

PISA, TIMSS e PIRLS em Portugal: Uma análise comparativa

RESUMO

Neste artigo apresentamos uma análise comparativa entre três questionários internacionais: o PISA, conduzido pela OCDE, e o TIMSS e o PIRLS, realizados pela IEA, tendo em conta os seus objetivos gerais e correlações entre os respetivos dados referentes a Portugal (por regiões da NUTS III). A revisão bibliográfica e documental mostrou como os resultados podem ser reformulados ou reinterpretados pelos diferentes atores sociais e que são vistos pelos decisores políticos e pelas organizações como uma ferramenta de comparação dos sistemas escolares, reveladora de pontos fortes e pontos fracos. A análise dos resultados gerais evidencia que Portugal tem vindo a melhorar os seus resultados no PISA, nos três domínios avaliados (leitura, matemática e ciências), que melhorou no TIMSS em matemática e piorou em leitura no PIRLS. A análise comparativa por regiões revelou que existe uma diferenciação dos resultados por regiões e que, enquanto o fator domínio não se apresenta como relevante, o fator objeto de avaliação em cada estudo parece determinante para os resultados dos questionários. Revelou, ainda, que há uma relação direta entre o PIB/habitante e os resultados dos questionários nas regiões com valores extremos de PIB/habitante, mas que essa relação não existe para a generalidade das regiões.

Palavras-chave: Regiões NUTS III; PISA; TIMSS; PIRLS

1. INTRODUÇÃO

Os estudos internacionais desenvolvidos a partir de questionários são cada vez mais numerosos e abrangem diferentes domínios. Os da educação dizem respeito a diversos níveis de ensino (da educação básica à universidade) e também aos adultos não escolarizados. Várias revistas e obras especializadas fazem eco destes estudos, insistindo sobre os seus objetivos, os métodos utilizados e os seus resultados. O *Programme for International Student Assessment* (PISA), conduzido pela Organização para a Cooperação e Desenvolvimento Económico (OCDE) e que incide tradicionalmente sobre a leitura, as ciências e a matemática, e mais recentemente inclui a resolução colaborativa de problemas e a literacia financeira, o *Trends in International Mathematics and Science Study* (TIMSS), que abrange a matemática e a física,

Vítor Rosaⁱ
Universidade Lusófona
de Humanidades e
Tecnologias, Portugal

João Sampaio Maiaⁱⁱ
Universidade Lusófona
do Porto, Portugal

Daniela Mascarenhasⁱⁱⁱ
Instituto Politécnico
do Porto, Portugal

António Teodoro^{iv}
Universidade Lusófona
de Humanidades e
Tecnologias, Portugal

e o *Progress in International Reading Literacy Study* (PIRLS), que incide sobre a leitura, ambos realizados pela *International Association for the Evaluation of Educational Achievement* (IEA), são, certamente, os mais conhecidos e tornaram-se uma referência para a governação nos campos da educação e da formação. Dado que todos estes estudos se centram essencialmente na avaliação dos alunos, foram designados pelo acrónimo ILSAs, que significa *International Large-Scale Assessments*. Vários documentos e bases de dados, de acesso livre, podem ser encontrados(as) na internet e, ao longo dos anos, todos eles se dotaram de recursos humanos, metodológicos e tecnológicos importantes.

Será que estes questionários são um instrumento útil para um diagnóstico dos sistemas educativos? Eles pretendem funcionar como um instrumento de controle dos níveis das competências adquiridas pelos estudantes. As referências utilizadas para definir a escala em que são medidos os níveis de competência são precisamente aquelas que fornecem a distribuição dos resultados do conjunto de alunos que participaram nos testes (a média é fixada em 500 e o desvio-padrão em 100). A comparação entre esses resultados permite dizer se Portugal tem o mesmo nível de outros países¹. Ao lermos a imprensa sobre estes questionários, verifica-se, muitas vezes, que incidem sobre as comparações internacionais, mas a apresentação dos níveis de literacia (de leitura, matemática, ciências, ou as mais recentes, financeira e de resolução de problemas em contexto colaborativo) é sumária, limitando-se a breves descrições.

Os trabalhos sobre a participação de Portugal no PISA, no TIMSS e no PIRLS são poucos e muito restritos nos seus campos de abordagem (Carvalho & Costa, 2009; Conselho Nacional de Educação [CNE], 2013; Marôco, 2020). Com este artigo, pretendemos alargar o espaço de informação sobre esses questionários e saber se é possível compará-los, tendo em conta os seus objetivos. Fazemo-lo a partir dos dados globais conhecidos dos países participantes, com um olhar particular para os dados de Portugal, mas numa perspetiva diferente dos autores atrás referidos. Enquanto, por exemplo, Marôco (2020) compara os resultados globais de Portugal nos estudos do PISA, do TIMSS e do PIRLS com os de outros países na sua evolução ao longo do tempo, neste artigo, além de fazermos referência à evolução dos resultados globais de Portugal nesses estudos, olhamos para os resultados discriminados por regiões da Nomenclatura das Unidades Territoriais para Fins Estatísticos (NUTS) III² em cada um dos estudos referidos, num período restrito de tempo, e comparamo-los entre si e com o Produto Interno Bruto (PIB) por habitante. Isto é, fazemos uma análise comparativa correlacional dos resultados regionais obtidos pelos estudantes portugueses naqueles estudos e que não encontrámos em nenhum outro artigo.

Assim, resumindo, temos como objetivos comparar os resultados de Portugal nos estudos do PISA, do TIMSS e do PIRLS no que respeita às respetivas evoluções e a algumas das suas características e fazer uma análise correlacional entre esses resultados, tendo como base as regiões da NUTS III.

1. De notar que a IEA não tem a OCDE como referência nos seus estudos e que o PISA inclui vários países que não são membros da OCDE.

2. As NUTS estão organizadas em três níveis. O NUTS III é constituído pelas 25 sub-regiões (23 no continente e 2 nas Regiões Autónomas dos Açores e da Madeira), que correspondem a entidades intermunicipais e em que se dividem as sete regiões do NUTS II (5 no Continente e as 2 Regiões Autónomas dos Açores e Madeira). O nível NUTS I é constituído por três unidades: Continente, Região Autónoma da Madeira e Região Autónoma dos Açores.

2. ASPETOS GERAIS DO PISA, DO TIMSS E DO PIRLS

2.1. PISA

A missão da OCDE é a de promover políticas que melhorem o bem-estar económico e social em todo o mundo. Convencida da importância da educação no desenvolvimento económico dos países e no bem-estar dos indivíduos, não poderia deixar de se interessar por este domínio, tendo-se tornado uma das principais organizações internacionais de análise da educação em todo o mundo (Rizvi & Lingard, 2006; Teodoro, 2010, 2019, 2020). A adesão de países a esta organização tem vindo a crescer ao longo dos vários ciclos do estudo. No quadro da sua reflexão, “e motivados pela vontade de aplicar o rigor da pesquisa às políticas de educação” (Schleicher, 2018, p. 17), a OCDE lançou o programa PISA. Com ele, e avaliando os alunos de 15 anos de idade³ e que frequentam, pelo menos, o 7º ano de escolaridade, pretende verificar se as escolas de cada país participante preparam os seus jovens para exercerem o papel de cidadãos na sociedade contemporânea. Não pretendem avaliar os currículos diretamente, mas as competências que os alunos adquiriram para a vida ativa. Este estudo internacional pretende avaliar como os alunos conseguem mobilizar as suas competências em três domínios de literacia: leitura, matemática e ciências. Procura estudar também a resolução colaborativa de problemas e a literacia financeira.

Três dimensões foram tomadas em conta para elaborar o material de avaliação do PISA: os tipos de tarefas de leitura, os tipos de texto e o contexto. As tarefas de leitura (questões) definem a natureza das capacidades a avaliar e situam-se a um “nível macro”. Podem ser agrupadas em três categorias: 1) encontrar informações em materiais variados (informar-se); 2) interpretar o que se lê (interpretar); 3) refletir sobre o que se leu e reagir (refletir). Os conteúdos são organizados em categorias. Por exemplo, as categorias referentes à matemática são as seguintes: transformações e relações; espaço e forma; quantidade; incerteza e dados. As categorias não têm sido alteradas desde 2000. No entanto, e sem modificar o espírito, a sua apresentação evoluiu para melhor colocar em evidência a importância dos conteúdos ensinados e a adequação existente entre as categorias e os objetivos definidos pelos programas curriculares dos países participantes.

Querendo inscrever as questões de avaliação da literacia no mundo real e na vida dos alunos, o PISA distingue e hierarquiza os tipos de situações nas quais os problemas são encontrados. Até 2012, o quadro de referência apontava para que as situações se podiam enquadrar numa grande variedade de contextos:

- Situação próxima do aluno: a sua vida pessoal e familiar;
- Situação afastada do aluno: a sua vida escolar, o mundo do trabalho e os lazeres;
- Situação ainda mais afastada do aluno: a comunidade local e a sociedade nacional ou mundial;
- Situação extremamente afastada do aluno: as situações de natureza científica.

3. Os alunos desta idade encontram-se no final da escolaridade obrigatória na maioria dos países da OCDE.

No quadro de referência de 2012, os tipos de situações tornam-se em categorias de contexto: pessoais, profissionais, sociais e científicos.

A Tabela 1 apresenta, como exemplo, as categorias do PISA-2015 para a literacia matemática.

Tabela 1

Síntese das categorias PISA-2015 em literacia matemática

Domínios de Avaliação	Categorias de cada um dos domínios avaliados
Processos	Formular Empregar Interpretar
Conteúdos	Transformações e relações Espaço e forma Quantidade Incerteza e dados
Contextos	Pessoais Profissionais Sociais Científicos

Fonte: Adaptado de OCDE (2013).

Em cada ano do questionário são avaliados os três domínios (leitura, matemática e ciências), mas um deles é considerado como “maior” (M), ao qual são consagrados dois terços da duração dos testes; os dois outros domínios são considerados “menores” (m). Na Tabela 2, e a sombreado, salienta-se o enfoque principal colocado nos diferentes tipos de literacia avaliados nas edições do PISA (2000 a 2018).

