where assessments can help us.

Yet, *how do we know that we found it*? How do we know that we get the answers right? When you devise your national curriculum in Portugal, how do you know that what you teach is really what matters? A lot of hypotheses are going in, and that's the same with international assessments.

When you have an examination system, how do you know that what you test is what really matters for people in their life? It's a challenge of external validity and it's another challenge where global comparisons can be very instructive.

And finally, what do we do with the results? What can we learn from this?

I'm sure this is also going to be the focus this afternoon basically: the kinds of lessons that you can learn.

As the Minister said, this is not about copying and pasting education systems. Nobody can copy solutions. But international assessments can help us to understand the drivers of success, what makes different education systems successful. And from that we can think about the way we can configure those drivers in our own national context.

So this is the kind of questions I want to answer with you today.

I'd like to start with something that isn't just about assessments, that's about our comparisons of higher education, because that's where we see the most dynamic changes.



Every dot on this chart is one country and the chart is twodimensional. On the horizontal axis we see how successful countries are in producing high qualifications, university graduates. On the vertical axis you can see what countries spend per student. That is, success and cost.

The size of the dot shows you where the money comes from. The larger the dot, the more money is mobilized from the private sector.

As you can see, the United States takes the lead in terms of graduate output, very far to the right: lots of people get high qualifications. It is also number one in terms of spending per student: lots of money is available for each student. In addition, the dot is very big because a lot of money gets mobilized from the private sector.

Finland is moderately successful in graduate output and moderately expensive, in the middle of countries. On the other hand, the dot is very

small, which means parents don't have to pay, it's all provided by the government.

Germany, for example, is pretty much to the back: University education does not have such a strong tradition.

Portugal is at the end of this scale in terms of output, not that expensive, but showing a lot of public expenditure.

The only thing I haven't said so far is that this was the picture in 1995, about ten years ago.

Now, I want to show you how quickly the world has changed. If you look at the picture above (1995), you'll see how far the United States is ahead of Portugal. Five years later the world looks different. By 2000, you can see many, many countries moving very quickly to catch up.



The United Kingdom moved to the front. Australia is also a very successful country in moving forward. But look at Finland! Do you remember where Finland was five years before? From the middle of countries, it moved to number one, in five years of time.

This shows that improvement by national standards is not enough. Every country is improving, every country is moving to the right, and actually Portugal is in a great place in points, as well. However, you also see that some countries are moving forward much faster than others. This is why international comparison is so important.

But the world did not stop in the year 2000, as some people had predicted, things continued to move. If you look at 2001, 2002, 2003, 2004, 2005, 2006, what you see is a rapidly changing world. In ten years time the global education has fundamentally changed.



Do you remember the United States was number one in 1995? By 2006 it is an average performer. Has the United States declined in its graduate output? No, it didn't, it actually even improved. But everybody else improved so much faster. You can see Australia, Finland, United Kingdom, Germany and Portugal.

All I want to show you in this chart is really the rapidly changing world in which we live, and it is no longer a question that every country makes progress. Every country can claim its progression, but the pace of change is very different across different countries. And the world will continue to change.

Now I want to show you the comparison between China, the European Union and the United States in terms of the future supply of college graduates.



Take 2006. Basically, you can see that China, the United States, and Europe are more or less comparable in terms of the number of people with a higher level of qualifications.

But look at what the picture is like the next year, 2010. You can see how countries like China, which are just starting to build up their education system, have been progressing. If you look at projections for 2015 and projections for 2020, you can see a rapidly improving education pattern, something that we couldn't expect by now. Europe has also improvements, the United States also a bit, but the world in the future will look very, very different in terms of the supply of qualified people. Of course this is all about quantity; we don't know how good these people actually are.

The only thing this shows is that the world in which we will be living will look very different. Globalisation will look very different. Nowadays, it is easy for the Western world: it is basically competing with countries (like China and India) that provide low skills at low cost. But the picture above tells us everything is going to change. In the future, competition will be with countries (like China and India) that provide better and better skills at increasing pace.

