

Avaliações internacionais e desempenho dos
alunos portugueses



Seminários e
Colóquios

Conselho Nacional de
Educação

**Avaliações internacionais
e desempenho dos alunos
portugueses**

Avaliações internacionais e desempenho
dos alunos portugueses

Avaliações internacionais e desempenho dos alunos portugueses



Conselho Nacional de
Educação

As opiniões expressas nesta publicação são da responsabilidade dos autores e não refletem necessariamente a opinião ou orientação do Conselho Nacional de Educação.

Título: Avaliações internacionais e desempenho dos alunos portugueses
[Textos do Seminário realizado no CNE a 25 de março de 2013]

Autor/Editor: Conselho Nacional de Educação

Direção: Ana Maria Bettencourt (Presidente do Conselho Nacional de Educação)

Coordenação: Manuel Miguéns (Secretário-Geral do Conselho Nacional de Educação)

Coleção: Seminários e Colóquios

Organização e apoio à edição: Carmo Gregório e Teresa Catarina Ribeiro

Composição e montagem: Paula Antunes

Edição Eletrónica: julho de 2013

ISBN: 978-972-8360-77-1

© CNE – Conselho Nacional de Educação

Rua Florbela Espanca – 1700-195 Lisboa

Telefone: 217 935 245 Fax: 217 979 093

Endereço eletrónico: cnedu@cne.min-edu.pt

Sítio: www.cnedu.pt

SUMÁRIO

ABERTURA	7
Ana Maria Bettencourt	9
Bártolo Paiva Campos	15
Avaliações internacionais e a melhoria da qualidade e da equidade dos sistemas educativos	17
Hans Wagemaker	19
Os resultados dos alunos portugueses no TIMMS em Matemática e Ciências e as suas implicações para o ensino, para a formação de professores e para o sistema educativo	53
João Pedro da Ponte	55
Pedro Reis	61
Os resultados dos alunos portugueses no PIRLS em Leitura e as suas implicações para o ensino, para a formação de professores e para o sistema educativo	69
Inês Sim-Sim	71
Um outro olhar sobre o TIMMS 2011	91
Jorge Miguel Marques da Silva	93
Filipe Oliveira	95
Pedro Fevereiro	103

SEMINÁRIO

Os testes internacionais, o sistema de avaliação externa e a promoção das aprendizagens	109
Helder Sousa	111
Ana Sousa Ferreira	119
José Augusto Pacheco	135
Leonor Santos	151
Rui Santos	163
Síntese e conclusões	169
Bártolo Paiva Campos	171

ABERTURA

SEMINÁRIO

Agradeço a todos a disponibilidade para, mais uma vez, refletirmos em conjunto sobre o impacto das avaliações internacionais na melhoria da qualidade e da equidade dos sistemas educativos.

Agradeço a todos os ilustres participantes nos painéis a disponibilidade para esta colaboração.

Um agradecimento especial ao Senhor Hans Wagemaker por apresentar aqui o trabalho da IEA (*International Association for the Evaluation of Educational Achievement*) sobre a avaliação internacional dos resultados de aprendizagem e, em particular, dos alunos portugueses do 4º ano de escolaridade, nos domínios da Literacia (*PIRLS Progress in International Reading Literacy Study*), da Matemática e das Ciências (*TIMSS – Trends in International Mathematics and Science Study*).

O Conselho Nacional de Educação é um órgão independente que tem por missão o acompanhamento das políticas educativas. Desde 2010 que, para além de pareceres e recomendações, e usando do direito de iniciativa própria, publica o Estado da Educação. Nestes relatórios, procuramos retratar o desenvolvimento do sistema, desde a educação pré-escolar até ao ensino superior, selecionando os indicadores mais pertinentes para a análise da qualificação dos portugueses, da rede de oferta e acesso à educação em todos os níveis, dos resultados de aprendizagem, do abandono precoce e da qualidade dos percursos educativos, nomeadamente apreciando a evolução do desvio etário nos vários ciclos do ensino básico e no ensino secundário. Os relatórios incluem igualmente recomendações do Conselho.

Da análise realizada ressalta uma evolução significativa, nos últimos quarenta anos, em matéria de democratização do acesso. A última década registou progressos assinaláveis ao nível da qualidade e da equidade do sistema.

¹ Presidente do Conselho Nacional de Educação

No Estado da Educação 2012 incluímos uma síntese dos dados da última edição do TIMSS e do PIRLS relativos aos alunos portugueses. Foi com o maior interesse que tivemos conhecimento dos nossos resultados obtidos em testes internacionais, realizados por uma instituição com a competência da IEA. Depois de várias provas internacionais, iniciadas na década de noventa do século passado, em que os nossos resultados eram negativos, é confirmada a tendência para uma evolução positiva, iniciada com os últimos testes PISA.

Considero ser nossa missão contribuir para o debate sobre esta evolução, os fatores que a podem explicar e sobre a sua sustentabilidade.

A importância da melhoria das aprendizagens e competências alcançadas nos primeiros anos de escolaridade tem sido objeto de grande preocupação por parte do CNE. Temos dedicado a maior atenção ao problema do desvio etário e à acumulação de dificuldades pelas nossas crianças desde os primeiros anos de escolaridade. Temos igualmente apelado à necessidade de uma intervenção ao primeiro sinal de dificuldade e ao reforço dos apoios como estratégia de prevenção do insucesso escolar e de melhoria das aprendizagens e dos percursos escolares.

A iniciativa que hoje levamos a cabo é mais um testemunho do interesse do CNE pela problemática da avaliação e mais uma oportunidade para o aprofundamento do debate sobre a qualidade da nossa educação. É, também, mais um momento de reflexão sobre o potencial das avaliações internacionais.

A avaliação é, aliás, uma das dimensões estruturantes da atividade do CNE, sendo inclusivamente um dos objetos de análise da 1ª Comissão Especializada, cujo coordenador, o conselheiro Bártolo Paiva Campos, me acompanha nesta abertura.

Permitam-me que, neste início dos trabalhos, partilhe convosco uma preocupação e dois desafios relacionados com as problemáticas do seminário:

A preocupação que vos queria transmitir diz respeito aos meios para a sustentabilidade da melhoria dos resultados obtidos pelos alunos portugueses.

Neste momento de graves dificuldades, sabemos como o futuro do país depende da prioridade política atribuída ao sector da educação. O CNE tem apelado de forma repetida para a necessidade de, mesmo em tempos de crise, a educação ser considerada como o investimento decisivo para o futuro do país.

Há poucos dias foi publicado um documento da Comissão Europeia sobre *O Impacto da Crise no Financiamento da Educação na Europa*. Embora algumas das análises aí publicadas tenham de ser lidas com algum cuidado, por se basearem em informação recente a necessitar, em alguns casos, de consolidação, os dados divulgados apontam para uma situação preocupante em Portugal, em comparação com outros países da União Europeia. Estamos no grupo dos países que mais reduziu o financiamento da educação desde 2010, o que se deve a fatores como a diminuição do número de professores, o aumento do número de alunos por turma e o agrupamento de escolas dos ensinos básico e secundário em unidades orgânicas de maior dimensão. E, se a extensão do orçamento não é por si só sinónimo de aumento da qualidade, na maioria dos países em que existe qualidade educativa estes dois fatores estão associados.

Não podendo descurar a necessidade de prosseguir esforços para melhorar a eficiência do sistema e a rentabilidade dos meios existentes, é com preocupação que encaramos a possibilidade de virmos a diminuir, ainda mais, a fatia do Orçamento do Estado dedicada à educação. Não podemos esquecer o passado sombrio do nosso país nesta área e os baixos níveis de qualificação que ainda persistem na população portuguesa, apesar dos progressos verificados. A dimensão do nosso atraso, sem paralelo na União Europeia, não permite que se abrande a recuperação em curso.

A tendência de queda mais acentuada do investimento em educação nos países em crise poderá ter como consequência o aprofundamento da clivagem já existente entre grupos de países europeus, em matéria de

desenvolvimento e competitividade, e contribuir ainda mais para a falta de coesão europeia.

Como vos disse atrás, a qualidade e a equidade da nossa educação têm vindo a registar melhorias significativas. São notáveis os progressos verificados no acesso desde a educação pré-escolar até ao ensino superior. A qualidade dos nossos técnicos e cientistas é conhecida em Portugal e além-fronteiras. A sustentabilidade destes progressos é essencial para que Portugal se afirme e vença a grave crise económica em que se encontra.

Passando aos dois desafios que vos queria deixar:

O primeiro desafio diz respeito à necessidade de valorizar a melhoria dos nossos resultados e o papel dos professores. Numa altura em que as nossas escolas e os professores vivem momentos de grande instabilidade e de fraca autoestima é essencial tomarmos consciência de que melhorámos e acreditar que somos capazes de progredir ainda mais, corrigir erros e encontrar as soluções para os problemas.

É sabido como o papel dos professores é o mais decisivo para a qualidade das aprendizagens e como a sua valorização é essencial. A recente democratização do acesso à educação em Portugal relativamente aos outros países da União Europeia, é geradora de situações de grande complexidade. O enquadramento exigente do trabalho dos alunos na escola e o apoio para que superem dificuldades são decisivos. A divulgação de soluções positivas é da maior justiça e pode ser a base para frutuosa intercâmbios entre escolas e professores. O CNE tem procurado conhecer e divulgar a ação de professores e escolas situadas em zonas sensíveis que são capazes de acrescentar valor e romper com expectativas negativas.

É essencial construir conhecimento sobre a ação das escolas e dos professores portugueses e empenhar toda a sociedade na sua valorização. Mas é igualmente necessário criar condições atrativas de trabalho que, por sua vez, possam contribuir para assegurar a permanência no sistema de um corpo docente altamente qualificado e motivado.

Um segundo desafio diz respeito à necessidade de conhecer melhor e debater os resultados obtidos pelas avaliações internacionais. Com práticas e

normas tão enraizadas na sociedade, as avaliações internacionais têm a vantagem de nos fazer descentrar de hábitos e crenças por confronto com outras realidades e outros resultados.

Numa área em que, com frequência, impera o senso comum na definição de propostas para a solução dos problemas educativos, o estudo de situações muito diversas é enriquecedor. Temas como, por exemplo, a autonomia das escolas, a sua responsabilização pelos resultados, a existência de metas claras, o agrupamento dos alunos e as vias para uma maior equidade têm sido objeto de estudos internacionais, podendo constituir um apoio fundamental à decisão política.

Das avaliações internacionais podemos igualmente colher ensinamentos sobre a importância da estabilidade das políticas, outra das recomendações persistentemente defendida pelo CNE.

Neste seminário apostámos em propor a análise das implicações dos testes internacionais na promoção das aprendizagens, na formação de professores e no sistema educativo. Para além do contributo do senhor Hans Wagemaker, contamos com a colaboração de ilustres investigadores de áreas essenciais como as Ciências da Educação e as Ciências Exatas e de responsáveis por estas áreas no Ministério da Educação.

A todos agradeço a disponibilidade e o contributo.

SEMINÁRIO

Depois desta intervenção mais política, vou sugerir algumas questões que podem organizar a nossa reflexão ao longo do dia.

A missão da 1.^a Comissão, que coordeno, é o acompanhamento global das políticas educativas; por isso, logo que saíram os resultados destes estudos internacionais, nos pareceu importante organizar uma reflexão alargada sobre os mesmos.

Como sabemos, a eficácia dos sistemas educativos tem sido analisada à luz de dois grandes grupos de indicadores. Durante um certo tempo, foi privilegiado o primeiro grupo, que é o dos indicadores de escolaridade. Nestes, a preocupação centra-se no acesso e na permanência na escola e na conclusão de um determinado número de anos de escolaridade. E, assim, foram realizados muitos estudos que relacionam o aumento do número médio de anos de escolaridade de uma população ou de um grupo etário com o crescimento do PIB, com os salários e com outros indicadores sociais.

A partir de um certo momento, sobretudo graças a estes estudos comparativos internacionais, facilmente se verificou aquilo que se poderia pensar, mesmo sem os estudos, isto é, que ao mesmo nível de escolaridade em vários países, não correspondia o mesmo nível de aprendizagem obtida. A aprendizagem conseguida começou, então, a ser considerada um melhor indicador da eficácia dos sistemas educativos. É justamente a este nível que se situa o presente seminário, ao nível da comparação das aprendizagens conseguidas, e não do número de anos de escolaridade realizado por uma população ou por um grupo etário.

Para a organização do debate, ousou propor três questões sem impedimento de surgirem outras.

A primeira questão é a de saber se os progressos que estes estudos evidenciam são progressos reais, consolidados, ou se são resultados

¹ Conselho Nacional de Educação

ilusórios, efémeros, devidos, por exemplo, a questões de amostragem ou de pouca dificuldade das provas. Esta primeira questão é se, de facto, houve progresso e onde é que o progresso se verificou. Claro que há sempre maneira de comparar os resultados com uma meta que se deseja; mas há também maneira de avaliar os resultados em função de um ponto de partida, de onde se veio. E conforme a nossa focalização, assim a nossa atitude, o nosso pessimismo ou o nosso otimismo.

A segunda questão que proponho é a de saber, havendo progresso, quais são os fatores do mesmo. É evidente que, em última análise, o progresso, se existe, é devido a um desempenho diferente dos professores e dos alunos. A questão é saber o que é que conduziu a um desempenho diferente dos professores e dos alunos. A identificação de causas prováveis, ou com alta probabilidade de estarem relacionadas com estes resultados, tem importância para a decisão de continuidade das políticas: as que convém continuar, aperfeiçoando, eventualmente, ou as que convém, como agora se diz, descontinuar.

A terceira questão refere-se à sustentabilidade destes resultados. Em função da segunda questão, convém examinar se há condições para continuar a contar com os prováveis fatores de progresso e garantir, assim, não só o não-recuo, como ainda a melhoria no nível das aprendizagens conseguidas pelos nossos alunos.

E num clima de crise financeira, a questão que se pode colocar é se se pode aumentar a eficiência, isto é, se com menos meios se conseguem os mesmos ou melhores resultados. Porque, com menos meios conseguir piores resultados, não é habilidade nenhuma. Bom trabalho.

Avaliações internacionais e desempenho
dos alunos portugueses

Avaliações internacionais e a melhoria da qualidade
e da equidade dos sistemas educativos

Presidente da Mesa - Maria Arminda Bragança

SEMINÁRIO

I understand you've already had an opportunity to listen to what PISA does. I think it's important to understand that what we do is very different. And those differences are important in terms of understanding what we can do and how we can work with the data.

This morning I will talk very briefly about some of the outcomes related to TIMSS and PIRLS for Portugal. Some of you already had the opportunity to see these results, so I would like to use some of the information from the data from Portugal, and talk a little bit about the information in general that exists, and also consider some questions related to equity, particularly in Portugal's context.

In the time that we have available, I can't possibly cover all the issues that arise when we go through the data. These studies take 4 or 5 years respectively to conduct, and the more comprehensive analysis that countries undertake usually takes the 4 or 5 intervening years between cycles.

What we've seen over the last years is a change of the demand for the types of information we used to make judgments about the quality of our educational systems. And the shift, the transition has been from a focus on the inputs out of the equation to what happens when children are actually in school.

And it's not just what happens, but it's also the nature of what happens, that is at the heart of the issue of quality in education. And quality is defined in a variety of ways, not always in terms of performance but it also takes into consideration questions of equity and efficiency.

The other big driver in terms of this transformation of focus, I think, as we are all aware, is globalization.

Students that are being produced in Portugal are not competing just for jobs and employment working in companies locally, but they are expected to be

¹ Diretor Executivo da International Association for the Evaluation of Educational Achievement (IEA)

able to compete internationally and help develop the local economy to be competitive internationally. So this is a different demand.

One of the things that often arise when we talk about issues of educational outcomes is, why we should bother with international comparisons? This it is argued is because many countries already work with national assessment systems that address a number of the issues and concerns related to quality and equity.

The following quotation however, from some of the early scholars from IEA captures the essence of what we're trying to achieve.

“If custom and law define what is educationally allowable within a nation, the educational systems beyond one's national boundaries suggest what is educationally possible”(Foshay 1962:p. 2).

I think it is about trying to find and define the limits of possibilities of schooling. The opportunity to learn from other nations or sub-national jurisdictions about what is possible in education and how we might achieve either the same results or hopefully better results for our students is what drives international assessments.

IEA's work spans more than 50 years now. While we're here to talk about TIMSS and PIRLS today, we have a number of other activities that are going and one of the things that you'll see in relation to TIMSS is our TIMSS advanced assessment, which focuses on advanced Mathematics and Physics.

We are also developing what we describe as TIMSS numeracy, which is intended for those countries which children are earlier in a stage of acquire the foundational skills in Mathematics. It's not just developing countries, it's also countries that for one reason or another may be actually rich in terms of GDP but have systems that have languished.

In the area of reading we also have a parallel assessment that we call pre-PIRLS. Pre-PIRLS has been around for a few years now and we administered it successfully to a number of countries during the last cycle.

We currently have a Citizenship study starting this year. We've just completed a study in the preparation of Mathematics teachers. This year

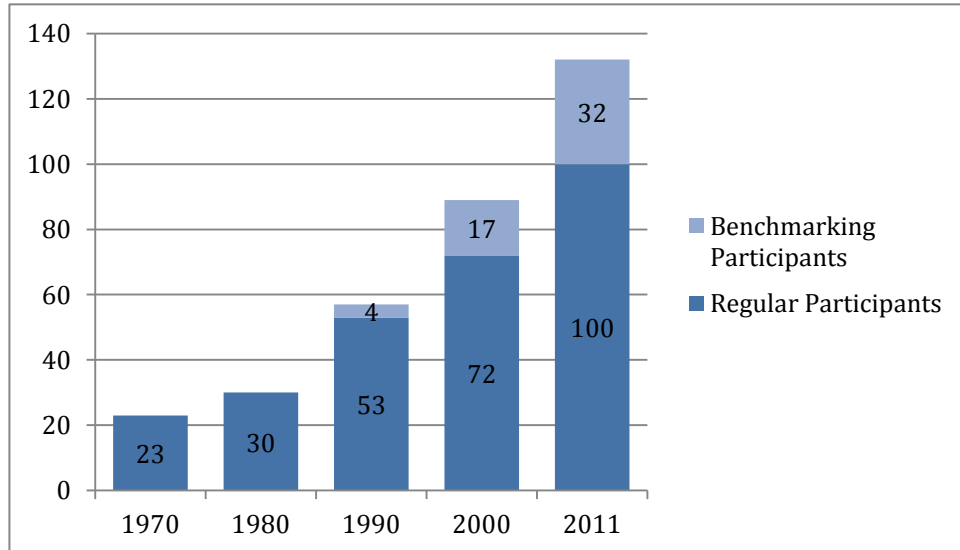
again we will be starting a new project in early childhood and we also have ongoing a study which is halfway through which has as its focus digital literacy.

A characteristic of these developments is that there's much greater diversity in terms of the countries that we work with. In the early years of IEA we were working largely with the countries of what is now the European Union. There were some exceptions, but it was a rather homogenous group.

We are now working with countries that have huge linguistic diversity, huge cultural diversity, and huge, global, economic diversity, which place new demands on the sort of work that we are trying to do.

To give you an idea of what has happened in the last four decades, in the 1970's when we talked about international assessment and this pre-dates the work of PISA, there were about 23 countries.

Cumulative growth in unique IEA participants over the decades (beginning of each decade)



You can see not only the growth in terms of unique countries that have participated but also the phenomenon that begins in the 1990's with a number of what we called sub-national or benchmark participants. They may be regions in Spain, they are states of the United States, even countries like the Emirates of the UAE. There are a couple of the Emirates participating independently, as well as a number of Canadian provinces which participated independently. And we are seeing that these autonomous, administratively autonomous, regions, also want to understand the development of their local systems in a global context.

Given the context of diversity and the breath of interests and our attempt to try and understand not only the outcomes of schooling but also the antecedents of learning, what approach does IEA take in its assessments?

IEA takes a grade-based approach, as I mentioned this is different from what PISA does. And the reason we do that is that, our argument is that children learn by going to school, not by getting older. Perhaps it is a bit glib but the reality is that in many countries that we work with, 15 year-olds are distributed across 4 or 5 grades in some cases.

The modal grade differences among some PISA countries are up to 2 years. So most of the children in New Zealand for example, have had 2 years more schooling than the 15 year-olds in Finland. So when New Zealand says, "we did pretty well in PISA", I remind them, "yes, but our kids have had 2 years more schooling (on average), they should be doing OK."

So, in that context and that age distribution is an issue. For us the challenge is, because we are curriculum focused, to make sure that we capture the curriculum emphasis in each country.

The other feature of the TIMSS design for example is that we offer a trend assessment. What this allows you to do is at successive participation intervals to look at a same cohort over time. For example 4 years later, the grade 4 students in 1995 are the grade 8 students in 1999, and so on. This permits an examination of comparisons among countries but also within the same country.

The target grades for TMSS and PIRLS are not arbitrarily chosen. They represent important transition points in the educational system: end of primary, beginning of lower secondary, and end of secondary, in the case of TIMSS. These grades also represent theoretically important developmental stages. In the case of Reading (PIRLS), in the developed world, grade 4 represents a point at which in most countries, reading instruction, that is, formal Reading instruction finishes.

While there are often remedial reading programs that go into the later grades and additional development often occurs through language programs, the reading process, the acquisition of those fundamental skills, usually ends at the end of grade 4.

Underpinning all of this is a curriculum model. We start with the intended curriculum, that is, what's dictated by policy, what is described in terms of curriculum statements. We try to capture this through background questionnaires, first from policy makers, from national research coordinators and to a certain extent from teachers.

We then are interested in the implemented curriculum. What happens in the classroom and how is it that those broad curricular statements are implemented in the classroom? What do teachers actually do?

Finally, at the last level, we are interested in the achieved curriculum, which is what can students demonstrate in terms of their knowledge attitudes, skills and dispositions towards learning. This is captured not only by the content of assessments but also by the background questionnaires.

Of course, one of the key players in this, in addition to the teachers, and in addition to all the resources that the State provides, is parents. We are interested in what parents do prior to children starting school, the extent to which they are involved in the acquisition of early skills in the learning process.

So we augment the content of the test that we develop with a comprehensive set of questionnaires which goes to, as I said, the research coordinators and the policy makers, the principals of schools, the teachers in the classroom,

the children in the selected class and the parents of the children that we work with.

Schools are selected on a random basis, and the sample intended to cover (represent) the whole system. If you start to exclude parts of your system we begin to annotate and describe what's missing, so that we understand exactly who is it we are talking about. And I think the last thing on this slide is the issue of trend design.

PIRLS 2011 - 19th

Country	Average Scale Score
Hong Kong SAR	571
Russian Federation	568
Finland	568
Singapore	567
Northern Ireland	558
United States	556
Denmark	554
Croatia	553
Chinese Taipei	553
Ireland	552
England	552
Canada	548
Netherlands	546
Czech Republic	545
Sweden	542
Italy	541
Germany	541
Israel	541
Portugal	541
Hungary	539
Slovak Republic	535
Bulgaria	532
New Zealand	531
Slovenia	530
Austria	529
Lithuania	528
Australia	527
Poland	526
France	520
Spain	513
Norway	507
Belgium (French)	506
Romania	502
PIRLS Scale Centerpoint	500

The assessments we offer were designed from the beginning to measure trend. In the case of TIMSS 2011 the design allows us to be able to link back to the 1995 assessment. This design requires a discipline in terms of retaining certain parts of the assessment, keeping it secure so that it can be used again. But also requires planning ahead in terms of what the rotation of booklets is, and the consequent development of new stimulus materials.

This design allows us to measure change and to modify or change our assessment with changing curricular expectations for the learning in these key subject areas.

The demands of schooling in terms of the knowledge and skills that students are expected to learn in Mathematics, Science and Reading are as you will be aware, are not stable and evolve over time.

At the end of this development and assessment process, we report not only the average scores that everybody talks about, but we also talk about the sub-domains of Mathematics and Reading. We describe performance in terms of achievement benchmarks, what students can actually do in terms of demonstrating their skills. And we look at the relationship between some of the antecedence of learning and the learning outcomes.

We produce an encyclopedia that describes the broader political and social and educational context in which this learning takes place. These are extremely useful documents in terms of interpreting not only your results, but also understanding the results of other countries.

The thing that always struck me, I think in 2003 in the Reading assessment, when I looked at the performance of some of the lower performing countries, was to see the limited aspirations stated in the curriculum about what children should learn by way of Reading. If you have low expectations, if you have low aspirations, it's pretty certain that that is what you are going to get. And so, it (the encyclopedia) provides another insight into the potential for the reform process.

We also produce thematic reports from time to time, address particular issues of interest. Gender is one of them. We have a new one coming up, on

high performing students, and I'll show you the graphs a bit later on about what this might look like.

Now, for TIMSS 2011, there was an opportunity to look at the same students for Math's, Science and Reading. So the question of what proportion of those students in participating countries, were high achieving for all three subjects? Similarly, what proportion of students within each country had high scores, and achieved at the high benchmarks in one subject. I would also refer you to the technical reports for the study, so that you can really understand the quality of the data, not only for your country but for every other country and make appropriate judgments about how that information might be used.

International databases are produced together with tools to allow you to analyze this data correctly. The challenge of analyzing, working with this data is non-trivial. That is, these studies are based on complex samples, and normal statistical software does not handle these complex sampling designs. These studies require analysts to be very aware of the sampling design and the weighting strategy that's applied to the data. The software we provide allows you to do this without fear of making an error.

Of course at the end of this process, there is always a portion of the assessment that is made available as released materials. The idea here is that, for example, in Reading there's a reader that we use as an example of authentic reading text. Many countries use this as a classroom resource later on just for augmenting their classroom libraries.

Examples of the questions are used in many countries as instructional tools to give them some guidance on how to start thinking differently about their instructional programs. There is considerable evidence that these resources are penetrating teacher's training providers.

So, to understand all this activity, I think what we try and convey is that the assessment is not the goal, the test is not the goal. We can develop tests. What we are trying to achieve is to provide countries with the sort of information that allows them to come to understand the functioning of their educational system, and perhaps provide some insights into the way in

which they can improve educational performances. The Goal is ultimately educational reform and improvement.

I think lastly we like to think of these studies (TIMSS and PIRLS) as an opportunity to develop local capacity. We argue that TIMSS and PIRLS and the other assessments need to be understood as both process and product. They represent a tremendous opportunity for people to learn what we think of as the cutting edge of assessment technology from the development of frameworks to item writing, (where we run workshops), to data management and analysis.

In 2011, we had a huge amount of press coverage. I guess most of that focused on the international averages. This is part of the kit that we provided for countries just to give them a quick summary we provided the media. I have it demonstrated at the bottom of the table where the countries bellow the scale center point, not because it is not interesting, it's just that I want to go on from here.

TIMSS 2011 G4 Math - 15th

Country	Average Scale Score
Singapore	606
Korea, Rep. of	605
Hong Kong SAR	602
Chinese Taipei	591
Japan	585
Northern Ireland	562
Belgium (Flemish)	549
Finland	545
England	542
Russian Federation	542
United States	541
Netherlands	540
Denmark	537
Lithuania	534
Portugal	532
Germany	528
Ireland	527
Serbia	516
Australia	516
Hungary	515
Slovenia	513
Czech Republic	511
Austria	508
Italy	508
Slovak Republic 27	507
Sweden	504
Kazakhstan	501
TIMSS Scale Centerpoint	500

TIMSS 2011 Science - 19th

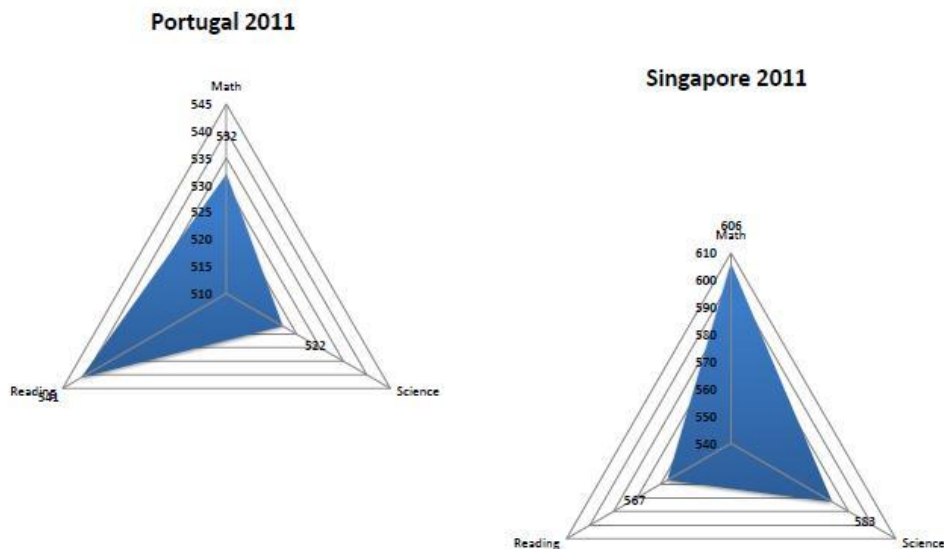
Country	Average Scale Score
Korea, Rep. of	587
Singapore	583
Finland	570
Japan	559
Russian Federation	552
Chinese Taipei	552
United States	544
Czech Republic	536
Hong Kong SAR	535
Hungary	534
Sweden	533
Slovak Republic	532
Austria	532
Netherlands	531
England	529
Denmark	528
Germany	528
Italy	524
Portugal	522
Slovenia	520
Northern Ireland	517
Ireland	516
Croatia	516
Australia	516
Serbia	516
Lithuania	515
Belgium (Flemish)	509
Romania	505
Spain	505
Poland	505
TIMSS Scale Centerpoint	500
New Zealand	497

I think what Portugal should do is celebrate its performance. If we look at PIRLS, particularly if you discount some of the Asian countries, in terms of Europe, you (Portugal) have a great performance. An average score of 542 is a significant level of performance.

But what you see at the very top, when you look at Korea and Singapore at 600, and in the case of Science, Korea and Singapore nearly at 600, Finland at 570, - these countries are all reaching this level of achievement in the same period of schooling, - 4 years of schooling. So there are always possibilities for improvement.

But as I say, this performance (for Portugal) is a very significant level of achievement.

Another way of looking at this is to look at the relative strengths and weaknesses for Portugal, and again I'm comparing it to Singapore here, which was at the top, generally at the top-performing countries. Reading is obscured here, but again, a very strong performance.



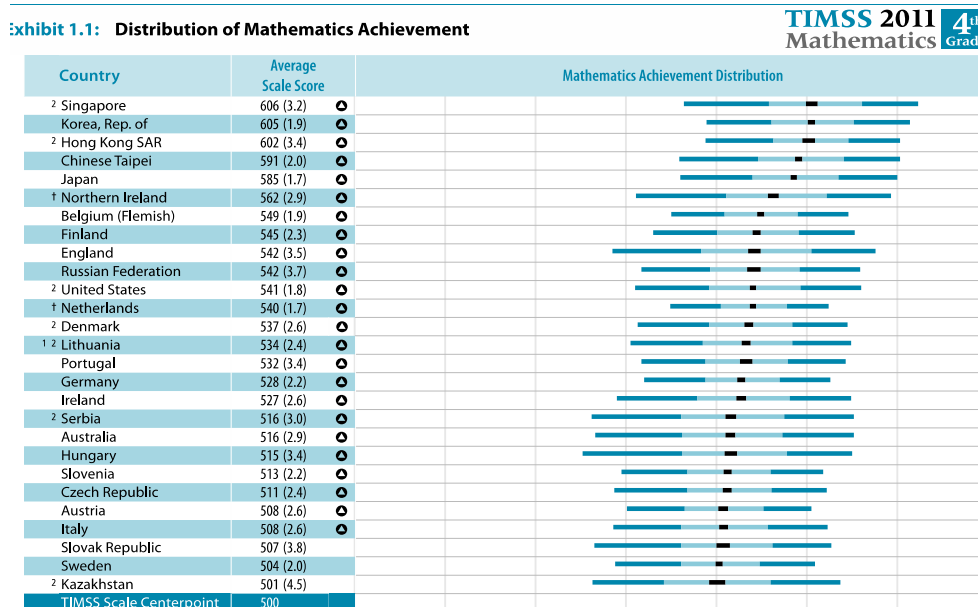
Singapore, on the other hand stands out for Maths and Science, but if you think that their assessment is in English and they're still achieving at 575, again, it's an extraordinary performance. Again, there are possibilities for improvement. New Zealand doesn't get a mention in this part.

When I was asked to come to the meeting I was asked first to talk a little bit about, addressing issues of equity and what could we learn from the data that might address these issues It's interesting when we are working with the artifacts, in fact, when we are looking at this data, when we are working with the artifacts of a policy, we are looking at how countries dictate how schools should be resourced and how this might impact learning outcomes.

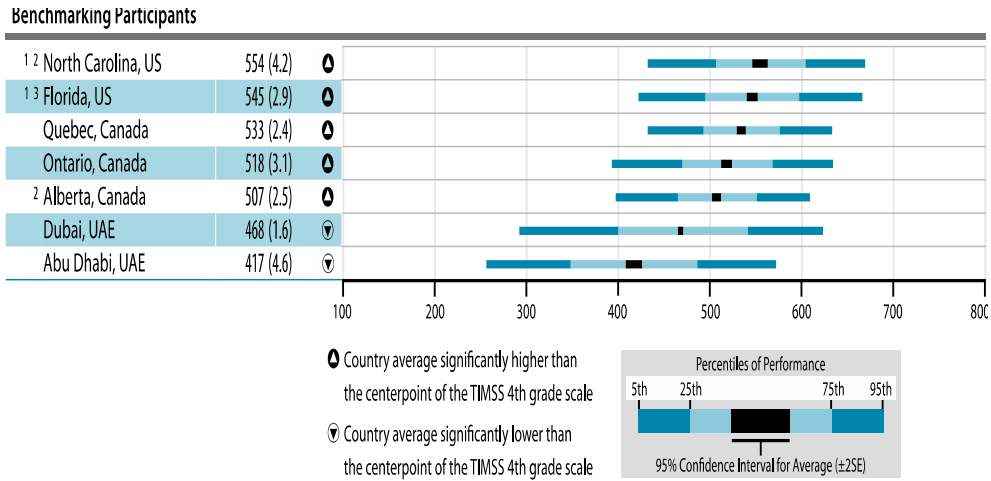
We're looking at the outcome of policies related to teaching (instructional programs) and resourcing. How do we know when we achieve equity? It appears it is often in terms of differences between groups and subgroups, this type of thing.

The first opportunity we have to begin to look at issues of equity, is in the first table, the first exhibit on our international report. So we never report those scores like I produced just a little while ago in our international reports. We begin to produce them like this, and you'll see on the right hand side of the table the range, and I'll blow this up a little bit.

I'll just take it from the bottom so you can see what this represents. It represents the spread of scores within a country. The wider they spread, the greater I guess you could say the inequity there is. And if you look at Abu Dhabi it's wide. If you look at countries like New Zealand, it's very wide.



TIMSS 2011 Grade 4



Other countries are much more effective in terms of educating the total cohort to a certain standard. I'll take that to the next step. Before I go into that, the other way, as I mentioned, we look at the sub-domains. I don't know if you can read all of these from the back of the room. But the idea here is that we look at where the strengths and weaknesses lie within the system in terms of the Maths content.