Tabela 2

Enfoque nos tipos de literacia avaliados nas edições do PISA

Anos	Literacia de Leitura	Literacia Matemática	Literacia Científica
2000	M	m	m
2003	m	M	m
2006	m	m	M
2009	M	m	m
2012	m	M	m
2015	m	m	M
2018	M	m	m

Portugal participou em todas as edições (2000, 2003, 2009, 2012, 2015 e 2018). Na edição do PISA-2018, numa lista ordenada de 79 participantes, Portugal ficou em 24^o lugar na literacia científica, 24^o na literacia de leitura e 22^o na literacia matemática, com 492 pontos em cada domínio, ficando acima da média da OCDE em todos os domínios⁴. Desde a edição do ano 2000, Portugal tem registado uma tendência de melhoria dos resultados nos três domínios analisados, conforme se mostra na Tabela 3 e nos pequenos gráficos (*sparkline*). No entanto, de 2015 para 2018 houve uma pequena descaída, não estatisticamente significativa, a leitura e a ciências.

4. Os resultados obtidos são expressos numa escala de 0 a 1000, sendo a média de 500.

Tabela 3

PISA, literacias de leitura, matemática e ciências (média e erro-padrão da média), 2000-2018

Ano	Literacia de Leitura (Médias)	Literacia de Leitura (Erro-padrão)	Literacia de Matemática (Médias)	Literacia de Matemática (Erro-padrão)	Literacia de Ciências (Médias)	Literacia de Ciências (Erro-padrão)
2000	470	4.5	454	4.1	459	4.0
2003	478	3.7	466	3.4	468	3.5
2006	472	3.6	466	3.1	474	3.0
2009	489	3.1	487	2.9	493	2.9
2012	488	3.8	487	3.8	489	3.7
2015	498	2.7	492	2.5	501	2.4
2018	492	2.4	492	2.7	492	2.8

Fonte: OCDE (1999, 2003, 2013, 2019a, 2019b); Lemos (2014).

2.2. TIMSS

O TIMSS, promovido sob a égide da IEA, realiza-se de quatro em quatro anos e é uma avaliação internacional da literacia em matemática e em ciências, dos 4º, 8º anos e, recentemente, do 12º ano de escolaridade. Procura avaliar o desempenho dos alunos em diversos contextos de aprendizagem, tentando medir o que designam por eficácia escolar ou competências académicas. Distingue vários níveis de currículos, interessa-se pelos conteúdos de ensino e a sua presença nos programas escolares, pelos contextos familiares e escolares, e analisa a aprendizagem dos alunos dentro de cada contexto. Isto é, o TIMSS diferencia-se do PISA porque se centra nos currículos, recursos e ambientes familiares e escolares e no sucesso dos alunos na escola (Martin et al., 2013), enquanto este pretende avaliar essencialmente competências para a vida.

Este questionário internacional inscreve-se na história da IEA, que se inicia em 1964 com o *First International Mathematics Study* (FIMS) e depois com o *Second International Mathematics Study* (SIMS), em 1982. O TIMSS, de 1995 a 1999, designou-se *Third International Mathematics and Science Study*, assumindo o acrónimo por que é conhecido, para se tornar, em 2003, o *Trends in International Mathematics and Science Study*.

O número de países que participam no TIMSS é irregular e muitos só participam em um ou dois níveis de escolaridade. Em 2015, 57 países participaram no estudo: 48 para o 4º ano de escolaridade, 40 para o 8º ano de escolaridade e 9 (EUA, França, Itália, Líbano, Noruega, Portugal, Rússia, Eslovénia e Suécia) para o nível terminal científico (12º ano) (Conseil National d'Évaluation du Système Scolaire [CNESCO], 2016). Para o nível terminal, o estudo designa-se *TIMSS Advanced*. Ele avalia a matemática e a física.

Integrando o conjunto dos 46 países aderentes, Portugal participou na edição de 1995, a primeira, avaliando os alunos dos 3º e 4º anos de escolaridade (1º ciclo do ensino básico) e dos 7º e 8º anos (3º ciclo do ensino básico). Depois de um hiato de 16 anos, correspondendo a três edições do TIMSS, voltou a participar em 2011 e em 2015 (Marôco, 2016b, 2016c), mas somente com alunos do 4º ano de escolaridade. Em 2015, envolveu cerca de 5000

alunos provenientes de 158 escolas. Portugal ficou em 13^o lugar a matemática e em 32^o lugar a ciências. Face aos resultados de 2011, subiu dois lugares no *ranking* (em matemática), mas desceu 14 pontos a ciências (em 2011 ocupava o 19^o lugar) (cf. Tabela 4). No TIMSS *Advanced*, e no reduzido universo de países participantes, Portugal ficou classificado em 5^o lugar a matemática e em 4^o lugar a física.

Tabela 4

TIMSS, médias para a matemática e para a ciência, 4.º ano de escolaridade, Portugal, 1995 e 2015

Anos	Média para a Matemática (4.º ano de Escolaridade)	Erro-Padrão	Média para a Ciência (4.º ano de escolaridade)	Erro-Padrão
1995	475	3.5	480	4.0
2011	532	3.4	522	3.9
2015	541	2.2	508	2.2

Fontes: Marôco (2016b); Martin e Kelly (1997).

Em 2019, decorreu o eTIMSS8, versão digital do estudo, que incluiu, também, os alunos do 8^o ano de escolaridade.

O Instituto de Avaliação Educativa (IAVE), organismo público de regime especial sob a tutela do Ministério da Educação que assegura a coordenação e a aplicação dos estudos internacionais em Portugal, compilou, recentemente, um conjunto de itens utilizados em edições anteriores do estudo (IAVE, 2019). Devido ao processo de renovação cíclica, esses itens deixaram de integrar o teste, pelo que foram libertados pela IEA com o propósito de mostrar ao público a natureza das questões colocadas aos alunos.

2.3. PIRLS

Com início em 1995, e com uma periodicidade quinquenal, o PIRLS é outro estudo internacional promovido pela IEA, que avalia o desempenho em literacia de leitura dos alunos a frequentar o 4^o ano de escolaridade e, como o TIMSS, centra-se nos currículos e outros fatores familiares e académicos e no sucesso escolar dos alunos. Como este, procura apreender também os elementos de contexto cultural, socioeconómico, pedagógico, familiar e individual, que podem estar associados às *performances* observadas, fazendo análises intra e interpaíses. Para o efeito, recolhe informações sobre as escolas, as classes e os contextos educativos através de questionários dirigidos aos diferentes interlocutores da escola: alunos, encarregados de educação, professores e diretores. O protocolo de avaliação (textos e itens) é baseado no cruzamento de duas dimensões: os processos de compreensão e os objetivos de leitura. Os questionários sobre os hábitos de leitura e atitudes sobre a escrita são submetidos aos alunos e aos encarregados de educação. Outros questionários procuram avaliar as práticas pedagógicas e a formação dos professores e dos diretores de escola.

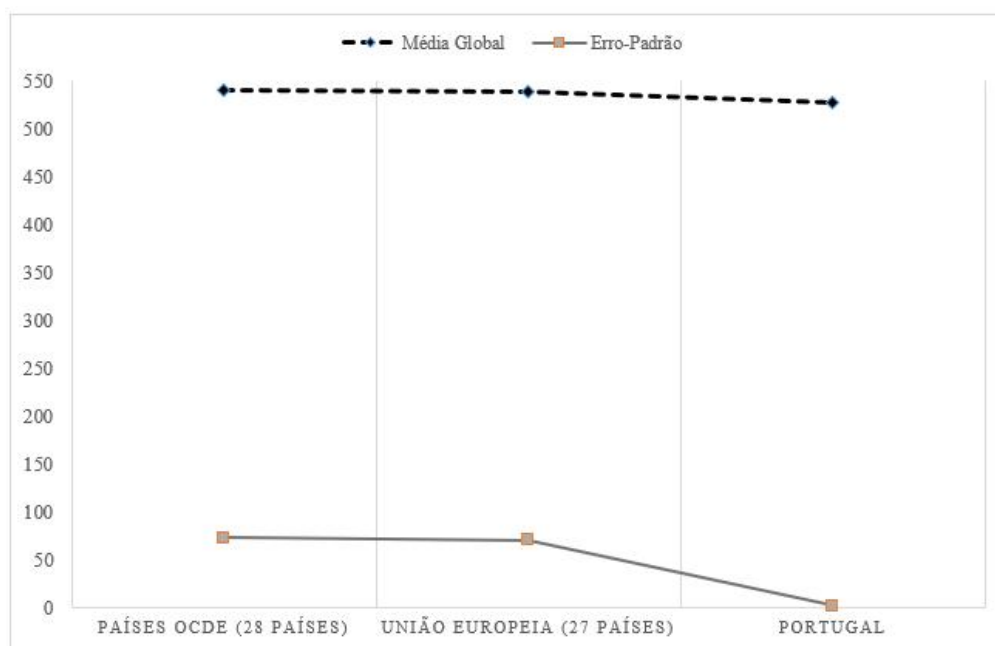
A amostragem, tal como para o TIMSS e o PISA, segue um protocolo rigoroso para este tipo de questionários internacionais, garantindo a

sua representatividade e que assegura a comparabilidade dos dados com outros países.

Portugal participou na primeira edição, em 2011. No PIRLS-2016, introduziu-se uma novidade, procurando-se avaliar a competência de leitura *online* dos alunos (ePIRLS). Se olharmos para os dados médios dos países da OCDE e da União Europeia (UE), verifica-se que Portugal ficou abaixo 13 e 12 pontos, respetivamente (cf. Figura 1).

Figura 1

Média global e erro-padrão da média, no PIRLS, segundo o agrupamento dos países, 2016



Fonte: IEA.

O grupo de especialistas do PIRLS definiu uma escala de competências em quatro níveis. O primeiro (baixo) corresponde a uma pontuação global entre 400 e 475. O segundo nível (intermediário) corresponde a uma pontuação entre 476 e 550. Os limites do terceiro nível (elevado) são 551 e 624, e uma pontuação global igual ou superior a 625 corresponde ao nível 4 (avanzado).