But this is again about quantity. However, this impact is very clear and puts pressure in all of us to increase the pace of change, to increase the pace of improvement.

What, then, should education systems deliver? Assessments can have an important influence on the kind of skills that are being valued by society. You can argue that you can't improve what you cannot measure. What is measured gets valued.

Let's have a look at this.

When school was invented and it was all easy: there was one truth and you just had to explain it to a lot of people.

In the industrial age, schooling was also quite simple: it was about making people compatible with the norms of the industrial society. Basically, schooling was about giving young people a set of skills that would last for their lifetime. Very easy to do! It's a reproduction of subjectmatter contents. That was the education of the industrial society.

And then the world as become more complicated, demand has increased, supply has increased, but the model of education hasn't changed very much in terms of skills.

However if you look at the challenges that education faces today it's all about creating people that can continue learning throughout life. Take science for example. What do you teach young people? Half of the science that people will need in their lifetime doesn't exist by the time they are at school. Where is the relevant content?

Let me show you these data for the United States. I do not have these data for Portugal.

This picture tells you that the nature of skills that make people successful is changing very rapidly too. It's another dimension that international assessments can talk about.

We can start with *routine manual skills*. In the 1960's and 1970's lots of factories were created and thereafter fewer and fewer is the need for people who earn their money with routine manual work. That's very clear because those kinds of jobs can be automated, machines can do them better than people. They can be digitised, computers can do this. They can be outsourced: in Vietnam and China people can do that cheaper.

If you look at *non-routine manual work* (work that you do with your hands but in ways that are harder to explain to a computer), you can also notice a decline but things are flattening out, non-routine manual work is quite safe today, because there are lots of jobs that you can't easily automate, digitise and outsource. Someone is going to clean up this room

later on; you can't put a computer at it. You can't outsource your hairdresser to China. There are lots of jobs that are not very demanding but they are quite safe.

This is not what matters now because assessment is not about manual skills. Assessment is about cognitive skills.

So, what about *routine cognitive skills*? What I mean by this is the capacity to reproduce what you have learned: you learn something and later in your life you use that knowledge.

What have we seen in the labour market about reproduction of subject-matter content? Actually, when you look at the picture above, you can see that the decline in the demand for skills no longer occurs in the manual sector but in terms of routine cognitive skills.

Now, why is that important? Because those are the things that are easier to teach and the easiest to test. Assessment has largely been about those kinds of skills (routine cognitive skills). You teach somebody something and you test whether they have learned it. Very easy to do, multiple-choice tests can do that very well.

Yet, the picture shows that our modern world is less and less in need of routine cognitive skills. Very important lessons we can derive from those kinds of labour market statistics that shape the view of comparative assessment... You can see those lines all going down, they are decreasing in demand.

What is really rising in labour market's demand is actually *non-routine analytic skills*. It's not about the capacity to reproduce what you have learned but the capacity to extrapolate from what you have learned. To apply your knowledge in a novel situation that you haven't seen before, that maybe nobody else has seen before. A very important dimension and we see it is rising.

And that's something, I must say, on which we are putting lots of emphasis in PISA. PISA contains lots of tasks that require students to extrapolate from what they know and apply their knowledge in a nonfamiliar situation. This is one view assessment can take. It is another way through which assessment can influence our thought about education.

Some people say that's unfair, some people say PISA is unfair to students because they're testing young people with things that haven't been taught in school. And that's true. That's a true point you can make in any country. But if you take that line of argument you can also argue that life is very unfair to people, because in real life you are going to see lots of things that you haven't seen before, and the test of truth is not whether you remember what you have learnt in school, but whether you can solve those kinds of novel situations.

The US Department of Labour estimates that today's generation will have had 12 to 14 jobs by the age of 40. Lots of changes, lots of new situations people haven't been prepared to think about. And I think we are actually doing that quite well. That's the kind of view, the kind of impact that PISA seeks to do and they are doing that quite well.

There is another class of competences, which we call *non-routine interactive skills*. This is about the capacity of people to relate well to other people – not people you know, not people who are similar to you (that's easy!), but people in heterogeneous groups. Basically you can see this section is steep rising in the market: How do you manage to solve conflicts? How do you collaborate? How do you work in teams? These are competences with very steep increase in the demand. They are also very hard to assess, and assessments like PISA do very poorly on them.