Portugal in terms of the strengths and weaknesses I guess is slightly less strong in Number, (and these are strategically significant differences,) but stronger in Geometry and measurement and data display.

Achievement in Mathematics Content Domains

TIMSS 2011
Mathematics 4th Grade

Country	Overall Mathematics Average Scale Score	Number		Geometric Shapes and Measures		Data Display	
		Average Scale Score	Difference from Overall Mathematics Score	Average Scale Score	Difference from Overall Mathematics Score	Average Scale Score	Difference from Overall Mathematics Score
² Singapore	606 (3.2)	619 (3.4)	13 (0.8) ●	589 (3.6)	-17 (1.5) ☹	588 (3.4)	-18 (1.7) ☹
Korea, Rep. of	605 (1.9)	606 (2.0)	1 (1.6) ●	607 (1.7)	2 (1.4) ●	603 (1.9)	-2 (2.0) ☹
² Hong Kong SAR	602 (3.4)	604 (3.3)	3 (1.0) ●	605 (3.4)	3 (0.9) ●	593 (3.6)	-8 (2.1) ☹
Chinese Taipei	591 (2.0)	599 (2.0)	8 (1.2) ●	573 (2.1)	-19 (1.3) ☹	600 (2.6)	9 (1.6) ●
Japan	585 (1.7)	584 (1.6)	-1 (0.9) ☹	589 (2.0)	4 (1.1) ●	590 (2.9)	4 (2.9) ●
[†] Northern Ireland	562 (2.9)	566 (2.9)	4 (1.6) ●	560 (3.3)	-2 (2.1) ☹	555 (3.0)	-8 (1.5) ☹
Belgium (Flemish)	549 (1.9)	552 (2.2)	2 (1.4) ●	552 (2.0)	3 (1.0) ●	536 (3.0)	-13 (2.0) ☹
Finland	545 (2.3)	545 (2.3)	0 (0.9) ●	543 (2.9)	-2 (2.2) ☹	551 (3.5)	5 (3.2) ●
England	542 (3.5)	539 (3.7)	-3 (1.1) ☹	545 (3.9)	3 (1.6) ●	549 (4.6)	7 (2.9) ●
Russian Federation	542 (3.7)	545 (3.3)	3 (1.4) ●	542 (4.3)	-1 (1.5) ☹	533 (4.1)	-9 (2.3) ☹
² United States	541 (1.8)	543 (2.0)	2 (0.9) ●	535 (2.2)	-6 (0.8) ☹	545 (1.8)	4 (1.1) ●
[†] Netherlands	540 (1.7)	543 (1.7)	3 (1.1) ●	524 (2.9)	-16 (2.6) ☹	559 (2.9)	19 (1.7) ●
² Denmark	537 (2.6)	534 (2.4)	-3 (0.9) ☹	548 (3.0)	11 (2.0) ●	532 (3.0)	-5 (1.5) ☹
[†] Lithuania	534 (2.4)	537 (2.4)	4 (1.1) ●	531 (3.0)	-3 (1.9) ☹	526 (3.0)	-7 (2.0) ☹
Portugal	532 (3.4)	522 (3.7)	-10 (1.6) ☹	548 (4.4)	16 (2.2) ●	548 (2.8)	16 (2.0) ●
Germany	528 (2.2)	520 (2.3)	-8 (0.7) ☹	536 (2.6)	8 (1.1) ●	546 (2.8)	18 (1.6) ●
Ireland	527 (2.6)	533 (2.6)	5 (1.4) ●	520 (3.1)	-7 (1.6) ☹	523 (2.8)	-4 (2.0) ☹

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS 2011

We also, as I mentioned report not only the content domains but also the cognitive domains. That is, what can children do with the information that they learn in school, what can they demonstrate in terms of what they know, how can they apply that knowledge and how can they reason with the Mathematical knowledge that they have acquired in the learning process. If you look at Portugal again, actually there it is pretty much the same.

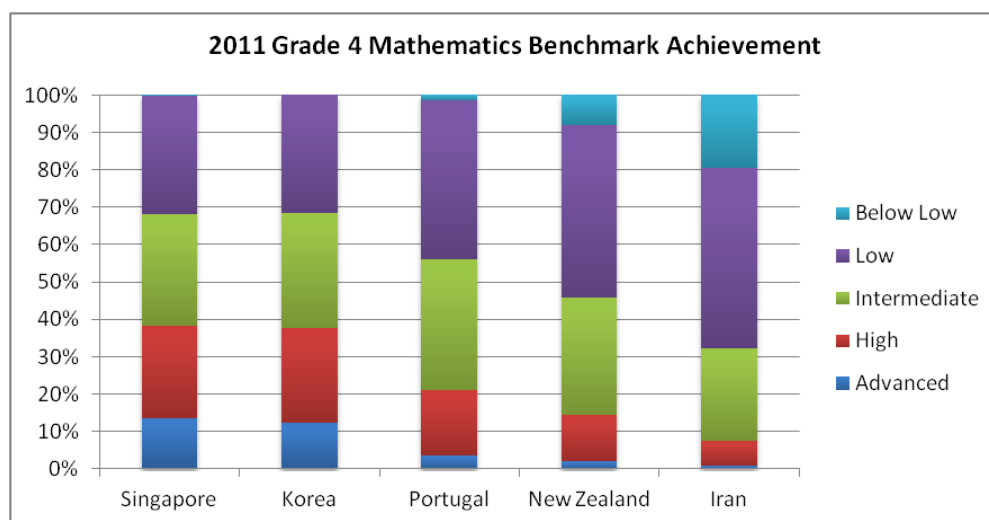
Achievement in Mathematics Cognitive Domains

TIMSS 2011 4th Grade
Mathematics

Country	Overall Mathematics Average Scale Score	Knowing		Applying		Reasoning	
		Average Scale Score	Difference from Overall Mathematics Score	Average Scale Score	Difference from Overall Mathematics Score	Average Scale Score	Difference from Overall Mathematics Score
² Singapore	606 (3.2)	629 (3.5)	23 (1.4) ⬆	602 (3.4)	-4 (1.1) ⬇	588 (3.7)	-18 (1.2) ⬇
Korea, Rep. of	605 (1.9)	614 (2.0)	9 (1.6) ⬆	600 (2.2)	-5 (2.1) ⬇	603 (2.3)	-2 (1.5)
² Hong Kong SAR	602 (3.4)	619 (3.2)	17 (1.2) ⬆	597 (3.2)	-4 (0.8) ⬇	589 (3.4)	-13 (1.4) ⬇
Chinese Taipei	591 (2.0)	599 (2.1)	8 (1.6) ⬆	593 (2.0)	2 (1.0) ⬆	577 (2.5)	-14 (2.0) ⬇
Japan	585 (1.7)	590 (1.7)	5 (1.0) ⬆	579 (1.6)	-6 (1.1) ⬇	592 (2.0)	6 (1.0) ⬆
† Northern Ireland	562 (2.9)	580 (3.4)	17 (1.7) ⬆	565 (2.9)	2 (2.0)	538 (3.3)	-25 (2.1) ⬇
Belgium (Flemish)	549 (1.9)	564 (1.9)	15 (0.9) ⬆	546 (2.2)	-3 (1.1) ⬇	532 (2.7)	-17 (1.6) ⬇
Finland	545 (2.3)	548 (2.6)	2 (1.3)	544 (2.7)	-2 (1.8)	546 (2.2)	0 (1.1)
England	542 (3.5)	552 (4.3)	10 (2.7) ⬆	542 (3.7)	0 (1.5)	531 (3.7)	-11 (2.2) ⬇
Russian Federation	542 (3.7)	541 (3.4)	-1 (1.7)	539 (3.9)	-3 (1.1) ⬇	548 (3.6)	6 (1.5) ⬆
² United States	541 (1.8)	556 (2.1)	15 (0.9) ⬆	539 (2.1)	-2 (0.7) ⬇	525 (2.2)	-15 (0.9) ⬇
† Netherlands	540 (1.7)	537 (2.0)	-3 (1.4)	540 (1.6)	0 (0.9)	543 (2.6)	3 (1.6) ⬆
² Denmark	537 (2.6)	531 (2.6)	-6 (1.4) ⬇	539 (2.9)	2 (1.7)	543 (2.7)	6 (1.4) ⬆
^{1 2} Lithuania	534 (2.4)	525 (3.0)	-9 (1.4) ⬇	540 (2.5)	7 (0.8) ⬆	536 (2.5)	3 (1.4)
Portugal	532 (3.4)	531 (3.5)	-2 (1.4)	534 (3.9)	2 (1.3)	531 (4.1)	-2 (2.0)
Germany	528 (2.2)	524 (2.3)	-4 (1.0) ⬇	528 (2.3)	0 (1.1)	532 (3.0)	4 (2.1) ⬆
Ireland	527 (2.6)	539 (3.1)	12 (1.5) ⬆	529 (2.7)	1 (1.4)	510 (3.1)	-18 (2.2) ⬇

Now, as I mentioned, one way of looking at the question of equity is to look at the performance distribution. In fact the average scores in terms of policy any review, are of very limited value. I mean what they tell you is that you are doing relatively well or relatively poorly but they don't provide you any insight into what should be better, what could be better., A more informative indicator is how do you understand the score distribution., Who lives at the bottom 10% of the distribution, who lives at the top 10%.

Percentage of G4 mathematics population reaching international benchmarks – TIMSS 2011



So one of the exhibits that we produce is the international benchmarks. We try to link score points to descriptions of what children can actually do. And if you look at the difference between this advanced benchmark and the low benchmark, you can see that at the low end this is a very basic level of performance.

One of the important things to note is that while we have four benchmarks, we have 5 zones.

But before I go into that I'll just talk about this. As I mentioned, this is a new report that we are just producing, this is not public yet, but what we are looking at high-performing students in all the countries; and what proportion in each countries of students perform at the advanced benchmark level in all subjects and what proportion perform each in single subjects.

And again, for Portugal this is a great performance. But we can always do better, and this for example, is what happens in Singapore. And one of the

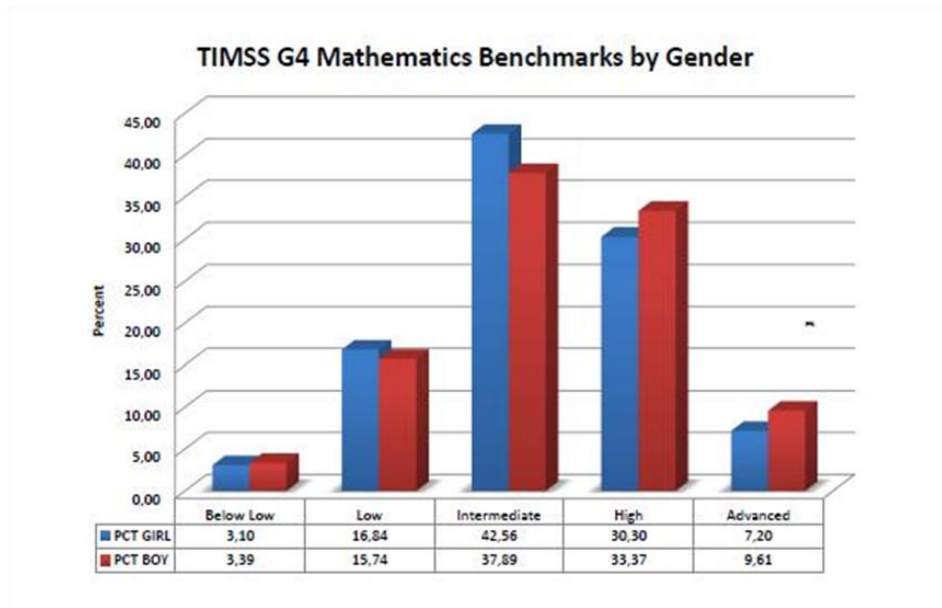
interesting things here for Mathematics, they have almost $\frac{3}{4}$ of their students achieving at a high benchmark. It is an extraordinary performance; it's more than one standard deviation above the scale center point, on average.

So, anyway, getting back to the 5 zones, here is a display of comparison of some countries in terms of the proportion of students in each of the zones, of the benchmark achievement.

Singapore, even with this high achievement (is difficult to see) has about 1% of the students who fall below that low benchmark, below the low benchmark, Korea has nobody in that zone. So this is one of the systems that I say performs well in terms of equity or in terms of efficiency, maybe. The system somehow manages to bring all their children up to this performance mark.

When we look at New Zealand, I think we have 16% of kids who fall below this low benchmark. Now when I look at Portugal it is a small percentage, is not large, but it's in an area, from the policy point of view that, I will say, should be targeted. Who are these people and what are the explanations for their position in the system? If you want to talk about equity, even if you don't want to talk about efficiency, that's an area where you can pay attention.

Given that you have these benchmarks, what can you achieve by trying to begin to understand who these people are?

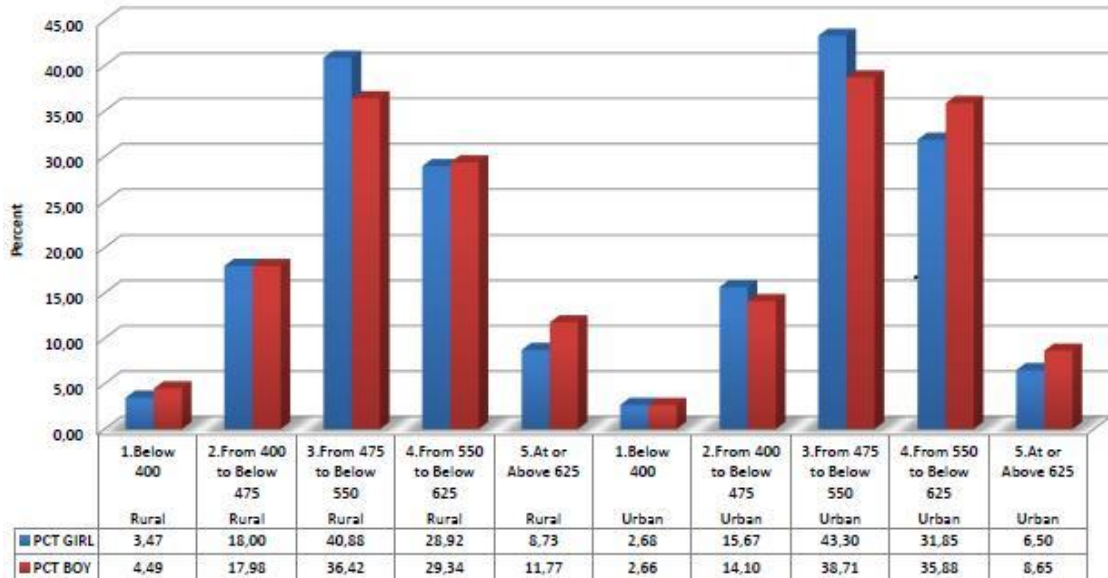


Here I've just broken it down by gender for Mathematics. The below low are more girls than boys; these differences by the way are probably not significant, and I couldn't put in the standard of errors, but for the purposes of illustration you can get the point. You need to begin to decompose the average and as you begin to decompose the benchmarks you can begin to understand the nature of the challenge that you have from a policy point of view.

What you are seeing at grade 4 is that the boys, at that point at least, are achieving at a higher level of performance than the girls. I don't think that stays. And again I'm not sure of what the standard error is, it may not be that significant.

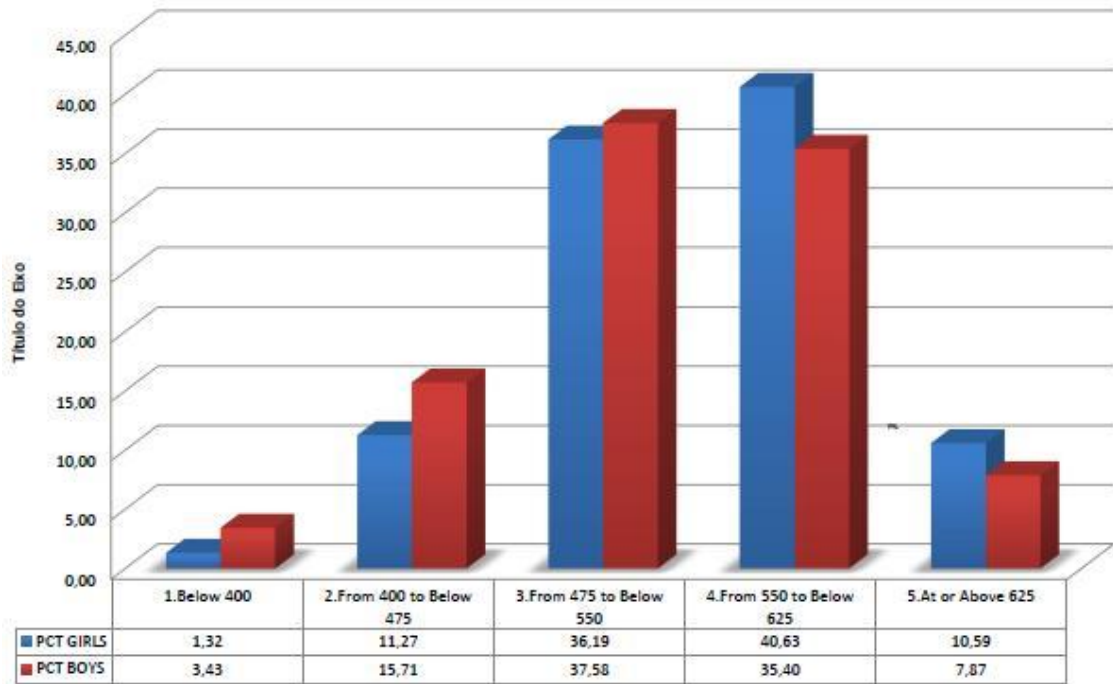
We can take this the next step and say, "OK, given that you have these gender differences, are there any differences between the rural and the urban? Is there something systematic about living in an urban community that stimulates better performances in comparison to a rural community? And of course, I should say from the beginning, these are just relationships. There is probably an underlying reason for all of this, it may be poverty, but again, systematically beginning to decompose these benchmarks begins to give you an understanding about the possible policy remediation.

TIMSS 2011 G 4 Mathematics Rural -Urban Benchmark Performance by Gender



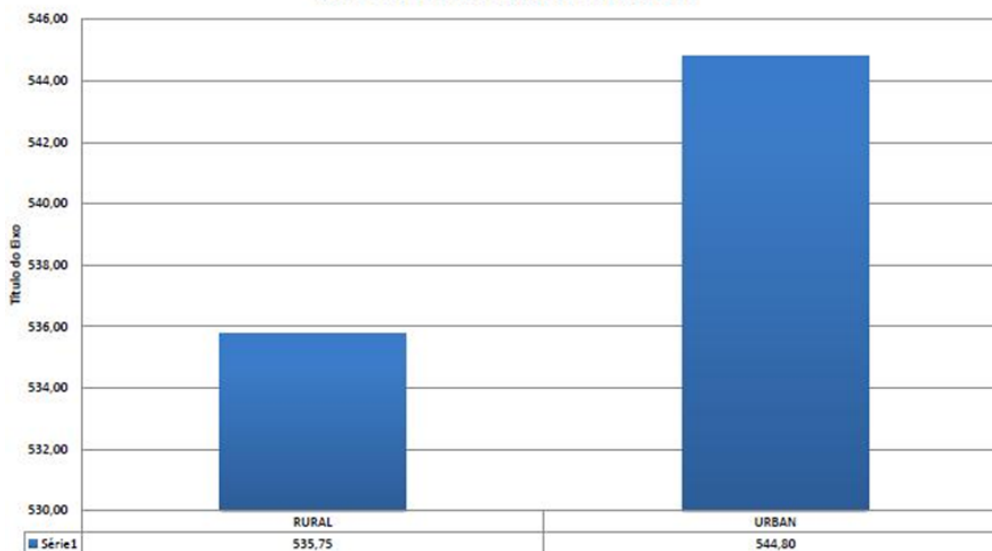
If you look at the difference between, let me see, rural and urban, on the left hand side are the rural, there is a slightly different profile between that and the urban areas. One I was looking at before, at the high end, the differences between the students in the high benchmark in urban areas and in rural areas, is likely to be significant.

PIRLS 2011 Reading Benchmarks by Gender

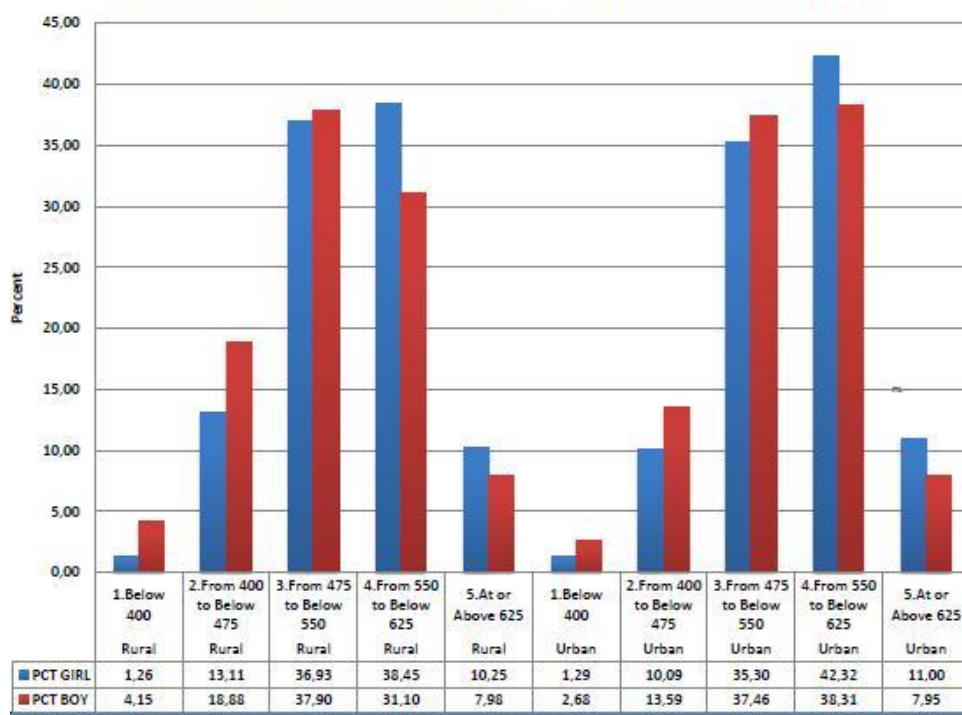


Reading. We have the same challenge here. In this case, the girls are the better readers, systematically. And this is showing up, and I think these differences are significant, statistically significant. And already at grade 4 you've got these differences appearing for boys, here. And from that point on, shifting to the right, there is improvement.

PIRLS 2011 Rural Urban Differences



PIRLS 2011 Benchmark Performance - Gender and Urbanization



Urban/rural differences, just like that, again, I think these are statistically significant differences. And you break it down by gender, and again you get a different profile, where urban students, particularly girls; have an advantage here. Again, it's a different way of understanding, or beginning to understand performance differences in terms of achievement.

Achievement in Mathematics Content Domains by Gender

TIMSS 2011 4th Grade Mathematics

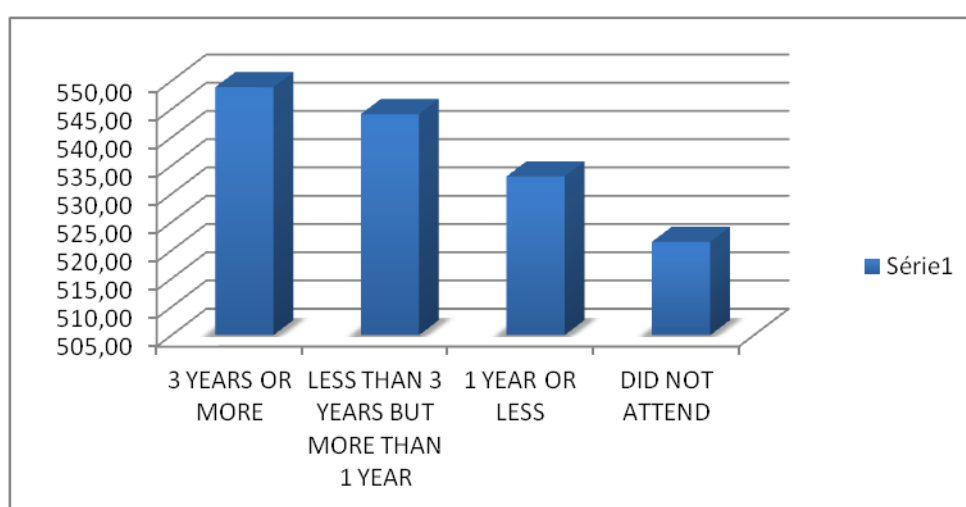
Country	Number		Geometric Shapes and Measures		Data Display	
	Girls	Boys	Girls	Boys	Girls	Boys
Armenia	485 (3.3)	483 (3.6)	426 (4.4)	422 (5.0)	392 (6.2)	381 (5.2)
Australia	505 (3.6)	511 (4.1)	532 (3.5)	536 (4.1)	512 (4.0)	519 (3.8)
Austria	502 (2.7)	511 (3.1)	507 (3.5)	516 (4.7)	510 (3.0)	520 (4.1)
² Azerbaijan	494 (5.8)	488 (5.4)	440 (7.7)	433 (7.8)	414 (6.6)	402 (6.9)
Bahrain	440 (3.8)	438 (3.9)	426 (5.7)	417 (4.5)	448 (5.6)	436 (4.8)
Belgium (Flemish)	547 (2.5)	556 (2.4)	549 (2.4)	555 (2.1)	533 (4.0)	539 (4.0)
Chile	457 (3.4)	466 (3.3)	449 (3.6)	462 (3.5)	463 (3.6)	468 (3.9)
Chinese Taipei	599 (2.7)	600 (2.5)	576 (2.9)	570 (2.6)	605 (2.8)	596 (3.8)
² Croatia	484 (2.0)	498 (2.7)	487 (2.9)	493 (3.4)	485 (3.3)	491 (3.7)
Czech Republic	502 (3.0)	515 (3.0)	511 (3.4)	515 (3.7)	512 (4.3)	526 (4.2)
² Denmark	530 (2.7)	538 (2.9)	546 (3.5)	550 (3.6)	530 (4.3)	533 (4.2)
England	536 (4.3)	542 (3.8)	544 (4.6)	547 (4.1)	551 (6.3)	547 (4.9)
Finland	538 (2.5)	552 (2.8)	544 (3.5)	543 (3.7)	549 (4.7)	553 (4.7)
¹ Georgia	474 (2.9)	472 (3.9)	416 (4.7)	408 (5.2)	441 (4.1)	425 (4.9)
Germany	514 (2.7)	526 (2.7)	532 (3.1)	540 (2.9)	546 (3.9)	545 (3.6)
² Hong Kong SAR	600 (3.2)	608 (4.0)	600 (3.3)	609 (4.1)	593 (3.5)	593 (4.8)
Hungary	513 (3.4)	517 (3.8)	521 (4.2)	520 (4.1)	510 (4.9)	510 (4.8)
Iran, Islamic Rep. of	439 (5.3)	440 (5.1)	434 (5.9)	435 (5.8)	397 (6.1)	396 (6.7)
Ireland	530 (3.8)	535 (3.3)	519 (4.4)	521 (3.8)	524 (3.6)	522 (5.0)
Italy	505 (3.1)	515 (3.0)	508 (3.5)	517 (3.5)	491 (3.1)	498 (4.1)
Japan	581 (2.0)	587 (2.2)	588 (2.4)	591 (2.6)	591 (2.3)	588 (4.4)
² Kazakhstan	511 (4.5)	518 (4.3)	486 (5.4)	496 (5.8)	477 (6.6)	474 (6.2)
Korea, Rep. of	600 (2.3)	610 (2.2)	606 (3.0)	608 (2.0)	607 (3.9)	599 (3.5)
¹ ✖ Kuwait	348 (4.5)	315 (6.6)	340 (4.6)	298 (7.5)	364 (4.3)	327 (7.4)
¹ ² Lithuania	536 (2.9)	539 (2.9)	531 (3.7)	530 (3.3)	528 (3.0)	524 (4.1)
Malta	493 (1.9)	502 (2.8)	484 (2.1)	489 (2.0)	497 (2.5)	499 (3.6)
✖ Morocco	344 (3.8)	337 (4.7)	352 (4.7)	348 (5.3)	278 (5.7)	264 (5.5)
[†] Netherlands	538 (2.3)	549 (2.1)	523 (2.7)	525 (4.1)	557 (2.9)	562 (4.5)
New Zealand	481 (3.2)	485 (3.1)	482 (2.9)	484 (3.1)	496 (3.3)	487 (3.6)
[†] Northern Ireland	566 (3.3)	567 (3.8)	561 (3.8)	559 (4.3)	558 (3.8)	552 (4.1)
[†] Norway	484 (3.2)	493 (4.0)	505 (3.9)	508 (3.8)	496 (4.3)	492 (4.7)
✖ Oman	394 (3.4)	373 (3.6)	390 (3.7)	363 (3.6)	396 (3.4)	366 (3.7)
Poland	474 (2.7)	486 (2.7)	472 (2.9)	478 (3.4)	486 (5.4)	491 (3.7)
Portugal	519 (4.5)	525 (3.6)	546 (4.5)	550 (4.9)	544 (3.5)	552 (3.7)
² Qatar	421 (4.9)	413 (3.9)	411 (4.8)	388 (4.7)	425 (5.3)	409 (5.7)

As I mentioned before, we also looked at the content domains, and we also looked at these by gender as a matter of routine. And if you go to Portugal, down on the bottom, in terms of content domains, there are no significant differences. But you see the patterns in other countries.

Again, in terms from a policy perspective, you begin then to question what the issues about our instructional practices are that may, or systematically advantage one group or another, assuming that we discount a number of generic explanations which here make no sense.

Content domains, we've got something here, which I can't explain, but I understand that it's not an unexpected outcome here in Portugal. That is the boys seem to be better at knowing. I don't know what that is; maybe someone here can tell me what the possible explanation is for that. But again, this sort of decomposition can begin to provide some sort of policy guidance into potential remediation programs.

PIRLS 2011 & Attendance at Pre school

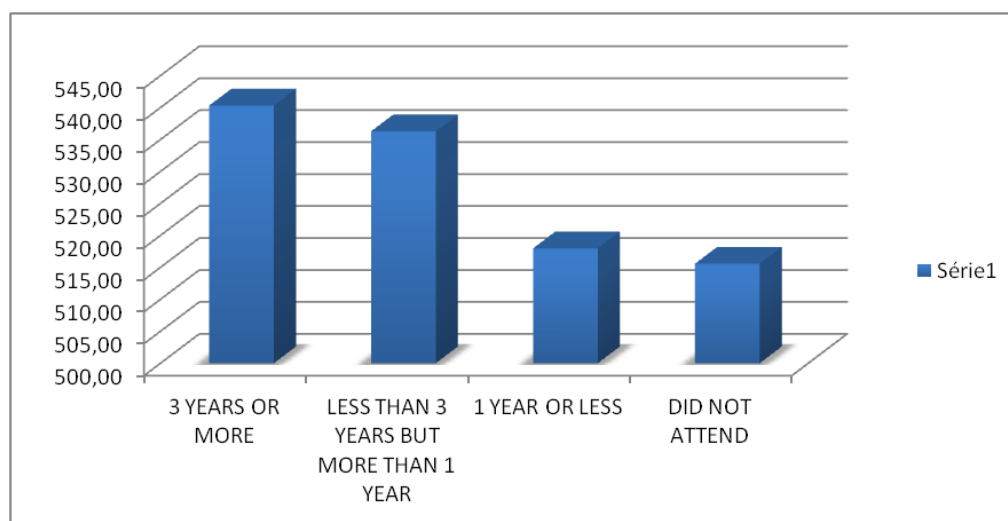


Aside from the question of what happens in the school system, there has been increasing attention and interest in the relationship between early childhood experiences and early parenting experiences and subsequent learning outcomes. And here, in the PIRLS and TIMSS 2011, we collected information from parents and asked them about their children's' early childhood experiences.

And here you can look at the relationship here between performance and participation in early childhood. Again, an underlying factor may in part be in play. It may be poverty, people who don't go, can't afford to go, that type of thing. But nevertheless, this sort of relationship (increasing performance with increased attendance at pre school) was present in all countries, while

maybe poverty is operating in all the countries, I suspect is more than that. I suspect that there is something about that pre-school experience which is important for later learning.

Math 2011 & Preschool attendance



For Mathematics it's the same, and in fact, from memory, I think the performance in Mathematics is assisted, maybe I'm wrong, but it had a greater impact on learning to read than early reading experiences.

In terms of thinking about on how to derive benefit from the assessment I would argue what is required is a much greater interrogation of data, and further decomposition of this information., And of course, beyond this sort of thing, much more sophisticated modeling is required.

There are clear questions that arise.

What policy (ies) could the government generate to address what might be perceived to be inequities which are attributable to the absence of pre-school experiences? So a challenge for all the countries, and I'll finish here, is to

think about how do we use this information? If all that we do is to produce a report and participate, we are wasting our time.

In order to maximize the impact, we really have to understand what it is that we (IEA) are trying to do. The detail of these assessments is important. There is much more than those international rankings, and it drives us (IEA) to distraction to think that most of the attention we get is about who came first and who didn't. This rather than talking about what are the things that make a difference in terms of understanding learning outcomes in our countries,

I think it is also important to think globally. It's important to think about what is it that makes it possible for countries like Singapore, Korea, Taiwan, Japan to perform at the levels they do, given that they are operating in the same period of schooling? It may be that some of the things that they do are not possible within Portugal or New Zealand or Germany or elsewhere. They may not even be desirable. But these are important questions to ask, and it's important to take the question of difference to the next step. Differences that we are talking about aren't explanations; these are the things that have to be explained. And this requires the integration of the information that we bring through participation in the assessments.

Think of international large-scale assessment as process and product. It's an opportunity for local capacity development. It's an opportunity for training and for the development of a culture of assessment and analysis. I may be talking to the converted here, that this is already built into your national systems. You have sophisticated people who work with us, universities are committed, as I understand to working with some of this data, but that's not true for many countries.

Many countries think this as an activity, and they don't, in my view, derive the full benefit, what is possible. We need to think as this as a strategic investment. It's about having a vision for education, not about participating once and thinking "now we solved this problem, we know what's happening."

Things change, and what's really apparent in the last 30 or 40 years is not just these educational systems that change our demands also change, and

context changes. And there are some rapid changes in context. And I'll give you a couple of examples.

When the European Union was created I remember reading an article in *The Economist* that talked about the estimated net migration gain in the United Kingdom, when the borders were opened. They thought of getting 60.000 people, in the first year they got 600.000 people.

These are all second language learners, many of them, and I would say most of them at that time were not from families with professional backgrounds. These are people who are looking for a second chance, better lives. While they are motivated, they were not endowed in terms of their educational background.

In Morocco between 1999 and 2003, under *Education for All*, they put 30% more children in school. It's a huge mass of children coming in, and if you believe that the kids who are already at school are more likely to be advantaged, what's coming in are children who are also less advantaged and with that come some of the challenges that those students bring. In Morocco's case, I would have expected average scores to probably decline, in reality. If they had achieved stability, it would have been a good outcome.

So how do we think about performance, equality and success? It has to be tempered knowledge of a rapidly changing broader context I mean, even the current economic crisis has seen changes in migration patterns from Britain back to Poland, and I'm sure this is happening here, in many directions. Again, this has an impact on schooling and learning outcomes, and the challenges that schools and teachers face.