Relativamente ao nível de desempenho dos alunos portugueses e da UE, nos dois ciclos de avaliação PIRLS (2011 e 2016), a Tabela 5 mostra as percentagens acumuladas referentes a cada nível em Portugal e na UE. Verifica-se que, em Portugal, apenas 9% e 7% conseguiram um nível avançado em 2011 e 2016, respetivamente, ou seja, que compreenderam os textos propostos, implicando inferências para explicar a ligação entre as intenções, as ações, os acontecimentos, exprimir uma preferência e justificá-la. Constata-se também que a esmagadora maioria dos alunos em Portugal (98% em 2011; 97% em 2016) foi capaz de localizar e restituir um detalhe ou uma informação explicitamente mencionada e diretamente acessível no texto; isto é, somente 2% e 3%, respetivamente, não o conseguiram. De notar que as percentagens não estão muito afastadas das verificadas para a UE.

Tabela 5*Nível de desempenho dos alunos no PIRLS, em %, 2011 e 2016*

PIRLS 2016		Avançado (625)	Elevado (550)	Intermediário (475)	Baixo (400)
Portugal	2016	7%	38%	79%	97%
	2011	9%	47%	84%	98%
União Europeia	2016	12%	48%	82%	96%

Fonte: IEA.

De 2011 para 2016, os alunos (rapazes e raparigas) do 4º ano pioraram o seu desempenho. Dos 4558 alunos e 218 escolas, Portugal desceu 13 pontos (de 541 para 528 pontos), em relação às provas realizadas em 2011 (cf. Tabela 6), não se tendo verificado diferenças significativas entre os participantes de cada sexo. Relativamente à literacia *online* (ePIRLS), Portugal ficou seis pontos abaixo dos resultados obtidos no PIRLS: os alunos portugueses obtiveram 522 pontos, ocupando a 12ª posição.

Tabela 6*Resultados do PIRLS em Portugal, 2011 e 2016*

PIRLS – Resultados	2011	2016
Portugal	541	528
Raparigas - Pontuação Média	548	529
Rapazes - Pontuação Média	534	527
Raparigas (% de Alunos)	49	49
Rapazes (% de Alunos)	51	51

Fontes: IAVE (2017); Mullis et al. (2012).

- Um estudo efetuado por Lafontaine (2008) permitiu apontar algumas características pedagógicas que distinguem os países com pontuações mais baixas:
- A tendência para se ensinar em vez de se promover a capacidade de compreensão;
- O ensino das estratégias de compreensão está longe da prática do quotidiano das aulas;
- O pouco tempo consagrado ao ensino formal da leitura no 4º ano de escolaridade;
- Os livros para crianças não são considerados como um material didático essencial;
- A irregularidade da leitura de livros mais extensos;
- A predominância das avaliações tradicionais do tipo questionário de escolha múltipla e questões abertas que exigem respostas escritas curtas.

Sobre a realidade portuguesa, e de uma forma mais abrangente, Benavente (2016) destaca vários constrangimentos a nível escolar: aulas expositivas, número elevado de alunos por classe, necessidades dissociadas da realidade das crianças e dos jovens, concorrência económica entre escolas, dispensa de milhares de professores e de funcionários, reformulação

curricular e programática, programas extensos e desadequados, obstáculos à integração de crianças e jovens com necessidades educativas específicas, menorização de outras disciplinas (desporto, educação artística, educação cívica, educação ambiental). As políticas e as práticas são irregulares e mudam consoante os governos, os atores educativos e as escolas.

3. METODOLOGIA

Neste artigo, usamos dois tipos de análise metodológica. Nos três primeiros pontos da secção 3. Resultados, analisamos e comparamos a conceção de literacia em cada um dos estudos envolvidos, os respetivos conteúdos dos protocolos e tipo de tratamento de dados e os resultados globais obtidos pelos alunos portugueses nos três estudos e a sua evolução. Usamos uma abordagem qualitativa de carácter interpretativo, seguindo as perspetivas de IAVE (2019), Mullis et al. (2009) e Rémond (2006), entre outros.

Nos dois pontos finais da secção 3. Resultados, fazemos uma análise quantitativa dos dados numéricos dos estudos em causa, como fizeram vários autores, como, por exemplo, Marôco (2018) e Mullis et al. (2012) e os autores da maioria dos capítulos de *Monitoring Student Achievement in the 21st Century. European Policy Perspectives and Assessment Strategies*, de Harju-Luukkainen et al. (2020). Apresentamos uma análise comparativa correlacional entre os resultados médios (RM) de cada um dos três domínios do PISA-2015, dos dois domínios do TIMSS-2015, dos dois domínios do TIMSS *Advanced*-2015, do único domínio do PIRLS-2016 e do ePIRLS-2016, por regiões da NUTS III, e, ainda, uma análise comparativa correlacional entre estes RM e os valores médios do PIB/habitante dessas regiões em 2015. São, pois, estas as 10 variáveis que comparámos correlacionalmente. Dentro dos estudos, dado que o PISA abrange três domínios, o TIMSS e o TIMSS *Advanced* dois cada um, e o PIRLS e o ePIRLS um cada, e que determinámos a correlação entre estas nove variáveis, duas a duas, calculámos trinta e seis correlações. Entre os RM de cada domínio de cada estudo (nove) e os valores médios da variável correspondente ao PIB/habitante, determinámos as nove (9×1) correlações possíveis. O número de valores possíveis de cada das 10 variáveis são sempre 25, correspondentes às 25 regiões da NUTS III. No tratamento estatístico destas variáveis e respetivas comparações para determinação das correlações e do valor de foi usado como suporte o *software* IBM SPSS® *Statistics*.

Embora já se tenha acesso aos RM do PISA de 2018 por NUTS III (Lourenço et al., 2019), para reduzir os efeitos de fatores externos aos resultados, reduzimos essa análise a um período temporal curto, o mais recente possível, mas que abrangesse todos os estudos. Esta opção levou-nos a escolher os anos de 2015, abrangendo o PISA e o TIMSS, e de 2016, que inclui o PIRLS e o ePIRLS. Assim, a análise comparativa que apresentamos compreende os últimos resultados conhecidos do TIMSS, do PIRLS e do ePIRLS e o penúltimo do PISA; no entanto, quando se justifica, os resultados do PISA-2018 são também referidos nessa análise. Em relação ao PIB/habitante, para que haja coerência nas comparações com os RM é o ano de 2015 que é considerado, pois é o que corresponde ao maior número de RM comparados.

Nessa análise correlacional, assumimos que os valores apresentados nos relatórios nacionais sobre o PISA de 2015, de Marôco (2016a), sobre o TIMSS e o TIMSS *Advanced* de 2015, também de Marôco (2016b, 2016c), e sobre o PIRLS e o ePIRLS de 2016, do IAVE (2017), são fidedignos e que as amostras referidas e usadas nos estudos citados são representativas de cada região da NUTS III em cada um deles.

Dado que os RM, por regiões da NUTS III, de quaisquer dois domínios de qualquer dos estudos, e o PIB/habitante são métricos e referentes às mesmas regiões, esses resultados são emparelhados, tendo sentido calcular se há correlação linear entre eles, usando um teste paramétrico. Para tal, usou-se o coeficiente de correlação linear de Pearson, r , que é o usado pelo SPSS em todas as situações com condições semelhantes às nossas.

Para avaliarmos se as correlações lineares calculadas tinham ou não significado estatístico relativamente às respetivas populações, usamos o teste de hipóteses, $H_0: \rho = 0$; $H_1: \rho \neq 0$, sobre a variável habitualmente usada em correlações lineares, $t_c = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$, que segue a distribuição *-Student* com $n-2$ graus de liberdade, sendo n = número de regiões da NUTS III (25) e r = coeficiente de correlação de Pearson entre as amostras em causa. Note-se que é a expressão que o SPSS usa com as variáveis com características semelhantes às nossas, e que está associada ao valor de prova, p . Para distinguirmos se a correlação calculada tem ou não significado estatístico para as respetivas populações, definimos o habitual nível de significância em estudos educativos: $\alpha = 5\%$. Assim, se o valor de prova, p , for inferior a 0,05 há correlação estatisticamente significativa, positiva ou negativa, entre as populações correspondentes aos dados emparelhados dos resultados dos dois domínios dos estudos em causa, e se p for superior ou igual a 0,05 não há correlação estatisticamente significativa, positiva ou negativa, entre as populações correspondentes aos dados emparelhados dos resultados comparados.

4. RESULTADOS

4.1. AS DEFINIÇÕES DE LITERACIA SEGUNDO O PISA, O TIMSS E O PIRLS

Dado que é a literacia, nas três áreas referidas, que está em foco no PISA, no TIMSS e no PIRLS, importa esclarecer o que se entende por este conceito nestes três questionários e quais as suas diferenças e semelhanças.

Em relação ao PISA, no relatório sobre Portugal no PISA-2018, Lourenço et al. (2019) fazem um resumo da respetiva posição. Referem que a *literacia científica* é encarada como “a capacidade de um indivíduo se envolver em questões relacionadas com as ciências e de compreender as ideias científicas como um cidadão reflexivo sendo capaz de explicar fenómenos cientificamente, avaliar e conceber investigações científicas, interpretar dados e evidências cientificamente” (Lourenço et al., 2019, p. 32). Em relação à *literacia de leitura*, dizem que é a “capacidade de um indivíduo compreender, utilizar, avaliar, refletir e de se envolver na leitura de textos escritos, com a finalidade de atingir os seus objetivos, de desenvolver os seus conhecimentos e o seu potencial e de participar na sociedade” (Lourenço et al., 2019, p. 20).

E, no que diz respeito à *literacia matemática*, consideram que é “a capacidade de um indivíduo formular, aplicar e interpretar a matemática em contextos diversos. Inclui raciocinar matematicamente e utilizar conceitos, processos, factos e ferramentas da matemática para descrever, explicar e prever fenómenos” (Lourenço et al., 2019, p. 39).