I think it's true also for national examination systems, national tests. I'm sure your curriculum talks like every other curriculum about the importance of interpersonal competences, but at the end of the day you give students an individual test, an individual examination to test their individual skills, so reality and ambition are often far away. With all this, what I really want to show is the fact that in the course of ten, fifteen years the world will have fundamentally changed in terms of the demands for skills, the kind of competences that are valued. This is what comparisons can tell us. They can tell us how the world is changing in terms of quantity (we've seen that before), but they also can tell us how the world is changing in terms of quality, in terms of the nature of skills that matter.

We can translate this into the following questions: How do we stand on this kind of comparisons? We have already seen the kind of competences that matter, so, how do people come out of this?

This is basically where PISA comes in, and this map just shows the kind of countries that are currently comparing.

In China we have just covered about half of the country, India is just beginning to use PISA assessments. With this we get a quite good picture of the industrialised world: those countries, like Portugal, that have made a contribution in the world economy, the OECD countries, South American countries, Asian countries. We can actually say a lot about how well they do in terms of the kind of knowledge and skills that school systems provide for young people, people at the age of 15.

And, of course, the first thing people do is try and look at results. That takes me to the second part: *How do we benchmark quality*?

Here is a very, very important question. This is one way where assessment can have a deep impact on a national debate. Not just through the results but through the way in which you define quality.

Let's take something that is familiar to everyone! Mathematics.

You can of course think about Mathematics as a world of formulas, of equations, of abstract thinking, but there's also the real world and the connections to be built between them. How do you ensure that Mathematics doesn't remain just a world of formulas, but a language through which students learn to structure, to understand the real world, and to mould the real?

In the PISA context, we start with a real situation, we confront students with something that they may have seen or may not have seen and then expect them to build a model of the reality. It's the first step.

Many students – that's a particular challenge for Portuguese students – many students in Portugal struggle with this. They know about the mathematical world but when they see a real problem they do not know why this could be a mathematical problem. They have difficulties in translating the real world into the world of Mathematics. If you struggle with those kinds of skills, what's the use of Mathematics? What value does an algebra equation

have if you can't use it as a language to understand the real world, to structure the real world, to mould the real world? Portuguese pupils do well in "The Mathematical world" but using Mathematics is quite another story.

In PISA we don't tell students whether this is an algebra problem, a geometric problem, even whether this is a mathematical problem. We expect them to be able to use what they have learned, to apply what they have learned in that kind of knowledge situation. Having a mathematical number, a mathematical result is not the end of the story. What does that result mean in the context of the real problem?

This is one view of Mathematics but there are others. You can create other models. That is another way in which educational assessment can have an impact on the debate. Assessment can formulate the paradigms that you take in different subjects, and those paradigms influence the results. They influence the results very, very strongly.

Portugal is not some sort of an extreme case. We can take the case of the Russian Federation. If we had only tested within "The Mathematical world" – traditional school tests –, the Russian Federation would come out number one or number two. When you look at international tests that have done so, this comes out very strongly. When you compare them with PISA they are number 23 among 40 countries.

The way in which you look at the world makes a difference: it shapes the outcomes that you see. It's not that PISA tells you how good or bad an education system is, but it tells you how good an education system is in a particular view of the world. So that's one of the important ways in which results can shape the impact.

You can do the same thing in Science. In Science, as well, we don't look just at how well students know something of Physics, Chemistry, Biology. We want to see to what extent they can identify Science in a real situation, whether they can scientifically explain what they see, whether they can use evidence. And of course this has to do with knowledge.

I don't deny knowledge is a very important part, but knowledge also has several components. There is of course the *knowledge of* Science, of Physics, Chemistry, and Biology. What about the *knowledge about* Science, Science as a discipline? What is evidence? What is a hypothesis? What is to draw conclusions? What is to make judgements? What distinguishes a question that is scientifically investigable from one that is just a popular belief? That's also Science, so knowledge about Science is an equally important part; it shapes competences in the same ways.