And I think the last thing I'd probably want to say is what we do represents an opportunity. We provide a template which we require everybody to follow, but beyond that, this can all be augmented, both in terms of time, and it can be augmented in terms of additional target investigations., For example looking at increasing the sample size, because you want to pursue differences between certain populations, whether there are differences between private and public schools, differences among language minorities or ethnic minorities.

There are opportunities to, depending on the resources within the country to add an additional assessment using some national component which focuses more clearly on national issues. Questionnaires can also be modified to include variables that we don't collect internationally but might be of local interest.

Again, this is part of the strategic investment that countries need to make. They need to think about this not as a routine data gathering exercise, but as something that can be directed nationally in a way that is likely to make the most difference in terms of understanding educational systems.

If you want to learn more about what we've got, these are some brochures that explain the other assessments that you're not involved with. There's a new one here on the right, this is a proposal, but there is a new brochure here that talks about early childhood which has just come out and they are all available from our website. So thank you very much.

DEBATE

Question

One of the things that strikes me was what you said about the relation between expectations and final results. In fact I was thinking, and I would like to have a comment from you, about the relation between over-expectations, possible over-expectations and the same possibilities. There should be equilibrium between things. And the idea I got from you was, well, if you have more expectations that surely will push up the results.

Hans Wagemaker

It's clear that you have to be realistic about some of these things. But I think it's also equally clear that when you read a statement for one country, and I'll talk about Reading, when goes something like 'can read about 100 words with some degree of understanding', and our framework, our assessment of framework on Reading is demanding the ability of our students to integrate information and to understand authors intent and to extract information from the text, then, unless the system is at least beginning to teach children or bring them up to a point where they can begin to do some of these things, the chances of scoring well on that assessment after 4 years is going to be difficult.

Now, how do they transition from the 100 words scenario to this (PIRLS assessment) point is a challenge that is faced by a number of countries. I can talk about South Africa and also of some of the countries in the Middle East. Incidentally I did look at Portugal in Reading and its statement isn't exactly expressed in the way that we express our learning outcomes in Reading, but it's not so different. I think it's reasonably demanding in terms of expectations of what kids should be doing.

Also it's clear you're right, if you raise expectations too high, and you don't have teachers who can deliver them, then it's not going to work either. I think however that we underestimate what children can achieve.

Question

How consistent are the TIMSS and PISA standards?

Hans Wagemaker

In terms of performance rankings? I actually haven't look at Portugal's performance in PISA, but I wouldn't expect them to be automatically consistent for the very reason that there really are different assessment models. We have the curriculum-based model; we have grade 4 and grade 8. PISA assesses 15 year-olds end of the compulsory school learning process.

At least when they started, PISA talked about it being totally independent of curriculum. The challenge for me in PISA in interpreting this, is given that they aren't linked to curriculum who's 15 year-olds drive the assessment? Is it the 15 year-olds of New Zealand or the 15 year olds from Finland?

How do you design an assessment that targets learning outcomes for 15 year-olds when there's no backdrop. How do you come to that agreement?

But I think it was originally conceived, and I know this, I was involved in the original development of PISA when I worked for the government. My concern was that when you abandon the curriculum you abandon an anchor which provides some sort of comparability.

When you abandon the grade, you've raised that challenge of having to interpret the outcomes in terms of where most 15 year-olds are found. In some countries 15 year-olds are in 2 grades, in some they are in 4 grades, in some countries they are in grade 11, in some they are in 9.

Question

You also have data about the span of the grades you've studied?

Hans Wagemaker

Yes, and they are all, with one or two exceptions in grade 4 and grade 8. What we published at the beginning was country participation, their rank, the grade, the average age and also their ranking on the Human Development Index, so that countries can understand that you're comparing the very richest with some of the very poorest of countries. So it's another context for interpreting that information.

Question

You have evidence about the factors that are behind the difference in results in the urban and rural children?

Hans Wagemaker

I don't have it here, and it's probably different for different countries. It may as well be related to poverty, it may be related to particular communities. In some countries it will be language community differences. As I said, I took the relationship between two variables. But to understand that, I think it's exactly the right question – you have to dig deeper into the data. The differences are not an explanation, so you have to start interrogating that, and I don't know enough about Portugal to be able to do that.

Question

Coming back to the previous question, when you say that there is a difference between the rural and urban children, you mean a difference because they are in rural areas and in urban areas, taking aside all others you made a selection from and so on? Or are you just comparing rural and urban?

Hans Wagemaker

I had not taken any of that into account. There are students in rural communities and in urban communities and they perform differently. Within those communities boys and girls perform differently, have different profiles.

But the question you then have to answer or ask is, the sort of things that you are alluding to, is it because there of differences in parental education? Is it differences that are related to language; is it differences that are related to both? In some countries they would be highly mobile communities, its unique to each country. And it may even be unique within countries, there may be different types of rural experiences that can help us understand those performance differences.

That is to say that what we produce is a national report that doesn't even go to the rural/urban, because it's not well defined internationally. Because of that we look at this in a national data base., Portugal happens to have this

information so I was able to get it, but it's a construct that needs to be unpacked as well.

What does it mean to be in a rural community, a rural community in New Zealand, or in the United States or whatever? How big are these communities? That is not a constant either, so I think there's systematic differences that needs to be understood, and that's the point.

Question

Do you have any key points that may explain why we have improved our results in TIMSS and PIRLS? Because as you say, it's complex, the analysis is complex, but if you have some ideas about what might have happened for the increasing results?

Hans Wagemaker

Only from what I've learned talking to a few people here. But I understand that Portugal, I've been told very recently invested 6% of GDP, which is probably above average for the OECD in terms of education. There has been an investment and a focus on learning outcomes in the last, let's say, 20 years. It may be that it reflects the political changes.

What might have been measured was a legacy of the past, of the previous political government here, the dictatorship. Who knows? It may be under the new more democratic form of organization in Portugal, that there are improved systems, better management, all of these things, better teacher preparation? I don't know, I was hoping that you could tell me.

Question

I would like to ask you how should Portugal take care if future results go down?

Well, I think that depends on how much they go down, hopefully they won't, but I think the same questions need to be asked. What is different about 2015 and 2016 in the students who now represent a different cohort of experience? What is different, what has changed in that period? What broader societal changes have occurred?

We work with the data that we get at the end of this period. We insure that all the countries carry out these studies to the same standard. We monitor the sampling; we monitor the translations and the quality in the translations. We have some quality procedures during and after that we ensure that, for example, when we find countries that manipulate data, that we don't publish it. When we find examples where countries suddenly remember that they have parts of the population that they have forgotten about, you know, accidentally...

We've had countries where we detected deliberate attempts to manipulate the data, and so we employ quality control monitors to go to sample schools to make sure. So what we can say at the end of this is that on the basis of the information that we have with us now, we're confident that the study was carried out in a manner that would give us reassurance that it actually reflects the performance in a country at that time.

But things do change and the financial crisis may well have an impact. I understand that here you've got increased class sizes. In some countries there are school closures that are having that exact same effect. The removal of support services, all these things are changing some very fundamental conditions for learning. On the other hand, maybe what the counselor said, maybe we'll learn to do more with less.

Question

With your long experience in this kind of studies are there, in your opinion any aspects, particular aspects that you would like to do it, in TIMSS and PIRLS, to change and you don't because of any reasons? Are there any changes you would like to introduce in TIMSS and PIRLS or not, and why?

Hans Wagemaker

We are making several changes. One of the challenges we face almost immediately is the technological transition, to capture all this information technologically and delivering assessments that aren't simply a replication of the paper and pencil on the computer. But it allows us to do different types of assessment, to work in different ways.

I mean, an ideal model might be what is called adaptive testing where students are exposed to successively more difficult questions depending on the previous success. That demands huge resources in terms of not only in terms of getting the technology, the electronic technology, but also in terms of the development of resources.

For PIRLS, the next cycle, we are planning to have what we call Internet based Reading. We talked about originally a sort of a digital literacy but there is a question of as to whether or not the experience of reading in the internet and that internet environment represents a different reading challenge, and so that's something that we definitely investigating, and we'll run that in the next cycle.

The technology opens up whole new possibilities. One of the disappointments I think of technology in schools is that it has been adapted largely for administrative purposes. There are instructional programs more so now than there were 30 years ago, but nevertheless, I still think that the opportunities, if you think about Physics or Chemistry or Mathematics where the opportunities for, for example, rotating molecules or manipulating molecules in a virtual environment as a teaching tool, and possibly as an assessment tool, all this sort of things are theoretically possible.

We are limited in two ways, one by resources, it's expensive to do this and we are probably at the limit of what we can charge countries to do this. And secondly we work with countries that range from Singapore, Taiwan, Korea, Portugal, Norway, the developed economies of the world, but also with Yemen and Syria and Lebanon and South Africa, and Botswana and Ghana. And the reality of delivering some of this in a technology rich environment is probably not there yet.

SEMINÁRIO

Os resultados dos alunos portugueses no TIMMS,
em Matemática e Ciências,
e as suas implicações para o ensino,
para a formação de professores e
para o sistema educativo

Presidente da Mesa - Maria Calado Gomes

SEMINÁRIO

Nesta intervenção passo em revista os resultados dos alunos portugueses do 4.º ano de escolaridade em Matemática no TIMSS, divulgados em dezembro de 2012, dando igualmente conta de alguns resultados de contexto que me parecem particularmente interessantes, procurando interpretá-los à luz das realidades portuguesas. Considero que, no seu conjunto, estes resultados são francamente notáveis e merecedores de uma reflexão aprofundada por todos aqueles que se interessam seriamente pelo ensino desta disciplina escolar.

Portugal (com 532 pontos) surge pela primeira vez muito bem colocado no ranking da disciplina de Matemática – 15.º lugar mundial e 8.º europeu, só atrás de Singapura (606 pontos), Coreia (605), Hong Kong (602), Taiwan (591) e Japão (585), países asiáticos com sociedades e sistemas educativos muito diferentes do nosso, e Irlanda do Norte (562), Bélgica (parte flamenga) (549), Finlândia (545), Inglaterra (542), Rússia (542), Estados Unidos (541), Holanda (540), Dinamarca (537) e Lituânia (534).

Resultados inferiores aos portugueses tiveram, por exemplo, países como a Alemanha (528), Irlanda (527), Sérvia (526), Austrália (516), Hungria (515), Eslovénia (513), República Checa (511), Áustria (508), Itália (508), Eslováquia (507) e Suécia (504). Abaixo da média internacional de 500 pontos, ficaram ainda Malta (496), Noruega (495), Nova Zelândia (486), Espanha (482), Roménia (482), Polónia (481), Turquia (469) e vários outros países.

Os resultados dos alunos portugueses traduzem uma subida de 442 para 532 pontos, ou seja, uma subida de 90 pontos em relação à avaliação anterior realizada em 1995, o que não tem paralelo a nível mundial – o segundo país com maior subida (58 pontos) é a Inglaterra, que passou de 484 para 542 pontos. Alguns países tiveram mesmo descidas, como a República Checa (-30 pontos, de 541 para 511), Áustria (-23 pontos, de 531 para 508), Holanda (-9 pontos, de 549 para 540), Hungria (-6 pontos, de 521 para 515 pontos). É de registar que esta subida dos alunos portugueses, ao contrário do que se

¹ Instituto de Educação da Universidade de Lisboa

poderia esperar, não se deveu à melhoria dos resultados dos melhores alunos, mas sim a uma melhoria a todos os níveis, com uma redução da dispersão dos resultados (passou-se de um desvio padrão de 3,9 para 3,4 pontos).

Por temas matemáticos, o desempenho dos alunos portugueses é especialmente forte nos temas “Formas geométricas e medida” (548 pontos) e “Representação de dados” (idem) e inferior, mas ainda assim bastante positivo, no tema “Números” (522 pontos). No que respeita aos níveis cognitivos matemáticos considerados na elaboração da prova, não se registam diferenças significativas nos resultados dos alunos portugueses, que obtiveram no Conhecimento (o nível mais simples) 531 pontos, na Aplicação (nível que já traduz alguma complexidade) 534 pontos, e no Raciocínio (o nível mais exigente) 531 pontos.

Nos itens cujos resultados são divulgados, os alunos portugueses têm, de um modo geral, excelentes desempenhos. Assim, no tema “Números”, num problema verbal envolvendo a adição de dois números naturais de três dígitos (questão de aplicação, dificuldade baixa), Portugal surge em 5.º lugar (89% correto). Noutro problema verbal envolvendo a adição de numerais decimais (com uma casa decimal) (questão de aplicação, dificuldade intermédia), Portugal surge em 7.º lugar (84% correto). Num problema verbal envolvendo adição de medidas de tempo e conversão entre horas e minutos (questão de aplicação, dificuldade elevada), registou-se o pior resultado, surgindo Portugal em 31.º lugar (47% correto). Em contrapartida, num problema numérico de raciocínio com vários passos (questão de raciocínio, dificuldade elevada), Portugal surge em 11.º (40% correto).

No tema “Formas geométricas e medidas”, num problema para determinar o número de cubos numa pilha com alguns cubos escondidos (questão de aplicação, dificuldade intermédia), Portugal surge em 7.º lugar (84% correto). Noutro problema para completar uma figura de tal modo que exista uma reta de simetria e um certo número de lados (questão de aplicação, dificuldade elevada), Portugal surge em 10.º lugar (53% correto). Num problema em que eram dadas as figuras de dois sólidos comuns e se pedia para classificar quatro afirmações como verdadeiras ou falsas (questão de

conhecimento, dificuldade elevada), Portugal surge em 1.º lugar (70% correto).

No tema “Representação de dados”, num problema em que era necessário completar um gráfico de barras usando dados numa tabela (questão de aplicação, dificuldade baixa), Portugal surge em 15.º lugar (84% correto). Num problema em que se pedia para identificar o gráfico de barras que corresponde à informação mostrada num gráfico de setores (questão de raciocínio, dificuldade intermédia), Portugal está em 10.º lugar (82% correto). Finalmente, num problema de vários passos em que era necessário usar dados de um gráfico de barras (questão de raciocínio, dificuldade elevada), Portugal ficou em 7.º lugar (70% correto). Deste modo, na grande maioria dos itens divulgados, Portugal surge acima do 15.º lugar que obteve na média geral, sendo que em diversos itens ultrapassou mesmo os países asiáticos que melhores resultados globais obtiveram no teste.

É também interessante analisar os resultados obtidos em algumas variáveis relacionadas com o desempenho. Por exemplo, no que se refere aos alunos, os professores reportam que estes se envolvem na aprendizagem (Portugal é 8.º em 50 países) e que os alunos gostam de aprender Matemática (12.º). Em contrapartida, os próprios alunos revelam muita confiança nas suas capacidades na disciplina de Matemática (41.º).

No que se refere aos próprios professores (recorde-se que estamos a falar de professores do 4.º ano de escolaridade), estes consideram que promovem um ensino para envolver os alunos na aprendizagem (Portugal é 4.º em 50 países), mostram confiança no uso de estratégias de ensino (5.º) e consideram importante a colaboração para melhorar o ensino (14.º). Em contrapartida, na perspetiva dos professores, as atividades com computador não têm uma expressão muito forte na aula de Matemática (22.º) e não vêm como muito favoráveis as condições de trabalho, dados os problemas criados pelos alunos (41.º), nem mostram satisfação com a carreira (44.º). Deste modo, os professores evidenciam muita confiança nas suas capacidades profissionais, ao mesmo tempo que mostram grande descontentamento relativamente a questões de carreira e de condições de trabalho.

No que respeita ao sistema educativo, Portugal é o país onde mais horas semanais são estabelecidas para o ensino da Matemática (1.º em 50 países), mas não são muitos os alunos que trabalham com números antes de entrar no 1.º ciclo (18.º em 32 países). No que se refere à família, registam-se expectativas elevadas na realização de um curso de pós-graduação para os seus educandos (7.º em 32 países), o que evidencia vontade de investir em qualificações académicas, e registam-se recursos favoráveis em casa para a aprendizagem (15.º em 32 países).

Para estes resultados contribuíram, no meu entender, diversos fatores de política educativa nacional, com destaque para o programa de formação contínua para professores do 1.º ciclo, da iniciativa do Ministério da Educação, que teve a coordenação de Lurdes Serrazina, da Escola Superior de Educação de Lisboa, e que decorreu de 2005 a 2011. Este programa envolveu cerca de 15 000 professores (mais de 50% do total em exercício), num trabalho consistente realizado ao longo de todo o ano letivo (em alguns casos durante um ano, noutros casos durante dois), com trabalho de aprofundamento matemático e didático e supervisão por parte de um formador na sala de aula. Terão igualmente contribuído, embora de forma menos direta, o programa de Matemática do ensino básico homologado em 2007, pelos materiais de apoio e pela formação produzidos no seu âmbito e que abrangeram muitos professores em exercício, nomeadamente através do programa de formação contínua referido. Alguns alunos que fizeram a prova do TIMSS seguiram já o novo programa, nas mais de 400 escolas e agrupamentos que o começaram a usar por iniciativa própria a partir de setembro de 2009. Para estes resultados terá ainda contribuído o Plano de Ação da Matemática (em vigor a partir de 2006), proporcionando apoio continuado ao longo do tempo às escolas e agrupamentos, procurando corresponder às necessidades sentidas pelos professores.

Destes resultados decorrem diversas implicações. Assim, é de sublinhar, antes de mais, que vale a pena valorizar, de modo positivo, a Matemática no currículo escolar, como foi feito no período 2006-2011, com políticas específicas nesse sentido. Em segundo lugar, é necessário valorizar o programa de 2007, principal referência que permitiu este sucesso, com as orientações curriculares que lhe estão subjacentes, nomeadamente:

- Valorização do sentido de número, visualização espacial, pensamento algébrico e literacia estatística,
- Ênfase nas capacidades transversais de resolução de problemas, raciocínio matemático e comunicação matemática,
- Ensino baseado numa abordagem exploratória, proporcionando uma formalização progressiva dos conceitos matemáticos,
- Aprendizagem dos procedimentos matemáticos com compreensão.

Finalmente, os resultados do TIMSS 2012 sugerem a necessidade de apoiar iniciativas de formação contínua com objetivos e natureza idênticos ao programa de 2006-11, de promover a formação inicial articulada com a formação contínua e com as boas práticas existentes no terreno e de apoiar as escolas e os agrupamentos, valorizando as suas iniciativas e correspondendo às suas necessidades. Trata-se de aspetos que convém não perder de vista nas políticas educativas futuras.

SEMINÁRIO

Para mim é particularmente grato estar nesta sessão pelo facto de ter integrado a equipa nacional do projeto TIMSS de 1995, altura em que os resultados negativos de Portugal foram amplamente divulgados. Lamento que o mesmo não esteja a acontecer agora com os bons resultados alcançados por Portugal na edição do TIMSS de 2011.

Comparação dos desempenhos nacionais de 1995 e 2011

Irei centrar a minha análise em vários aspetos, mostrando os desenvolvimentos que ocorreram de 1995 a 2011 e tentando discutir convosco algumas das possíveis razões para a melhoria de desempenho do nosso país. Portugal faz parte do grupo de países que teve o maior incremento de desempenho entre 1995 e 2011. Aliás, é o segundo país em termos de melhoria: melhorou 70 pontos. Contudo, para além do desempenho bastante positivo de Portugal, os resultados divulgados também revelam áreas a melhorar e sugerem o reforço e a realização de intervenções específicas em determinados aspetos.

Constata-se que de 1995 para 2011 os alunos portugueses evoluem consideravelmente em todos os níveis de referência internacionais:

- a) 7% dos participantes atingem o nível avançado (quando em 1995 apenas 2% o conseguiam fazer), sendo capazes de aplicar o conhecimento e a compreensão das relações e dos processos científicos e evidenciando algum conhecimento do processo de investigação científica (*scientific inquiry*);
- b) o nível elevado é alcançado por 35% dos alunos (apenas 13% o tinham feito em 1995) que conseguem aplicar o seu conhecimentos e compreensão das ciências para explicar fenómenos em contextos do quotidiano e abstratos;

¹ Instituto de Educação da Universidade de Lisboa

c) o nível intermédio é atingido por 75% dos participantes (apenas 43% o tinham alcançado em 1995) que apresentam níveis básicos de conhecimento e compreensão de situações práticas de ciências;

d) o nível baixo é ultrapassado por 95% dos alunos portugueses (em comparação com 73% em 1995) que evidenciam alguns conhecimentos elementares de ciências da vida, físicas e da terra.

Na edição de 2011 do TIMSS, os alunos portugueses obtêm desempenhos superiores à média internacional nas diferentes áreas de conteúdo e competências avaliadas. No entanto, os níveis de desempenho são maiores a Ciências da Terra e a Ciências da Vida do que a Ciências Físicas. Por outro lado, os alunos portugueses evidenciam melhores desempenhos ao nível do Conhecimento e do Raciocínio do que ao nível da Aplicação.

O desempenho por género também é particularmente interessante. Verifica-se que a situação que existia em 1995 se mantém, ou seja, não há uma diferença significativa entre rapazes e raparigas ao nível das Ciências. No entanto, se analisarmos especificamente por níveis cognitivos existe uma diferença que é estatisticamente significativa ao nível da aplicação, precisamente a competência sobre a qual também teremos de refletir melhor, em que os rapazes apresentam um nível significativamente superior. A nossa reflexão deverá centrar-se não só nestes resultados mas também no que são os resultados de rapazes e raparigas nos exames finais do secundário, em que parece haver uma inversão de desempenhos: as raparigas acabam por atingir resultados médios mais elevados, apesar dos rapazes atingirem desempenhos superiores, mas em menor quantidade. As questões de género são uma área em que há muito pouca investigação em Portugal, nomeadamente no ensino das Ciências, mas que importa realmente estudar. Torna-se necessário compreender as causas destas diferenças de desempenho entre géneros.

Se analisarmos os sete exemplos de itens incluídos no relatório do TIMSS 2011, constata-se que as maiores evoluções ao nível de desempenho se verificam em tópicos curriculares que foram trabalhados no Programa de Formação Contínua de Professores do 1.º Ciclo em Ensino Experimental das Ciências. O facto deste programa de formação ter implicado os professores

participantes no planeamento e na realização de diversas atividades experimentais com os seus alunos no seu próprio contexto de trabalho terá assegurado um impacto positivo (visível através dos resultados do TIMSS 2011) nos conhecimentos e competências destes mesmos alunos.

Os melhores níveis de desempenho dos alunos portugueses verificam-se em tópicos curriculares que integram o currículo do 1.º Ciclo do Ensino Básico. A área menos representada no currículo – as Ciências Físicas – é aquela em que o desempenho dos alunos é inferior.

A análise pormenorizada dos resultados obtidos pelos alunos portugueses no TIMSS 2011 é uma tarefa decisiva para Portugal: é através desta análise de pormenor que conseguimos detetar lacunas ao nível da educação em ciências. A análise do tipo de respostas dos alunos perante determinados tipos de itens permitirá detetar as suas dificuldades e os aspetos que deverão ser melhorados ao nível da aprendizagem de conhecimentos e do desenvolvimento de competências.

A importância do contexto

Existem alguns dados de contexto que poderão ter contribuído para o bom desempenho dos alunos portugueses. É o caso da formação académica dos pais e do número de alunos que frequentaram o pré-escolar (que aumentaram consideravelmente de 1995 para 2011). Apesar de não se tratar de uma relação de causa-efeito, observa-se uma relação positiva entre a frequência do pré-escolar e o nível de desempenho obtido no TIMSS 2011.

Verifica-se que os melhores níveis de desempenho ocorrem em escolas em que a dimensão da biblioteca é de 500 ou menos livros ou em que não há biblioteca. Isto significa que os grandes centros escolares, melhor apetrechados (nomeadamente, com bibliotecas com mais de 500 livros), não são sinónimo de bom desempenho no TIMSS 2011. Assim, a variável professor parece emergir como a mais importante para o bom desempenho dos alunos.

Também se verifica que a existência de laboratórios de Ciências não se revela particularmente significativa para o bom desempenho dos alunos

portugueses. Esta situação poderá estar relacionada com o Programa de Formação de Professores do 1.º Ciclo em Ensino Experimental das Ciências que, ao longo da sua existência, insistiu muito no desenvolvimento de competências científicas através da utilização de materiais simples, ou seja, na promoção de competências de nível cognitivo elevado através do recurso a materiais simples em salas de aula normais (sem necessidade de laboratório).

Os professores

Relativamente aos professores dos alunos que participaram no TIMSS 2011, verifica-se que 31% frequentaram ações de formação sobre ensino da ciência nos últimos dois anos e que a maioria (76%) se sente muito bem preparada para ensinar todos os tópicos de ciências do programa. Contudo, apenas 64% dos professores se sentem muito bem preparados para ensinar os tópicos de Ciências Físicas. Logo, existe um défice de formação dos professores ao nível das Ciências Físicas que é necessário ultrapassar, principalmente porque essa é a área de conteúdo na qual os alunos obtêm um desempenho inferior.

Os professores de 71% dos alunos apresentam um nível de confiança bastante elevado nas suas capacidades gerais de ensino das ciências. Contudo, manifestam falta de confiança (1) na explicação de conceitos ou princípios através de experiências científicas, (2) na proposta de atividades que sejam desafiantes para os alunos com melhor nível de desempenho e (3) na resposta às perguntas dos alunos sobre Ciência. Trata-se de capacidades de ensino das ciências às quais as instituições de formação inicial de professores e o Ministério da Educação (através da realização de programas de formação contínua) deverão prestar particular atenção.

No que se refere ao tempo que os professores consideram dedicar ao ensino das ciências, os professores portugueses ocupam o primeiro lugar a nível internacional. Contudo, para além do que poderá ser o carácter irrealista desta avaliação, importa refletir sobre a qualidade do trabalho que é realizado em sala de aula. Os resultados obtidos pelos alunos portugueses (por exemplo, o desempenho inferior ao nível da Aplicação de

Conhecimentos) sugerem a necessidade de melhorias em termos de práticas de sala de aula, nomeadamente no respeito à realização de atividades mais centradas nos alunos que reforcem a sua autonomia, as suas competências intelectuais e o seu conhecimento científico.

Segundo os professores, a percentagem de alunos a quem foram ensinados os tópicos TIMSS das diferentes áreas de conhecimento varia entre os 96% para as Ciências da Vida, os 75% para as Ciências Físicas e os 88% para as Ciências da Terra. Curiosamente, verificamos que na Coreia e no Japão (países com ótimos níveis de desempenho) os tópicos TIMSS apenas foram abordados, segundo os professores, por 50% e por 38% dos alunos, respetivamente. Isto significa que nestes países, os alunos conseguem obter bons níveis de desempenho apesar dos seus professores não terem abordado vários tópicos TIMSS, o que poderá indiciar um ensino mais centrado no desenvolvimento de competências e não tanto em conteúdos específicos.

Outro aspeto extremamente interessante é a relação entre a experiência dos professores e o nível de desempenho atingido pelos seus alunos. Constata-se uma relação positiva entre a experiência dos professores portugueses e o nível de desenvolvimento dos seus alunos. Contudo, o maior nível de desempenho é obtido por alunos de professores com menos 5 anos de experiência. Claro que, nos últimos anos, são as instituições de ensino particular (muitas vezes frequentadas por alunos de estratos socioculturais privilegiados) que contratam professores recém-licenciados. Contudo, esta situação mostra, pelo menos, que nos últimos anos se tem realizado um bom trabalho ao nível da formação inicial de professores de 1.º ciclo do Ensino Básico.

Implicações e recomendações ao nível do ensino

A leitura e a análise dos resultados obtidos pelos alunos portugueses no TIMSS 2011 sugerem diversas implicações e recomendações para o Sistema Educativo:

- A necessidade de recurso a atividades e metodologias capazes de reforçar o gosto pela aprendizagem das Ciências, especialmente das raras. O

aumento da percepção da relevância da educação em Ciências é algo que também temos de trabalhar bastante, através do estabelecimento de ligações com situações do dia-a-dia, contextualizando o currículo.

- A necessidade de melhorar a qualidade do tempo dedicado ao ensino das Ciências: é extremamente importante ter um ensino mais desafiante para os alunos de todos os níveis de desempenho – tanto para aqueles que têm um fraco desempenho, como para aqueles que têm um desempenho muito bom.
- A necessidade de se implementar uma maior diferenciação pedagógica e de se realizar um ensino mais contextualizado, mais centrado no aluno e menos no professor, mais promotor do raciocínio e da aplicação de conhecimentos, mais investigativo e menos prescritivo, mais centrado nos processos da Ciência.
- Uma maior e melhor exploração de tópicos de Ciências Físicas.
- A necessidade de se continuar o trabalho de desenvolvimento pessoal e profissional dos professores do 1.º ciclo do Ensino Básico, iniciado com o Programa de Formação de Professores em Ensino Experimental das Ciências. Analisando o desempenho dos alunos portugueses nos diferentes itens do TIMSS 2011 detetam-se evidências de um impacto positivo deste programa na qualidade do ensino das ciências.
- É preciso, também, que os documentos de orientação curricular não se restrinjam a listagens de conteúdos e que procurem estimular um trabalho de qualidade no que respeita ao desenvolvimento de competências, nomeadamente, ao nível da aplicação de conhecimentos.
- A importância de um trabalho precoce com os alunos com maiores necessidades educativas (mal se detetem determinadas dificuldades). Países como a Finlândia conseguem obter bons níveis de desempenho global graças ao trabalho feito com os alunos mais fracos.

É de assinalar o efeito extremamente contraproducente que os exames no final do 1.º ciclo do Ensino Básico – centrados na Matemática e na Língua Portuguesa – já estão a ter no ensino das Ciências. Com a realização destes exames, a mensagem que se passa para os professores e os familiares é a de que as áreas da Matemática e da Língua Portuguesa são as mais importantes.

Logo, cada vez mais tempo e recursos são deslocados para estas áreas em detrimento das outras áreas, nomeadamente a de Estudo do Meio. Os professores sentem que não têm tempo para a realização de atividades de tipo investigativo, centradas nos alunos, acabando por favorecer o ensino factual e expositivo das ciências (pouco promotor de competências cognitivas mais elevadas).

Constata-se que o trabalho de investigação realizado na área da Educação em Ciências teve repercussões diretas (efeitos bastante positivos) ao nível da formação contínua de professores. Logo, os resultados alcançados no TIMSS 2011 devem ser estudados de forma aprofundada por esta comunidade de Educação em Ciências com o objetivo de se corrigir o que seja necessário e de se manter as estratégias que se tenham revelado positivas.

Acrescentaria, também, algo que considero decisivo para nós como país: a necessidade de se divulgar o bom desempenho dos alunos portugueses no TIMSS 2011, para reforçar o autoconceito dos professores, dos alunos e das famílias. É necessário mostrar que as escolas portuguesas têm realizado um trabalho bastante interessante, ao qual é necessário dar continuidade.

SEMINÁRIO

Os resultados dos alunos portugueses no PIRLS, em Leitura,
e as suas implicações para o ensino,
para a formação de professores e para o sistema educativo

Presidente da Mesa - José Augusto Bernardes

SEMINÁRIO

Gostaria de começar por agradecer ao Conselho Nacional de Educação o convite para vos apresentar a minha análise sobre os dados do PIRLS e, concomitantemente, partilhar a reflexão sobre um tema que me é muito caro: a leitura, os leitores e o ensino desta competência fulcral no desenvolvimento do indivíduo e da sociedade.

De uma forma geral, a participação em avaliações internacionais, porque envolve muito trabalho, dispêndio de energias e de recursos nacionais, inclusivamente financeiros, deve, em minha opinião, reverter em lucros substanciais para o país participante. Daí que tivesse escolhido como eixo central da minha comunicação *o que podemos aprender com os resultados do PIRLS*.

Concebi a estrutura da comunicação em três partes e, revisitando os idos de 1974, escolhi três D, para designar cada uma dessas partes: Dados, Dúvidas e Dividendos.

Na primeira parte — os Dados — abordarei alguns resultados do PIRLS relativos a Portugal, centrando-me na questão, *como leem as nossas crianças?* Procurarei comparar os dados atuais com os de países que integram o mesmo espaço geográfico e económico e, num olhar retrospectivo de 20 anos atrás, com os resultados do Reading Literacy (1991), promovido pelo IEA, e que está na base do PIRLS. Em seguida levantarei algumas Dúvidas que me ocorreram na leitura dos mesmos. Na secção final remeterei para o que considero os Dividendos nacionais na participação do estudo em análise.

Como ponto de partida, considero que a análise dos resultados deste tipo de estudos deve ter como bússola a avaliação da qualidade das aprendizagens do país, tendo como referência outros países, e não a visão de competições tipo olimpíadas. O fulcro principal são as aprendizagens escolares, tentando compreender como é que as aprendizagens se materializam no que pode ser objeto de comparação entre sistemas educativos. E devo referir, à partida,

¹ Escola Superior de Educação de Lisboa

que quando se trata de um estudo com características linguísticas, como é o caso, há que tomar em linha de conta limitações muito fortes que as diferenças linguísticas comportam. Na potencialização dos ganhos de participação gostaria de salientar a necessidade de interpretar os dados à luz de variáveis específicas do sistema educativo e da população nacional. Os dados são tratados em termos internacionais, mas há informação de pormenor que é de importância capital e que não pode ser ignorada ou menosprezada; aí se incluem variáveis da população, espelhadas ou não nas características da amostra desenhada e usada que inclui quer alunos, quer professores, e que se repercute nas especificidades da organização pedagógica. Tal significa que o aprofundamento dos dados deste tipo de estudos requer o cotejo com outros dados de todo o país, tais como as provas de aferição ou os atuais exames nacionais.

Um ganho adicional da participação em estudos internacionais tem que ver com desejáveis melhorias no processo e no sistema de avaliação e, conseqüentemente, no sistema de ensino. No que se refere especificamente ao processo de melhorias na avaliação da compreensão de leitura, quero salientar especificamente o GAVE, com todas as suas provas de aferição, que beneficiou das participações nacionais quer no *Reading Literacy*, quer no PISA. Julgo, portanto, que tendo em vista o sistema de avaliação, e conseqüentemente a formação de professores, será da máxima conveniência olhar microscopicamente para os dados obtidos nos estudos internacionais. Finalmente, porque de pesquisa se trata, os resultados e a experiência de participação nestes estudos podem fornecer âncoras para posteriores investigações. Todos estes aspetos são dividendos não desprezáveis se quisermos recapitalizar a energia e recursos investidos em participações internacionais.