A IEA, responsável pelo TIMSS e pelo PIRLS, não aborda diretamente o conceito de literacia e localiza-a nas *performances* académicas. No caso do TIMSS, a literacia matemática e científica é encarada como a capacidade do indivíduo para usar e aplicar o conhecimento em cada uma destas áreas, mas, como já se referiu, essa capacidade está centrada nas competências curriculares e nas competências cognitivas *conhecer, aplicar e raciocinar* (IAVE, 2019). No entanto, não encontramos qualquer definição de literacia nos diferentes relatórios do TIMSS nem qualquer preocupação em fazê-lo. Por exemplo, no Relatório de 1995, no volume referente a Conceção e Desenvolvimento, a primeira referência a literacia é feita na página 11, dizendo somente que o “TIMSS planeia relatar medidas de literacia em matemática e em ciências” (Martin & Kelly, 1997, p. 11); nas restantes vezes que a palavra “literacia” aparece é para dizer se se trata de matemática ou ciências ou para a relacionar com a população em causa. A avaliação da literacia de leitura feita pelo PIRLS assenta numa noção abrangente do que é saber ler, uma noção que inclui a capacidade de refletir sobre o que se lê e de fazer uso disso para alcançar objetivos individuais e sociais. Segundo Campbell et al. (2000), a IEA define literacia de leitura como “a aptidão para compreender e utilizar as formas de linguagem escrita, que a sociedade tem necessidade e que são necessárias ao indivíduo” (p. 3). Ou seja, para o PIRLS, o aluno lê para aprender, mas, tal como no PISA, deve também participar na vida coletiva, chamada “comunidade de leitores” (Mullis et al., 2009, p. 11; Rémond, 2006, p. 72).

Ora, em termos gerais, a OCDE (1999) assume, para os estudos do PISA, que literacia é a capacidade de “compreender, de utilizar e de analisar os textos escritos, mas também de refletir sobre o seu propósito, por forma a realizar os seus objetivos, de desenvolver os seus conhecimentos e o seu potencial e assumir um papel ativo na sociedade” (p. 20). Portanto, apesar de a OCDE e a IEA apontarem para focos diferentes nas suas avaliações, a posição da OCDE sobre o que é literacia não difere muito da IEA. De facto, também para Rémond (2006), estas duas definições, elaboradas num intervalo de um ano, apontam numa perspetiva semelhante, pois ambas refletem “uma conceção interativa da leitura, onde o leitor apela aos seus conhecimentos e às suas estratégias para construir a significação do texto e utilizar a escrita” (p. 72).

4.2. O CONTEÚDO DOS PROTOCOLOS E DO TRATAMENTO DE DADOS NO PISA, NO TIMSS E NO PIRLS

Uma complexa metodologia condiciona a construção e a organização do material, elaborado em inglês, para formar uma “versão fonte”, servindo de base comum às traduções. Vários itens foram criados (141 para o PISA-2000), mas os alunos apenas veem uma parte do material produzido, sob o princípio de “cadernos rotativos”. Os protocolos do PIRLS recaem sobre textos longos e integrais, de onde são elaboradas várias questões. São constituídas diversas versões do teste. No PIRLS-2016, por exemplo, foram utilizados 16 cadernos

de teste, com diferentes combinações de itens que avaliam diferentemente os processos e as finalidades de leitura. Utilizando uma *multiple matrix sampling test design* (distribuição dos itens pelos cadernos de teste), assegura-se a “estimação da literacia de leitura para grupos grandes” (IAVE, 2017, p. 5). Devido às limitações da amostragem, os estudos como o PIRLS recorrem à estimação de valores plausíveis para o desempenho dos alunos. Considera-se a aptidão dos alunos em função das respostas aos itens do teste. O PISA, como outras avaliações da OCDE, incide sobre um grande número de exercícios, na medida em que procuram cruzar vários fatores para obter uma declinação dos itens, correspondendo às características definidas (OCDE, 1999, 2003). No caso do TIMSS, e consoante o ano em causa (4.º, 8.º ou 12.º), os domínios específicos avaliados em matemática e em ciências são: números, formas geométricas, medidas, álgebra, probabilidades, ciências da vida, ciências físicas e da terra, biologia, química, física, entre outros. Para além dos conteúdos, o quadro de referência define os processos cognitivos na resolução dos itens (aplicar, conhecer e raciocinar).

Os relatórios internacionais publicados pela OCDE e pela IEA colocam a ênfase na comparação de resultados entre os países e são representados como se traduzissem uma escala de competências que induz um *ranking*.

A Tabela 7 resume alguns aspetos dos questionários internacionais e a participação portuguesa.

Tabela 7*Questionários internacionais e a participação de Portugal*

	PISA	PIRLS	TIMSS	TIMSS Advanced
Pilotagem do Inquérito	OCDE <i>Organismo Intergovernamental</i>		IEA <i>Cooperativa Internacional de Instituições de Investigação</i>	
Data de Criação	2000	2001	1995	1995
Periodicidade	Todos os 3 anos	Todos os 5 anos	Todos os 4 anos	Irregular
Número de Países/Regiões*	72 (34 da OCDE)	50 (26 da OCDE)	48 (26 da OCDE)	9 (7 da OCDE)
Público-Alvo	Alunos de 15 anos de idade que frequentam, pelo menos, o 7.º ano de escolaridade	4.º ano de escolaridade	4.º, 8.º anos de escolaridade	12.º ano de escolaridade
Domínios Avaliados	Literacia de Leitura Literacia Matemática Literacia Científica	Literacia de Leitura	Matemática e Ciências	Matemática e Física
Contexto dos Itens	Competências úteis ao futuro cidadão	Competências de leitura	Saberes disciplinares	
Exercícios Libertados	Poucos	Poucos	Muitos	
Participação de Portugal	2000, 2003, 2006, 2009, 2012, 2015, 2018	2011, 2016**	1995, 2011***, 2015***, 2019	2015

* A participação dos países é variável durante as várias edições.

** Em 2016, realizou-se também o PIRLS *Online* (ePIRLS).

*** Em 2011 e 2015, Portugal só participou com alunos do 4.º ano de escolaridade.

4.3. ALGUMAS COMPARAÇÕES DE RESULTADOS ENTRE O PISA, O TIMSS E O PIRLS

Desde o primeiro ciclo do PISA, em 2000, Portugal tem registado uma melhoria significativa dos resultados obtidos nos três domínios analisados. Em 2018, os alunos obtiveram 492 pontos nas três literacias (leitura, matemática e científica), e, segundo Lourenço et al. (2019), as regiões de Coimbra, Leiria, Área Metropolitana do Porto e Médio Tejo são as que obtêm melhores resultados nos três domínios, com a região do Ave a sobressair também na leitura e na matemática. As regiões com resultados menos conseguidos nos três domínios foram a Região Autónoma dos Açores, o Baixo Alentejo e o Alto Alentejo. Estes resultados representam uma manutenção dos valores de 2015 ou uma descida não significativa, pois em 2015 os alunos obtiveram 498 pontos em leitura, 492 pontos em matemática e 501 pontos em ciências. As unidades territoriais com resultados mais conseguidos e menos conseguidos em 2018 também são diferentes das correspondentes em 2015.

No relatório de Marôco (2016a), referente ao PISA-2015, é mencionado que são as unidades territoriais do litoral, em particular o Alentejo Litoral, que apresentam os melhores resultados médios dos alunos. Em contrapartida, são as unidades do interior norte (Alto Tâmega, Terras de Trás-os-Montes e Tâmega e Sousa) e do Baixo Alentejo onde se verifica que o desempenho dos alunos em leitura foi mais débil.

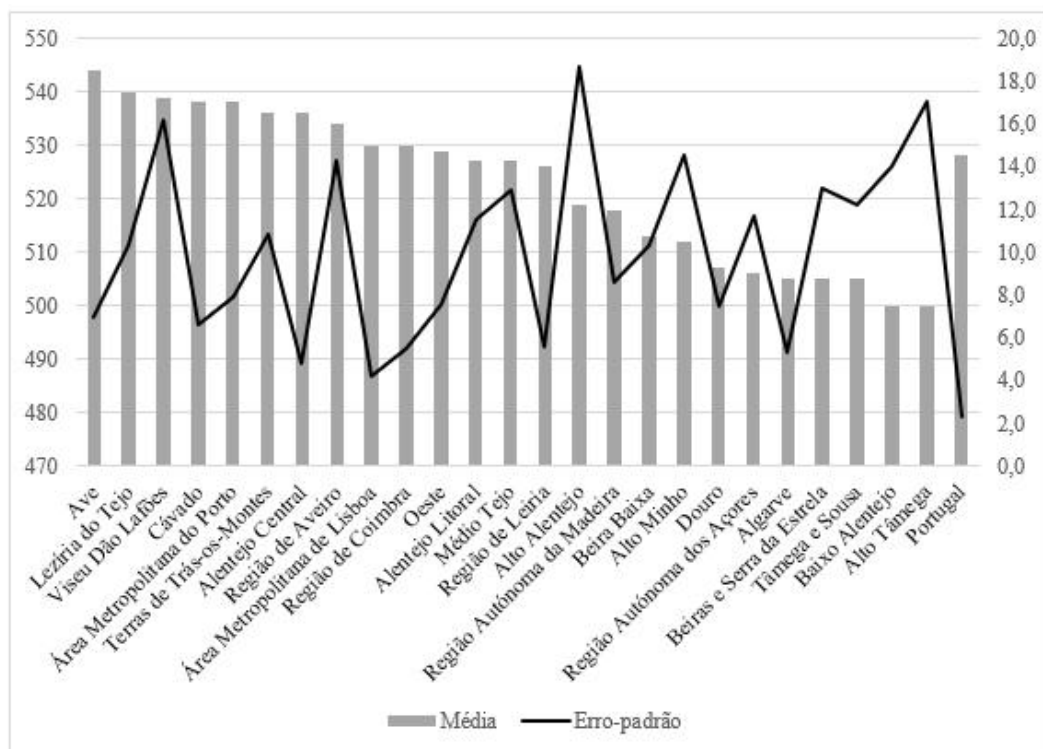
A família, nas suas diferentes dimensões (económica, cultural e social), afeta o desempenho escolar dos seus filhos (Ferreira, Flores, & Casas-Novas, 2017). A OCDE criou um indicador, denominado “Estatuto Socioeconómico e Cultural” (ESCS), que agrega várias características do agregado familiar (nível de escolaridade, estatuto profissional, bens materiais, indicadores culturais). No estudo de Ferreira et al. (2017) refere-se que “desde 2003, Portugal tem vindo a melhorar a condição socioeconómica dos seus alunos” (p. 25). No entanto, ela é ainda relativamente baixa quando comparada com a de outros países (Espanha e Polónia, por exemplo). O desemprego ou subemprego a que as famílias foram expostas a partir de 2008, devido à crise económica, poderá ter contribuído para os efeitos negativos na *performance* dos alunos. Os autores referem que “Portugal tem feito uma trajetória decrescente, sendo que a correlação entre ESCS e scores PISA tem vindo a reduzir-se ao longo do tempo, estando em 0,4 em 2015” (Ferreira et al., 2017, p. 46), isto é, em Portugal, a influência das condições socioeconómicas das famílias nos resultados do PISA tem vindo a reduzir-se. Numa análise aos dados do PISA-2012, Ferrão (2015) afirma que “os alunos mais desfavorecidos socioeconomicamente [apresentam] um imenso cenário de desigualdade de oportunidades, com início na escola, a partir dos 7 anos de idade” (p. 113).