To what extent attitudes, interests, support, taking responsibility should be part of Science is a very debated question. For some countries school is about knowledge, it is not about influencing young people. Other people say 'well, if you look at environmental questions today, this is not just about knowledge, this is about people taking responsibility, people using Science to shape their lives, people making trades'. And of course whatever you do, you do it in a specific context. You cannot ignore this as well. Take the languages, one other language, one other view on the world.

We could talk about many of those kinds of paradigms here, but what I want to analyse is how assessments can actually shape the way we look at the world, and therefore they have a deep impact. And in some countries this impact is more important than the impact of the results. I just came back from Japan, talking at the Japanese Parliament and they are not interested in where they come out. They know they are doing well by international standards in terms of the ranking, but they are very interested in how other countries model the world; how they define the skills they need to succeed, how they measure them, how they teach them. This is also very important and of course we shouldn't neglect it.

And then there comes the results. That's what everybody tries to look at and all the media like to map countries.

Look at this slide on the average performance in science! Red are the countries bellow the average, the average is yellow, and green are the countries that are doing well by OECD standards.

Here you can see how Finland comes out on top of everyone else, but you also see some countries in Asia doing well. Canada in North America coming out with good results. We can basically see there is a range of differences here, and here you have the position of Portugal. You can look at this in two ways, you can always ask yourself, 'is the glass halffull?' or 'is the glass half-empty?' 'Why is Portugal where it is?'

First of all, Portugal is by OECD standards one of the most rapidly improving education systems – you've seen that for higher education but I could have shown you the same chart for school education. School today is very different from what it was twenty years ago. However, the kinds of strengths of Portuguese students are perhaps not ideally valued by the PISA systems for the reason I've shown you before. This is a challenge for Portuguese students: they are often good at the reproduction of subject matter competence, but they have greater difficulties in using their knowledge. Once again, that's what assessments can do, that's another form of impact. They can show you where they are, where you are in terms of quality of outcome. They can also show you where you can be, and that's the important dimension of comparisons.

When you see yourself in a national perspective you do not know how good things can be in terms of quality, equity, efficiency. When you look at yourself in the mirror of what everybody else achieves you can actually look at this in a different perspective. And I'd like to show you this also in a bit more nuanced way.

Let us start with France!

French students are average performers overall. They are sort of fragile, sort of so-so, not particularly good, not particularly bad.

Students can *identify scientific issues*, they figure out what science is, it's positive. They are very bad when it comes to *explaining what they see scientifically*. Their knowledge base is not really well developed. But they can actually *use scientific evidence*, they have very good strengths here. So basically they can think scientifically.

French students have quite good knowledge about Science, about the fundamental paradigms of Science, but they have poor knowledge of Geography, poor knowledge of Biology and Physics. So what you see here is that an average score tells you something, but the real story starts when you look at below the relative strengths and weaknesses of countries.

When you look at this pattern you can ask yourselves, "Does every country look like that? Is every country similar?"

Let's take a typical Eastern European country, Russia or Czech Republic for instance.

If I contrast France with the Czech Republic, you can see it's almost exactly the opposite. On average, the French and the Czech Republic do quite similar. But the kind of relative strengths and weaknesses of their education system is very different. This tells you that what you have in your curriculum really maps the outcome, how you teach, what you teach. It does make a difference.

Maybe not in average, but in terms of the relative strengths and weaknesses, if you compare France and Czech Republic, I think I know where I would hire an engineer and where I would not, in terms of the relative strengths and weaknesses of the education system. Czech students are not very strong about understanding what Science is, but they do have a very solid quality of knowledge.

Where would you expect Portugal? There you are, a more mixed picture, but not as extreme.

International assessment, what is the impact? They tell you not just how good things are but they tell you what the relative strengths and weaknesses are.

How do you evaluate your national curriculum? Well, one way is to look at how systems come out across different student groups.

I will now introduce another dimension, which your minister spoke about as well and which is very, very important to the OECD. That's the dimension of equity.