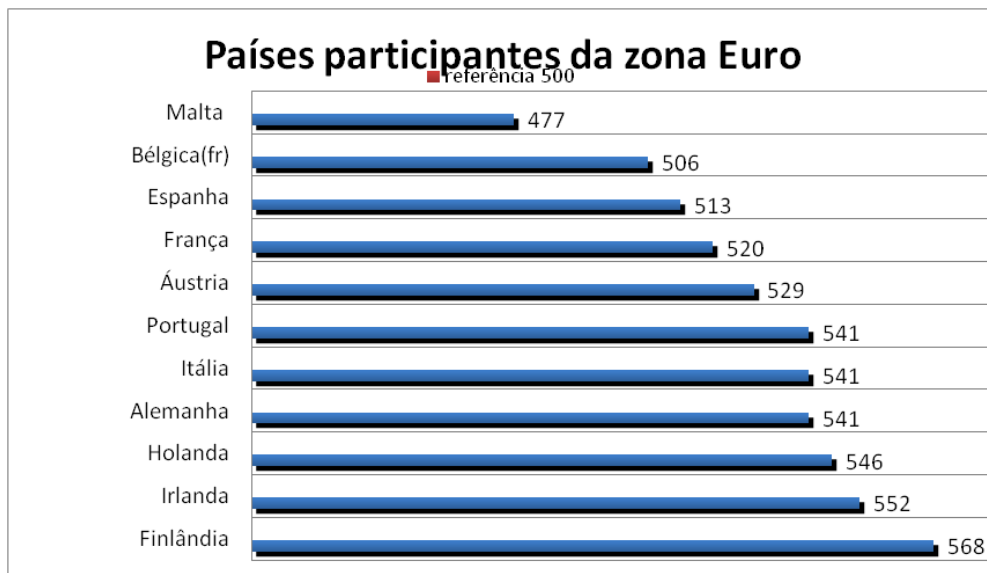
Os Dados – Partindo da minha declaração de interesses, que é a posição que assumo perante os estudos internacionais, analisemos os dados. Trata-se aqui do PIRLS, *Progress in International Reading Literacy Study*, no qual participaram alunos do 4.º ano de escolaridade de 45 países. A escolha do 4º ano não é mero acaso, nem sequer por ser fim de ciclo, o que não acontece na maioria dos sistemas educativos. A razão tem que ver com a etapa de escolaridade em que todos os mecanismos básicos de aprendizagem da

leitura devem estar consolidados. Os alunos participantes tinham no momento da recolha de dados 9/10 anos; na medida em que os dados foram recolhidos em Portugal no 1.º trimestre de 2010, os alunos portugueses tinham uma média de idades de exatamente de 10 anos. A amostra portuguesa contempla 148 escolas e inclui 4085 alunos.

Trata-se de um teste de compreensão de leitura, i.e., o estudo avalia exatamente a capacidade para compreender material escrito. A prova integrava diversos tipos de textos e, através das questões formuladas, procurava avaliar níveis de compreensão, tendo como parâmetros os objetivos/propósitos de leitura e as diversas categorias de processos envolvidos na compreensão. Foram também passados questionários de contexto aos alunos, aos pais e à escola. As coordenações nacionais contribuíram com informação para uma designada *Enciclopédia* sobre políticas educativas e organização do currículo para a leitura em cada país participante.

Os dados aqui mencionados foram obtidos no relatório internacional, tendo sido selecionados os que considere mais pertinentes para a situação em vista.

Para nos situarmos comparativamente, escolhi os resultados globais dos países que, como Portugal, integram a zona euro. De salientar que, à imagem do que sucede neste tipo de estudos internacionais, estamos a trabalhar numa escala de 0-1000, com valores típicos que caem na faixa 300-700 e com um ponto central de referência de 500.

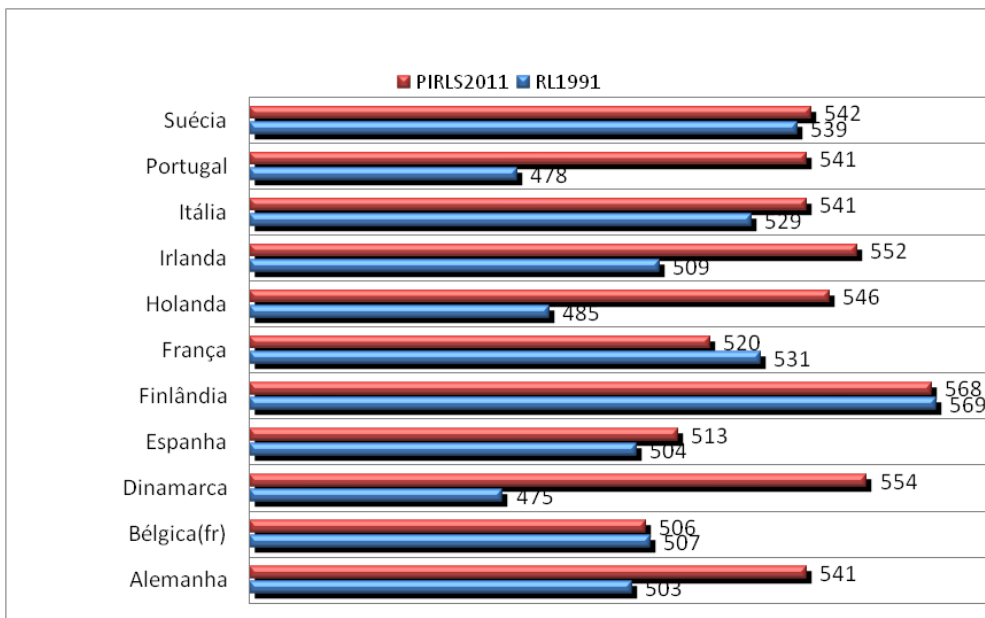


Quadro 1- Situação em relação a países da zona Euro

Numa comparação relativa com países que partilham a mesma moeda, verifica-se que Portugal, com uma pontuação de 541, claramente acima do valor de referência (500), ocupa um lugar central na zona Euro, *ex aequo* com a Itália e com a Alemanha. Com valores inferiores surgem Malta, a Bélgica francófona, a Espanha, a França e a Áustria. Pontuando acima de nós a Finlândia, a Irlanda e a Holanda.

Neste olhar comparativo, vale a pena equacionar os dados de 1991 em que também participámos. Era outro estudo, era uma outra prova. De qualquer maneira, foi esse estudo *Reading Literacy* que esteve na origem do PIRLS.

O que pretendo mostrar aqui é o que se passou com países participantes nos dois estudos; quais os países que melhoraram, os que pioraram ou os que se mantiveram, 20 anos depois desse primeiro estudo. De salientar que são estudos diferentes, embora com objetivos idênticos, mas em que é possível identificar diferenças e tendências no posicionamento relativo.

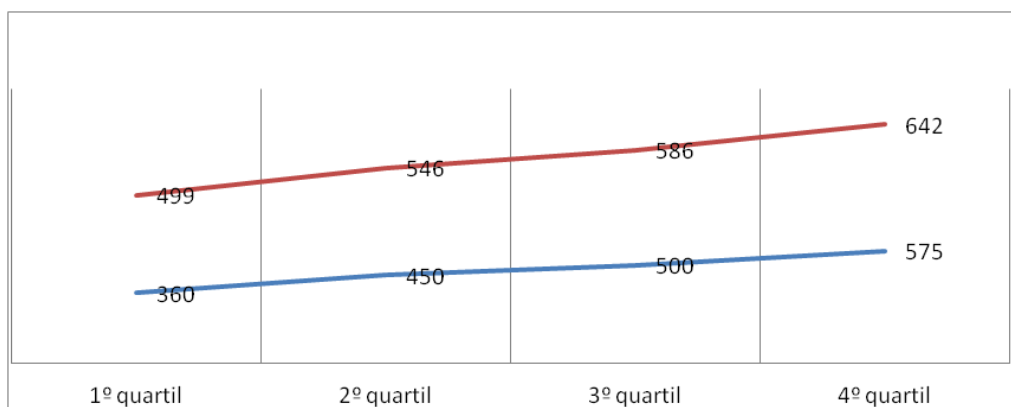


Quadro 2- Comparação de desempenhos 1991/2001 em países da UE

Como se constata, há países que ficaram praticamente na mesma, o caso da Finlândia, da Suécia ou da Bélgica, e há alguns países que melhoraram, designadamente a Holanda, a Dinamarca, a Alemanha e Portugal. Embora, como já referi, não se trate do mesmo teste e não se possa dizer que estes valores representam exatamente o mesmo, em termos de competência de leitura, os alunos portugueses parecem ter melhorado, pelo menos na sua posição relativa em relação a outros países da União Europeia.

Uma análise mais fina na comparação interna dos desempenhos nacionais no PIRLS e no *Reading Literacy* (1991) pode ser feita numa partição em quartis. Reafirmo que estamos nos dois casos a lidar com valores médios obtidos por procedimentos estatísticos em que o valor de referência é 500 em qualquer dos estudos. No que respeita ao PIRLS, uma média de 625 é considerada excelente, enquanto a de 550 é superior, a de 475 é intermédia e a de 400 é baixa.

Através do quadro 3, podemos verificar que, considerando sempre o valor central de 500, em 1991 só no 3º quartil esse valor é atingido, enquanto no PIRLS 2011 a média do 1º quartil (25% das crianças) cai nos 499 e o valor médio de desempenho das crianças no 3º quartil no PIRLS (586) não é atingido nem sequer no 4º quartil no *Reading Literacy* de 1991. Em 1991 50% dos alunos tiveram uma média entre 360 e 450. No PIRLS, 50% tiveram uma média entre 499 e 546. Finalmente, no grupo dos 25% de desempenhos superiores, obtiveram o valor de 575, em 1991, e de 642, em 2011.

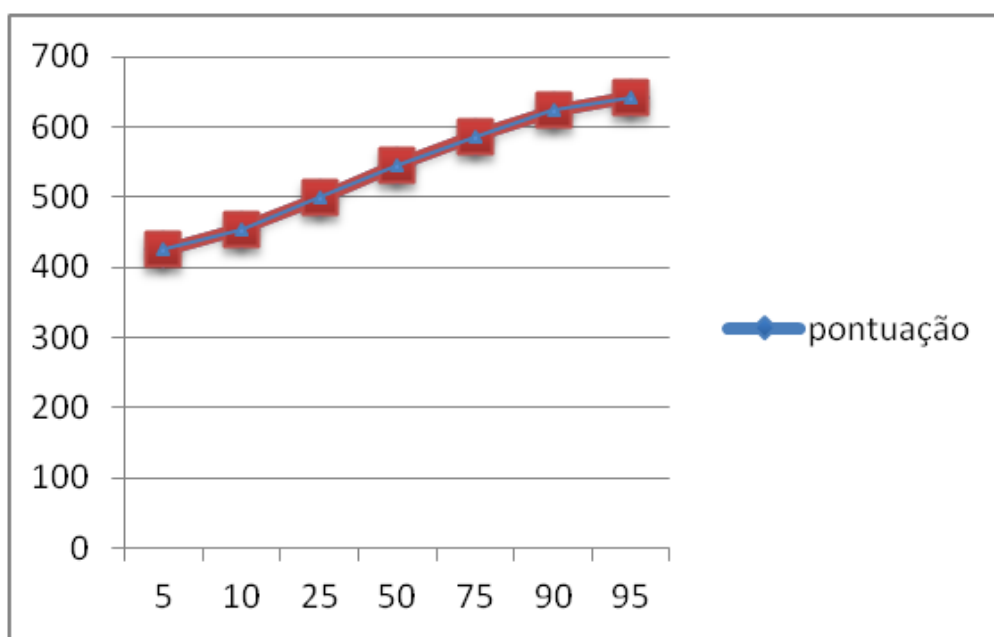


Quadro 3- Comparação de distribuições por quartis (1991/2001 __ RL, 1991_ PIRLS, 2011)

Em síntese, mesmo esquecendo a comparação com os outros países e debruçando-nos na comparação interna de desempenhos, verificamos que no PIRLS o valor central de referência (500) é atingido nos 25% de desempenhos inferiores (1º quartil).

No sentido de comparar o desempenho por quartis e deixar claro o que significam, em termos de compreensão de leitura na escala internacional, os valores atrás mencionados (excelência, nível superior, nível intermédio e nível inferior), é importante deixar-vos aqui com o significado de cada uma destas categorias. O nível de compreensão de leitura *excelente* (pontuações de 625 ou acima) significa que o leitor integra ideias e informação ao longo

dos diversos textos, identificando razões e explicações subjacentes. No nível superior (550), compreende inferências e interpreta com base em suportes textuais. No nível intermédio (475) extrai inferências diretas. E num nível inferior (400) o leitor identifica, localiza e obtém informação de diferentes partes do texto, portanto aspetos muito básicos de compreensão. É interessante associar os conteúdos de cada nível à percentagem nacional dos que o atingem. A figura 4 permite-nos essa interpretação.



Quadro 4- Partição em percentis no PIRLS
Percentil: 5-425; 10-454; 25-499;50-546;75-586;90-623;95-642

Como se verifica, só 5% dos alunos portugueses pontuam num valor médio de 425 (400 é indicativo de nível inferior de leitura). No percentil 10, a média é 454. No percentil 25, a média é de 499, acima do nível intermédio de leitura (475). No percentil 75 o valor médio é de 586, acima do valor superior de leitura (550); no percentil 90 o valor médio 623 significa que 10% dos alunos portugueses quase atingem o nível de excelência de leitura (625), com 5%, percentil 95, claramente excelentes leitores, com desempenho médio de 642 pontos.

Se abordarmos os resultados do PIRLS em termos de padrões internacionais (*benchmarks*), verificamos que 9% dos alunos portugueses participantes no PIRLS resolveram itens de dificuldade muito elevada, compreendendo de forma excelente o que leram. Por sua vez, 47% atingiram um nível superior de compreensão, 84% resolveram itens de dificuldade intermédia e 98% resolveram itens de baixa dificuldade. Numa perspetiva comparativa dentro da zona euro, aos 61% de respostas corretas dos alunos portugueses correspondem 68% dos finlandeses e 46% de alunos de Malta. Salientamos um outro dado que vale a pena olhar comparativamente a percentagem de alunos com desempenho total demasiado baixo, com pontuações obtidas nas questões de escolha múltipla e devidas ao acerto por mero acaso. Em Portugal foram identificados 1% de participantes nestas circunstâncias, enquanto em Malta se identificaram 6% dos alunos com desempenhos demasiado baixos e na Finlândia não há nenhum aluno nesta situação. Tal facto pode ajudar a explicar a constância dos bons desempenhos da Finlândia.

Uma outra perspetiva de análise são os resultados obtidos, tendo por base os propósitos de leitura, ou seja, ler para recriação, o que eu chamei ficção narrativa, porque são sempre histórias, ou ler para obter informação, textos informativos ou com instruções. Os valores nacionais globais em termos de ficção narrativa apontam para um desempenho de 538 *versus* 544 para textos informativos e com instruções (cf. Quadro 5).

Média Global 541	Processos de compreensão		Propósitos de leitura	
	Compreender inf. explícita inferências diretas (539)	Interpretar, integrar, avaliar (542)	Ficção narrativa (538)	Textos informativos/ Instrucionais (544)
raparigas	547	549	548	549
rapazes	532	535	528	539

Quadro 5- Resultados por categorias de textos e processos e sexo

Numa primeira leitura, parecia que, ao contrário de 1991, os alunos portugueses liam melhor narrativas do que textos informativos. Quando,

porém, se faz uma partição por sexo, rapazes e raparigas, o que se verifica é que para as raparigas não há diferenças entre ler ficção narrativa ou ler textos informativos. Elas leem tão bem uns quanto os outros. Por sua vez, os rapazes leem muito melhor textos informativos do que ficção narrativa, ainda que sempre com desempenhos inferiores às raparigas. A diferença global é, portanto, justificada pelo grande desequilíbrio na compreensão de textos narrativos por parte dos rapazes. Em síntese, regista-se consistentemente uma diferença positiva estatisticamente significativa no desempenho feminino.

Uma outra forma de análise dos resultados do PIRLS diz respeito aos processos envolvidos na compreensão de leitura. Quando se aborda um texto, os processos envolvidos na sua compreensão são gradativamente diferentes, podendo exigir apenas um nível básico ou requerer a mobilização de processos de nível superior. Nesta análise foram contemplados de forma ascendente de dificuldade os seguintes processos: (i) localizar e retirar informação explícita no texto; (ii) compreender inferências diretas; (iii) interpretar e integrar ideias e informação; (iv) analisar e avaliar conteúdos, linguagem e elementos textuais. O quadro 5 espelha os resultados atingidos por processos, tendo sido anexados dois a dois por níveis de complexidade. A categoria “compreender informação explícita e inferências diretas” engloba os dois níveis de processos básicos, enquanto a categoria “interpretar, integrar e avaliar” contempla os dois níveis de processos superiores.

Os resultados da população portuguesa revelam-nos que a diferença de desempenhos por categorias de processos não é estatisticamente significativa, quando cotejada com a média global obtida pelos mesmos alunos. Ou seja, quando analisamos o desempenho global este é tão bom em níveis básicos de compreensão quanto nos níveis mais altos. Porém, numa análise mais fina, e tomando em linha de conta uma partição por sexos, as diferenças também aqui surgem estatisticamente significativas. Os rapazes, com efeito, obtêm resultados muito inferiores às raparigas, ressaltando como urgente a necessidade de um maior cuidado no ensino da compreensão de leitura aos rapazes portugueses nesta etapa da escolaridade, para não

corremos o sério risco de os perder para a leitura à medida que a complexidade de textos aumentar.

Em síntese, parece evidente que, no PIRLS de 2011, os dados da população nacional do 4º ano de escolaridade, pelo menos na amostra testada, apontam para níveis de desempenhos de compreensão de leitura equiparáveis, e por vezes superiores, aos nossos parceiros europeus e substancialmente em muito melhor posição do que em estudo idêntico em 1991.

As Dúvidas - Ao mesmo tempo que julgo da máxima importância realçar as melhorias reveladas, gostaria também de partilhar convosco algumas dúvidas que o estudo internacional não esclarece e que será de toda a importância clarificar numa análise nacional dos dados. Todos os dados atrás discutidos são diretamente provenientes do relatório internacional, sendo completamente lacunar a informação nacional, quer sobre a particularidade da definição da amostra, quer sobre os procedimentos de adaptação de materiais. E é com base na falta de disponibilização da informação nacional que equaciono as dúvidas que gostaria de partilhar. Começemos pelos materiais. Referi atrás que tratando-se de um teste que avalia competências linguísticas, há que salvaguardar as especificidades e os controlos linguísticos.

Quando traduzimos um texto destinado a avaliar capacidade de leitura, é preciso ter em linha de conta o vocabulário escolhido na tradução. Basta que o vocabulário usado tenha sido mais simples do que na língua original para que a condição de compreensão de leitura seja afetada. E mais importante do que o vocabulário são as estruturas sintáticas. Se na tradução do original, neste caso do Inglês para o Português, não tiver sido tomada em linha de conta a paridade na complexidade frásica nas duas línguas, o texto traduzido pode reverter em maior ou menor facilidade de compreensão. Habitualmente para acautelar distorções desta ordem usam-se procedimentos de controlo de tradução e de contra-tradução, sempre com suporte de especialistas na língua. Se foram ou não acautelados estes procedimentos não é possível detetar nem através da descrição nacional do estudo, nem pelo cotejo de textos, na medida em que não são disponibilizados os mesmos. Seria da maior vantagem que nacionalmente fossem esclarecidos os procedimentos assumidos. A não existência desta

informação disponível, levanta-me, em termos científicos, dúvidas muito concretas.

Um outro aspeto que é interessante, e que limita o âmbito da avaliação de compreensão de leitura, diz respeito ao tipo de textos. Não se trata propriamente de dúvida em termos de investigação, mas de uma limitação que é importante revelar. Quando o material é linguístico, e estamos em presença de várias línguas, é necessário encontrar uma franja ou uma margem em que todos os países, cultural e linguisticamente diferentes se encontrem... Neste estudo, por exemplo, foi assumido que poesia não constava da avaliação da compreensão, porque dadas as características da poesia, o seu aspeto metafórico e os jogos de sonoridade da língua (rima, ritmo...) podem ficar muito enviesados quando se traduz. O controlo do possível enviesamento tem um efeito na limitação da compreensão do desempenho dos alunos, na medida em que não se testa a compreensão de poesia. Seria da máxima importância a realização de estudos paralelos com população nacional idêntica, com metodologias similares ao PIRLS e com tipologias textuais particularizadas.

Uma outra dúvida que quero partilhar convosco tem que ver com a fiabilidade e o rigor metodológico da amostra usada em Portugal. Dito de uma outra forma, a necessidade de alguns esclarecimentos sobre a representatividade da mesma, na medida em que só foi possível aceder à informação sobre a definição e construção da amostra pelo que o estudo internacional nos revela. Num primeiro olhar sobre os resultados dos alunos portugueses, um dos aspetos que me chamou a atenção foi a pouca dispersão, medida pelo desvio padrão dos valores globais na prova. Como se sabe, o desvio padrão (DP) é o indicador de variação que nos permite aferir a maior ou menor dispersão em relação à média de desempenho dos alunos. Pelo meu conhecimento da população nacional, esperaria, de facto, uma dispersão, leia-se um desvio padrão, mais alto, dadas as grandes assimetrias do país e com possíveis reflexos na variabilidade dos desempenhos atingidos. Tomando de novo em linha de conta os países da zona euro, o país com menor dispersão foi a Holanda, com um DP de 54 e o de maior dispersão Malta com um DP de 97, o que revela uma variação significativa em termos dos desempenhos da população testada. Quanto a Portugal, o DP

é de 66, com uma dispersão igual à da Itália e Alemanha e inferior à da Irlanda, Espanha e França. Este facto motivou-me para procurar mais dados sobre a construção da amostra.

O que nos diz o relatório internacional, é que na amostra portuguesa foram seleccionadas 148 escolas, 132 escolas públicas e 16 escolas privadas. É também afirmado que foram excluídas as escolas pequenas. Nada porém é esclarecido sobre a dimensão, em número de alunos e, conseqüentemente, de docentes, no conceito de “escolas pequenas”. Sabendo nós que as escolas pequenas existem no nosso país em zonas rurais, e que habitualmente os desempenhos afetados pela ruralidade tendem a corresponder a valores mais baixos de resultados escolares, em que o peso do fator social do meio de pertença não é despreciando, levanto a questão do efeito dessa exclusão nos resultados nacionais dos alunos.

Uma outra dúvida que o relatório internacional não esclarece diz respeito às variáveis de estratificação da amostra. No referido relatório são mencionadas 12 regiões, não se percebendo, contudo, que tipo de regiões e como é que certas variáveis geográficas e sociais foram contempladas. Também não é possível discernir se a partição em regiões, quaisquer que elas tenham sido, contemplaram também as escolas privadas ou se serviram apenas para a identificação das escolas públicas. Com vista a um conhecimento mais focado, seria da máxima importância que os dados nacionais contemplassem análises mais finas e com base em variáveis já muito experimentadas na população portuguesa. Daí a importância de se saber quais as 12 regiões contempladas. Uma outra questão relacionada com as regiões e com particularidades da estratificação prende-se com o modo como foram seleccionadas as escolas dentro de cada região. Como sabemos, mesmo pertencendo à mesma área geográfica as escolas apresentam diferenças significativas; na escolha das escolas, para garantias de representatividade da população, seria importante conhecer os controlos de estratificação dentro das regiões geográficas. Ainda sobre a seleção da amostra, o relatório internacional também não explica como foram escolhidas as turmas dentro de cada escola. São pormenores da maior importância quando se pretende compreender e interpretar com transparência o valor do desempenho dos alunos nacionais.

Na tentativa de um maior esclarecimento sobre a representatividade e fiabilidade da amostra, procurei informação suplementar sobre o contexto familiar dos alunos testados, obtida através do questionário respondido pelos pais. No sentido de caracterizar a população testada foram identificados os recursos disponíveis no agregado familiar de cada aluno. De acordo com uma classificação prévia, foram estabelecidas três níveis de recursos: (a) muitos recursos; (b) alguns recursos; (c) poucos recursos. Os recursos identificados referem-se ao nível de diploma de um dos pais (universitário; secundário ou médio; menos que secundário), à profissão dos mesmos, ao número de livros existentes em casa (infantis e não infantis) e à existência ou não de quarto próprio para a criança e à disponibilidade de serviços de internet. A síntese desses dados é apresentada no quadro 6. A correlação com o desempenho no PIRLS é a esperada; quanto mais recursos disponíveis, melhores os desempenhos conseguidos. A distribuição global de recursos pela população na amostra (16% com muitos recursos; 75% com alguns recursos e 9% com poucos recursos) requer uma análise mais fina e algum cotejo com outros dados nacionais para confirmar a representatividade da amostra. O indicador mais fácil de obter comparações prende-se com o nível de escolarização dos adultos. Nesta amostra, 25% das crianças apresentam um dos progenitores, segundo o pai respondente, com a titularidade de um diploma universitário. Tentámos obter dados que nos confirmem se na população nacional o grau de escolaridade corresponde a essa titularidade, tendo em comparação a população da mesma faixa etária que os pais em questão, i.e., entre os 30 e os 44 anos. Nos dados da PORDATA de 2010, apenas 11,8% da população portuguesa total é titular de diploma académico de nível superior; por sua vez, em 2008, os dados do *Eurostat Labour Force Survey* apontam para uma variação entre os 19,8%, da população portuguesa entre os 30- 34 anos, 14,9% entre os 35-39 anos e 12,2% entre os 40-44 anos com diploma académico de nível superior. Os dados mais próximos dos revelados no PIRLS (25%) parecem estar no CENSOS 2011, em que 24,63% da população portuguesa entre os 25-44 anos surgem com a frequência do ensino superior. Apenas uma nota, tendo nós uma faixa etária dos pais que certamente não desce abaixo dos 30 anos, e embora frequência de ensino superior não signifique titularidade de diploma, é provável que os valores do CENSOS se aproximem dos 25% do

PIRLS, ao contrário dos dados revelados por outras bases de dados. É muito provável que na amostra testada 25% os pais garantissem ser titulares de diploma de ensino superior, a questão permanece em aberto sobre a veracidade das respostas e sobre a generalização à população.

Nível de recursos	Qualificação de recursos	PIRLS
Muitos recursos 16% alunos	Pai/mãe diploma universitário 25% Pai/mãe profissão diferenciada 33% + 100 livros em casa 21% +25 livros infantis 63% Quarto próprio e internet 63%	578
Alguns recursos 75%	Pai/mãe diploma secundário/médio Pai/mãe pequeno negócio/terciária Entre 25-100 livros em casa Entre 10-25 livros infantis Quarto próprio ou internet	541
Poucos recursos 9%	Pai/mãe menos que secundário/médio Pai/mãe profissão indiferenciada Menos de 25 livros em casa 10 ou menos livros infantis Nem quarto próprio nem internet	508

Quadro 6- Recursos familiares e resultados no PIRLS

Os dividendos – Procurar dividendos nos resultados do PIRLS significa aqui, por um lado, procurar fatores responsáveis pelos resultados atuais e, por outro, identificar o que os atuais resultados nos podem ensinar para o futuro. Por outras palavras, o que contribuiu para chegarmos onde chegámos e que políticas, ou que particularidades das políticas educativas devem ser seguidas num futuro imediato ou próximo.

Tentemos identificar fatores que podem ter afetado positivamente os resultados. Parece evidente, quer pelo senso comum, quer pelos dados que a investigação vai recolhendo, que os alunos aprendem porque os professores os ensinam, independentemente da maior ou menor disponibilidade de

recursos. A valorização dos docentes é uma forma eficaz de afetar as aprendizagens.

Partindo dos dados disponibilizados pelo relatório internacional, procuremos caracterizar os docentes das crianças participantes no estudo. Começamos pela caracterização da formação académica e profissional dos docentes, cotejando momentos diferentes da nossa história recente. Voltemos à comparação com 1991, aquando da realização do já mencionado *Reading Literacy*. Nesse momento, 63% dos docentes possuía até 11 anos de escolaridade, a escolaridade primária, secundária e média. Portanto quatro anos de primária, cinco de liceu e dois de Magistério Primário e apenas 37% dos docentes tinha frequentado a escola por mais de 11 anos (quatro de primária, sete de liceu e dois de Magistério Primário). O PIRLS 2011, revela-nos que 91% dos alunos têm professores com o nível de licenciatura, 3% com o nível de pós-graduação e 6% sem formação de nível superior. Este último valor parece muito estranho e poderá referir-se ou a docentes de escolas privadas ou a antigos professores que não realizaram complementos de formação para equiparação a licenciatura. É claramente evidente que o nível de escolaridade dos atuais docentes é superior ao do passado. Talvez que mais anos de escolaridade possam ter influenciado o nível de docência e quiçá as aprendizagens dos alunos, refletidas na diferença de desempenhos nos dois estudos (1991 e 2011).

Foquemo-nos agora na formação profissional nos cursos de formação inicial de professores. Em 1994¹, ou seja, utilizando dados de 1991, decidi caracterizar os docentes cujos alunos tinham participado no *Reading Literacy* e que responderam a um questionário de contexto e, em simultâneo, na medida em que tinham passado a existir as Escolas Superiores de Educação que ministravam cursos de bacharelato para professores do então ensino primário, realizar junto de todas estas instituições um inquérito para caracterizar a formação oferecida para o ensino da língua portuguesa, designadamente a formação específica em didática do Português. Responderam a esse inquérito 19 instituições de ensino superior, públicas e privadas. Posteriormente, em 2000, todos os

¹ Sim-Sim, I. (1994). Os professores que hoje temos e os professores que vamos ter amanhã. In *Inovação*, vol.7, 2, 145-171.

cursos de formação inicial para professores do Ensino Básico eram obrigatoriamente do nível de licenciatura. Por sua vez, o novo regime de formação de professores exige atualmente a titularidade do grau de mestre para lecionar o 1º ciclo do Ensino Básico. O quadro 7 mostra as horas de formação em 1994, em 2000 e em 2011 especificamente para o ensino do Português e da didática da leitura. Os dados de 2000, já com a obrigatoriedade da licenciatura, não correspondem a nenhum inquérito realizado junto das instituições de ensino superior, mas espelham um exemplo recolhido da informação pedida à ESE de Lisboa que traduziu os ECT em horas presenciais, quer na formação específica em Português, quer em didática do Português. O mesmo se passa com a informação de 2011.

	1994 Bacharelato/Licenciatura para variante do 2º ciclo Inquérito a todas as escolas de formação R - 19)	2000 Licenciatura em 1º ciclo do Ensino Básico (ESE LX)	2011 Licenciatura em Ensino Básico+ mestrado em 1º ciclo EB (ESE LX)
Formação específica em Português*	S/I	180h	260,5h
Formação específica em didática do Português *	<ul style="list-style-type: none"> • < 90h - 5 instituições • 91h-200h -11 instituições 	97,5h	135h
Total de horas em Português e didática do Português	Entre 90 e 200h	277,5h	395h

Quadro 7- Formação profissional inicial para o ensino do Português
*S/I – sem informação; **3 instituições não responderam a esta questão

Pelo que se constata, em 1994, cinco das instituições (ESE e Universidades) afirmaram ter no respetivo plano de estudos menos de 90 horas em formação didática do Português, contra 91 a 200 horas de formação em 11 das referidas instituições. Por sua vez, em 2000, reporta-se o exemplo da ESE de Lisboa, com um total de 277,5h de ensino presencial em Português e na respetiva didática. Embora nenhum aluno que realizou o PIRLS tenha beneficiado da formação mencionada em 2011, podemos exemplificar o que

passa numa ESE em que a carga presencial de formação específica em Português e didática do Português sobe para 395h. É provável que o aumento da formação inicial em língua portuguesa e no seu ensino afetem o desempenho de leitura dos alunos.

Discutir a formação profissional implica não nos quedarmos na formação inicial obtida. A formação em exercício pode ter um efeito na melhoria das práticas de ensino. No que se refere aos efeitos das horas de formação contínua em ensino da leitura, tem sido consistentemente comprovado que menos de 16 de horas de formação não parece afetar a melhoria do ensino e, conseqüentemente, o desempenho dos alunos. Nos resultados do PIRLS, expostos no quadro 8, é possível verificar que, com efeito, os 45% dos alunos com melhores desempenhos (545) pertenciam a turmas em que os professores tinham usufruído nos dois últimos anos de 16 ou mais horas de formação contínua em ensino da leitura. Nas turmas em que os docentes não tiveram qualquer tempo de formação ou menos de 6 horas, os resultados são significativamente inferiores (537 e 536). O que podemos retirar destes dados é que para que o desenvolvimento profissional seja afetado pela formação contínua, e que tenha efeito nas aprendizagens, requer mais do que ações esporádicas e atomísticas, pelo menos no que à aprendizagem da leitura diz respeito.

Horas de formação contínua em leitura	Desempenho dos alunos no PIRLS
16 horas ou mais 45% alunos	545
Algum tempo, mas menos de 6 horas 36% alunos	537
Nenhum tempo de formação 19% alunos	536

Quadro 8- Horas de formação contínua e resultados no PIRLS

Concomitantemente com a formação académica e profissional, o PIRLS tentou perceber até que ponto o grau de satisfação profissional dos docentes podia influenciar a aprendizagem dos alunos. As respostas obtidas no questionário aos docentes revelaram que 36% dos alunos portugueses eram lecionados por professores satisfeitos com a sua profissão, 59% por

professores com alguma satisfação e apenas 5% das crianças tinham professores pouco ou nada satisfeitos. No quadro 9, verificamos que parece existir uma relação direta entre o grau de satisfação docente e os resultados dos desempenhos dos alunos.

Grau de satisfação dos docentes	Desempenho dos alunos no PIRLS
Satisfeito – 36% dos alunos	547
Alguma satisfação 59% alunos	539
Pouco ou nada satisfeito 5% alunos	527

Quadro 9- Grau de satisfação dos docentes e resultados no PIRLS

Para além da formação e satisfação docente, um outro fator que parece influenciar os desempenhos de aprendizagem está relacionado com a frequência da Educação Pré-escolar. No sentido de obter uma maior fiabilidade nas respostas, foi perguntado aos pais de cada criança se a mesma tinha frequentado a Educação Pré-escolar e durante quanto tempo, antes da entrada no 1º ano de escolaridade. De salientar que a Educação Pré-escolar foi objeto de uma considerável expansão nos últimos 20 anos, o que se reflete na frequência de uma grande maioria das crianças participantes no PIRLS. A relação entre essa frequência e os resultados do PIRLS está espelhada no quadro 10 e parece ser um fator poderoso no desempenho da leitura.

Frequência de pré-escolar	Desempenho dos alunos no PIRLS
3 ou + anos - 45% dos alunos	549
1-3 anos - 37% dos alunos	544
< 1 ano - 8% dos alunos	533
Nunca - 9%	522

Quadro 10- Frequência da Educação Pré-escolar e resultados no PIRLS

Contudo, para que a frequência do pré-escolar afete as posteriores aprendizagens de leitura é necessário mais do que a mera presença num jardim-de-infância. Está sobejamente confirmado pela investigação que esse

efeito beneficia principalmente com atividades linguísticas direcionadas para a emergência da leitura e da escrita. Nesse sentido foi perguntado também aos pais de cada a criança, se, quando o filho entrou no 1º ano de escolaridade, a criança realizava cinco determinadas tarefas que envolviam conhecimento precoce sobre leitura e escrita. Dos pais inquiridos, 12% revelaram que as crianças realizavam as cinco tarefas muito bem, 45% razoavelmente e 42% nada bem. A relação entre esses conhecimentos precoces e os desempenhos no PIRLS está patente no quadro 11.

Conhecimentos precoces sobre leitura e escrita	Desempenho dos alunos no PIRLS
Muito bons - 12% dos alunos	561
Razoáveis - 45% dos alunos	549
Poucos - 42% dos alunos	532

Quadro 11- Conhecimentos de leitura e escrita antes do ensino formal e desempenho no PIRLS

Em síntese, para além das características e do grau de satisfação profissional dos docentes é evidente o efeito da frequência do pré-escolar e da estimulação de conhecimentos precoces de leitura e escrita, antes do ensino formal da leitura, no desempenho dos alunos no PIRLS. O investimento, nos últimos vinte anos, em políticas educativas de formação de professores e de expansão e melhoria da qualidade da Educação Pré-escolar mostram reverter em dividendos mensuráveis nos resultados do PIRLS.