Numa análise dos resultados nacionais do TIMSS-2015 (4^o ano) por regiões da NUTS III, tomando como referência o estudo de Marôco (2016b), é possível verificar que quase um quarto dos alunos da região de Leiria alcançaram resultados de nível Avançado (625 pontos ou mais na escala de matemática). Em contraste, foi a região do Tâmega e Sousa, tal como aconteceu no PISA-2015, que obteve os dados mais fracos a nível nacional (500 pontos em matemática; 477 pontos em ciências). De uma forma global, as regiões do litoral centro e norte apresentam, no âmbito das duas literacias, resultados muito positivos.

No que respeita ao PIRLS, em 2016, Portugal, com 528 pontos, encontra-se no conjunto de participantes que registaram uma pontuação média significativamente superior ao ponto central da escala (500 pontos). No entanto, face a 2011, esta pontuação representa uma descida significativa de 13 pontos (541, em 2011, desce para 528, em 2016). Na escala ordenada dos resultados, entre 2011 e 2016, Portugal passou do 19º para o 30º lugar. A Figura 2 apresenta a distribuição dos resultados nacionais por regiões da NUTS III. Foi a região do Ave que obteve a melhor classificação no PIRLS-2016, com 544 pontos. É uma diferença de 16 pontos relativamente à média nacional (528). As regiões que registaram os desempenhos médios mais baixos foram: Alto Tâmega, Baixo Alentejo, Tâmega e Sousa, Beiras e Serra da Estrela, Algarve, Região Autónoma dos Açores e Douro. No entanto, todas elas registaram valores acima de ou iguais a 500.

Figura 2

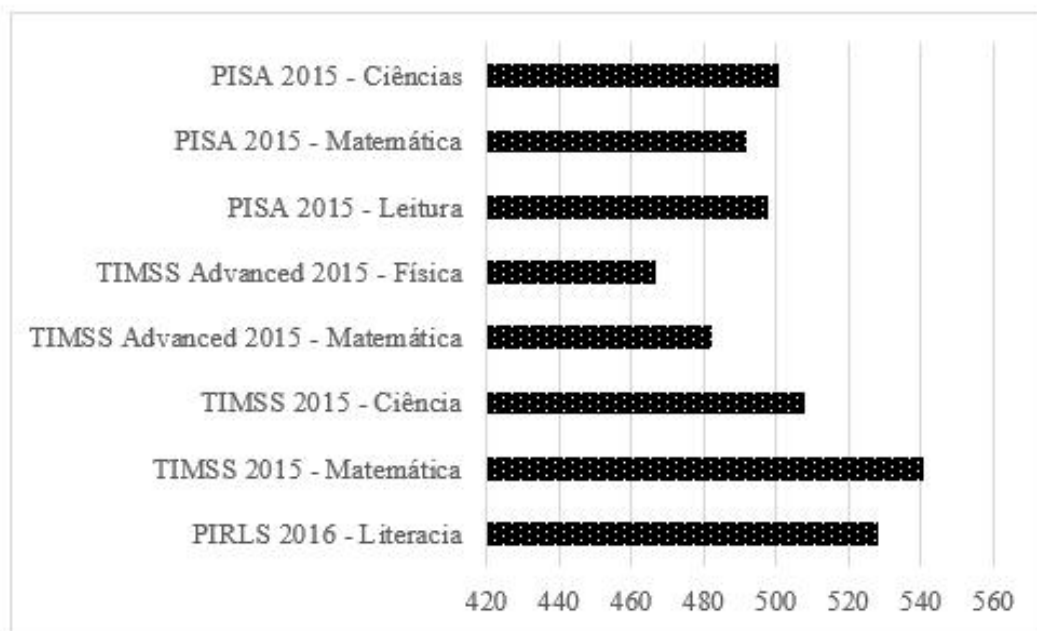
Resultados do PIRLS (média e erro-padrão da média), por regiões da NUTS III, 2016



Fonte: IAVE (2017).

Comparando os resultados de todos os estudos referidos, verificamos que Portugal apresenta resultados mais fracos em física e matemática no TIMSS *Advanced*-2015 (cf. Figura 3) do que nos restantes estudos.

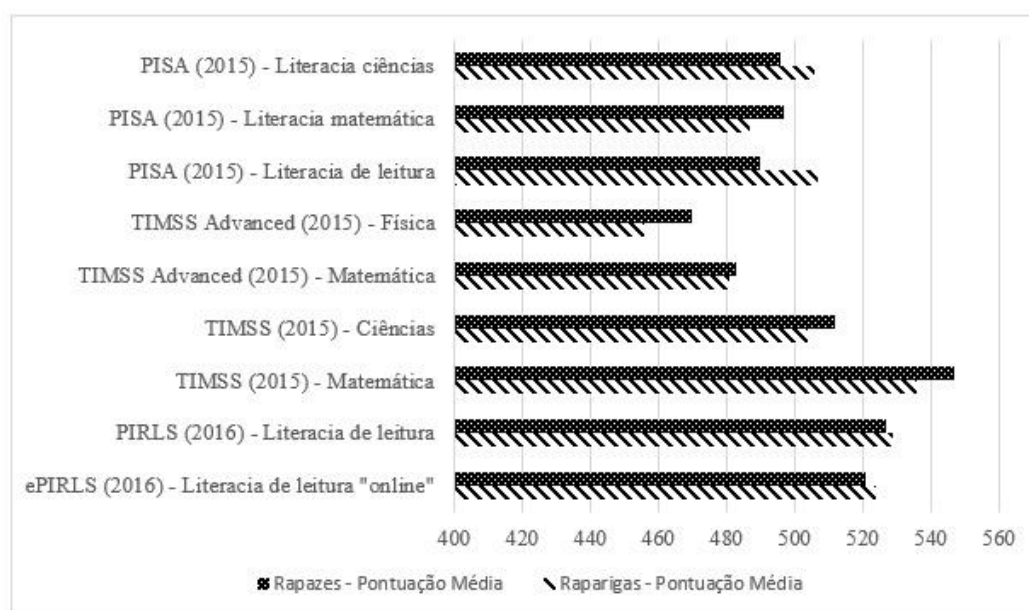
Figura 3
 Médias dos três questionários e domínios de avaliação, 2015-2016



Fontes: IAVE (2017); Marôco (2016b, 2016c).

Numa análise por género, também relativamente a todos os estudos referidos (cf. Figura 4), verifica-se que, em literacia matemática, de ciências e de física, os rapazes têm médias mais altas do que as raparigas e que estas têm médias mais elevadas na literacia de leitura.

Figura 4
 Médias dos três questionários, por domínios e por género, 2015-2016



Fontes: IAVE (2017); Marôco (2016b, 2016c).

Numa análise sobre as medidas e os impactos de política educativa no respetivo desempenho, em 2015, dos alunos do 4.º ano na área da matemática (referente ao TIMSS) e dos alunos de 15 anos no que se refere à literacia em ciências, Saraiva (2017) sublinha que

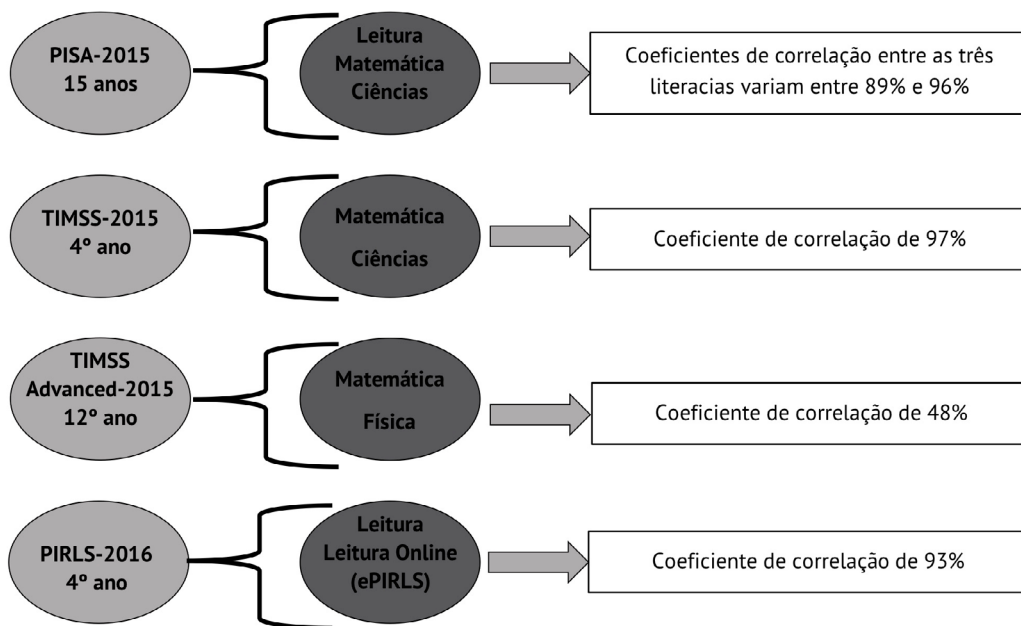
Ao nível das ciências, e ao contrário do que sucedeu em matemática, os relatórios nacionais do TIMSS (4.º ano de escolaridade) e do PISA (2015) revelam resultados discrepantes. Se no PISA a média dos alunos portugueses foi significativamente superior à obtida em anos anteriores e à média da OCDE, o mesmo não sucedeu no TIMSS ... Mas esta discrepância ficou quase “esquecida” no debate público, talvez porque as ciências permanecem afastadas dos “palcos” dos *rankings* nacionais, onde se mantém o protagonismo da matemática e do português. (p. 5)

4.4. ANÁLISE CORRELACIONAL ENTRE OS RESULTADOS DO PISA, DO TIMSS E DO PIRLS, POR REGIÕES DA NUTS III

Com o intuito de se saber se haveria alguma relação entre os resultados dos estudos que temos vindo a mencionar, comparámos os RM, por regiões da NUTS III, dos seguintes estudos do PISA, TIMSS e PIRLS: PISA-2015 (15 anos – leitura, matemática e ciências); TIMSS-2015 (4.º ano – matemática e ciências); TIMSS *Advanced*-2015 (12.º ano – matemática e física); PIRLS-2016 e ePIRLS-2016 (ambos no 4.º ano – leitura).

Fizemos uma análise correlacional entre esses RM, por regiões da NUTS III, calculando o coeficiente de correlação linear de Pearson entre os dados emparelhados de cada par desses estudos. Como há nove estudos/domínios, calculámos as 36 correlações possíveis entre dois desses estudos. Verificámos que havia seis correlações estatisticamente significativas, todas positivas (ver Figura 5): a) no PISA-2015, entre a leitura, a matemática e as ciências, com valores entre 0,89 e 0,96 (em qualquer dos casos verifica-se um $p < 0,001$); b) no TIMSS-2015, entre a matemática e as ciências, com 0,97 ($p < 0,001$); c) no TIMSS *Advanced*-2015, entre a matemática e a física, com 0,48 ($p = 0,016$); entre o PIRLS e o ePIRLS, referente à leitura, com 0,93 ($p < 0,001$).

Figura 5
Correlações significativas positivas dentro de cada estudo



Dessa análise, duas evidências se destacam:

- Entre os RM de dois domínios do mesmo estudo, há sempre uma correlação positiva, estatisticamente significativa, quaisquer que sejam os domínios em causa;
- Entre os RM de dois questionários referentes ao mesmo domínio ou à mesma faixa etária, mas de estudos diferentes, não há qualquer correlação estatisticamente significativa (ver Tabela 8).

Tabela 8
Correlações não significativas entre RM de estudos diferentes

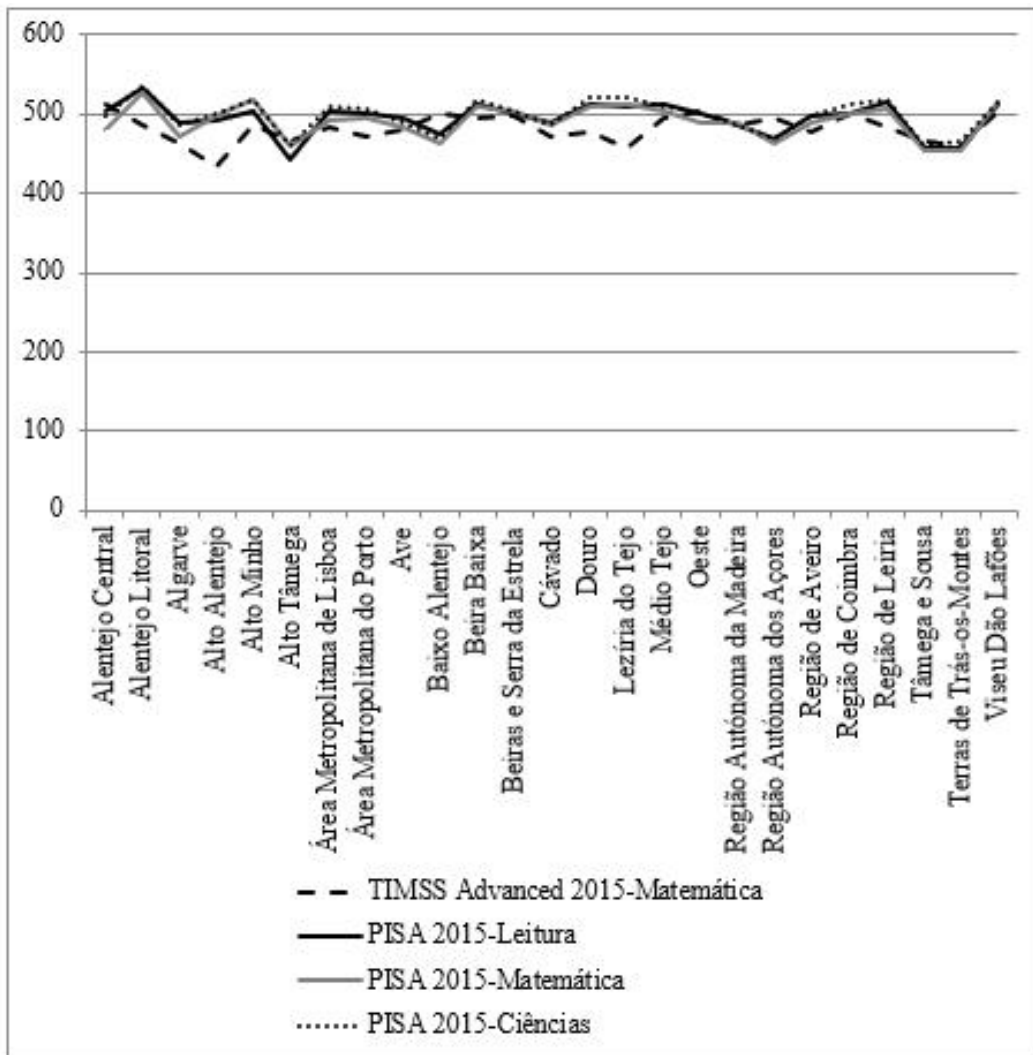
LITERACIAS	CORRELAÇÕES	PISA-2015	TIMSS-2015	TIMSS Advanced-2015	PIRLS-2016	ePIRLS-2016
MATEMÁTICA	PISA-2015	---	31%	15%		
	TIMSS-2015	31%	---	6%		
	TIMS Advanced-2015	15%	6%	---		
CIÊNCIAS	TIMSS-2015	30%				
LEITURA	PISA-2015				40%	30%
	PIRLS-2016	40%				
	ePIRLS-2016	30%				

Verificar-se uma correlação muito forte entre os RM do PIRLS e do ePIRLS não é de estranhar, pois correspondem ao mesmo ano e à mesma área, mas o mesmo já não se pode dizer em relação a correlações muito fortes entre os RM de estudos referentes a domínios diferentes, como, por exemplo, dentro do PISA. A Figura 6 mostra os RM do PISA-2015, de leitura, matemática e

ciências, e do TIMSS *Advanced*-2015, de matemática, por regiões da NUTS III. Nela podemos verificar que os RM dos três domínios do PISA-2015 seguem linhas muito próximas, o que confirma a correlação quase perfeita que existe entre eles, e o desfasamento da linha do TIMSS *Advanced*-2105, da matemática, em relação às outras, o que mostra a não existência de correlação significativa entre qualquer dos RM do PISA-2015 e os do TIMSS *Advanced*-2015 referido. De facto, o coeficiente de correlação entre os RM do TIMSS *Advanced*-2015, da matemática, e do PISA-2015, da matemática, é de 0,15 ($p=0,42$).

Figura 6

Resultados do PISA-2015, de leitura, matemática e ciências, e do TIMSS Advanced, de matemática, por regiões da NUTS III



A evidência a) – entre os RM de dois domínios do mesmo estudo, há sempre uma correlação positiva, estatisticamente significativa, quaisquer que sejam os domínios em causa –, isto é, o facto de as correlações entre os RM de dois domínios do mesmo estudo serem sempre positivas e estatisticamente significativas (em quatro das cinco correlações possíveis nesta situação o coeficiente de correlação é muito alto: $\geq 0,89$, o que corresponde a $p < 0,001$) diz-nos que, dentro de cada estudo, as regiões da NUTS III com bons RM num domínio também os têm no outro ou nos outros domínios do mesmo estudo.

Isto aponta para que, dentro de cada estudo, o resultado obtido por um aluno aleatoriamente escolhido dentro dos que preencheram um dos questionários em causa dependa muito menos do domínio estudado do que da região da NUTS III a que o aluno pertence. Assumindo que as amostras de cada estudo são representativas de cada região e de Portugal no seu todo e que os questionários são válidos e fidedignos, isto indicia que obter bons resultados em qualquer dos questionários, de qualquer domínio, é muito mais dependente da região a que pertence o aluno do que do domínio em causa.

Para aprofundar a evidência b) – entre os RM de dois questionários referentes ao mesmo domínio ou à mesma faixa etária, mas de estudos diferentes, não há qualquer correlação estatisticamente significativa –, especifiquemos os resultados obtidos. No que diz respeito à matemática, entre o PISA-2015 e o TIMSS-2015, entre este e o TIMSS Advanced-2015 e entre o último e o primeiro, os coeficientes de correlação obtidos foram, respetivamente, de 0,31 ($p=0,14$), de 0,06 ($p=0,78$) e de 0,15 ($p=0,48$). No que respeita ao domínio ciências, entre o TIMSS-2015 e o PISA-2015, obteve-se um coeficiente de correlação de 0,29 ($p=0,15$) e, em relação ao domínio leitura, o coeficiente de correlação entre o PIRLS-2016 e o PISA-2015 foi de 0,40⁵ e entre o ePIRLS-2016 e o PISA-2015 foi de 0,30 ($p=0,15$). Todos estes valores, como dissemos, não têm significado estatístico. A não existência de correlação estatisticamente significativa entre os RM de estudos diferentes, mesmo referentes ao mesmo domínio, aponta para que as diferenças de objeto de avaliação em cada questionário (competências para a vida no PISA e competências académicas no TIMSS e no PIRLS) possam ser fatores relevantes para os RM obtidos em cada questionário. O facto de também não haver correlação entre os RM dos diferentes questionários da IEA no que respeita ao mesmo domínio, isto é, entre os RM do PIRLS-2016 e do ePIRLS-2016 (4º ano) em comparação com os RM do PISA-2015 (15 anos), em relação à leitura, e entre os RM do TIMSS-2015 e do TIMSS Advanced-2015, em relação à matemática, poderá ser justificado pela diferença etária e de nível de escolaridade dos alunos em causa, já que dificilmente se poderá dizer que os objetos de avaliação (currículos, recursos e ambientes familiares e escolares e sucesso dos alunos na escola) no 4º ano e no 12º ano de escolaridade sejam semelhantes. Ou seja, o fator objeto de avaliação em cada estudo parece ser decisivo para que não haja correlação significativa entre os RM de um domínio, mas de estudos diferentes. Porém, a não existência de correlação linear entre os RM de questionários aplicados a alunos do mesmo ano de escolaridade ou nível etário, como é o caso do TIMSS, do PIRLS e do ePIRLS (todos aplicados a alunos do 4º ano e em anos consecutivos, portanto da mesma geração), já não aponta para a mesma justificação. De facto, as quatro correlações calculáveis entre os RM de cada um dos dois domínios do TIMSS e os RM do PIRLS e do ePIRLS estão compreendidas entre 0,25 ($p=0,26$) e 0,34 ($p=0,09$) e são, como dissemos, todas estatisticamente não significativas. Sendo estes estudos da responsabilidade da mesma organização, a IEA, portanto, com os mesmos princípios e objetivos genéricos, talvez só possa ser justificada pela diferença entre domínios (leitura no PIRLS e ePIRLS e matemática e ciências no TIMSS).