Finalmente, dada a importância revelada pela frequência da Educação Pré-Escolar nos resultados do PIRLS, gostaria de corrigir um erro presente no relatório internacional do PIRLS, provavelmente, por lapso da informação fornecida nacionalmente. Na sequência da questão 6 do *Questionário sobre Currículo*, é omitida a existência de orientações curriculares na Educação pré-escolar para a estimulação de domínios /atividades linguísticas. No sentido de repor a veracidade da informação, recordo que nas *Orientações Curriculares para a Educação Pré-Escolar* (1997) existe especificamente um *Domínio da linguagem oral e abordagem à escrita* que espelham a importância que política, pedagógica e investigativamente tem sido dada a estas áreas nas últimas décadas. Certamente que as relações evidenciadas

nos quadros 10 e 11 dos parágrafos anteriores não são alheias a estas opções na política curricular.

Gostaria, para terminar, de um olhar futuro para o que o PIRLS nos pode ainda fornecer para um conhecimento mais aprofundado da nossa própria realidade e das orientações de política educativa que nos levem a consolidar e melhorar desempenhos escolares dos alunos no que respeita ao domínio da leitura. Como referi, toda a informação recolhida para esta comunicação é proveniente do relatório internacional. Daí que tenham ficado por esclarecer muitos aspetos e algumas dúvidas que só estudos nacionais podem dar resposta. Seria da máxima importância a disponibilização de todos os dados nacionais, da informação específica da construção da amostra e dos processos usados na tradução/adaptação dos textos da prova. Só essa divulgação permite potenciar todo o esforço humano e financeiro despendido na participação deste estudo internacional, através da promoção de estudos específicos que olhem para os dados nacionais, salientando as características da amostra, no que respeita, por exemplo, ao rural e urbano, à dimensão das escolas, às características dos docentes, aos tipos de textos e questões. O que vos mostrei são dados globais, mas se queremos que esses resultados melhorem a formação dos docentes e, conseqüentemente as aprendizagens dos alunos, é necessário partir para uma pesquisa pormenorizada e ancorada na nossa realidade social e escolar. Num tempo em que se valorizam tanto as avaliações e os exames, será da máxima perspicácia aprender com o PIRLS sobre a construção de itens de avaliação de compreensão de leitura e deixar que essa aprendizagem contamine as avaliações nacionais, de classe e de escola. Se os dados do PIRLS ficarem arrumados num qualquer repositório, físico ou virtual, sem que deles se potencializem todos os dividendos possíveis, o país perderá não só a possibilidade de se conhecer, mas, principalmente, de melhorar.

Um outro olhar sobre o TIMMS 2011
Presidente da Mesa - Jorge Marques da Silva

SEMINÁRIO

Duas ideias principais presidiram à incorporação desta sessão no programa: por um lado, avaliar o impacto que os resultados dos alunos portugueses têm no desenvolvimento do sistema científico português; por outro lado, discutir o papel que a comunidade científica nacional pode ter no progresso dos resultados dos alunos portugueses em ciência e em matemática. Em particular, neste segundo aspeto, valerá a pena discutirmos o papel que a ciência e os cientistas podem ter na formação de professores, sobretudo no ensino e formação ao longo da vida, bem como na elevação da cultura científica dos portugueses (e os dados do TIMMS assinalam claramente a correlação entre os resultados escolares dos alunos e o ambiente científico-cultural em que estão inseridos).

O Conselho Nacional de Educação segue com atenção os programas de avaliação de processos e resultados em educação, muito em particular o programa nacional de Avaliação Externa das Escolas (AEE) em cuja avaliação tem, aliás, responsabilidades legalmente consagradas. Fá-lo pela convicção de que é necessário avaliar para conhecer e conhecer para intervir. Isto é, a avaliação deve consagrar-se como um processo estruturante para o desenho e implementação de políticas. Para cumprir eficazmente esse papel, porém, os processos de avaliação em educação têm que reunir várias condições. Desde logo, têm que ser competentes, têm que ter qualidade. Uma avaliação mal feita pode ter efeitos muito perniciosos. Na ausência de avaliação, um decisor procede com cautela, age com a precaução que a incerteza impõe. Na posse de resultados de uma avaliação errada, um decisor procede com convicção, na direção errada. Em meu entender, portanto, é preferível não avaliar, a avaliar mal. Importa, portanto, avaliar com qualidade. Neste contexto, valerá a pena analisar o TIMMS, para além dos resultados, pelo processo em si.

Outro aspeto relevante nos processos de avaliação em educação é a sua continuidade temporal, a possibilidade de fornecer séries temporais consistentes. Uma análise pontual dos sistemas, não sendo inútil –

¹ Conselho Nacional de Educação

proporciona uma radiografia com potencial diagnosticante – não permite, contudo, avaliar as tendências de evolução, essenciais para as definições de políticas educativas. Esta necessidade de proporcionar leituras continuadas - sem prejuízo da evolução dos sistemas educativos, na permanente busca da máxima qualidade – tem sido recorrentemente expressa pelo CNE, em particular nos pareceres que tem emitido sobre o programa de AEE. O TIMMS opera desde 1995, proporcionando aos países que nele participaram regularmente um riquíssimo manancial de informações sobre os seus progressos ou recuos em matéria de educação. Infelizmente, a participação de Portugal tem sido pontual e algo errática. Participou no TIMMS matemática em 1995, com alunos de 4º e 8º anos, e só voltou a participar em 2011, e apenas com alunos de 4º ano. Importa, parece-me, que Portugal selecione os programas de avaliação internacional em que quer participar e mantenha, aí, presença regular.

Os relatórios em discussão nesta sessão, o TIMMS Ciências e o TIMMS Matemática 2011, estão repletos de informação muito relevante sobre a performance, absoluta e relativa, dos alunos portugueses de 4º ano. Mas sobre os seus conteúdos, não me vou pronunciar, deixo-o para os nossos convidados.

Começo por agradecer ao CNE este gentil convite que me endereçou. Espero poder vê-lo como um reconhecimento - ou pelo menos como um sinal - de que este Conselho também considera importante ouvir os especialistas das disciplinas.

Digo-o no sentido dos especialistas científicos das disciplinas, porque há uma grande comunidade de matemáticos, de físicos, de biólogos e de outros cientistas que, não sendo especialistas das Ciências da Educação, tem, ao longo dos anos, manifestado um grande interesse por estes assuntos, cujo contributo é certamente valioso para completar a discussão destas importantes problemáticas.

Confesso que não tinha tido ainda acesso ao texto do Prof. Jorge Marques da Silva, mas fico bastante satisfeito porque ele coincide em vários pontos com o que eu gostaria de partilhar convosco.

Não resisto à tentação, já que há poucas oportunidades para os cientistas se manifestarem neste tipo de fórum, de trazer até vós alguns matemáticos profissionais que tiveram um papel muito significativo na Educação, quer através de reformas que lideraram, quer através de outras iniciativas, como a formação de professores.

O Professor Ron Aharoni é um matemático profissional que resolveu uma conjectura em Matemática muito difícil, chamada conjectura de Ryser, um problema de combinatória que esteve em aberto durante muitos anos. Para além deste currículo científico invejável, o Professor Ron Aharoni é simultaneamente considerado, a nível mundial, um dos maiores especialistas do ensino da Matemática no 1.º ciclo. Aliás, foi durante algum tempo professor do 1.º ciclo. Escreveu o livro *Aritmética para Pais* que é um *best-seller* a nível internacional e uma referência para muitos professores e encarregados de educação. Foi traduzido para Português há pouco tempo, pela Sociedade Portuguesa de Matemática.

¹ Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa

Vou completar com outro matemático muito importante, um matemático português do século XX, também de renome internacional, o Professor José Sebastião e Silva. Desenvolveu uma teoria estruturante da Análise Matemática, a teoria das distribuições. Infelizmente, fê-lo em simultâneo com outro matemático, o francês Laurent Schwartz, e foi a abordagem deste último que acabou por vingar. No entanto, nos meios cultos, é tido como um dos autores de uma das principais ferramentas de Análise Matemática dos séculos XX e XXI. Teve um papel muito importante no ensino secundário no nosso país, liderou uma reforma piloto que infelizmente nunca foi concluída, no seguimento do seu desaparecimento prematuro, em 1972. Deixou de certa maneira um processo inacabado, mas a verdade é que teve e tem uma enorme influência. Ainda hoje, muitas pessoas que frequentaram estas turmas-piloto, independentemente da sua atual profissão, referem a importância dessa iniciativa.

Um dia, se for convidado para falar especificamente dos matemáticos na Educação, certamente que começarei pelo próprio Euclides, porque, de facto, e contrariamente aos meus colegas das Ciências mais experimentais, os matemáticos são todos professores também. É muito raro nós vermos um matemático que é apenas investigador. Todos ou quase todos os matemáticos têm uma experiência letiva e, por essa razão, se vão interessando, em maior ou menor grau, por estes assuntos.

Relativamente ao TIMSS, vejo que há, neste momento, um grande interesse por este tipo de estudo: estruturado, com medidas precisas e rigorosas. Durante muito tempo, estes estudos foram muito criticados e desvalorizados. Considero muito positiva esta clara mudança de opinião, materializada nas intervenções de alguns oradores desta manhã. Desde o estudo de 1995, sempre que se trazia este assunto, as conclusões que o TIMSS revelava eram imediatamente postas de lado, como se o estudo tivesse todos os vícios e mais alguns, como se não se pudessem medir as coisas nestes termos. Mas vejo que alguma opinião está a mudar e que, neste momento, os cientistas da educação olham com muita atenção para este tipo de resultados. É importante porque, sendo a educação um sistema extremamente complexo, sem medidas rigorosas é muito difícil

conseguirmos perceber o alcance das políticas educativas e qual a evolução exata do sistema, no sentido de o tentar empurrar na boa direção.

É justamente esse o primeiro facto que tenho a lamentar. Para estudar sistemas complexos precisamos de séries temporais minimamente reguladas e, infelizmente, o que nós temos neste momento é um ponto em 1995 e um outro ponto em 2011, 16 anos depois.

Ora, se o PISA e o TIMSS se repetem com uma certa periodicidade, é por uma razão muito específica: só podemos ter uma ideia do que se está a passar no sistema se o acompanharmos ao longo do tempo. Com 16 anos de intervalo é muito difícil mas tem este interesse da radiografia. Dá-nos, no fundo, um vislumbre da situação atual. Mas saber o que é que levou a esta melhoria enorme entre 1995 e 2011 é, francamente, mais difícil.

Reparem que para estes sistemas temos dois tipos de problemas. Primeiro, temos de identificar variáveis que estejam interligadas, que pareçam funcionar em harmonia. A isto chama-se correlação. Mas ainda há uma dificuldade suplementar. Não é porque vemos duas variáveis que estão correlacionadas que obrigatoriamente uma é a consequência da outra. Aliás, os estatísticos gostam muito de uma frase feita: em Estatística, «correlação não é causalidade».

Com apenas dois pontos, um em 1995 e outro em 2011, tudo isto é extremamente complicado. Não sei se alguém na assistência tem ideia de quantos ajustamentos de programas, de currículos e outras reformas existiram, em 16 anos. Não sei se têm ideia de quantos responsáveis políticos exerceram diretamente a sua ação no sistema educativo. Até uma pergunta mais fácil: não sei se alguém sabe quantos ministros da Educação existiram entre 1995 e 2011. Eu fiz o exercício e posso-vos dizer que são muitos, foram nove. Assim, neste contexto, é muito difícil sabermos exatamente o que foi positivo para o sistema, o que foi negativo, e em que altura. E muitas vezes caímos no erro de considerar que foi apenas o passado muito recente que influenciou o sistema.

Há cerca de sete ou oito anos começou a aparecer a “moda” de se pensar que uma medida de política educativa tinha uma ação imediata no sistema. Era frequente ouvir argumentos como: «os resultados dos exames são estes

porque há seis meses fez-se isto ou aquilo, ou porque há um ano se implementou tal coisa, ou porque há um mês tal pessoa fez tal afirmação na televisão.»

Começámos com uma visão muito, muito restrita do que leva efetivamente os alunos a melhorar ou a piorar. Reparem que a intervenção do Professor Pedro da Ponte ligou essencialmente esta melhoria registada no TIMSS a eventos que se passaram todos a partir de 2007/2008: referiu o Plano de Ação da Matemática e referiu o Programa de que é coordenador. Fui entretanto recolher alguns dados, porque fiquei um pouco espantado com este argumento do Novo Programa de Matemática. Foi implementado em 2008/2009 em dez turmas-piloto. No país inteiro, em dez turmas apenas. Em 2009/2010 ainda não era obrigatório. Dos dados que consegui recolher, tenho a informação de que se candidataram sensivelmente um terço das escolas do país. O Programa só se tornou obrigatório em 2010/2011, e a sua implementação foi faseada a partir desta data. O TIMSS é de 2011, portanto, na melhor das hipóteses, os alunos que desde a primeira hora tiveram professores entusiasmados com o programa, no máximo tiveram dois ou três trimestres de contacto com este novo currículo.

Será realmente razoável achar que houve um salto significativo ao nível dos resultados do TIMSS por causa do novo Programa de Matemática? É a pergunta que eu vos deixo.

Curiosamente, penso que foram os políticos, e não os académicos, que estrearam estas linhas de argumentação imediatistas. Como os ciclos políticos duram quatro ou cinco anos e os políticos gostam, naturalmente, de demonstrar que tomaram medidas acertadas, tentam, por vezes, tirar conclusões de forma um pouco precipitada. É por esta razão que muitos defendem que em Educação deve existir um pacto de regime. Um ciclo político não é suficiente para se verem efetivamente resultados na educação.

Em intervenções anteriores foram dadas algumas opiniões – que não foram mais do que isso – sobre o porquê desta grande evolução. Eu também tenho algumas suspeições.

Devo dizer-vos que, para mim, o Professor David Justino é uma referência. Trouxe mais exigência e veio contrariar uma certa tendência que havia no

país contra a avaliação, contra a mensurabilidade da escola, dos alunos, das performances. Em particular, introduziu os exames nacionais no Ensino Básico, o que teve um efeito muito positivo. E todos estes eventos decorreram há cerca de dez anos.

Outro fator que penso que tenha tido boa influência foi aquilo a que chamo a «*perestroika*» do Gabinete de Avaliação Educacional (GAVE). Este organismo fornecia poucos dados de *feedback* aos professores. A partir de 2009, 2010 começaram a aparecer dados e penso que isso foi muito positivo, porque os sistemas melhoram quando os alunos e as escolas recebem indicações sobre a *performance* dos seus alunos.

Estamos agora a ver o que é uma série temporal mais completa relativa à Rússia e à República da Coreia onde se percebe um bocadinho melhor a evolução. Imaginem se nestes países só tivessem os dados de 1995 e de 2011, certamente na Rússia ninguém pensaria que tiveram ali uma quebra muito acentuada. A Coreia subiu e está a subir de maneira consistente ao longo dos anos, não é um “pulo”. Aliás, este é outro aspeto que queria partilhar convosco. Estes sistemas não são “sistema de pulo”. Não existem “guinadas”. Isso acontece quando há provas mal calibradas e quando os exames não são equivalentes entre duas aplicações. Mas não em provas como as do PISA ou do TIMSS que são feitas com todo o cuidado, com itens bem estudados e calibrados. Estes perfis são aquilo que nós chamamos em Matemática de baixas frequências. São ondas que oscilam suavemente no tempo mas que não variam violentamente em curtos lapsos de tempo. Na evolução da Rússia, temos separações de dois pontos, depois 14 para baixo, 4 para cima, e finalmente 20 para cima. Não há “pulos” de 80 ou 90 pontos separados por 4/8 anos.

Eu garanto-vos que não foi de 2006 até agora - não pode ter sido - que isso sucedeu em Portugal.

Gostaria agora de chamar a vossa atenção para um aspeto muito importante que se pode concluir destas aplicações do TIMSS e que, infelizmente, não é discutido, nem na imprensa nem neste tipo de fóruns. Trata-se de uma situação que começa a preocupar muita gente na Europa. A este respeito, poderei falar um pouco mais tarde da experiência que tenho num grupo de

trabalho da União Europeia, de que faço parte e onde costumo discutir estes assuntos com colegas de outros países europeus.

Neste momento, temos os países asiáticos a afastar-se de forma consistente dos restantes países ocidentais. Reparem que já não estamos com dois pontos separados por 16 anos. Já temos aqui uma sequência temporal mais completa. Já se consegue discernir uma tendência que, como expliquei, que não é “de guinada”. É algo de consistente, de suave no tempo. E o que nós vemos realmente é esta ilha dos países asiáticos a fugir.

Alguém disse, noutra comunicação, que nem vale a pena falar dos países asiáticos porque são “extraterrestres”. Eu lamento informar que não são extraterrestres. Vivem no mesmo planeta que nós, competem pelos mesmos recursos, competem na mesma economia global e competem connosco em todos os ramos da atividade humana.

Hoje, abrem-se os jornais e lê-se por todo o lado a palavra “competitividade”. É pena que na Educação se entreguem as armas logo de início dizendo “não faz mal, os países asiáticos são diferentes”. É verdade que são diferentes, mas as diferenças culturais não explicam que os dois blocos se comecem a separar tal como está a acontecer.

Reparem que em 1995 asiáticos já estavam realmente no topo, é verdade, mas ainda havia aqui um contínuo quando comparados com os países europeus: Hong Kong com 547, a Holanda e a República Checa com 549 e 541 pontos, respetivamente. Não se vê aqui uma clara separação.

Se observarmos, à medida que o tempo passa, começamos a assistir à abertura de um fosso. Aqui já são 24 pontos de diferença entre Taipé e a Holanda, ali são 19 e aqui já são 40 pontos. Portanto, está aqui a acontecer qualquer coisa.

Notem que todos estes dados são muito macroscópicos, constituem o resultado final, médio, dos diferentes países. Se formos estudar os *benchmarks* – que são, no fundo, permitam-me o abuso, “níveis de desempenho” –, a situação é ainda mais dramática. Gosto do resumo que o Professor Pedro da Ponte hoje fez do segundo *benchmark*: neste nível, os alunos resolvem apenas problemas de resposta imediata. Foi o que ele

escreveu. A descrição oficial é um pouco mais completa, mas este parece-me ser um bom resumo.

E o grave é o seguinte: não há um único país ocidental que tenha menos de 50% dos alunos estagnados neste nível de só responderem a perguntas imediatas. Eu posso mostrar aqui alguns dados muito interessantes: nos países asiáticos, praticamente 80% dos alunos conseguem ultrapassar este nível. Ou seja, onde nós temos mais de metade dos alunos bloqueados (60% no caso de Portugal), os países asiáticos têm apenas 20%.

Temos de ser realistas, 60% dos nossos alunos que terminam o 1.º ciclo só conseguem responder a questões triviais. Isto não é suficientemente dito e penso que tem de nos preocupar muito, independentemente dos grandes progressos que fazemos face aos nossos parceiros europeus e pura e simplesmente ignorando o resto do mundo. Temos que levar este problema muito a sério, enquanto país e enquanto União Europeia. Os países asiáticos conseguem colocar cerca de 40% dos seus alunos no nível 4, que é o *benchmark* mais elevado.

Mal passamos para a Europa este valor cai drasticamente. O país que tem a melhor performance é a Inglaterra com 18%. Em Portugal são 8%. Vi na comunicação social que 8% era excelente, porque em 1995 eram só 4% e portanto “duplicámos”. Penso que temos que ser um bocadinho mais sérios do que isto: 8% não é excelente, é mesmo muito fraco, há muito trabalho para fazer.

Aliás, o próprio relatório final do TIMSS chama a atenção para estes factos que aqui relato, o que é curioso. Cada parágrafo do relatório começa com um texto que no fundo resume o que é mais importante em cada capítulo. No capítulo dos *benchmarks* podem ler que os cinco países asiáticos se distanciaram completamente dos seus parceiros.

Esta realidade já se vê no 4.º ano, mas no 8.º ano ela é elevada para níveis insustentáveis. No 8.º ano, temos quase 50% dos alunos asiáticos no nível mais elevado. O que significa isto? Significa que o TIMSS já nem é adequado para estes alunos, tornou-se demasiado fácil. Portugal não participou, como foi dito, no estudo do 8.º ano, mas podemos ver os resultados dos países parecidos com o nosso, os países europeus.

Reparem que temos uma progressão suave dos diferentes indicadores, como não me canso de referir. A percentagem de alunos no nível mais avançado evoluiu, por exemplo na Ilha Formosa, da seguinte forma: 37%-38%,-45%-49%. Em Singapura essa série é a seguinte: 40%-42%-44%-40%-48%. Neste momento, como se pode ver, praticamente 50% dos alunos asiáticos estão naquele nível. Mal se passa para o mundo ocidental, estes valores passam a ter apenas um dígito: 9%, 8%, 7%, 8%, 10%, 5%...etc. Não tracei a linha porque considerei que nem era necessário, ela é demasiado evidente.

Não sei qual é a razão exata deste fenómeno, mas posso falar-vos um bocadinho da minha experiência na União Europeia, onde começam a aparecer várias pessoas a questionar o modelo que temos e que continuamos a apregoar, muito baseado no “ensino pela descoberta”, na “avaliação formativa”, no “ensino centrado no aluno”, numa quantidade de ideias que tornam o ensino, e a respetiva avaliação, pouco rigoroso e demasiado permissivo. A palavra *accountability* começa a aparecer muito nestes círculos. Os alunos devem ser responsabilizados pela falta de estudo, devem ser responsabilizados pela falta de resultados, dentro, obviamente, da especificidade de cada um. Ainda há pouco tempo, em mais uma sessão de trabalho em que no fim do *workshop* as pessoas vieram com as mesmas conclusões que tiram há mais de 20 ou de 30 anos, do professor não como um professor mas como um amigo ou” um facilitador de aprendizagens” e a aula como não sei o quê e já houve muitas vozes discordantes.

Não queria entrar nestas polémicas de metodologias de ensino, isso seria se calhar para um próximo encontro do CNE. Aquilo que eu quero dizer é que apesar destas ideias serem certamente bem-intencionadas, estamos provavelmente, em termos do mundo ocidental, a exagerar um pouco na eficácia destas abordagens.

Sou investigador no Instituto de Tecnologia Química e Biológica da Universidade Nova de Lisboa e professor na Faculdade de Ciências da Universidade de Lisboa. Na realidade, de educação sei pouco. Evitarei, portanto, imiscuir-me num campo em que tenho pouco conhecimento.

Há vários aspectos da análise que foi feita pelos anteriores intervenientes com os quais concordo. Evitarei tornar a esses aspetos, farei apenas alguns apontamentos, necessariamente pouco refletidos, dado o pouco tempo disponível para uma análise mais aprofundada, sobre o que apreendi relativamente a este processo de avaliação.

Em primeiro lugar tentei perguntar-me que experiências pessoais poderia recrutar para tentar interpretar alguns dos dados apresentados. Aquilo que verifiquei foi que embora não trabalhe em Educação, poderia recorrer a dados empíricos e pessoais sobre o sistema.

Sou professor há 28 anos na Faculdade, tendo sido anteriormente professor do ensino básico e secundário; sou, há 4 anos, auditor das provas nacionais de Biologia e Geologia, sendo, como presidente de uma Associação de Pais, promotor das Atividades Extra Curriculares de três escolas básicas, Vejo como se produzem os manuais (a minha esposa é coordenadora de livros escolares de ciências). Enfim, tudo isso junto acaba por me permitir pensar onde estou situado e como é que este sistema afeta a minha vida pessoal como interveniente no sistema, mas também como é que afeta a vida dos meus filhos, em particular a dos mais novos.

Gostaria de vos dizer, antes de mais, que uma das questões aqui abordadas me deixou perplexo. Foi apresentada, uma distinção entre a questão da formação na Língua e a formação em Ciências. Devo dizer que me custa compreender esta dicotomia. Esta minha falta de compreensão tem também

¹ Faculdade de Ciências da Universidade de Lisboa/Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa

como base o acompanhamento do meu filho que se encontra no 4ºano do ensino básico. Uma das dificuldades que ele tem em Matemática, não é a Matemática, não é o manuseamento dos números, mas sim a compreensão dos enunciados. Ou seja, muitas vezes, o problema não é a resolução da questão científica que está a ser apresentada, mas sim a compreensão propriamente dita do texto que apresenta o problema. Depois, a mesma dificuldade aparece muitas vezes na explicação dos resultados e na comunicação desses mesmos resultados.

Portanto, gostaria de deixar evidente a necessidade de, a determinada altura, articular os dois tipos de resultados e as duas formações, para que, inclusive em termos de análise, se perceba, por exemplo, se os alunos que têm melhor *performance* são aqueles que tanto compreendem a linguagem, como compreendem os conhecimentos científicos. Porque, a meu ver, quer ao nível da formação quer ao nível da análise dos resultados estes dois aspetos não podem ser observados em separado.

Reitero aqui a importância dada à formação no 1.º ciclo do Ensino Básico. É verdade que apenas abordamos aqui dados relativos a 2011 e ao 4.º ano do ensino básico, o final do 1.º ciclo. Mas, na realidade, é facilmente verificável que uma boa formação base dá uma capacidade aos alunos de progredirem mais facilmente em ciclos subsequentes.

Se for verdade que em Portugal os resultados da formação estão a evoluir, então isso vai permitir aos alunos obterem melhores resultados em ciclos posteriores e, eventualmente, uma maior facilidade e, no caso daqueles que quiserem prosseguir uma carreira como investigadores, conseguirem não só melhores resultados mas melhores *performances* científicas também. Portanto, esta maior importância dada à formação no 1.º ciclo parece-me de facto fundamental.

A formação e desempenho dos professores parece-me um aspeto a ter em conta na melhoria dos resultados apresentados. Isto por um lado, pela verificação do esforço que tem vindo a ser feito ao longo dos anos, mas também pelo contacto direto que tenho com os professores.

É verdade que a minha experiência direta ao nível do 1.º ciclo é numa situação quase privilegiada. Vivo em Oeiras e contacto com uma escola pública que do meu ponto de vista é excelente. O contacto que tenho tido com os professores demonstra que eles são maioritariamente excelentes, muito competentes e muito empenhados na relação com os alunos e na formação dos seus alunos. É, portanto, provável que o empenho na qualidade da formação dos professores tenha tido um impacto na melhoria da qualidade dos resultados apresentados neste estudo.

Há dois outros aspetos que acho interessantes que são a maior interação da escola com as famílias e com a sociedade. Estes aspetos são relevantes para um maior enquadramento dos alunos e também para uma facilitação da formação dos mesmos.

Outro aspeto que não sei qualificar como de positivo ou negativo é o aumento das expectativas das famílias relativamente à formação dos alunos a um nível tão baixo como o 4.º ano do Ensino Básico. Aumento este que, em alguns casos, induz uma maior tentativa de intervenção direta no processo educativo. O que tenho verificado, e que não sei se tem um impacto positivo ou negativo no processo formativo, é uma tensão muito grande em muitos pais, uma expectativa muito grande em relação às *performances* dos filhos desde os primeiros anos de formação. Pessoalmente, não considero esta disposição positiva. Acho que há pais que esperam demasiado das crianças numa fase muito precoce, mas não sei qual o impacto que esta disposição tem na *performance* final das crianças.

Dois comentários relativamente às relações entre o conhecimento, o raciocínio e a aplicação dos conhecimentos. Com base naquilo que observo na receção dos alunos no 1.º ano da Faculdade, a nível de Ciências, penso que continua a dar-se um peso excessivo à retenção estrita de conhecimento, ou seja, insistindo em que os alunos “saibam coisas”.

O conhecimento está em constante e rápida evolução. Eu vejo exigências excessivas na memorização de factos e dados. Talvez não a nível do 1.º ciclo, mas a nível do 2º e 3º ciclos do ensino básico e, sobretudo, do Ensino secundário. Por exemplo, exige-se aos alunos que saibam o nome de cada uma das proteínas que está envolvida num determinado complexo proteico,

quando esse nome passado três semanas pode vir a ser mudado. Ou seja, são completamente irrelevantes os nomes que atribuem a essas proteínas. Cheguei a experienciar o caricato de um professor do Ensino Secundário me perguntar se eu sabia qual era o nome da proteína eIF-4G que estava associada ao complexo proteico que medeia a ligação do RNA mensageiro ao ribossoma, de forma a dar-se início à sua tradução. Tal não tem qualquer significado, nem qualquer utilidade para a formação do aluno ao nível do Ensino Secundário.

O segundo comentário diz respeito à questão levantada pelo Professor Pedro Reis quando comparou a perspetiva da formação na Finlândia, que é tentar melhorar a *performance* dos alunos com menos sucesso *versus* uma perspetiva, que julgo existir no nosso país, que é fazer sobressair aqueles alunos que têm melhores resultados. Esta perspetiva ajusta-se, do meu ponto de vista, a uma tendência elitizante que faz parte de algum discurso político que se ouve em vários momentos, como, por exemplo, quando se afirma que são as elites que conseguem fazer evoluir os países, que o nosso país tem falta de elites, etc. Do meu ponto de vista, esta perspetiva elitizante não é a mais adequada.

Como comparar esta perspetiva com a da Finlândia, onde a ênfase não é tentar aumentar a performance daqueles que estão no primeiro quartil mas sim garantir que todos se enquadram nas perspetivas de uma formação integradora, o que parece garantir melhores resultados à partida?

Finalmente, qual é o possível impacto do processo formativo na qualidade da investigação? E aqui a pergunta que se pode fazer é: o que é um investigador ou o que é que eu quero de um investigador?

Um investigador tem que saber definir um modelo de estudo, tem que saber colocar hipóteses, desenhar experimentação para falsificar as hipóteses que coloca, organizar e interpretar os resultados, apresentar os resultados, questionar os resultados, os seus e os dos outros, redesenhar os seus modelos de estudo. Quando estou a trabalhar com investigadores quero que, pelo menos, saibam fazer isto. Já não estou a falar de inventividade, de capacidade de reconstruir estratégias, de trazer novas aproximações completamente díspares daquelas que estão a ser seguidas.

Como é que se obtêm este conjunto de competências e saberes? Seguramente não será privilegiando alguém com conhecimentos enciclopédicos. Até porque, continuo a dizer, o conhecimento está em rápida evolução. As aplicações estão em rapidíssima evolução. Assim, uma pessoa que se forma com base em conhecimentos enciclopédicos está à partida condenada. A formação tem que ter outra perspetiva, a formação de um investigador tem que ter outra perspetiva.

Neste contexto, convém levantar uma questão: existe alguma política de formação de investigadores ou existe alguma formação de investigadores? Penso que não.

Não existe nenhuma política estruturada para formar investigadores em Portugal. O que acontece em geral é: uma pessoa começa a fazer investigação, porque encontra uma oportunidade para isso e vai-se tornando um investigador. Porque, academicamente, aquilo que oferecemos é essencialmente uma formação mais tradicional, fundamental, com cada vez menos recursos e não para fazer investigação.

Mesmo na formação superior, na área das Ciências Exatas, não damos, regra geral, formação específica para fazer investigação. Logo, estas questões, colocadas desta maneira, têm um reflexo naquilo que é possível esperar quando se pergunta: “então a evolução detetada em Portugal na formação ao nível do 1.º ciclo terá impacto ou não depois na investigação que se faz em Portugal?”.

Foi antes de mim referido, e é constatável, que evoluímos muito em termos da nossa qualidade em investigação ou, pelo menos, em resultados quantitativos evoluímos muito. A questão, do meu ponto de vista, não é se evoluímos muito, mas se, comparativamente com os outros países, essa evolução é concordante ou não. Confesso que, neste momento, tenho algumas dúvidas que isso seja verdade.

Termino desta maneira, dizendo-vos que na realidade não tenho muitas certezas. Que à partida fico agradado com os resultados que nos foram apresentados, que do ponto de vista da observação da escola onde o meu filho se encontra parecem corresponder à realidade. Resta saber se isso é

SEMINÁRIO

generalizado e, como outros antes de mim referiram, se é possível manter esta tendência. Muito obrigado por me ouvirem.

Avaliações internacionais e desempenho
dos alunos portugueses

Os testes internacionais, o sistema de avaliação externa
e a promoção das aprendizagens

Presidente da Mesa – Maria Emília Brederode Santos

SEMINÁRIO

A abordagem do tema proposto foi conduzida tendo como fio condutor três perspetivas, interligadas e complementares:

- a valia dos estudos internacionais enquanto fontes de informação para o conhecimento da qualidade da aprendizagem dos alunos em Portugal;
- a utilização dos resultados dos estudos internacionais de avaliação de alunos, em articulação com os resultados da avaliação externa nacional, enquanto instrumento de avaliação e de regulação do sistema de ensino, adicionalmente como meio de validação, ainda que indireta, dos resultados gerados pelas avaliações externas de alunos de âmbito nacional;
- o futuro do sistema de ensino concebido a partir de evidências, entre outras, que emergem de uma análise integrada dos resultados dos estudos internacionais de avaliação de alunos e das provas de avaliação externa nacionais.

Os estudos internacionais de avaliação de alunos apresentam, em comparação com as avaliações nacionais, a particular vantagem de gerarem resultados alicerçados em provas não públicas, em que a estabilidade dos itens usados em estudos realizados em diferentes momentos retiram da análise global de resultados uma variável que, no caso das avaliações nacionais, está sempre presente: a produção de provas públicas, cujos itens se renovam a cada aplicação. Também o facto de se realizarem com uma periodicidade que oscila entre três (PISA) e cinco anos (PIRLS) permite a obtenção de informação espaçada no tempo, mais coerente com o propósito de uma avaliação de âmbito longitudinal de médio/longo prazo.

Os resultados gerados pelas avaliações nacionais, com periodicidade anual, estão limitados, no que se refere à informação que deles se pode extrair, pela natureza pública das provas. A pressão mediática e social para que desses resultados se façam inferências em relação ao estado da educação, assentes numa comparação longitudinal de resultados, exercício repetido de forma

¹ Diretor do Gabinete de Avaliação Educacional

acrítica e pontualmente em cada final de ano escolar, ignora as limitações e insuficiências atrás referidas.

Salvaguardando e considerando as limitações técnicas que decorrem da natureza específica das provas e do tratamento de resultados em contexto de estudos internacionais e de provas nacionais, ainda assim não deixa de ser possível contar com os resultados dos estudos internacionais como indicadores, mesmo que indiretos, de validação dos resultados dos exames nacionais e das provas finais, realizadas todos os anos no nosso país.

Senão, vejamos. O desempenho dos alunos portugueses nos estudos internacionais cujos resultados permitem estabelecer comparações longitudinais mostra progressos importantes, como se comprova nos casos do PISA 2009 e do TIMSS 2011. Contudo, como ponto prévio, é preciso ter em atenção que as diferenças no objeto de avaliação que subjaz a cada um dos estudos, o diferente público-alvo (alunos com 15 anos e a frequentar 4.º ano, respetivamente no caso do PISA e do TIMSS) e o diferente intervalo de tempo que serve de referência para a comparação longitudinal (9 e 16 anos, respetivamente) são variáveis indissociáveis de uma análise global que sustente inferências fiáveis e relevantes para o delinear de estratégias de melhoria da aprendizagem.