5. Este valor corresponde ao limite de uma correlação não significativa ($p=0,05$).

4.5. RELAÇÃO ENTRE OS RESULTADOS DOS TRÊS ESTUDOS COM O PIB POR REGIÕES DA NUTS III

Investigámos também se haveria alguma relação entre os RM de cada um destes estudos e o PIB/habitante de 2015 (dados obtidos na PORDATA – Base de Dados Portugal Contemporâneo)⁶, por regiões da NUTS III, dado que a maioria dos resultados estudados são deste ano.

Relativamente às correlações lineares entre o PIB/habitante e os RM em cada domínio de cada questionário, constatámos que, das nove correlações possíveis, a única estatisticamente significativa (0,47, o que corresponde a $p=0,017$) verificou-se entre os valores do PIB/habitante e os RM do PISA-2015 de leitura, o que parece apontar para que as diferenças regionais a nível de PIB tenham pouca influência nos resultados destes questionários, conforme Marôco (2018) refere a propósito do PIRLS-2016. Contudo, se olharmos para os valores extremos do PIB/habitante, verificamos que há algumas coincidências entre estes e os RM de cada estudo. A região que apresenta o valor máximo de PIB/habitante – Alentejo Litoral – é também a região que, no PISA-2015, obteve melhores RM em cada um dos três domínios. Em qualquer dos outros questionários isso não sucede. Se alargarmos a análise à segunda região (Área Metropolitana de Lisboa) com maior PIB/habitante, verificamos que não é a região com melhores RM em qualquer dos questionários TIMSS-2015, TIMSS *Advanced*-2015 e PIRLS-2016. Quanto à região com menor PIB/habitante – Tâmega e Sousa –, verificamos que é a região com RM mais baixo no PISA-2015, de matemática e de ciências, e no TIMSS-2015, de matemática e ciências. A segunda região com menor PIB/habitante (Alto Tâmega) é a região com RM mais baixo no PISA-2015, de leitura, no TIMSS *Advanced*-2015, de física, e no PIRLS-2016, de leitura. Somente o RM mais baixo no ePIRLS, de leitura, e no TIMSS *Advanced*-2015, de matemática, não é obtido por uma das duas regiões com menor PIB/habitante da NUTS III.

Sintetizando, a região com PIB/habitante mais alto é a região com os RM mais altos em três dos nove estudos/domínios e as duas regiões com PIB/habitante mais baixo são as regiões com RM mais baixos em sete dos nove estudos/domínios. Ou seja, embora a maioria⁷ das correlações, por regiões, entre o PIB/habitante e os RM de cada domínio dos estudos citados seja não significativa, o facto de a maioria dos resultados mais altos e mais baixos em cada domínio de cada estudo serem obtidos nas regiões, respetivamente, com PIB/habitante mais alto e com PIB/habitante mais baixo parece mostrar que há alguma relação entre o PIB/habitante das regiões e os RM dos diferentes estudos/domínios e que essa relação é essencialmente nas regiões com valores extremos de PIB/habitante. Se atendermos ao facto de, nos três domínios do PISA, os RM mais elevados serem na região de PIB/habitante mais alto e isso não se verificar em qualquer dos estudos TIMSS e PIRLS e ao facto de o PISA avaliar competências para a vida e os outros estudos competências académicas, podemos pôr a hipótese de isso apontar para que os filhos das famílias com mais rendimentos estejam mais preparados para a vida, embora não sejam necessariamente aqueles que têm mais sucesso escolar.

6. Em <https://www.pordata.pt/>

7. Note-se que, se não considerarmos a região do Alto Alentejo (a que tem PIB/habitante mais alto), a correlação entre o PIB/habitante e os RM do PISA-2015, de leitura, já não seria estatisticamente significativa, acompanhando, assim, o que sucede com os RM de todos os restantes estudos.

5. CONCLUSÕES

A análise dos resultados globais referentes a Portugal desde 2000 evidencia que estes têm vindo a melhorar no PISA, nos três domínios (leitura, matemática e ciências); que, no TIMSS, progrediram em matemática, mas que, em ciências, as médias oscilaram; e que no PIRLS, em leitura, pioraram. No entanto, os resultados obtidos pelos alunos não são iguais em todo o país, havendo diferenças entre as várias regiões da NUTS III. Na análise correlacional, por regiões da NUTS III, entre os RM em cada domínio de cada estudo, constatámos que:

- Entre os RM de dois domínios do mesmo estudo, há sempre uma correlação positiva estatisticamente significativa, quaisquer que sejam os domínios em causa;
- Entre os RM de dois questionários referentes ao mesmo domínio ou à mesma faixa etária, mas de estudos diferentes, não há qualquer correlação estatisticamente significativa.

Estes resultados fazem destacar dois fatores com relevâncias de sentidos opostos nos resultados das regiões da NUTS III:

- 1) O objeto de avaliação (competências académicas nos estudos da IEA – PIRLS, ePIRLS, TIMSS e TIMSS *Advanced*, e competências para a vida nos estudos da OCDE – PISA) parece ter uma importância relevante nos resultados dos questionários;
- 2) O domínio (leitura, matemática, ciências e física) que se avalia parece não ser determinante nos resultados de cada estudo.

A irrelevância do domínio diz-nos que as regiões têm bons ou maus resultados independentemente do domínio em causa, e a relevância do objeto de avaliação indica-nos que a competência na escola poderá estar pouco relacionada com as competências para a vida.

Olhando para a relação entre as questões socioeconómicas, se Portugal tem vindo a melhorar a condição socioeconómica dos seus alunos, ela é ainda baixa quando comparada com a de outros países europeus e repercute-se nos resultados escolares dos alunos. A escola reproduz as desigualdades sociais, como concluíram, entre outros, Bourdieu e Passeron (1970), e, como sublinha Ferrão (2015), a desigualdade dos alunos mais desfavorecidos socioeconomicamente começa logo a partir dos 7 anos de idade. A comparação que fizemos, por regiões da NUTS III, entre o PIB/habitante e os RM em cada domínio dos diferentes estudos mostra-nos uma relação direta entre o PIB/habitante e os RM dos estudos referidos nas regiões com PIB/habitante extremos: a região com PIB/habitante mais alto é a região com RM mais elevados nos três estudos do PISA; as duas regiões com PIB/habitante mais baixo são as que têm RM mais baixos em sete dos nove estudos/domínios aqui citados, incluindo os três do PISA. No entanto, essa relação não se verifica nas restantes regiões, isto é, não há correlações significativas entre o PIB/habitante e os RM em cada domínio de cada estudo, por regiões da NUTS III, se excluirmos as que têm valores extremos de PIB/habitante.

Os resultados apresentados neste artigo e correspondentes conclusões têm as suas limitações. A principal dessas limitações tem que ver com a assunção que fizemos sobre a representatividade das amostras de cada uma das unidades territoriais da NUTS III usadas em cada um dos estudos citados. De facto, o cuidado que as entidades competentes tiveram em usar amostras que representassem as diversas regiões da NUTS III é evidente nos relatórios nacionais sobre os diferentes estudos usados nas comparações correlacionais realizadas neste artigo (cf. IAVE, 2017; Marôco, 2016a, 2016b, 2016c). Contudo, alguns erros amostrais associados aos RM das regiões da NUTS III são elevados e, portanto, como refere Marôco (2016a), os resultados que obtivemos “deverão ser lidos com precaução” (p. 62).

Pretendemos, com este artigo, dar alguns contributos para a educação em Portugal e também noutros países, no sentido de melhor se conhecer os estudos do PISA, do TIMSS e do PIRLS, em particular numa ótica comparativa. As análises sobre o conceito de literacia, o conteúdo dos protocolos e do tratamento de dados usados em cada estudo já tinham sido desenvolvidas noutros artigos, mas não com o carácter comparativo que apresentámos e que se acrescenta ao saber já adquirido pela comunidade. As comparações entre os resultados globais dos ILSAs também são muito pouco frequentes e como, sobre Portugal, só conhecemos o estudo de Marôco (2020), publicado muito recentemente, entendemos que o nosso artigo acrescenta informação e outras comparações que permitem perceber melhor as diferenças que existem entre eles. Por fim, a análise correlacional entre os resultados do PISA, do TIMSS e do PIRLS, por regiões da NUTS III, e destes com o PIB/habitante, que apresentamos e que cremos não ter sido tratada em artigo algum, consideramos que é uma abordagem nova sobre a avaliação de currículos e de competências para a vida. Cremos que abre uma nova janela por onde outros autores podem vislumbrar novos caminhos de investigação, em particular sobre a relação entre os percursos académicos e os das competências para a vida e entre estas e as assimetrias socioeconómicas regionais das famílias.

AGRADECIMENTOS

Este trabalho foi financiado pela Fundação para a Ciência e a Tecnologia (FCT) [PTDC/CED-EDG/30084/2017].

REFERÊNCIAS

- Benavente, A. (2016, outubro). O ‘dia’ seguinte: O que a Troika fez à escola. *Le Monde Diplomatique* – edição portuguesa, 8-9.
- Bourdieu, P., & Passeron, J.-C. (1970). *La reproduction: Éléments pour une théorie du système d’enseignement*. Éditions Minuit.
- Campbell, J., Kelly, D., Mullis, I., Martin, M., & Sainsbury, M. (2000). *Framework and specifications for PIRLS assessment 2001: Progress in international reading literacy study*. Boston College.

Carvalho, L., & Costa, E. (2009). *Production of OECD's Programme for International Student Assessment (PISA)*. Rapport pour le projet Know&Pol. Université de Lisbonne, Faculté de Psychologie et de Sciences de l'Éducation.

Conseil National d'Évaluation du Système Scolaire. (2016). *Comparaison des évaluations PISA et TIMSS: Acquis des élèves: Comprendre les évaluations internationales* – Rapport scientifique. CNESCO.

Conselho Nacional de Educação. (2013). *Avaliações internacionais e desempenho dos alunos portugueses*. Lisboa, CNE.

Ferrão, M. (2015). Investigação em educação e os resultados do PISA. Análise estatística da retenção através do PISA 2012. In J. Justino & M. Miguéns (Org.), *Investigação em educação e os resultados do PISA* (pp. 104-117). Conselho Nacional de Educação.

Ferreira, A., Flores, I., & Casas-Novas, T. (2017). *Introdução ao estudo – Porque melhoraram os resultados PISA em Portugal? Estudo longitudinal e comparado (2000-2015)*. Fundação Francisco Manuel dos Santos.

Harju-Luukkainen, H., McElvany, N., & Stang, J. (Eds.). (2020). *Monitoring student achievement in the 21st century*. European policy perspectives and assessment strategies. Cham Springer. <https://doi.org/10.1007/978-3-030-38969-7>

Instituto de Avaliação Educativa. (2017). *Resultados globais PIRLS 2016 – ePIRLS 2016 – Portugal. Literacia de leitura & literacia online*. Lisboa, Portugal: IAVE.

Instituto de Avaliação Educativa. (2019). *TIMSS8 – Itens de matemática e de ciências* – I. IAVE.

Lafontaine, A. (2008). *PIRLS 2006. Progress in reading literacy study. Note de synthèse*. Université de Liège.

Lemos, V. (2014). *A influência da OCDE nas políticas públicas de educação em Portugal*. Almedina.

Lourenço, V., Duarte, A., Nunes, A., Amaral, A., Gonçalves, C., Mota, M., & Mendes, R. (2019). *PISA 2018 – Portugal. Relatório nacional*. IAVE. Disponível em: <http://www.cnedu.pt/pt/noticias/internacional/1517-resultados-pisa-2018>.

Marôco, J. (2018). O bom leitor: Preditores da literacia de leitura dos alunos portugueses no PIRLS 2016. *Revista Portuguesa de Educação*, 31(2), 115-131. <https://doi.org/10.21814/rpe.14105>

Marôco, J. (2020). International large-scale assessments: Trends and effects on the Portuguese public education system. In H. Harju-Luukkainen, N. McElvany, & J. Stang (Eds.), *Monitoring student achievement in the 21st century. European policy perspectives and assessment strategies* (pp. 207-222). Springer. https://doi.org/10.1007/978-3-030-38969-7_17

Marôco, J., Gonçalves, C., Lourenço, V., & Mendes, R. (2016a). *PISA 2015 – Portugal. Volume I: Literacia científica, literacia de leitura e literacia matemática*. Lisboa, Portugal: IAVE.

Marôco, J., Lourenço, V., Mendes, R., & Gonçalves, C. (2016b). *TIMSS 2015 – Portugal. Volume I: Desempenhos em matemática e em Ciências*. IAVE.

Marôco, J., Lourenço, V., Mendes, R., & Gonçalves, C. (2016c). *TIMSS Advanced 2015 – Portugal 1 – Desempenhos em matemática e em física*. IAVE.

Martin, M., Foy, P., Mullis, I., & O'Dwyer, L. (2013). Effective schools in reading, mathematics, and science at the fourth grade. In M. Marin & I. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science*

achievement at the fourth grade – Implications for early learning (pp. 109-178). TIMSS & PIRLS International Study Center, Boston College.

Martin, M., & Kelly, D. (Org.). (1997). *TIMSS – Technical report*, vol. I. IEA.

Mullis, I., Martin, M., Foy, P., & Drucker, K. (2012). *PIRLS 2011 international results in reading*. IEA.

Mullis, I., Martin, M., Kennedy, A., Trong, K., & Sainsbury, M. (2009). *PIRLS 2011 assessment framework*. TIMSS & PIRLS International Study Center, Lynch School of Education: Boston College.

Organização para a Cooperação e Desenvolvimento Económico. (1999). *Measuring student knowledge and skills: A new framework for assessment*. Paris, França: OCDE.

Organização para a Cooperação e Desenvolvimento Económico. (2003). *La lecture, moteur de changement: Performances et engagement d'un pays à l'autre: Résultats de PISA 2000*. OCDE.

Organização para a Cooperação e Desenvolvimento Económico. (2013). *Cadre d'évaluation et d'analyse du cycle PISA 2012*. OCDE.

Organização para a Cooperação e Desenvolvimento Económico. (2019a). *PISA 2018 Results. What students know and can do*, vol. I. OCDE.

Organização para a Cooperação e Desenvolvimento Económico. (2019b). *PISA 2018 Results. Where all student can succeed*, vol. II. OCDE.

Rémond, M. (2006). Éclairages des évaluations internationales PIRLS et PISA sur les élèves français. *Revue Française de Pédagogie*, 157, 71-84.

Rizvi, F., & Lingard, B. (2006). Globalisation and the changing nature of the OECD's educational work. In H. Laurder, P. Brown, J. Dillabough, & A. Halsey (Eds.), *Education, globalization and social change* (pp. 247-260). Oxford University Press.

Saraiva, L. (2017). A aprendizagem das ciências em Portugal: Uma leitura a partir dos resultados do TIMSS e do PISA. *Mediações – Revista OnLine da Escola Superior de Educação do Instituto Politécnico de Setúbal*, 5(2), 4-18.

Schleicher, A. (2018). *World class: How to build a 21st-century school system, strong performers and successful reformers in education*. OCDE Publishing. <https://www.oecd.org/education/world-class-9789264300002-en.htm>

Teodoro, A. (2010). *Educação, globalização e neoliberalismo. Os novos modos de regulação transacional das políticas de educação*. Lisboa, Portugal: Edições Universitárias Lusófonas.

Teodoro, A. (2019). The end of isolationism: Examining the OCDE influence in Portuguese education policies, 1955-1974. *Paedagogica Historica: International Journal of the History of Education*, 1-14. <https://doi.org/10.1080/00309230.2019.1606022>

Informação dos autores:

i Universidade Lusófona de Humanidades e Tecnologias, Centro de Estudos Interdisciplinares em Educação e Desenvolvimento (CeIED), Portugal. <https://orcid.org/0000-0002-5093-8115>

ii Universidade Lusófona do Porto, Centro de Estudos Interdisciplinares em Educação e Desenvolvimento (CeIED), Portugal. <http://orcid.org/0000-0003-1466-1013>

iii Instituto Politécnico do Porto, Escola Superior de Educação, Centro de Estudos Interdisciplinares em Educação e Desenvolvimento (CeIED), Portugal. <https://orcid.org/0000-0001-5854-536X>

iv Universidade Lusófona de Humanidades e Tecnologias, Centro de Estudos Interdisciplinares em Educação e Desenvolvimento (CeIED), Portugal. <https://orcid.org/0000-0001-7819-0498>

Toda a correspondência relativa a este artigo deve ser enviada para:

Vítor Rosa
ULHT/CeIED
Campo Grande, 376
1749-024 Lisboa, Portugal
vitor.rosa@ulusofona.pt

Recebido em 29 de julho de 2019
Aceite para publicação em 4 de junho de 2020

PISA, TIMSS and PIRLS in Portugal: A comparative analysis

ABSTRACT

In this article we present a comparative analysis of three international surveys: PISA, carried out by OECD, and TIMSS and PIRLS, undertaken by IEA, considering their general objectives and the data correlation on Portugal (by NUTS III regions). The literature and documental review has shown how results can be reformulated or reinterpreted by the different social actors, and that they are seen by the political decision makers and organizations as a tool for school system comparison, revealing strengths and weaknesses. The analysis of general results reveals that Portugal's results has been improving in PISA, in the three domains (reading, mathematics and sciences), that they have improved in TIMSS in mathematics, while the reading performance in PIRLS has declined. The comparative analysis by regions has shown that there is a difference in results by regions and that, while the domain factor does not seem relevant, the object of evaluation factor in each study seems decisive for the results of the surveys. This comparison has also revealed a direct link between the GDP/inhabitant and the results of the surveys in regions with extreme values of GDP/inhabitant, although this relation does not exist for most regions.

Keywords: NUTS III Regions; PISA; TIMSS; PIRLS

PISA, TIMSS y PIRLS en Portugal: Un análisis comparativo

RESUMEN

En este estudio presentamos un análisis comparativo de tres encuestas internacionales: PISA, realizada por la OCDE, y TIMSS y PIRLS, realizadas por la AIE, teniendo en cuenta sus objetivos generales y una correlación de los datos referentes a Portugal (por regiones de la NUTS III). La revisión bibliográfica y documental mostró cómo los resultados pueden ser reformulados o reinterpretados por los diferentes actores sociales y que son vistos por los responsables políticos y las organizaciones como una herramienta de comparación de los sistemas escolares, revelando sus fortalezas y debilidades. El análisis de los resultados generales muestra que Portugal ha mejorado sus resultados en PISA, en los tres dominios (lectura, matemáticas y ciencias), que mejoró en TIMSS en matemáticas y empeoró en lectura en PIRLS. El análisis comparativo por regiones reveló que existen diferencias de resultados por regiones y que, si bien el factor dominio no es relevante, el factor objeto de evaluación en cada estudio parece ser un factor determinante de los resultados de los cuestionarios. También reveló que existe una relación directa entre el PIB/habitante y los resultados de los cuestionarios en regiones con valores extremos de PIB/habitante, pero que esta relación no existe para la mayoría de las regiones.

Palabras clave: Regiones NUTS III; PISA; TIMSS; PIRLS