Enquanto no estudo PISA 2009 Portugal se apresenta como um «aluno» com um desempenho no limite inferior do «suficiente» (21.ª posição entre 32 países da OCDE) e com um *score* próximo do ponto médio da distribuição (489 vs 500), no caso do TIMSS (Matemática) regista-se, globalmente, uma prestação relativa bem mais positiva, com *scores* muito superiores ao valor médio de referência (532 vs 500 Matemática; 522 vs 500 Ciências). Ainda com um *score* muito positivo temos os resultados do PIRLS 2011 (541 vs 500).

Em sentido inverso, o *First European Survey on Language Competence* (adiante designado abreviadamente por *Survey Lang*), estudo promovido pela Comissão Europeia em que se avaliava o desempenho dos alunos nacionais no domínio das línguas estrangeiras no final do 9.º ano, mostra-nos um desempenho medíocre e preocupante: no final do 9.º ano, 69% dos alunos atinge um nível inferior ao esperado (58% média do estudo) e 31%

atinge ou supera o nível esperado (42% média do estudo), colocando o país na 10.^a posição entre 13 países em que o inglês é a primeira língua estrangeira de estudo.

Podemos assim concluir que em termos longitudinais os estudos PISA e TIMSS mostram uma progressão positiva em matéria de resultados, embora considerando intervalos temporais longos (especialmente no último caso), mas que em termos sincrónicos colocam Portugal numa posição em torno do ponto médio da escala de medida de cada um dos estudos. Em termos nacionais, também por força das alterações curriculares, a série de resultados que podem ser considerados numa análise cronológica não cobrem períodos tão longos, não indo além de cinco a sete anos. Mas, recorrentemente, têm vindo a mostrar, em termos médios e no essencial:

- por um lado, uma estabilidade de resultados ao longo do período considerado e, por outro lado, a ocorrência de resultados médios que se situam em torno do ponto médio das escalas de medida utilizadas (0 a 20 valores ou 1 a 5);
- dificuldades de aprendizagem que se concretizam com maior acuidade ao nível das chamadas *high skills*.

Estas constatações são concordantes com as evidências que os resultados dos estudos internacionais nos mostram: Portugal progrediu mas está ainda longe de alcançar *scores* que nos coloquem ao nível de países como a Finlândia (já para não falar de países asiáticos com resultados ainda superiores), o que também é um direto resultado da baixa percentagem de alunos portugueses que resolve com sucesso os itens que mobilizam operações mentais mais complexas, ou que exigem processos de resolução mais elaborados.

Os resultados dos estudos internacionais de avaliação de alunos, bem como a análise dos resultados da avaliação externa nacional, constituem um manancial de informação que deve sustentar uma avaliação do efeito de medidas de política educativa, informar sobre a qualidade do trabalho desenvolvido por cada professor e por cada escola, ao nível micro, e ajudar a delinear, se necessário, a adoção de outras medidas ou formas de atuação mais eficientes no apoio à melhoria da aprendizagem.

É certo que presentemente as evidências dos contactos com escolas, com diretores e professores no terreno, levam-nos a concluir que há uma insuficiente divulgação e sobretudo apropriação dos resultados e dos relatórios nacionais e internacionais por parte de muitos *stakeholders*, o que, de certo modo, não ajuda a uma reflexão e à construção de estratégias de atuação desenhadas localmente, privilegiando processos mais ou menos autónomos de decisão, no sentido da superação das fragilidades e dificuldades de todos conhecidas.

Ao nível da informação gerada pela avaliação externa nacional, a política de produção de relatórios nacionais e de relatórios técnicos, estes para uso exclusivo das escolas, iniciada há cerca de 4 a 5 anos, tem vindo a revelar-se muito promissora, com uma crescente difusão e utilização nas escolas. Já no que se refere aos relatórios produzidos por organismos internacionais, em concreto os relatórios PISA, *Survey Lang* e os mais recentes TIMSS e PIRLS, constituem ainda ferramentas cuja apropriação continua muito confinada ao mundo académico e da investigação.

Por exemplo, enfatizar o papel de uma sistemática produção de feedback aos alunos e às famílias, ou o trabalho colaborativo com as famílias, enquanto agentes centrais de uma possível mudança de paradigma nos processos de aprendizagem, na metodologia de estudo, na valorização da leitura, parecem longe de estar a funcionar.

A título de exemplo, se atendermos ao facto de que em Portugal os alunos em cujas famílias se promove o hábito de ler obtêm no PIRLS um score 22 pontos acima da média nacional (563 pontos vs 541, valor próximo do alcançado pela Finlândia – 568), se aceitarmos que a literacia de leitura constitui talvez a principal variável-chave na promoção do sucesso educativo, que é transversal a todas as áreas disciplinares, que é indispensável a uma correta leitura dos enunciados escritos e que a sua melhoria generalizada teria resultados extremamente positivos em todas as avaliações escritas que cada aluno realiza ao longo da vida, percebe-se o muito que há a fazer neste domínio.

Um outro aspeto em que se torna evidente intervir é o que se prende com a cultura de avaliação dominante, em que se privilegia o resultado pelo

resultado, descurando a riqueza informativa que o resultado de uma qualquer avaliação comporta. Com resultados nacionais desagregados por aluno e por item estão criadas as condições para identificar de forma fina o que falhou do lado da ação do professor, o que está por alcançar, por parte do aluno. E, a seguir, deveriam estar criadas as condições para que, de forma generalizada e sistemática, houvesse uma contínua correção de processos, situação que implica um olhar sobre as funções formativas e reguladoras que estão longe de estar apropriadas pela maioria dos professores, não tanto por sua direta responsabilidade, mas sobretudo pelas lacunas que a formação inicial e contínua propiciam nesta matéria específica.

Replicam-se, num processo de mimetismo geracional que não é exclusivo nacional, processos de avaliação apreendidos pelas vivências dos atuais professores enquanto alunos, o que também se aplica aos pais, cujo conhecimento que possuem sobre o que é avaliar e para que serve avaliar em nada difere do obtido pelas suas experiências pessoais na qualidade de alunos.

Nestas circunstâncias, em termos gerais, parece evidente um fraco aproveitamento das potencialidades da articulação entre uma avaliação externa, rigorosa e exigente, e processos de avaliação baseados na sala de aula, capazes de gerar o retorno que constitua um forte incentivo à superação do erro e à conseqüente melhoria e elevação da qualidade da aprendizagem.

Dito isto, reconhece-se ser essencial a melhoria e a diversificação dos processos de avaliação de sala de aula, sem nunca descurar a importância de ter sempre presente uma ação reguladora que a avaliação externa, realizada ao nível internacional ou nacional, deve desempenhar.

Três exemplos, a três níveis de análise territorial diferentes ilustram quão indispensável é a prática regular de processos de avaliação externa. Ao nível regional, os atuais resultados dos alunos da Região Autónoma do Açores nas provas finais de ciclo do 6.º e 9.º anos de escolaridade, em Português e Matemática (dados de 2012), ficam significativamente aquém dos resultados alcançados nas NUTS III do território continental com piores resultados e

ainda mais distantes do resultado médio nacional: em Português as diferenças são da ordem dos 3,5% a 10%, respetivamente; em Matemática as diferenças aumentam para valores da ordem dos 8,5% a 17%, respetivamente.

Ao nível nacional, talvez um dos exemplos mais elucidativos pode ser encontrado nos resultados de Portugal no estudo *Survey Lang*, atrás referido, numa disciplina nunca sujeita a qualquer regulação realizada através de provas de avaliação externa nacionais (com exceção dos testes intermédios aplicados em 2011, 2012 e 2013, mas sem qualquer impacto nos resultados do estudo).

Em termos internacionais, países como a Noruega, a Suécia e mesmo a Finlândia, considerados modelos de sucesso, com processos de avaliação de alunos muito focados na designada *classroom-based assessment*, têm vindo a adotar procedimentos de avaliação externa mais frequentes, em resposta a retrocessos observados nos seus resultados em estudos internacionais (por exemplo, a Finlândia passa de um score de 520 para 482 pontos no estudo TIMSS, 8.º ano, entre 1999 e 2011).

Em suma, o forte incentivo à promoção da leitura e ao estudo da Matemática e o impacto da avaliação externa no 1.º ciclo do ensino básico (provas de aferição e em 2013 provas finais) no ensino e na aprendizagem do Português e da Matemática, ou a total ausência de mecanismos de regulação externa, que no caso da disciplina de Inglês não permitem entrar na «caixa fechada» que é a sala de aula, podem constituir tópicos de reflexão incontornáveis no caminho para o diagnóstico e desejável superação das dificuldades de aprendizagem que alguns dos resultados da avaliação externa nacional têm evidenciado nos últimos anos.

Considera-se ainda que os bons resultados internacionais alcançados no final do 1.º ciclo não devem mascarar ou iludir o facto de ao longo de vários anos os resultados da avaliação externa nacional mostrarem fragilidades de aprendizagem que se mantêm e, em regra, se ampliam, em anos escolares subsequentes.

Ao nível das capacidades cognitivas mais complexas, mas também ao nível de algumas tarefas elementares, a análise sistemática de resultados da

avaliação externa altamente desagregados (por aluno e por item) tem vindo a demonstrar que há ainda um longo caminho a percorrer para alcançar, de forma generalizada, níveis de sucesso académico em que a simples obtenção de um resultado coincidente ou pouco superior ao definido como linha de corte seja motivo de plena satisfação para pais, alunos e professores.

SEMINÁRIO

Estudos internacionais: desempenhos contextualizados

Ana Sousa Ferreira¹

Introdução

A OCDE (*Organisation for Economic Co-operation Development*), a IEA (*International Association for the Evaluation of Educational Achievement*) e a Comissão Europeia promovem programas internacionais de avaliação do desempenho de crianças e jovens do mundo inteiro em matemática, ciências, leitura e línguas estrangeiras.

Portugal tem vindo a participar em diversos estudos internacionais: PISA - *Programme for International Students Assessment*; TIMSS - *Trends in International Mathematics and Science Study*; PIRLS - *Progress in International Reading Literacy Study*, ESLC - *European Survey on Language Competences*. Todos estes projetos avaliam a proficiência de alunos de vários países em diversos domínios.

Nestes estudos, os itens não são públicos, possibilitando-se deste modo a comparação de resultados ao longo das várias edições da prova e a identificação de tendências. A OCDE, a IEA e a Comissão Europeia disponibilizam, em cada ciclo de cada um destes estudos, alguns itens que deixam de fazer parte das provas e que ilustram o tipo de situações apresentadas aos alunos.

O PISA, desenvolvido pela OCDE², visa avaliar se os alunos de 15 anos, aqueles que na maior parte dos países participantes se aproximam do final da escolaridade obrigatória, estão bem preparados para enfrentarem os desafios da vida quotidiana. A prova do PISA é constituída por vários cadernos de teste que combinam itens dos diferentes domínios avaliados. Cada aluno responde, apenas, a um caderno de teste.

¹ ProJAVI

² Para mais informações consultar <http://www.projavi.mec.pt/> ou <http://www.oecd.org/pisa/>

Os testes PISA foram concebidos para avaliar as competências de leitura, matemática e ciências dos alunos de 15 anos, na resolução de situações relacionadas com a realidade, e ocorrem em ciclos de três anos. Em cada ciclo PISA, é selecionada, como domínio principal, uma das três áreas avaliadas (leitura, matemática, ciências). Em 2000, o domínio principal foi a leitura, domínio novamente avaliado em 2009. Em 2003 e em 2012 selecionou-se a matemática e, em 2006, as ciências. Em 2015, seguindo a rotatividade, o domínio principal será ciências.

Neste projeto, são também recolhidas informações que permitem contextualizar os resultados, através da aplicação de questionários aos alunos, aos pais e às escolas. O PISA permite, pois, identificar os fatores que influenciam os níveis de desempenho nos vários domínios de literacia, nomeadamente, elementos sociodemográficos dos alunos e da sua relação com a aprendizagem e, ainda, diversas características das escolas participantes, como a sua organização e recursos.

Portugal tem vindo a participar no projeto PISA desde o primeiro ciclo, em 2000, até ao último ciclo aplicado, em 2012.

O TIMSS é uma avaliação internacional do desempenho dos alunos dos 4.º e 8.º anos de escolaridade em matemática e ciências, desenvolvida pela IEA¹. O TIMSS *advanced* é uma versão do estudo que tem o objetivo de avaliar as tendências do desempenho dos alunos no final do ensino secundário em matemática e física.

A avaliação em matemática considera duas dimensões: uma relativa ao conteúdo – *Números, Formas Geométricas e Medida, Apresentação de Dados* – e a outra ao domínio cognitivo – *Aplicar, Conhecer e Raciocinar*. A avaliação em ciências segue o mesmo desenho, sendo contempladas as seguintes áreas: *Ciências da Vida, Ciências Físicas e Ciências da Terra*. Tanto para matemática como para ciências, o TIMSS tem, igualmente, em conta a abordagem de cada área específica e o domínio cognitivo nos currículos nacionais.

¹ Para mais informações consultar <http://www.projavi.mec.pt/> ou <http://timssandpirs.bc.edu/>

A prova é constituída por um conjunto de cadernos, numa composição de itens de matemática e ciências que abrangem os três processos cognitivos (*aplicar, conhecer e raciocinar*). Cada aluno responde a um único caderno de prova. Neste estudo, a par dos testes, são também aplicados questionários visando recolher informação de contexto que permite descrever as situações e os fatores que influenciam a aprendizagem da matemática e das ciências.

Desde 1995, os testes TIMSS são aplicados de quatro em quatro anos com a finalidade de gerar informação de qualidade sobre os resultados do desempenho dos alunos e sobre os contextos em que estes aprendem. Em 2011, a aplicação dos dois estudos da IEA, o TIMSS e o PIRLS, coincidiu, o que irá permitir relacionar a literacia de leitura dos alunos do 4.º ano de escolaridade com o seu desempenho em matemática e ciências.

Portugal participou no primeiro ciclo do TIMSS em 1995 e voltou a participar na prova do 4º ano em 2011.

O PIRLS é uma avaliação internacional sobre a compreensão da leitura dos alunos do 4.º ano de escolaridade, desenvolvida pela IEA¹. Esta prova foi desenhada de modo a contemplar duas finalidades de leitura, a *literária* - ler como experiência literária - e a *informativa* - ler para adquirir e utilizar informação -, assim como os processos de compreensão da leitura - *Reter, Fazer inferências diretas, Interpretar e integrar e Avaliar* -.

A prova é constituída por um conjunto de cadernos, numa composição de itens que envolvem as diferentes finalidades e processos de compreensão da leitura. Cada aluno responde apenas a um caderno de prova. Neste estudo, a par dos testes, são aplicados questionários visando recolher informação de contexto que permite descrever as situações e os fatores que afetam a literacia de leitura. Desde 2001, os testes PIRLS são aplicados de cinco em cinco anos com a finalidade de gerar informação de qualidade sobre os resultados do desempenho dos alunos em leitura e sobre os contextos em que estes aprendem. Em 2011, foi criada uma versão para aplicar em países cujos alunos estão aquém dos níveis de leitura estabelecidos para o PIRLS (prePIRLS).

¹ Para mais informações consultar <http://www.projavi.mec.pt/> ou <http://timssandpirls.bc.edu/>

Portugal participou pela primeira vez no PIRLS em 2011.

O ESLC é uma avaliação das competências dos alunos europeus em línguas estrangeiras, que permite disponibilizar informação comparável sobre a proficiência em línguas e sobre boas práticas no ensino de línguas estrangeiras na Europa. Com este estudo pretende-se também construir um indicador europeu para a monitorização dos progressos alcançados na aprendizagem das línguas estrangeiras nos Estados-membros. Este estudo foi promovido pela Comissão Europeia¹ e foi desenvolvido pelo *SurveyLang*, um consórcio internacional constituído por várias organizações. A prova foi definida para avaliar a proficiência dos alunos em *ouvir, ler e escrever* numa língua estrangeira. Estes alunos têm, pelo menos, um ano de aprendizagem da língua estrangeira e frequentam, preferencialmente, o 9.º ano de escolaridade. A avaliação abrange as duas línguas estrangeiras mais ensinadas em cada país, de entre o inglês, o francês, o alemão, o italiano e o espanhol. Os resultados foram analisados de acordo com o Quadro Europeu Comum de Referência para as Línguas (QECR).

Cada aluno responde a uma prova adaptada a um nível de proficiência estabelecido *a priori* pelo seu resultado num *routing test*. A par dos testes, são aplicados questionários às escolas e aos professores de línguas estrangeiras. Esta informação de contexto permite descrever as situações e os fatores que influenciam o ensino e a aprendizagem das línguas estrangeiras em cada país participante.

Portugal participou no primeiro ciclo do ESLC em 2011.

A informação gerada por estes estudos permite identificar tendências de resultados, referenciados a padrões internacionais, em vários sistemas educativos, o que os potencia como instrumentos relevantes para a análise e para a definição de políticas educativas.

¹ Para mais informações consultar <http://www.projavi.mec.pt/> ou <http://www.surveylang.org/>

Desempenhos contextualizados

Que tipos de informação nos fornecem os estudos internacionais?

Os diversos estudos internacionais já referidos representam uma fonte de informação sobre sistemas educativos de inegável qualidade, quer ao nível da construção dos itens de avaliação, quer ao nível do desenho do plano de amostragem, possibilitando a avaliação das políticas educativas dos vários países participantes. A OCDE, a IEA e a Comissão Europeia planeiam e documentam cuidadosamente todo o processo de desenvolvimento das provas internacionais, de forma a garantir a comparabilidade dos dados, definindo procedimentos padronizados e implementando um exigente controlo de qualidade. Por exemplo, é realizada uma extensa série de verificações para garantir a comparabilidade das traduções dos itens de avaliação e dos questionários, é fornecida documentação detalhada sobre os padrões de amostragem, são minuciosamente verificadas as taxas de participação das escolas e dos alunos e foi desenvolvido um ambicioso programa de garantia da qualidade para monitorizar a recolha dos dados.

O plano de amostragem considerado nos vários projetos internacionais de avaliação do desempenho dos alunos em que Portugal participa considera uma amostragem estratificada, em duas etapas. Na primeira etapa, é selecionada aleatoriamente uma amostra de escolas; na segunda, recolhe-se uma amostra aleatória de alunos das escolas previamente selecionadas. Todos os processos de amostragem são monitorizados pelas instâncias internacionais responsáveis pelos estudos, garantindo-se que todos os países participantes cumprem os requisitos estabelecidos. Genericamente, nos vários estudos, têm sido usadas como variáveis definidoras dos estratos a região (por exemplo, NUTS II ou NUTS III), a natureza institucional da escola (pública *versus* privada) ou a tipologia das áreas urbanas (área predominantemente urbana, área medianamente urbana e área predominantemente rural).

Estes projetos podem, pois, contribuir para o desenvolvimento da investigação científica na área da educação, proporcionando aos investigadores o acesso a um amplo conjunto de dados, comparáveis entre países e abrangendo um alargado leque de informação.

A análise dos resultados dos estudos internacionais permite comparar o desempenho entre países, ou, num mesmo país, entre regiões (por exemplo, NUTS II ou NUTS III), por sexo, segundo a origem imigrante ou de acordo com o estatuto socioeconómico e/ou cultural dos pais.

Por outro lado, o desempenho dos alunos pode ser analisado por valores de referência internacionais (*benchmarks*), ou por conteúdos e domínios cognitivos, podendo, assim, conduzir à reflexão sobre como melhorar as práticas em sala de aula, a aprendizagem ou a formação de professores.

Por exemplo, podemos realizar uma comparação entre os resultados obtidos pelos vários países participantes no TIMSS & PIRLS 2011, analisando a Figura 1:



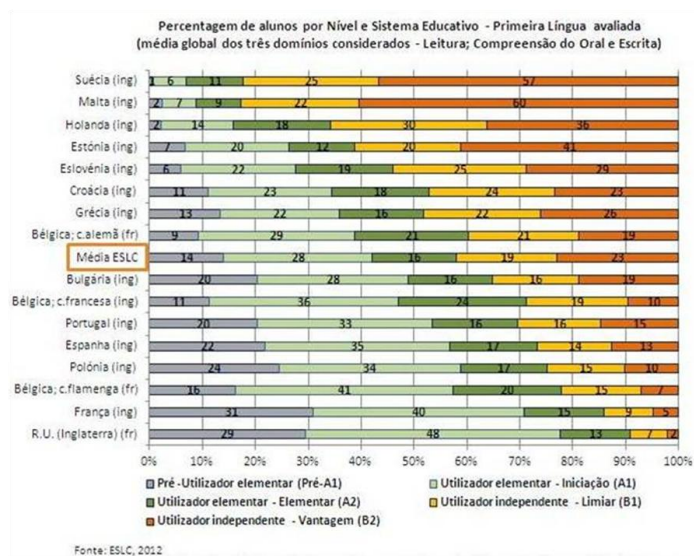
Figura 1. Desempenho dos vários países participantes no TIMSS & PIRLS 2011

O desempenho dos alunos portugueses do 4.º ano no TIMSS & PIRLS 2011 situa-se entre os 15 melhores países participantes em matemática e entre os 19 melhores países participantes em ciências e leitura.

Que contributo estes estudos internacionais podem dar ao nosso sistema educativo?

A análise sobre os países com melhor desempenho nos estudos TIMSS & PIRLS 2011 e as características dos seus sistemas educativos pode conduzir a pistas importantes para os decisores políticos e outros intervenientes no sistema educativo aperfeiçoarem a qualidade do nosso sistema de ensino-aprendizagem.

Ou, por exemplo, podemos realizar uma comparação entre os resultados obtidos pelos vários países participantes no ESLC 2011, analisando a Figura 2:



Nota: A indicação dada à frente do nome de cada país – (ing) ou (fr) – significa, respetivamente, inglês ou francês, e refere-se à língua avaliada nesse país como Primeira Língua.

Figura 2. Desempenho dos vários países participantes no **ESLC 2011** relativamente à primeira língua avaliada.

O desempenho dos alunos portugueses do 9.º ano de escolaridade no **ESLC 2011** situa-se na 11ª posição (em 16 participantes) em inglês (primeira língua).

Que contributo este estudo internacional pode dar ao nosso sistema educativo?

Se analisarmos quais os países com melhor desempenho no ESLC 2011, poderemos analisar as características dos seus sistemas educativos bem como as suas práticas de ensino e de aprendizagem das línguas estrangeiras e, desta forma, contribuir para a tomada de medidas eficazes na melhoria do desempenho dos nossos alunos em línguas estrangeiras.

Como já referimos, o desempenho dos alunos nos vários estudos internacionais pode ser analisado por valores de referência internacionais (*benchmarks*), ou por conteúdos e domínios cognitivos, podendo, assim, conduzir à reflexão sobre como melhorar as práticas em sala de aula, a aprendizagem ou a formação de professores.

Apresentamos, neste trabalho, apenas alguns exemplos desse olhar que pode ser aplicado aos resultados destes estudos.

No TIMSS 2011, o desempenho médio dos alunos portugueses do 4.º ano em matemática é de 532 pontos, numa escala de 0-1000, com um ponto médio de referência de 500 e um desvio padrão de 100.

Este estudo estabeleceu 4 níveis de desempenho para a matemática, a partir de 4 valores de referência (*benchmarks*): *Avançado*¹ (*scores* globais não inferiores a 625), *Elevado*² (*scores* globais com valores no intervalo [550, 625[), *Intermédio*³ (*scores* globais com valores no intervalo [475, 550[) e *Baixo*⁴ (*scores* globais não inferiores a 400). Estes quatro pontos na escala internacional originaram uma descrição de quatro níveis de desempenho em matemática, permitindo comparar o desempenho dos alunos portugueses com os alunos de Singapura, como pode ver-se no Quadro 1.

¹ Advanced international benchmark

² High international benchmark

³ Intermediate international benchmark

⁴ Low international benchmark

Quadro 1. Comparação entre os resultados no TIMSS 2011 – Matemática dos alunos portugueses e dos alunos de Singapura por valores de referência (*benchmarks*).

Matemática (532)		Portugal	Singapura	Benchmarks
TIMSS 2011 Benchmarks Internacionais	Avançado			
	Aplica conhecimento e compreensão em situações relativamente complexas e explica o raciocínio.	8%	43%	625
	Elevado			
	Aplica conhecimento e compreensão para resolver problemas.	40%	78%	550
	Intermédio			
	Aplica conhecimentos básicos em situações de resolução imediata.	80%	94%	475
	Baixo			
	Tem alguns conhecimentos matemáticos básicos.	97%	99%	400

Isto é, quase todos os alunos portugueses do 4.º ano (97%) atingiram o nível mais baixo de *benchmark*, 80% alcançaram o nível Intermédio, o que significa que 4/5 dos alunos portugueses são capazes de aplicar conhecimentos básicos em situações diretas, e 40% atingiram o nível Elevado, pelo menos 550 pontos (valor superior à média nacional de valor 532) o que corresponde a serem capazes de aplicar conhecimento e compreensão para resolverem problemas; apenas 8% dos alunos atingiram o nível Avançado, o nível de maior complexidade da escala, que envolve compreensão e explicação do raciocínio. Note-se que, apesar de os resultados dos alunos portugueses terem melhorado claramente quando comparados com os resultados obtidos em 1995, existe ainda uma diferença evidente relativamente aos resultados dos alunos de Singapura nos níveis *Intermédio*, *Elevado* e *Avançado*.

Não devemos, pois, esquecer que a avaliação do desempenho é uma componente indispensável do sistema de ensino-aprendizagem e os resultados destes estudos, referenciados a padrões internacionais, devem conduzir a uma reflexão que contribua para ajudar os professores a ensinarem melhor e os alunos a aprenderem melhor.

Exemplifiquemos, então, a análise do desempenho dos alunos portugueses por conteúdos e domínios cognitivos, visando, assim, melhorar as práticas em sala de aula, a aprendizagem e a formação de professores.

Consideremos os resultados do estudo TIMSS 2011 – Ciências, por conteúdos e domínios cognitivos, apresentados seguidamente, nos Quadros 2 e 3.

Quadro 2. Resultados no TIMSS 2011 – Ciências dos alunos portugueses, por conteúdos.

TIMSS 2011 - Ciências	Ciências (522)	
		PT
	Ciências da Vida	520
	Ciências Físicas	517*
Ciências da Terra	531*	

* As diferenças entre o desempenho médio global em ciências e na área específica de conteúdo são significativas ($p < 0.05$).

As áreas específicas de conteúdo de ciências avaliadas para o 4.º ano no TIMSS 2011 são *Ciências da Vida*, *Ciências Físicas* e *Ciências da Terra*. Os alunos portugueses do 4.º ano mostraram um desempenho médio significativamente superior à média global nacional (522 pontos) em *Ciências da Terra* (531). No que concerne às *Ciências da Vida* (520), os alunos portugueses tiveram um desempenho médio abaixo da média global nacional e, nas *Ciências Físicas* (517), os alunos portugueses tiveram um desempenho médio significativamente inferior à média global nacional (522 pontos).

Quadro 3. Resultados no TIMSS 2011 – Ciências dos alunos portugueses, por domínios cognitivos.

TIMSS 2011 - Ciências	Ciências (522)	
		PT
	Conhecer	528*
	Aplicar	515*
Raciocinar	524	

* As diferenças entre o desempenho médio global em ciências e no domínio cognitivo específico são significativas ($p < 0.05$).

Uma análise do desempenho de acordo com os domínios cognitivos que especificam os processos mentais mobilizados pelos alunos — *Conhecer*, *Aplicar* e *Raciocinar* — mostra que, em Portugal, a diferença do desempenho médio por domínio cognitivo relativamente à média global nacional é significativamente superior para o domínio do **Conhecimento**. Também no domínio do **Raciocínio**, os alunos portugueses do 4.º ano apresentam desempenhos médios acima da média global nacional. Já no que toca ao domínio da **Aplicação**, os desempenhos médios dos alunos portugueses encontram-se significativamente abaixo da média global nacional.

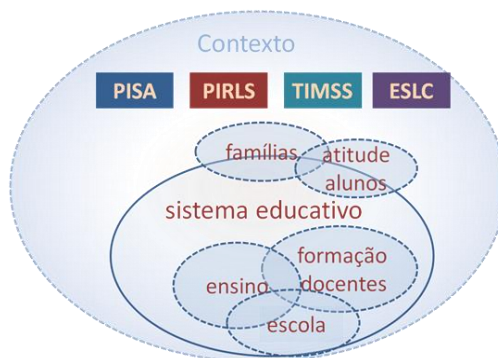


Figura 3. Desempenhos contextualizados nos vários projetos internacionais

Os questionários de contexto aplicados nestes estudos permitem aliar à informação sobre o ensino e a formação de professores a informação sobre o contexto da escola (recursos/professores) e das famílias e alunos, recolhendo dados sobre o *status* socioeconómico das famílias (formação e profissão dos pais, número de livros em casa, recursos educativos disponíveis), as características das escolas (recursos materiais, projeto educativo), a atitude dos alunos e o tempo de aprendizagem (conforme Figura 3).

No estudo PIRLS 2011, analisou-se o desempenho dos alunos por sexo e observou-se que o desempenho médio das raparigas em leitura é superior ao desempenho médio dos rapazes em quase todos os países (regiões) participantes. A média internacional é superior em 16 pontos para as raparigas (520 relativamente a 504 de média dos rapazes). Em *Portugal*, as raparigas tiveram uma pontuação média significativamente mais elevada (548) do que os rapazes (534) ($p < 0.05$).

A análise do desempenho dos alunos portugueses pode ser efetuada segundo muitos outros olhares, usando a informação de contexto recolhida por qualquer dos estudos internacionais em que Portugal participa. Os dados recolhidos e os resultados obtidos por estes estudos são tornados públicos pelos consórcios internacionais após a divulgação do relatório internacional e, como já foi referido, permitem aos investigadores da área da educação o acesso a um alargado leque de informação, contribuindo para que estes instrumentos de avaliação internacional possam ser reguladores do sistema de ensino-aprendizagem.

Relação com o currículo

Os estudos ESLC, TIMSS e PIRLS são avaliações que têm como referente os currículos dos países participantes, enquanto o projeto PISA é uma avaliação que não assenta nos currículos dos países participantes, mas sim nas competências que um aluno de 15 anos deve ter quando se prepara, na maioria dos países, para entrar no mundo do trabalho.

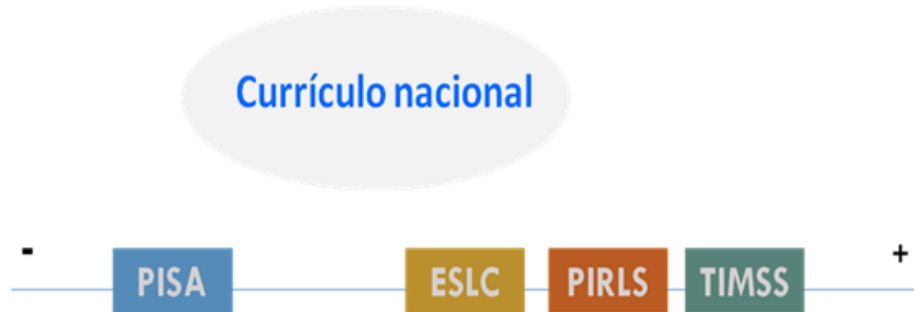


Figura 4. Projetos internacionais referentes ao currículo dos países participantes

Portugal, ao participar nestes estudos internacionais, acrescenta um procedimento de avaliação externa ao sistema educativo. Para algumas áreas disciplinares e para alguns anos de escolaridade existem pontos de encontro entre as provas nacionais e internacionais que podem ser explorados, como podemos observar na Figura 5.

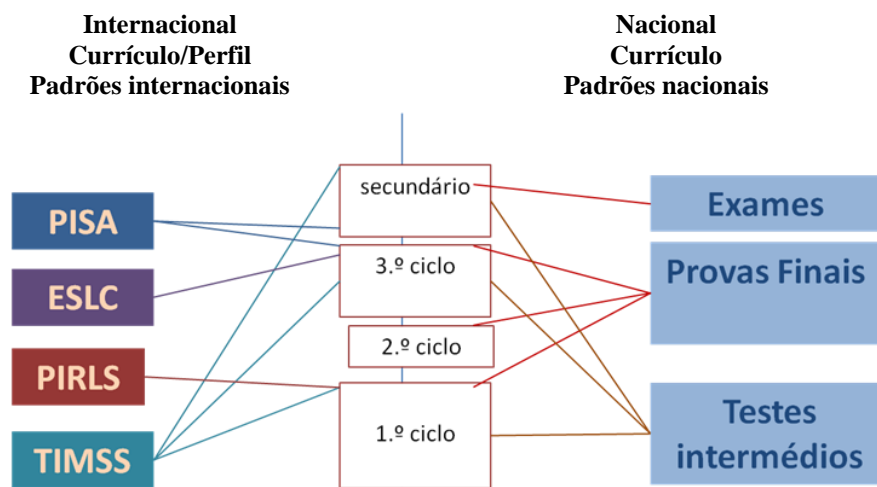


Figura 5. Provas nacionais e internacionais relacionadas com o currículo nacional

O facto de as provas de aferição de matemática serem constituídas por itens cuja tipologia se assemelha à dos itens TIMSS e que abordam conteúdos que se cruzam com os do TIMSS é outro fator que apoia uma comparação de resultados.

As provas de aferição são classificadas segundo 5 níveis de desempenho em *matemática* - A, B, C, D e E -, correspondendo os três primeiros a um desempenho positivo. Em 2011, 79,5% dos alunos do 4.º ano de escolaridade obtiveram um dos níveis A, B ou C (GAVE, Relatório de Provas de Aferição – Matemática 1.º ciclo 2010/2011), sendo a distribuição global por níveis a que se apresenta na Figura 6. A média nacional foi de 68%, com um desvio padrão de 21%.

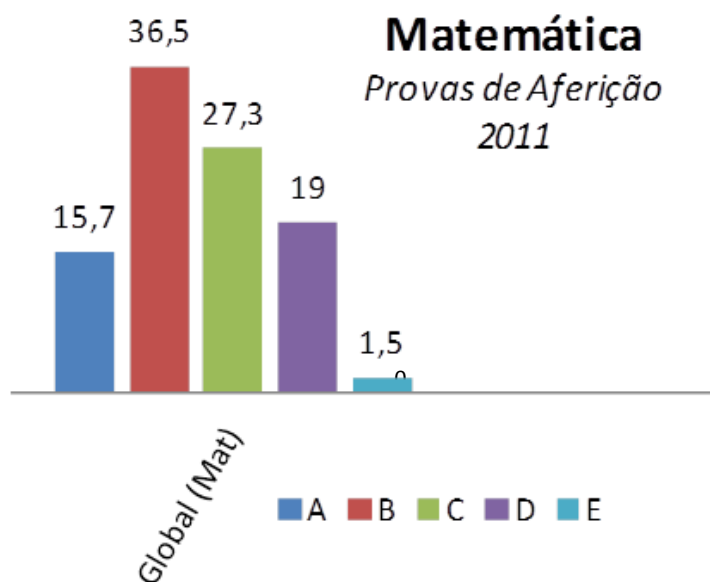


Figura 6. Níveis de desempenho em matemática no 4.º ano (Provas de aferição 2011).

Conclusão

Os estudos internacionais em que Portugal participa fornecem informação fiável e contextualizada sobre o desempenho dos nossos alunos.

Estes resultados podem, pois, contribuir não só para a regulação do sistema educativo, mas também para a melhoria das práticas em sala de aula e da aprendizagem, promovendo a avaliação formativa e o aprofundamento das temáticas da avaliação (itens libertos). O ProjAVI, pretendendo promover a divulgação destes projetos internacionais, disponibiliza informação sobre estes estudos e os resultados no seu sítio, com o endereço <http://www.projavi.mec.pt>.

Referências bibliográficas

FIRST EUROPEAN SURVEY ON LANGUAGE COMPETENCES: FINAL REPORT. (SurveyLang & Comissão Europeia, 2012).

MULLIS, I. *et al.* (2012). *PIRLS 2011 International Results in Reading*. (IEA, International Report).

MULLIS, I. *et al.* (2012). *TIMSS 2011 International Results in Mathematics*. (IEA, International Report).

MULLIS, I. *et al.* (2012). *TIMSS 2011 International Results in Science*. (IEA, International Report).

ProjAVI, *PIRLS 2011 - Principais resultados em Leitura*. (Nota de Leitura, acedido em 22 de abril de 2013, em: <http://www.projavi.mec.pt/np4/179/>).

GAVE, Relatório de Provas de Aferição – Matemática 1.º ciclo 2010/2011 (acedido em 23 de abril de 2013, em: <http://www.gave.min-edu.pt/np3/24.html>).

ProjAVI, *TIMSS 2011 Principais resultados em Ciências*. (Nota de Leitura, acedido em 22 de abril de 2013, em: <http://www.projavi.mec.pt/np4/179/>).

ProjAVI, *TIMSS 2011 Principais resultados em Matemática*. (Nota de Leitura, acedido em 22 de abril de 2013, em: <http://www.projavi.mec.pt/np4/179/>).

SEMINÁRIO

Páginas da Internet com informação relevante

- <http://www.projavi.mec.pt/>
- <http://www.oecd.org/pisa/>
- <http://timssandpirs.bc.edu/>
- <http://www.surveylang.org/>

Introdução²

No quadro de políticas de partilha de conhecimento e de reformas viajantes (Steiner-Khamsi, 2012), que fazem da globalização uma nova forma de governamentalidade educacional (Azevedo, 2007; Rizvi & Lingard, 2010), analiso a promoção das aprendizagens no sistema educativo português em função de razões externas, sobretudo dos testes à larga escala (Pereyra, Kotthoff & Cowen, 2011), e de razões internas, que se prendem com a análise dos efeitos da avaliação externa de escolas (Pacheco, Seabra, Morgado & Hattum-Janssen, 2013). Além disso, argumento que a avaliação das aprendizagens confronta-se com uma situação dilemática, entre o formativo e o sumativo (Pacheco, 2012), tornando-se necessário explicitar de que modo o currículo nacional/supranacional/transnacional se conjuga em planos, programas e metas, por um lado, e em que sentido as políticas de *accountability* (Afonso, 2012) se tornam em mecanismos de regulação pós-burocrática, por outro, com incidência na standardização de resultados.

1. Da governamentalidade educacional

Na síntese teórica que faz dos sistemas de governação, Azevedo (2007, p. 106) escreve sobre a regulação transnacional da educação, afirmando que “o rendilhado do sistema educativo mundial, tal como uma teia de aranha, é uma construção imparável, conflituosa, contínua, silenciosa e, paradoxalmente, quase impercetível para os cidadãos”, mas assente em padrões de institucionalização e em ideologias dominantes que estabelecem formas de governação através da mudança conceitual (Pacheco, 2011). Institui-se, assim, uma legitimidade discursiva que se torna realidade através de uma governação regulada por organismos transnacionais, verdadeiros berços de políticas globais, baseada na abordagem do conhecimento

¹ Instituto de Educação da Universidade do Minho. jpacheco@ie.uminho.pt

² Este trabalho é financiado por Fundos FEDER através do Programa Operacional Fatores de Competitividade – COMPETE e por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia no âmbito do projeto PTDC/CPE-CED/116674/2010.

(Afonso & Costa, 2011), ou nas políticas de partilha de conhecimento (Steiner-Khamsi, 2012), e nas políticas de “accountability” (Afonso, 2012). Alguns desses conceitos são, por exemplo, qualidade, eficácia, eficiência, *performance*, competência, boas práticas, testes e resultados da aprendizagem (Taubman, 2009; Rizvi & Lingard, 2010). Neste sentido, Maroy (2012, p. 91) fala da “accountability” educacional como uma governamentalidade global, cuja fórmula instrumental se processa pela implementação de uma cultura de avaliação, ancorada em sistemas padronizados de garantia da qualidade, em comparações nacionais, em modelos das boas práticas, em *standards* e na relação causal *input-output*.

Neste processo global de governação, as políticas de transferência em torno do conhecimento instituem-se como um referencial coercivo em que as reformas são implementadas mediante “travelling policies - one does not know where they come from, or go to; they are at the same time nobody’s and everybody’s reform” (Steiner-Khamsi, 2012, p. 11) – e na base de uma linguagem técnica que privilegia critérios de convergência. Daí que as reformas educativas se convertam em reformas viajantes, sendo mais caracterizadas pelos aspetos que as unificam e tornam semelhantes do que pelos aspetos que as diferenciam. Aliás, esta tendência é explicada por Waldow (2012) como sendo um processo de globalização que caminha em direção à homogeneização e similaridade, à partilha de *standards* e à intenção de criar um certo grau de uniformidade. Porém, e como vivemos em tempos de uma governação pós-burocrática, a estandardização de processos é substituída pela estandardização de resultados (Maroy, 2012) e a regulação normativa, na procura de uma conformidade de regras, é suplantada pela imposição cognitiva (Correia, Pereira & Vaz), de que os testes à larga escala são um exemplo no domínio do controlo das aprendizagens.

*The Programme for International Student Assessment (PISA*¹), da OCDE (Organização para a Cooperação e Desenvolvimento Económico), e de

¹ Cf. <http://www.oecd.org/pisa/>

Trata-se de um programa orientado para a avaliação de competências básicas dos alunos iniciado em 2000, com a periodicidade de três em três anos, abrangendo cada vez um maior número de países. Sobre dados dos relatórios, Cf. Gabinete de Avaliação Educacional, do Ministério da Educação e Ciência.

acordo com as ideias expressas por vários especialistas (Pereyra, Kotthoff & Cowen, 2011), é um evento global, para além de um circo mediático, que estabelece não só uma governação por comparação de resultados, mas também uma competição baseada em *rankings* de países, não estando ligado ao que os alunos aprendem, a partir de currículo nacional, nem tão pouco aos métodos de ensino/aprendizagem utilizados nas salas de aula, aliás, ao contrário do TIMSS, PIRLS e (EICL)¹. Tendo-se tornado num modelo de governação global, o PISA é um instrumento político (Popkewitz, 2011) que pretende tornar-se num instrumento de melhoria das escolas (Lundgren, 2011), mediante testes e questionários, com primazia para a avaliação de competências em função da preparação dos jovens para o mercado do trabalho e para a cidadania produtiva, procurando respostas para estas questões: “Are students well prepared to meet the challenges of the future? Can they analyse, reason and communicate their ideas effectively? Have they found the kinds of interests they can pursue throughout their lives as productive members of the economy and society?” (OECD, 2012).

Neste caso, o PISA avalia a competência de jovens de 15 anos em termos de uso de conhecimentos ligados ao quotidiano, isto é, é um instrumento de mensuração da literacia “which refers to the capacity of students to apply knowledge and skills in key subject areas and to analyse, reason and communicate effectively as they pose, interpret and solve problems in a variety of situations” (OECD, 2009a), convertendo-se num poderoso instrumento político de monitorização curricular em torno de três questões essenciais: i) identificação de disciplinas nucleares (língua materna, matemática e ciências) e de temas abrangentes, caso da cidadania e literacia financeira; ii) contributo para a melhoria organizacional das escolas; iii) incentivo para a valorização da aprendizagem ao longo da vida.

O impacto do PISA verifica-se mais a nível macro e meso do que a nível micro, sendo um instrumento de monitorização que proporciona informação para a tomada de decisão política a nível nacional e contendo análises que incidem sobre os fatores de aprendizagem, ainda que os resultados não

¹ *Programme for International Student Assessment Progress in International Reading Literacy Study (PIRLS); Trends in International Mathematics and Science Study (TIMSS); European Indicator of Language Competence (EICL)*. Cf. <http://www.projavi.mec.pt/np4/projavi>

possam ser inferidos para qualidade do ensino e para o progresso dos alunos e das escolas. Ao não incidir nos conteúdos lecionados¹, o PISA torna-se num instrumento de avaliação normativa, em que o conteúdo da avaliação se organiza em *standards*, e onde a escola é isolada do contexto, como se os fatores de aprendizagem fossem intrínsecos aos alunos e aos seus países de origem. Tendo escassa utilidade para avaliar a eficácia do sistema educativo, o PISA apresenta os resultados dos alunos numa relação direta entre escola e aprendizagem (“what happens in school has a direct impact on learning” (OCDE, 2009b) e numa relação indireta entre aprendizagem e contexto (“socio-economic background has only a moderate impact on performance (*ibid.*)²), pelo que, nas palavras de Lundgren (2011, p. 28), “PISA is maybe no longer a comparative projet. It is a model for the governing of national school development in a global word”. E esta governação está relacionada com a estandardização curricular, servindo os resultados dos alunos como referencial de avaliação da própria escola quer ao nível do currículo e das formas de diversificação da oferta e da diferenciação pedagógica, quer ao nível das dinâmicas de gestão e liderança,

¹ Cf. OCDE (2009a): “Rather than examine mastery of specific school curricula, PISA looks at students’ ability to apply knowledge and skills in key subject areas and to analyse, reason and communicate effectively as they examine, interpret and solve problems. Every PISA survey tests reading, mathematical and scientific literacy in terms of general competencies, that is, how well students can apply the knowledge and skills they have learned at school to real-life challenges. PISA does not test how well a student has mastered a school’s specific curriculum”.

² Para além dos dois relatórios citados, o ciclo de avaliação de 2009 contém mais dois relatórios: PISA 2009 Results: *Learning to learn. Student Engagement, Strategies and Practices (Volume III)*; PISA 2009 Results: *Overcoming social background equity in learning opportunities and outcomes (Volume II)*. Relativamente a Portugal, e considerando o documento *OECD Reviews of Evaluation and Assessment in Education PORTUGAL* (Santiago, Donaldson, Looney & Nusche, 2012), verifica-se o seguinte: “Regarding the PISA relationship between socio-economic background and performance (*i.e.* between the PISA index of economic, social and cultural status and the performance of 15-year-olds), the following indications emerge: (*i*) Portugal is not statistically different from the OECD average in terms of the percentage of variance in student performance explained by student socio-economic background (strength of the socio-economic gradient), but the impact of socio-economic background on learning outcomes tends to be larger than in most OECD countries; and (*ii*) Portugal is significantly below the OECD average in terms of the score point difference associated with one unit increase in the PISA index of economic, social and cultural status (slope of the socio-economic gradient) (OECD, 2010d) – and there was no significant change between 2000 and 2009 in this indicator”.

permitindo, aliás, uma leitura económica através da relação investimento/produzidade¹.

2. *Standards* curriculares

Como constata Teodoro (2010, p. 58), a OCDE tem “uma agenda global baseada na comparação e, sobretudo, na competição de *performances* dos sistemas educativos”, em que os inquéritos internacionais, caso do PISA, “pouco (ou nada) preocupados com os contextos socio-históricos geradores desses resultados tornaram-se uma das principais tecnologias de governação”. Mesmo remetendo para plano secundário a contextualização dos processos de aprendizagem, há, porém, um *efeito procedimental* no que diz respeito à organização curricular e que se prende com os *standards* de avaliação e com o peso dos testes na aprendizagem.

Em termos curriculares, o PISA é um eficaz instrumento de regulação da aprendizagem centrada nas abordagens *test-driven curriculum*², *competency-based curriculum*³, *goals-based curriculum* e *learning outcomes-based curriculum* (entendidas no prolongamento da abordagem *behavioral performance-based objectives*⁴), contribuindo, também, para uma regulação política baseada na competitividade e mercadorização, como

¹ A este respeito, leiam-se os seguintes artigos publicados no Boletim Económico do Banco de Portugal: Manuel Coutinho Pereira, *Uma análise da evolução do desempenho dos estudantes portugueses no Programme for International Student Assessment (PISA)*, 2011; Manuel Coutinho Pereira e Hugo Reis, *Diferenças regionais no desempenho dos alunos portugueses: evidência do programa PISA da OCDE*, 2012.

² Cf. Janet Penner-Williams, *Tested curriculum*, In Craig Kridel, *Encyclopedia of curriculum studies*, 2010, vol. 2, pp. 878-879: “The tested curriculum consists of that portion of the curriculum over which a student is tested via national norm-referenced achievement tests, state criteria-referenced tests, and teacher-made tests (...) The tested curriculum then becomes the measure of the school’s success (...) If the tested curriculum comprises curriculum based on high standards, it should result in schools with high student achievement”

³ “In a competency-based curriculum, emphasis is not placed upon the learner’s accumulation of memorized knowledge or behavior, but instead on their proficiency in a particular realm”. Cf. Laurel K. Cheavly, *Competency-based curriculum*, In Craig Kridel, *Encyclopedia of curriculum studies*, 2010, vol. 1, p. 131.

⁴ “Behavioral performance-based objectives in curriculum refer to lesson objectives for students set by the teacher that are precise and observable (...) Learning objectives are normally set for each lesson by the teacher and they should drive the instruction and the assessment of student learning”. Cf. Janet Penner-Williams, *Behavioral performance-based*, In Craig Kridel, *Encyclopedia of curriculum studies*, 2010, vol. 1, pp. 131-132.

se a racionalidade escolar fosse determinada pela racionalidade económica. Os testes são o ponto essencial destas abordagens, tornando-se marcadores do processo de desenvolvimento do currículo, em geral, e da aprendizagem, em particular. Por isso, os resultados dos testes tendem a definir a realidade educacional (Taubman, 2009), em que o valor dos números, e da sua comparabilidade a nível nacional e internacional, substituem os alunos, conferindo ao mesmo tempo mais visibilidade aos resultados do que aos processos de aprendizagem. Agindo a partir de uma avaliação que determina o currículo, os professores “are increasingly under pressure to shift the balance from the feedback role of assessment towards its accountability role as policy or curriculum driver” (Young, 2013, p. 113)¹.

A linguagem dos testes, ou da proficiência e da performatividade da aprendizagem, processa-se na base de *standards*, que podem ser utilizados de diferentes modos na organização curricular, por exemplo, providenciando informação sobre o desempenho dos alunos, clarificando posições em tabelas comparativas de grupos, escolas e países e permitindo leituras sumativas e/ou formativas sobre as aprendizagens alcançadas pelos alunos.

Com efeito, o *standard* pode estar relacionado quer com objetivos comportamentais, metas curriculares e resultados de aprendizagem, quer com competências, sobretudo se estas disserem respeito a capacidades de utilização do conhecimento em contextos de ação. Em linguagem avaliativa, e segundo Stake (2006), o *standard* é a quantidade do atributo necessário para a descrição de um critério, fixando o nível de realização do critério e estabelecendo padrões de mensurabilidade mediante indicadores, que funcionam como referências do que é observável. A noção de *standard* aplica-se ao objetivo, à meta e à competência, sendo sempre um processo de referencializar aspetos que contribuem para a mensuração do conhecimento.

No que respeita à organização curricular ao nível do currículo nacional e dos planos que o materializam, o *standard* é uma lógica que serve de referência para o estabelecimento de resultados esperados, isto é, de

¹ O autor faz a distinção entre “assessment as feedback on pupil progress” e “assessment as a driver for the curriculum and pedagogy”.

horizontes de aprendizagem de base que podem ser expressos em objetivos mínimos (em Portugal, com a designação de objetivos curriculares, na década de 1980), ou em competências transversais (na década de 2000), ou em metas curriculares (na década de 2010).

Os *standards* curriculares são um fator de homogeneização de programas e práticas de avaliação (Maradan, 2008), no quadro de políticas educativas centradas nos resultados. Esta cultura de prestação de contas das escolas torna-se mais visível quanto mais for implementada uma prática de formulação e avaliação de indicadores de desempenho. Neste sentido, o currículo corresponde, cada vez mais, ao estabelecimento de resultados esperados, à listagem do que é esperado, à identificação de procedimentos pretendidos, à listagem daquilo que o aluno deve saber e fazer, à mensurabilidade do que deve saber, à identificação do que é fundamental saber e à demonstrabilidade do que realmente sabe. O denominador comum a esta organização curricular diz respeito ao carácter finalizante do resultado de aprendizagem, o que faz com que a sua formulação corresponda a enunciados gerais e sintéticos sobre o que é esperado que o aluno venha a saber e fazer no final de um processo. Se no ensino superior, com o Processo de Bolonha, se utilizam resultados da aprendizagem, nos ensinos básico e secundário faz-se uso, atualmente, das metas curriculares, entendidas como *standards de conteúdos* (Scallon, 2009) e como objetivos terminais (Pacheco, 2011).

Assim definidos, os *standards* estabelecem, por um lado, o que se ensina e o que se aprende em contexto escolar e, por outro, os degraus desejáveis de proficiência. Além disso, “standards serve to standardize work, making it commensurable, erasing differences among individuals, populations, histories, locations” (Taubman, 2009, p. 117). O estabelecimento de metas de aprendizagem (mensuráveis por níveis, ciclos e anos), o que está em sintonia com as políticas de educação que têm vindo a ser implementadas, servirá de referencial para a validação de modalidades de avaliação ligadas às classificações (avaliação sumativa), aos critérios (avaliação aferida) e à comparabilidade nacional e internacional (avaliação normativa).

Na justificação do projeto das metas de aprendizagem, os seus autores, citados no Parecer do Conselho Nacional de Educação (2010, p. 7) afirmam

que “as metas de aprendizagem são entendidas como “evidências de desempenho das competências que deverão ser manifestadas pelos alunos”, sustentadas na aquisição dos conhecimentos e capacidades inscritos no currículo formal, “constituindo por isso resultados de aprendizagem esperados” e “serão sempre expressas em termos do desempenho esperado por parte do aluno”. Equivalendo a resultados esperados da aprendizagem, as metas de aprendizagem aproximam-se mais dos objetivos que das competências, pois, o que a escola confere aos alunos é a aquisição, desenvolvimento e domínio de habilidades e capacidades, nem sempre suscetíveis de serem avaliadas em contextos de ação escolar. Por isso, poder-se-á dizer que as metas representam a reintrodução dos objetivos, como afirma Scallon (2009, p. 126): “à luz de muitos exemplos dados pelos autores, os *standards* de *conteúdos* assemelham-se aos objetivos de aprendizagem que se nos tornaram familiares”, podendo funcionar, de igual modo, como *standards* de *performance* (por vezes, denominados descritores). Se forem considerados objetivos finais e intermédios, as metas de aprendizagem tão-só equivalem a um documento estratégico de regulação do currículo nacional, sendo uma um das mais fortes componentes do currículo nacional, pois os planos curriculares são de natureza normativa e os programas tendem para uma manifesta desvalorização.

3. Entre o sumativo e formativo

A linguagem da eficiência curricular, visando a *performance* dos resultados, tem raízes nas décadas de 1960 e 1970, com a valorização da abordagem *competency-based curriculum*, associada também à pedagogia por objetivos, cuja legitimação no dizer de Taubman (2009), se constrói em função de objetivos militares. Por conseguinte, competências, objetivos e metas são facetas de uma standardização curricular, diferenciando-se, essencialmente, pelo domínio da aplicabilidade (competência), da orientação pedagógica (objetivo) e da terminalidade (meta). Apesar de distintos, mas também com muitos aspetos em comum, objetivo, competência e meta têm contribuído para a valorização dos testes, para a avaliação das escolas pelos resultados, mesmo que seja tido em conta o valor acrescentado da escola, e para a

sumativização das aprendizagens. É sintomático que, no Relatório da OCDE, *Reviews of Evaluation and Assessment in Education PORTUGAL* (Santigao, Donaldson & Nusche, 2012), seja referido que os alunos têm escassa participação na avaliação das aprendizagens e que a avaliação formativa foi substituída pela avaliação sumativa. Esta obsessão pelos resultados, ainda segundo o referido relatório, faz com que os indicadores da qualidade do sistema educativo fiquem quase exclusivamente não só na avaliação externa e no *ranking* de escolas, que pode ser elaborado a partir dos seus resultados, como também no peso excessivo da avaliação sumativa, já que a prática da sala de aula é dominada pelo teste (incluindo os tempos da preparação, realização e correção) e a qualidade do ensino/aprendizagem corresponde à qualidade dos resultados. Advogam, ainda, os autores do referido Relatório, que é preciso colocar os alunos no centro da aprendizagem, tornando exequível a diferenciação dos percursos de formação e garantindo os apoios educativos como forma de melhoria das aprendizagens.

Como refere Looney (2011), a integração das avaliações sumativa e formativa tem sido uma ambição adiada devido aos obstáculos técnicos existentes. Um primeiro obstáculo diz respeito à forma como são implementadas. Enquanto a avaliação sumativa surge na parte final de um processo de ensino/aprendizagem, a avaliação formativa ocorre ao longo desse processo, não sendo possível dissociá-las se se pretende melhorar a qualidade das aprendizagens. Deste modo, a efetividade da avaliação formativa (*ibid*, pp. 8-9) depende das “culturas das salas de aula”, do “feedback”, do “questionamento” e da “autoavaliação e avaliação por pares”, reconhecendo-se que nem, sempre se torna fácil a sua implementação, sendo caracterizada mais pela “exceção que pela regra”, sobretudo em sistemas onde “os professores são encorajados a ensinar para os testes” (*ibid*, p. 10).

Um outro obstáculo depende do modo como a escola está organizada curricularmente ao nível do tempo do trabalho docente, com tendência para a fragmentação da aprendizagem em disciplinas, e do trabalho discente, com a exigência da certificação como forma de progressão. Por fim, e não menos importante, o problema dos rácios professor/turma e professor/aluno e do

aumento crescente do número de alunos por turma. Relendo-se os normativos, seria desejável que a função formativa fosse, *de facto*, implementada em contexto escolar, e que a retenção fosse apenas uma medida pedagógica de última instância, pois a sumativização da avaliação, bem como a correspondente política de accountability, “destroys not only the joy of learning, but learning itself” (Ravitch, 2007, citada por Taubman, 2009, p. 127).

Os efeitos dos estudos internacionais, caso do PISA, tal como foi referido no ponto um deste texto, são de natureza procedimental, já que introduzem sistemas de monitorização a nível macro, de que a avaliação externa de escolas é uma evidência, observando-se que poucos ou reduzidos são os efeitos na avaliação das aprendizagens, a não ser que se considere que reforçam a avaliação sumativa em detrimento da avaliação formativa. No entanto, não se trata de uma conclusão que possa ser generalizada, como o confirmam os resultados da Finlândia. Sendo um dos países com melhores resultados nos vários estudos internacionais PISA, os resultados são atribuídos à i) elevada crença na escolarização, ou seja, “the expansion of schooling, the modernisation of the occupational structure and the construction of the welfare state” (Simola & Rinne, 2011, p. 231); ii) ao elevado *status* social dos professores nas escolas compreensivas – “teaching is clearly the number-one career choice and overtakes traditionally favourite professions such as medicine, law, psychology, engineering and journalism” (*ibid.*, p. 235); iii) à descentralização e autonomia das escolas; iv) à efetividade dos apoios pedagógicos e conseqüente valorização da avaliação formativa.

Ideias para terminar

Mais do que uma conclusão, pois este texto surge na sequência de outros textos que tenho escrito sobre organização curricular e avaliação, pretendo sublinhar ideias que podem servir para uma discussão em torno das aprendizagens. Reconhecendo-se a necessidade de os alunos terem sucesso escolar, mesmo que esteja em causa a sua comparabilidade a nível internacional, pois os educratas tendem a correlacionar resultados/recursos e

ensino público/privado, para defenderem o desinvestimento na educação e a sua mercadorização, de reconhecer que a sofisticação técnica dos estudos internacionais tem um impacto significativo nos sistemas nacionais de educação, reforçando a abordagem *test-driven curriculum* e fazendo com que os *standards* de avaliação sejam equivalentes aos *standards* curriculares. Ainda que tais resultados estejam ligados “à tarefa explícita de restabelecer o orgulho nacional” (Hargreaves, 2004, p. 29), os seus efeitos observam-se no reforço do controlo curricular, em que o currículo nacional se conjuga, acima de tudo, em planos e metas curriculares, e na valorização dos testes, com tendência para que os resultados, e não os alunos, sejam colocados no centro da aprendizagem. Independentemente dos resultados e dos seus valores estatísticos, de mensuração da meritocracia, a linguagem dos *standards* e do “accountability” faz com que “many schools have shifted the general nature of classroom experience from activem relevant, and creative curriculum to drilling specific skills in the critical áreas” (Breault¹, 2010, p. 5).

Esta tendência origina quer uma aprendizagem apressada, em detrimento de uma aprendizagem ampla e profunda (Hargreaves & Fink, 2007)², quer a adoção de lógicas produtivistas nas escolas, em que o desempenho escolar obtido na avaliação se sobrepõe ao processo de aprendizagem, e onde a avaliação sumativa enfraquece a avaliação formativa, quer ainda a redução do tempo de aprendizagem: “The time of learning is the time of providing the right answer. It is the time of stimulus and response. The time of learning is terminable. The time of learning is the time of short-term memory” (Taubman, 2009, p. 184).

¹ Cf. Donna Adair Breault, *Accountability*, In Craig Kridel, *Encyclopedia of curriculum studies*, 2010, vol. 1.

² Ainda neste mar de obsessão pelos resultados, o que interessa são os resultados e não os processos de aprendizagem, como se fosse possível “continuar a alimentar os alunos à força e depois esperar obter *foie-gras*”. Face à aprendizagem apressada, mais guiada por um tipo de currículo Karaoke acelerado, ou seja, de seguimento de guiões concebidos por outros, Hargreaves e Fink contrapõem a noção de aprendizagem ampla e profunda através do conhecer lento, onde “as águas profundas são mais calmas” (*Ibid.*, p. 63), o que leva a uma luta contra “as obsessões políticas e burocráticas que privilegiam a prescrição de competências básicas, os níveis de sucesso orientados para as metas pré-definidas e os *rankings* de escolas” (*Ibid.*, p. 60).

Por outro lado, as escolas são confrontadas com fatores de segregação social, manifestando os professores a tendência para perfilharem discursos de desculpabilização socioeconómica perante os resultados dos alunos. Como se observa pelo caso da Finlândia, a qualidade dos resultados não depende somente do peso dos testes sumativos, estando relacionados, de igual modo, com a valorização social da escola.

Referências bibliográficas

AFONSO, Almerindo (2012). Para uma concetualização alternativa de *accountability* em educação. *Educação & Sociedade*, 33, 119, 471-484.

AZEVEDO, Joaquim (2007). *Sistema educativo mundial. Ensaio sobre a regulação transnacional da educação*. Vila Nova de Gaia: Fundação Manuel Leão.

KRIDEL, Craig (Ed.). (2010). *Encyclopedia of curriculum studies*. London: Sage, Vol. 1-2.

CONSELHO NACIONAL DE EDUCAÇÃO (2010). *Parecer sobre metas de Aprendizagem*. Lisboa: CNE.

CORREIA, José Alberto, Pereira, Luísa Álvares, & Vaz, Henrique (2012). Políticas educativas e modos de subjetivação da profissão docente. *Cadernos de Pesquisa*, 42, 146, 368-387.

HARGREAVES, Andy & Fink, Dean (2007). *Liderança sustentável*. Porto: Porto Editora.

HARGREAVES, Andy (2004). *O ensino na sociedade do conhecimento*. Porto: Porto Editora.

LOONEY, Janet W. (2011). *Integrating formative and summative assessment: progress toward a seamless system?* OECD Education Working Paper, n. 58. Paris: OECD.

LUNDGREN, Ulf. P. (2011). PISA as political instrument: one history behind the formulating of the PISA programme. In M. A. Pereyra, H. Kotthoff & R. Cowen (Eds.), *PISA under examination. Changing*

knowledge, changing tests, and changing schools (pp. 17-30). Rotterdam: Sense Publishers.

MARADAN, Olivier (2007). L'espace curriculaire entre horizon et placher. In F. Audigier, & F. Tutiaux-Guillon (Dir.), *Compétences et contenus. Les curriculums en questions* (pp. 65-84). Bruxelles: De Boeck.

OECD (2009a). PISA 2009 Results: *What Students Know and Can Do. Student performance in reading, mathematics and science* (volume I). Disponível em <http://www.oecd.org/pisa/pisaproducts/48852548.pdf>, acesso a 19 de março de 2013.

OECD (2009b). PISA 2009 results: What makes a school successful? Ressources, policies and practices (Volume IV). Disponível em http://www.oecd-ilibrary.org/education/pisa-2009-results-what-makes-a-school-successful_9789264091559-en, acesso a 19 de março de 2013.

OECD (2012). *PISA 2012 Assessment and Analytical Framework- Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Disponível em http://www.oecd-ilibrary.org/education/pisa-2012-assessment-and-analytical-framework_9789264190511-en, acesso a 19 de março de 2013.

PACHECO, José Augusto (2011). Discursos e lugares das competências em contextos de educação e formação. Porto: Porto Editora.

PACHECO, José Augusto (2012). *Avaliação das aprendizagens. Políticas formativas e práticas sumativas*. Disponível em: <http://hdl.handle.net/1822/21170>.

PACHECO, José Augusto, Seabra, Filipa, Morgado, José Carlos, & Hattum-Janssen, Natascha (2014). *Avaliação Externa. Para a referencialização de um quadro teórico sobre o impacto e efeitos nas escolas do ensino não superior* (em publicação).

PEREIRA, Manuel Coutinho (2011). Uma análise da evolução do desempenho dos estudantes portugueses no Programme for International Student Assessment (PISA). *Boletim Económico do Banco de Portugal-Inverno*, 131-146.

PEREIRA, Manuel Coutinho, & Pires, Hugo (2012). Diferenças regionais no desempenho dos alunos portugueses: evidência do programa PISA da OCDE. *Boletim Económico do Banco de Portugal-Inverno*, 59-83.

PEREYRA, Miguel A., Kotthoff, Hans-Georg, & Cowen, Robert (Eds.).(2011). *PISA under examination. Changing knowledge, changing tests, and changing schools*. Rotterdam: Sense Publishers.

POPKEWITZ, Thomas (2011). PISA: Numbers, standardizing conduct, and the alchemy of Scholl subjects. In M. A. Pereyra, H. Kotthoff & R. Cowen (Eds.), *PISA under examination. Changing knowledge, changing tests, and changing schools* (pp. 31-46). Rotterdam: Sense Publishers.

RIZVI, Fazal, & Lingard, Bob (2010). *Globalizing education policy*. London: Routledge.

SANTIAGO, Paulo, Donaldson, Graham, Looney, Anne, & Nusche, Deborah (2012). *OECD Reviews of Evaluation and Assessment in Education PORTUGAL*: Paris: OECD.

SCALLON, Gérard (2009). *L'évaluation des apprentissages dans une approche par compétences*. Bruxelles: De Boeck.

SIMOLA, Hannu, & Rinne, Risto (2011). Education politics and contingency: belief, status and trust behind the Finnish PISA miracle. In M. A. Pereyra, H. Kotthoff & R. Cowen (Eds.), *PISA under examination. Changing knowledge, changing tests, and changing schools* (225-244). Rotterdam: Sense Publishers.

STAKE, Robert E. (2006). *Evaluación comprensiva y evaluación basada em estándares*. Barcelona: Editora Graó.

STEINER-KHAMSI, Gita (2012). Understanding policy borrowing and lending. Building comparative policy studies. In G. Steiner-Khamsi & F. Waldow (Eds.), *World yearbook of education 2012. Policy borrowing and lending in education* (pp. 5-17). London: Routledge.

TAUBMAN, Peter M. (2009). *Teaching by numbers. Deconstructing the discourse of standards and accountability in education*. London: Routledge.

TEODORO, António (2010). *Educação, globalização e neoliberalismo. Os novos modos de regulação transnacional das políticas de educação*. Lisboa: Edições Universitárias Lusófonas.

WALDOW, Florian (2012). Standardisation and legitimacy. Two central concepts in research on educational borrowing and lending. In G. Steiner-Khamsi & F. Waldow (Eds.), *World yearbook of education 2012. Policy borrowing and lending in education* (pp. 413-427). London: Routledge.

YOUNG, Michael (2011). Overcoming the crisis in curriculum theory: a knowledge-based approach. *Journal of Curriculum Studies*, 45, 2, 101-118.

SEMINÁRIO

Estudos internacionais: o caminho percorrido

Num olhar retrospectivo cobrindo o início dos anos 90 do séc. XX até à atualidade podemos afirmar que sempre que saem os resultados de estudos internacionais, seja o TIMSS (iniciado em 1995 e com uma periodicidade de quatro em quatro anos), seja o PISA (iniciado em 2000, com uma periodicidade de três em três anos) existe sempre polémica, assente em pontos de vistas ideológicos sobre a avaliação, muitas vezes contraditórios. Em particular, a comunicação social atribui uma elevada importância a estes resultados, embora de curta duração, ocupando artigos de opinião, com títulos sonantes.

Sendo um fenómeno internacional, argumentos são então esgrimidos. A favor deste tipo de estudos são apontados atributos qualitativos tais como, rigor, equidade, credibilidade, comparabilidade. Baseiam-se em valores numéricos, objetivos e rigorosos. Colocam os alunos em igualdade de situação. São aplicadas técnicas rigorosas e adequadas aos valores de que se dispõem. Colocam-se à disposição não só análises estatísticas globais, mas também outras, procurando estudar a existência ou não de correlações, que poderão ser mais informativas para um uso eficaz da informação. São dadas a conhecer as situações de diversos países, as suas dificuldades, permitindo compreender a complexidade dos problemas.

Em contraponto, questiona-se o significado atribuído aos valores obtidos através de tratamentos estatísticos mais ou menos elaborados e a sua utilidade pedagógica. O que nos dizem sobre a aprendizagem e os seus processos? Assume-se ou é-se convidado a assumir, que é possível medir os desempenhos dos alunos e que estas medidas têm um valor em si próprias e independente do contexto onde se produz. Mais, para que esta medida seja rigorosa atribui-se a responsabilidade de a determinar a técnicos, a especialistas.

¹ Instituto de Educação da Universidade do Minho (mlsantos@ie.ul.pt)

A comparabilidade é, igualmente, questionada. Críticas vão no sentido de questionar até que ponto não se está a usar como referência um currículo idealizado, mais do que recolher informação sobre se os alunos de um dado sistema educativo estão ou não a aprender o currículo que o seu sistema lhes oferece:

Quando damos a todos, num grupo de alunos oriundos de vários sistemas, um conjunto comum de itens de desempenho, fica assim definido um currículo idealizado que serve de modelo, em contraponto com o desempenho que os alunos têm dentro de cada sistema. (Keitel & Kilpatrick, 1999, p. 72).

Ora, ainda segundo estes mesmos autores, ao não se considerar que os currículos dos vários países diferem entre si, os estudos internacionais continuam a comparar "maçãs com laranjas" (Keitel & Kilpatrick, 1999, p. 72)!

A rentabilização dos custos decorrentes da participação nestes estudos nem sempre é convincente. Por exemplo, no que respeita a Portugal, surge de forma inequívoca a ineficácia pedagógica do sistema de retenções (evidenciada aliás anteriormente), mas sobre a qual nada se procura alterar.

Em síntese, a comunicação social, elabora grandes machetes, incentiva discussões acesas e superficiais, que mais se traduzem em desinformação da sociedade em geral, e que rapidamente desaparecem (a comunicação social amplia esta visão tecnocrata da avaliação enquanto medida, aceitando o valor como verdade absoluta ignorando o contexto onde foi produzido). Contudo, é também possível afirmar que, ao longo destes últimos 20 anos, esta polémica parece ter vindo a diminuir de intensidade. Diversas razões poderão estar na base desta redução. Talvez por se verificar um fenómeno de habituação, talvez por habitualmente não existirem diferenças muito marcantes nos resultados (verificam-se sobretudo pequenas oscilações) e/ou talvez por "o esforço da melhoria que decorre dos resultados das avaliações poderá ter contribuído para questionar cada vez menos o valor e a pertinência das comparações internacionais (Bettencourt: 2010, p. 10).

Os resultados dos estudos internacionais e os seus impactos

Não sendo o único fator, podemos afirmar, sem correr muitos riscos de errar, que os resultados dos estudos internacionais influenciam as políticas educativas, fazendo surgir a necessidade e/ou legitimando mudanças nessas mesmas políticas. Vejamos o que aconteceu em Portugal mais recentemente, a partir do início do séc. XXI, sintetizado na figura 1.

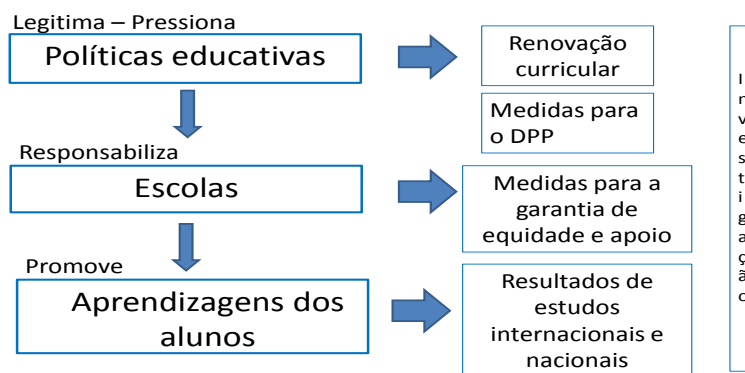


Fig. 1 Níveis de intervenção da política educativa

A existência de estudos internacionais torna claro o que são as tendências curriculares em cada época. O que é mais relevante na aprendizagem dos alunos em cada área curricular. Onde se encontra o enfoque dessa aprendizagem. Note-se que, muitas vezes, quem elabora os itens considerados nos estudos internacionais, propostos aos alunos, são educadores de renome internacional em cada área, como é, por exemplo, o caso da Matemática. Estas tendências curriculares são naturalmente moldadas pelo conhecimento que se vai construindo através da investigação na área. Assistimos, assim, a *processos de renovação curricular*, que não acontecem apenas em Portugal, mas também na grande maioria dos países da Europa, para apenas dar um exemplo.

A acompanhar estes processos de renovação curricular, no que a Portugal diz respeito, pôde assistir-se a um conjunto de medidas de apoio ao *desenvolvimento profissional dos professores* - conhecidas através de programas de formação contínua (em Matemática, Língua Portuguesa, Ensino Experimental das Ciências), outros como o Plano de Leitura, o Plano da Matemática, a elaboração de materiais de apoio aos professores, nomeadamente tarefas de aprendizagem para os alunos.

Ao nível médio, houve uma atenção especial às condições de trabalho e à responsabilidade da escola na aprendizagem dos alunos. São exemplo disso, a sensibilização para formas mais promissoras de responder às dificuldades dos alunos (avaliação formativa no quotidiano da sala de aula); o apoio a projetos de escola para responder a situações particulares (como é o caso do Plano da Matemática I e II); e a criação de condições favoráveis ao trabalho colaborativo entre os professores e co docência (oferecendo condições materiais para a sua concretização).

Embora em educação não se possam, nem se devam, estabelecer relações lineares de causa e efeito, os últimos resultados dos estudos internacionais apontam que estamos no caminho certo. Por exemplo, se recorrermos aos resultados do PISA 2009 (OECD, 2010), no que respeita à Matemática, Portugal é o 4.º país que mais progrediu em literacia matemática, verificando-se uma melhoria significativa (um aumento de 21 pontos, quando comparados com os resultados obtidos em 2003 e 2006 – de 466 para 487). Tal facto leva a uma aproximação à média da OCDE (496, que embora menor do que a de 2003 - 500 - não apresenta uma diferença estatisticamente significativa). Outros dados relevantes dizem respeito à redução da percentagem de resultados de nível 1 ou abaixo de 1 e ao aumento de níveis mais elevados, surgindo pela primeira vez alunos com resultados no nível 6. A melhoria dos resultados obtidos pelos alunos portugueses acontece igualmente na leitura (aumento de 17 pontos) e em Ciências (aumento de 19 pontos).

No TIMSS 2011, os resultados em Matemática dos alunos do 4º ano posicionam Portugal no 15º lugar (Mullis, Martin, Foy, & Arora, 2012). Se atendermos aos resultados médios destes alunos, pelo nível das questões a

que responderam, podemos observar que Portugal apresenta em qualquer um dos quatro níveis um valor médio sempre superior à média (Quadro 1).

Quadro 1. Percentagem de alunos por nível de complexidade dos itens

Nível ¹	Média de todos os países participantes	Portugal
Avançado	4%	8%
Elevado	28%	40%
Intermédio	69%	80%
Baixo	90%	97%

Ainda do relatório do TIMSS 2011 pode ler-se sobre a formação frequentada pelos professores em Portugal e a média dos países que participaram neste estudo (Quadro 2). Da leitura deste quadro pode afirmar-se que, à exceção da área da avaliação, em todos os outros temas considerados a percentagem dos professores que afirmaram ter estado envolvidos é, em média, superior à média do total dos países intervenientes neste estudo.

Quadro 2. Percentagem de professores que afirmaram ter tido formação, em 2009 e 2010, por domínio de formação

	Didáticas e pedagogia	Matemática	Currículo	Avaliação	Tecnologia
Média	46%	44%	41%	37%	33%
Portugal	54%	58%	61%	25%	36%

¹ Nível avançado - aplica compreensão em situações relativamente complexas e explica raciocínios
 Nível elevado - aplica conhecimentos compreensão na resolução de problemas
 Nível intermédio - aplica conhecimentos básicos em situações diretas
 Nível baixo - Tem alguns conhecimentos básicos de matemática

Em termos de resultados nacionais, não nos é possível comparar os resultados obtidos pelos alunos em provas a nível nacional, quer nos exames, quer nas provas de aferição, dado estes testes não terem sido construídos de forma a serem comparáveis. Assim, apresentamos, a título de exemplo, alguns elementos de uma escola da região de Lisboa, no que respeita à evolução das suas classificações finais ao longo do 3.º ciclo de escolaridade (Viana, 2013).

Quadro 3. Comparação dos resultados obtidos no 7º e, dois anos depois, no 9º ano

Nível	Ciclo sem PM		Ciclo com PM		Ciclo com PM		Ciclo com PM(*)	
	2003/04	2005/06	2006/07	2008/09	2007/08	2009/10	2008/09	2010/11
	7ºano	9ºano	7ºano	9ºano	7ºano	9ºano	7ºano	9ºano
1 ou 2	38 %	53 %	23 %	23 %	25 %	22 %	21%	22 %
3	50 %	31 %	31 %	41 %	50 %	49 %	46%	42 %
4 ou 5	12 %	16 %	46 %	36 %	25 %	29 %	33%	36 %

(*) Neste ciclo, devido às restrições impostas pelo Ministério, apenas duas turmas com mais dificuldades tiveram PM no 8º ano.

Da leitura do Quadro 3 (Viana, 2013, p. 10), é possível afirmar-se que a tendência de aumento de classificações inferiores ao longo do ciclo, que se verifica no grupo de alunos não abrangidos pelo Plano da Matemática (PM), não acontece nos grupos de alunos abrangidos pelo PM. É, ainda, de assinalar que ao longo de todo o ciclo, a percentagem de alunos com nível inferior a 3 é claramente menor, quando comparada com o outro grupo.

É do conhecimento geral que as mudanças educativas não produzem efeitos imediatos. Assim, e dado o Plano da Matemática ter sido uma medida que existiu ao longo de seis anos, é igualmente interessante verificar os seus

efeitos pelo número de anos a que os alunos nela estiveram envolvidos (Quadro 4).

Quadro 4. Análise comparativa da avaliação interna no 9º ano

	Ano letivo	Nível 1 ou 2	Nível 3	Nível 4 ou 5
Sem PM	2005/06 (*)	53 %	47 % (**)	
1 ano do PM	2006/07	50 %	26 %	25 %
2 anos do PM	2007/08	34 %	48 %	18 %
3 anos do PM	2008/09	23 %	41 %	36 %
3 ou 4 anos do PM	2009/2010	22%	49%	29%
2 ou 3 anos de PM	2010/2011	22%	42%	36%

Da leitura do Quadro 4 (Viana, 2013, p. 10), emerge de forma inequívoca, que a permanência de uma medida prolongada no tempo traz efeitos positivos na aprendizagem dos alunos a Matemática (50% obtêm nível 1 ou 2 com apenas um ano no PM, passando este valor para 22% ou 23%, quando o período de tempo é de três ou quatro anos). Verifica-se, igualmente, uma tendência para que o aumento de anos com PM seja acompanhado de um aumento na percentagem de alunos com nível 4 ou 5.

Assim, embora estes dados se refiram apenas a uma escola do país, podemos afirmar que a tendência para a melhoria dos resultados dos alunos a Matemática acompanha a tendência igualmente verificada nos estudos internacionais.

A perplexidade do presente

Tendo em conta o que acabei de descrever, é difícil de entender o que está a acontecer em Portugal na atualidade. Senão vejamos:

Valorização dos resultados dos estudos internacionais

Ao contrário do que atrás afirmei, os últimos resultados obtidos pelos alunos portugueses no PISA praticamente não foram noticiados. Se algum comentário houve por parte do Ministério da Educação e Ciência (MEC), este passou despercebido. Mais recentemente, os resultados positivos dos alunos portugueses no TIMSS mereceram uma nota justificativa no *site* do MEC onde constava, entre os três argumentos avançados, a referência à avaliação de Manuais Escolares. Mas acontece que esta medida não abrangeu um só aluno participante neste estudo (alunos a frequentar o 4º em 2010). Ora, a primeira vez que os alunos portugueses foram abrangidos pelos Manuais Escolares avaliados foi em 2010/2011, alunos a frequentar os 1º, 3º, 5º e 7º anos de escolaridade, altura em que se deu a generalização do Programa de Matemática para o Ensino Básico (DGIDC, 2007). Note-se que os alunos que participaram no TIMSS 2011 podiam já estar a trabalhar com este programa, quer por terem professores que estiveram envolvidos no Programa de Formação Contínua para Professores de Matemática, quer através do Plano de Matemática II e apoio à antecipação da concretização deste programa de Matemática. Contudo, estas medidas não foram referidas pelo MEC.

Novas mudanças curriculares

Surgem Metas Curriculares para a Matemática e para a Língua Portuguesa, em 2012, embora as metas de aprendizagem nestas disciplinas estivessem ainda no seu primeiro ano de aplicação. Na sua sequência, um ano mais tarde, um novo programa de Matemática para o ensino básico é homologado, apesar do anterior não ter sido avaliado, nem sequer completado o seu primeiro ciclo para o 3º ciclo de escolaridade.

Novas condições de trabalho nas escolas

Ao nível das escolas, muitas são as alterações que se vão introduzindo, para as quais é difícil apontar razões que não de natureza economicista. É, por exemplo, o caso da redução do número de professores; do aumento do número de alunos por turma; da eliminação de condições favoráveis ao trabalho colaborativo entre professores; da criação de sérias perturbações no funcionamento de escolas onde existem projetos específicos (TEIP), através da alteração das condições de contratação de professores.

Avaliação das aprendizagens

A nova legislação para a avaliação dos alunos introduz mudanças também elas preocupantes. O Despacho normativo n.º 24-A/2012, de 6 de dezembro, regulamenta a avaliação de desempenho dos alunos do ensino básico, surgindo na sequência do Decreto-Lei 139-2012, de 5 julho, que estabelece os princípios orientadores da organização, gestão e do desenvolvimento dos ensinos básico e secundário e a avaliação e certificação dos alunos. Estas novas diretrizes merecem um espaço próprio de análise que não procurarei aqui desenvolver. No entanto, ficam algumas notas de preocupação, emergentes da leitura do Despacho normativo n.º 24-A/2012:

- A importância anteriormente atribuída à avaliação formativa desaparece. Este termo não é possível ser encontrado. Ora a investigação diz-nos quão importante é esta modalidade de avaliação para a aprendizagem dos alunos e mesmo para os resultados em provas de avaliação externa. De uma meta-análise realizada a partir de um número bastante significativo de investigações, cujo objeto de estudo era a prática de avaliação formativa, Black e Wiliam (1998) afirmam que os alunos que frequentam aulas em que a avaliação é formativa não só aprendem mais, como obtêm melhores resultados em avaliações externas, exames, do que aqueles que frequentam aulas em que apenas existe avaliação sumativa.
- Algumas das medidas de promoção do sucesso escolar enunciadas neste despacho normativo traduzem uma visão muito problemática do apoio às aprendizagens. É, por exemplo, o caso da "constituição temporária de

grupos de homogeneidade relativa em termos de desempenho escolar, em disciplinas estruturantes" (Artigo 20.º, c), assente na ideia de que a formação de grupos homogêneos de alunos tem efeitos na aprendizagem. Tal pressuposto não é, contudo, confirmado pela investigação. Se é certo que os melhores alunos progridem mais depressa, tal não acontece com aqueles que apresentam maiores dificuldades de aprendizagem. A resposta encontrada para os alunos com retenções é igualmente muito questionável: "Sempre que se verifiquem retenções, deverão os alunos ser acompanhados pelo serviço de orientação escolar, de modo que possam ser propostas as medidas mais adequadas ao seu percurso escolar, nomeadamente percursos curriculares alternativos, programas integrados de educação e formação, cursos de educação e formação ou cursos vocacionais" (Artigo 24.º). Será que o pressuposto que está na base desta medida assenta na ideia preconizada nos finais do séc. XVIII, em 1799, num discurso proferido por Antoine Tracy, membro do conselho do ensino público francês (in Kahn, 2010, p. 39):

“Em toda a sociedade civilizada, há necessariamente duas classes de homens, uns que tiram a sua subsistência no trabalho físico, outros do produto de certas funções, nas quais o trabalho da mente é mais significativo do que a do corpo. A primeira, é a classe dos trabalhadores, a segunda é a classe dos intelectuais”.

Por último, a ideia de reduzir o tempo escolar para o ensino e a aprendizagem, de forma a se poder introduzir as provas de equivalência à frequência, parece corresponder a uma viragem do entendimento da essência da escola. A principal função da escola não é a educação, mas sim a preparação para o exame!

Em síntese, será que olhando para a situação atual de Portugal não somos levados a repensar a relação existente entre os elementos constitutivos do tema desta mesa redonda: os testes internacionais, o sistema de avaliação externa e a promoção das aprendizagens?

Referências bibliográficas

- BETTENCOURT, A. M. (2010). Prefácio. In Conselho Nacional de Educação (Ed.), *Impacto das avaliações internacionais nos sistemas educativos* (pp. 9-13). Lisboa: ME.
- BLACK, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- DIGDC (2007). *Programa de Matemática do Ensino Básico*. Lisboa: DGIDC/ME.
- KAHN, S. (2010). *Pédagogie différenciée*. Bruxelles: de boeck.
- KEITEL, C. & Kilpatrick, J. (1999). Racionalidade e irracionalidade dos estudos comparativos internacionais. *Educação e Matemática*, 55, 71-80.
- MULLIS, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International results in mathematics*. Boston: International Study Center, Boston College. (disponível em <http://timssandpirls.bc.edu/timss2011/international-results-mathematics.html>)
- OECD (2010). *PISA 2009 Results: What students know and can do. Students performance in reading mathematics and science* (Vol. I). (disponível em <http://www.oecd.org/pisa/pisaproducts/pisa2009/pisa2009keyfindings.htm>)
- VIANA, J. P. (2013). Os efeitos do Plano da Matemática. *Educação e Matemática*, 121, 10-11.

SEMINÁRIO

Avaliações internacionais e legitimação social do sistema educativo

Rui Santos¹

É sempre um prazer colaborar com este Conselho, de que fui membro durante alguns anos e onde me envolvi em algumas discussões que, sem entrar em pormenores, já aqui foram referidas e que são transversais a estes assuntos.

Mas hoje estou aqui como professor universitário, com um interesse de cidadão e com participação em alguns projetos de investigação sobre questões educativas e escolares – não sendo no entanto essa a minha área de estudo prioritária. Devo desde já esclarecer a condição de relativa menoridade da minha posição relativamente a tudo que me antecedeu.

Tendo aceitado o convite, pensei: com o que é que eu posso contribuir? – estando longe, à data, de saber que seria o último interveniente neste painel, o que ainda por cima me acresce a responsabilidade. O que me ocorreu como ideia diretriz, que aliás toda esta jornada confirmou, é que o efeito mais importante destas avaliações de larga escala internacional e comparativa, que esteve presente, mais ou menos conscientemente em ato da primeira à última comunicação, é talvez o seu efeito de legitimação.

Permitam-me que recue um pouco e vos explicito o meu olhar, como sociólogo que sou: as organizações que existem na sociedade, e particularmente as que dependem da dotação de recursos por parte dessa sociedade, necessitam de se legitimar.

As empresas que funcionam num ambiente de mercado de alguma maneira legitimam-se pelas quotas de mercado que detêm, pelos lucros que auferem, pelos dividendos que distribuem aos acionistas (embora mesmo aqui nem

¹ Departamento de Sociologia / CESNOVA. Faculdade de Ciências Sociais e Humanas da Universidade Nova de Lisboa

tudo, de longe, se resume a isso). Há neste caso algumas formas muito tangíveis de mensuração do sucesso ou do fracasso dessas formas organizacionais. Mas há na sociedade uma vasta gama, não menos importante, de organizações que dependem fortemente de uma legitimidade institucional para a sua dotação de recursos, mormente tratando-se de recursos públicos.

Em ambientes de grande incerteza e complexidade, como já aqui foi referido, com resultados de mensurabilidade difícil e frequentemente questionável, e em que as relações de causalidade das formas e os procedimentos organizacionais como os resultados, como quer que sejam medidos, com são elas próprias complexas e difíceis de determinar, a maneira de essas organizações se legitimarem tende a ser a conformidade normativa. Por outras palavras, a aproximação, patenteada pelos seus discursos e pelas suas práticas, a normas consideradas socialmente legítimas; tais organizações tendem a adotar e a encenar formas de racionalidade que sejam socialmente percebidas como “corretas”, a desenvolver rituais de procedimentos e formas discursivas conformes a essas normas, em detrimento da difícil avaliação dos resultados.

Regressando ao nosso tema de hoje, não foi por acaso que a linha de investigação sociológica que teorizou e se vem dedicando a estudar estes processos, o chamado “novo institucionalismo organizacional”, foi lançado por uma investigação sobre escolas públicas norte-americanas nos anos 70 do século passado.

Este tipo de questionamento parece-me especialmente relevante para o nosso caso nacional porque, devido à história recente, o nosso sistema educativo passou, como é sabido, por uma expansão enorme num período relativamente curto: simultaneamente de capacidade, de volume e diversidade de recrutamento, bem como de recursos afetados pela sociedade ao sistema. Isso criou uma série de tensões, às quais a sociedade foi solicitando respostas e que fossem criadas maneiras de aumentar a eficiência do sistema de ensino.

Mas para se medir a eficiência, é preciso primeiro medir a eficácia, ou seja, conseguir uma mensuração de resultados obtidos relativamente aos

desejados, para depois se ver a relação entre os recursos fornecidos e os resultados apresentados pelo sistema educativo e por cada uma das organizações que o integram. Isso foi muito difícil e continua a ser muito difícil.

Por outro lado, o sistema educativo também sofre, desde há muitos anos, de um défice – ou de uma crise, se assim o quiserem chamar – de legitimidade. Em parte, isso deve-se a dificuldades internas devidas à sua própria expansão, à sua própria gestão da diversidade e dos meios para lhe fazer face. Em parte também, compreende-se, creio, por frustração de expectativas. Quando nos anos 70 havia poucas pessoas com ensino secundário ou superior na população portuguesa, a correspondência entre a obtenção de um diploma e a obtenção de uma determinada recompensa no mercado de trabalho era linear, com uma distância relativamente curta. A expectativa de que, por isso, o aumento dos níveis educacionais iria melhorar e harmonizar as oportunidades de vida para os cidadãos era, por isso, muito fortemente fundamentada numa experiência anterior.

Acontece que, à medida que a população se escolariza e que os diplomas se vão tornando menos raros, as recompensas obtidas por esses mesmos diplomas, traduz-se na desvalorização das oportunidades prováveis de cada diplomado de um determinado nível, salvo evolução igualmente forte da procura e da capacidade de incorporação de conhecimento pelas sociedades e pelas economias que, no nosso caso, foi relativamente lenta, embora importante, e muito liderada pela administração e pelos serviços públicos, sistema de ensino incluído. Isto gerou frustrações que o senso comum, combinado com a evidência de alguns desajustamentos e dificuldades do sistema, muito facilmente transformou num discurso de “agora não se aprende nada” (já toda a gente ouviu este discurso nos autocarros, para não dizer à mesa lá de casa). Isto, no essencial, identifica um enorme problema de legitimidade.

O sistema educativo como instituição, as organizações que corporizam essa instituição, desde o topo das organizações estatais responsáveis pelas políticas educativas até à base das escolas e dos grupos profissionais organizados que desempenham essas políticas, têm vindo crescentemente a

sentir a necessidade de se (re)legitimarem, de lutar contra esse déficit de legitimidade.

Foram feitos progressos significativos nas práticas de avaliação do nosso sistema educativo e na incorporação dos resultados escolares nessa avaliação, como já aqui foi referido. Eu lembro-me de, nesta mesma sala, ter assistido não há muitos anos a uma exposição de resultados de uma avaliação externa de escolas em que havia um sentimento de grande otimismo, e de ter cometido a grande inconveniência de perguntar: “então e os resultados?”. Porque grande parte dos parâmetros avaliados, o grande peso da avaliação, incidia, como eu dizia antes, na conformidade a normas procedimentais: tinha isto, tinha aquilo, o corpo docente era qualificado, fazia isto, fazia aquilo... mas os resultados não estavam lá.

Não estavam lá também porque não havia infraestrutura de disponibilização e de análise comparável desses resultados. Quando hoje, como há pouco foi dito, já temos modelos de avaliação externa que incluem uma ponderação dos resultados, tendo em conta a composição socioeconómica e cultural do contexto das escolas, demos já um passo enorme no caminho de organizações que se legitimam fundamentalmente por rituais de conformidade, a organizações que começam a ser obrigadas a questionar os seus próprios resultados.

Mas isso internamente continua a ser complicado. Não preciso de lembrar a ninguém as controvérsias, que aqui foram sendo recordadas pontualmente ao longo da jornada, em torno da fiabilidade da avaliação das aprendizagens, da validade ou mesmo da seriedade dos exames. Os resultados pioram? É déficit das aprendizagens ou é exigência desadequada. Os resultados melhoram? É resultado de boas políticas educativas ou é facilitismo.

Enfim, estas flutuações em termos de discursos públicos e de opinião pública só são, de alguma maneira, colmatadas se tivermos elementos externos que permitam uma aferição comparativa das aprendizagens, que seja independente de nós próprios. Chegados aqui, podemos dizer que os resultados que foram hoje apresentados e amplamente comentados têm um importante efeito legitimador. Não tanto pelo seu valor absoluto ou pela sua

posição num escalonamento momentâneo, mas pelo que representem de progressão. Não é a mesma coisa dizer “estamos muito abaixo do que gostaríamos e não progredimos nada”, ou “estamos *ainda* muito abaixo do que gostaríamos – como foi dito –, mas progredimos *X*, comparativamente com os outros países”. Este é por si só um efeito bastante benéfico e positivo.

Passando, muito rapidamente, a questões mais críticas: este tipo de formas de avaliação internacional tende (passo o palavrão!) ao isomorfismo, ou seja, leva as organizações a procurarem encontrar pontos de referência – um dos princípios da ideia de *benchmarking* – e a adequar-se à formas organizacionais, procedimentais e inclusive discursivas mais legítimas, porque conotadas com maior eficiência em termos de resultados. E aqui que se enquadram alguns elementos de crítica que já foram hoje aqui formulados.

Com efeito, ouvimos hoje apelos – e não estando aqui para ajuizar se com ou sem razão, à partida inclinar-me-ia a dizer que com razão – para que elementos das formas de evidenciar e mensurar aprendizagens que estão incluídos nestes instrumentos internacionais sejam, como de facto têm vindo a ser, incorporados nos objetivos de aprendizagem, nas práticas pedagógicas, e nas práticas de avaliação. Deste modo, tende-se a padronizar as nossas metas de aprendizagem, práticas de ensino e de avaliação pelos pressupostos subjacentes aos estudos internacionais, estabelecendo quais devem ser as aprendizagens e como devem ser mensuradas. Com efeitos, imaginemos, possivelmente positivos e que permitirão explicar alguma da evolução registada.

Mas agora a questão é em que medida essa padronização é ela própria positiva. Temos tido suficiente reflexão crítica sobre os pressupostos dos indicadores de aprendizagem que estes estudos revelam, a sua adequação aos propósitos do nosso sistema educativo? Pessoalmente, desconheço.

Em termos de representação internacional, está aqui quem certamente saberá responder melhor do que eu, e estou certo de que os nossos representantes nas organizações que conduzem estes estudos têm feito o melhor esforço nesse sentido. Mas a incorporação desses pressupostos nas práticas

escolares e nas aprendizagens, tem vindo a ser feita respeitando os requisitos dessas mesmas aprendizagens e a sua relação com outras aprendizagens que a escola também está obrigada a facultar? Em suma, se não sei adiantar respostas, há pelo menos questionamentos a colocar.

Queria acabar, porque também já esgotei o meu tempo, com a ideia de que eu acho que estes resultados são globalmente muito positivos *se* a progressão for sustentada, precisamente porque essa progressão nos dota de meios de legitimação de um sistema educativo (e preferiria falar de um sistema educativo do que de políticas educativas específicas, cuja “rotatividade” corre o risco de retirar coerência ao próprio sistema) que carece muito de legitimidade social e que quero crer que assim a irá ganhando.

Avaliações internacionais e desempenho
dos alunos portugueses

Síntese e conclusões

SEMINÁRIO

Nesta discussão toda a preocupação em atribuir eventuais resultados positivos a este ou àquele governo, a este ou àquele partido, ou em excluir resultados positivos, por causa deste ou daquele partido, não permite serenidade nenhuma nem é útil, é mesmo uma perda de tempo, com o perigo de se cair no ridículo, o que é ainda pior.

A primeira questão em debate era a de saber se houve progresso. A maioria dos intervenientes considerou que houve progressos. Julgo que só houve uma voz que tentou dizer que está tudo quase na mesma, e que não há nada a celebrar. E pode ser que tenha razão. Costumo dizer que a identificação das verdadeiras razões não resulta de um voto democrático.

O progresso verifica-se só na amostra ou também na população? Esta é outra questão também interessante. Foram emitidas algumas dúvidas sobre a qualidade da amostra, mas não me parece que tenham sido invocados argumentos suficientes para questionar a sua representatividade.

Pode, no entanto, levantar-se uma questão que não surgiu no debate: a que se deve o facto de serem muito diferentes destes os resultados nas provas nacionais, em que participam todos alunos? A este propósito, há uma dúvida que tem surgido na praça pública, que é a de saber se o que foi avaliado nas provas internacionais corresponde ao que é relevante saber. O que está implícito é que para haver progresso basta elaborar um teste fácil e não em função daquilo que é relevante saber. Isto é, as provas nacionais mediriam o que é relevante saber, nesta idade, o que os testes internacionais não fazem.

Se houve progresso, a outra questão é a de saber se o progresso é consolidado. Julgo que aqui é que está a maior dúvida. Temos de esperar, temos de ver as próximas provas. A este propósito, só uma pequena nota. Houve uma referência aos resultados do PISA, em que se registou um pequeno avanço, e aos destes estudos, em que o progresso foi maior. Convém ter presente que no PISA são avaliados os alunos de uma determinada idade, 15 anos, mesmo que estejam no 7.º ano de escolaridade,

¹ Conselho Nacional de Educação

em que ainda não aprenderam o que vai ser ensinado no 8.º e no 9.º, o que pode justificar menor progressão. Porque, quando se analisam os resultados dos alunos que têm 15 anos e estão no 9.º ano, os resultados já são muito melhores. Estes estudos que aqui nos reúnem abrangem, se não erro, todos os que estão no 4.º ano, mesmo que tenham uma idade superior à normal para isso.

Passemos ao segundo grupo de questões, o dos fatores de progresso. Obviamente que não podemos, com os dados e os estudos que existem, indicar causas. A maioria, se não a totalidade, dos que estão aqui presentes sabe que uma correlação não é uma causa, mesmo que não sejamos cientistas. Os fatores de progresso referidos constituem, de facto, hipóteses, mais ou menos fecundas, que podem originar estudos mais aprofundados.

Não saber quais são as causas não significa que não as haja. Mas não tem sentido afirmar que os fatores apresentados pelos outros não são causas e fazer crer que aqueles que nós apresentamos o são. Também são meras hipóteses. Por exemplo, o impacto positivo neste progresso da generalização das provas nacionais efetuada por um ministro, também é apenas uma hipótese, e não uma causa.

Ao longo do dia foram enunciadas muitas hipóteses de causas possíveis. Na maioria dos casos, foram indicadas diferenças em alguns fatores entre os anos noventa e a atualidade e não apenas a partir de 2006.

Obviamente que há uma grande diferença entre os anos noventa e o tempo presente nas condições socioeconómicas da maioria da população. Normalmente estas condições são consideradas um fator importante nos resultados escolares conseguidos. O desenvolvimento socioeconómico está relacionado com os recursos existentes nas famílias e com o nível de escolaridade dos pais, aspetos normalmente considerados como suscetíveis de criar um ambiente mais favorável à aprendizagem.

Verifica-se também uma diferença no nível de qualificação ou no grau académico dos professores, que poderá, ou não, ter influência no seu desempenho. Parece igualmente haver uma diferença na sua preparação em didáticas específicas de cada disciplina, cuja importância tem vindo a ser

valorizada, não bastando uma didática geral comum a todas as disciplinas e a todas as idades. Há diferença, terá havido influência?

Houve um maior investimento a partir dos anos noventa na formação contínua dos professores, pode não ter tido impacto nenhum, mas há uma diferença entre o princípio dos anos noventa e agora. A partir de um certo momento, a formação contínua investiu cada vez mais na mudança das práticas na sala de aula e na sua articulação com os programas. Terá tido influência?

Foi, também, a partir de meados dos anos noventa que foi aumentando progressivamente a frequência da educação pré-escolar. Entre as crianças que agora estavam no 4.º ano, há mais com frequência de maior número de anos desta educação do que no início dos anos noventa. Não teve efeito nenhum? É possível, mas também é possível que tenha tido. É uma questão de fazer os estudos necessários.

Também foi referida a progressiva emergência de uma cultura de avaliação, com os seus aspetos negativos, que são evidenciados, mas também com os seus aspetos positivos, devido sobretudo ao acento colocado na centralidade dos resultados. Lembro-me que, 20 anos atrás, a maior parte das escolas nem sabia quantos alunos reprovavam, quantos é que eram repetentes, quais os resultados. Foi graças às diferentes metodologias de avaliação das escolas e à divulgação dos resultados de provas nacionais que tal começou a constituir uma preocupação e progressivamente promoveu uma focalização nos resultados e, portanto, na aprendizagem. Isto é, como dizia esta manhã, até um dado momento falava-se mais da educação para todos. Hoje, até o Banco Mundial, na sua estratégia de 2010-2020, já não fala em educação para todos, fala em aprendizagem para todos. Também esta ideia foi entrando e talvez possa ter tido influência.

Claro que há a ausência de resultados nos níveis mais elevados. Já no plano europeu, quando nos comparamos com as metas colocadas, dizemos muitas vezes: «ainda estamos distantes daquelas metas, ao contrário do que acontece com outros países...». Neste caso a comparação é feita com países que têm pontos de partida diferentes. Mas há outra perspetiva possível, que

é a de considerar o progresso em função do nível de partida. Julgo que é razoável ter isso em conta.

Neste seminário, quando se tratou de explicar a ausência de níveis mais elevados, houve quem admitisse que isso acontecia porque não se responsabilizavam os alunos. Penso que sim, que a aprendizagem depende muito do desempenho dos alunos. Mas eu julgo ter percebido naquele « não se responsabilizam os alunos » que devem desresponsabilizar-se os professores. Desde que os professores sejam muito sábios e digam aquilo que sabem, cumprem o seu papel. Transmitiu o que sabia. Se o aluno não aprendeu, não cumpriu o papel dele. É um pouco simplista.

A terceira questão que na abertura propus para reflexão foi a da sustentabilidade destes resultados. Saber se estes resultados são sustentáveis é ainda mais difícil do que identificar as causas do progresso registado. Houve quem dissesse que não vão ser, porque as políticas e os fatores a que são devidos estão a desaparecer. Isso é uma hipótese. Obviamente que os resultados dos próximos exames nacionais do 4.º ano vão ser objeto de comparação e, nessa altura, o que será preciso comparar será o tipo de provas. E se são muito diferentes, porque é que são diferentes ? Julgo que será uma questão interessante.

Finalmente, uma conclusão a que muitos chegaram e para que muitos foram apontando, é a de que são necessários estudos mais específicos a partir dos dados que existem. Não bastará ao ministério disponibilizar os dados, é necessário ainda fomentar os estudos e, eventualmente, constituir uma agenda da investigação. Será importante dizer: « queremos que se estude esta questão e financiamos ». A fixação de uma agenda de investigação é importante.

De qualquer modo, embora a investigação seja importante, para agir não se pode estar à espera dos seus resultados. É preciso correr sempre o risco de utilizar boas hipóteses, atribuindo os resultados a este ou àquele fator, e investir nas políticas que colocam esses fatores em ação.

Para terminar, gostaria de assinalar que não se falou muito, durante o dia, sobre a utilização dos resultados destas provas para melhorar, para conseguir melhores resultados ainda. Referiu-se, é verdade, que era preciso

insistir mais na Matemática, ou na aplicação das Ciências, e que havia resultados menos bons a que era necessário dar atenção. Mas a preocupação com a melhoria talvez não tenha sido o objetivo das intervenções. Claro que estava implícita quando foi feito o enunciado dos fatores que eram bons para o progresso, que seriam bons para não recuar e, eventualmente, para melhorar.

Resta-me em nome da Comissão que coordeno e em nome do Conselho, se a Senhora Presidente me permite, agradecer aos intervenientes pela riqueza das intervenções que fizeram e pelos debates que as mesmas animaram e valorizaram. Julgo que cada um leva a sua síntese sobre o que poderá contribuir para que mais alunos em Portugal aprendam mais.