

Avaliação internacional em larga escala: a participação de Portugal no TIMSS, PIRLS e ICILS

The international large-scale assessment: issues from Portugal's participation in TIMSS, PIRLS and ICILS

Vítor Rosa

Universidade Lusófona, CeIED

vitor.rosa@ulusofona.pt

Resumo

Nas últimas décadas a avaliação em larga escala adquiriu uma grande importância no âmbito educacional. Os estudos internacionais desenvolvidos a partir de questionários são cada vez mais numerosos e abrangem diferentes domínios. Portugal tem vindo a participar em vários desses estudos, permitindo obter informações sobre o sistema educativo e o contexto socioeconómico das famílias e pessoais dos alunos. Esses dados permitem influenciar na definição e implementação de políticas educacionais. Neste artigo apresentamos uma análise sobre três estudos internacionais no âmbito da educação, em que Portugal participa: Trends in International Mathematics and Science Study (TIMSS), o Progress in International Reading Literacy Study (PIRLS) e o International Computer and Information Literacy Study (ICILS), conduzido pela *International Association for the Evaluation of Educational Achievement* (IEA). Procuramos analisar os resultados gerais de Portugal, recorremos a uma metodologia interpretativa, com base na análise dos relatórios e bases de dados produzidos(as) por diferentes organizações, nomeadamente a IEA. Apesar das diferentes oscilações, nos diferentes ciclos dos estudos, os alunos portugueses têm vindo a apresentar resultados positivos.

Palavras-chave: Avaliação; educação; TIMSS; PIRLS; ICILS

Abstract

In the last few decades, the large-scale evaluation has acquired an importance in the educational field. International studies developed from surveys are increasingly numerous and cover different domains. Portugal has been participating in several studies, allowing to obtain information about the educational system and the socioeconomic context of the families and students. These data are reflected in the definition and implementation of educational policies. In this article we present an analysis of three international studies in the field of education, in which Portugal participates: Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS) and the International Computer and Information Literacy Study (ICILS), carried out by *International Association for the Evaluation of Educational Achievement* (IEA). In this article we present an analysis of the general results of Portugal and we use an interpretative methodology, based on the analysis of reports and databases produced by different organizations, and undertaken by IEA. Despite the different fluctuations, in the different study cycles, Portuguese students have been showing positive results.

Keywords: Evaluation; education; TIMSS; PIRLS; ICILS

Introdução

Os estudos internacionais de avaliação, baseados em amostragem de larga escala, produzem informações e diversos indicadores sobre os conhecimentos e as competências dos alunos de diferentes sistemas educativos. Essas avaliações, no âmbito da educação, adquiriram uma grande importância nas últimas décadas. Governos, de vários quadrantes políticos, passaram a utilizar os resultados destes estudos, com o objetivo de melhorar os investimentos e obter melhores resultados escolares.

Embora as avaliações internacionais em grande escala, conhecidas usualmente por ILSA, sejam atualmente consideradas por muitos uma regularidade do panorama educativo, elas são um fenómeno relativamente recente. As suas origens podem ser rastreadas até ao inquérito-piloto da *International Association for the Evaluation of Educational Achievement* (IEA), relativo à avaliação do desempenho dos alunos realizada na década de 1960 (Rosine & Postlethwaite, 1994). Desde então, registaram-se desenvolvimentos significativos. O número de organizações responsáveis pelo desenvolvimento, gestão e condução das ILSA cresceu de um para sete grandes atores: IEA, Conferência dos Ministros da Educação dos Estados e dos Governos da Francofonia (CONFENEM), Banco Interamericano de Desenvolvimento (IDB), Organização para a Cooperação e Desenvolvimento Económico (OCDE), Consórcio da África Austral e Oriental para a Monitorização da Qualidade Educacional (SACMEQ), Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) e o Banco Mundial (Lietz, Cresswell, Rust, & Adams, 2017; Wagemaker, 2014). Atualmente, estima-se que cerca de 70% dos países do mundo participam nas ILSA (Lietz, Cresswell, Rust & Adams, 2017).

As metodologias de medição evoluíram consideravelmente. Nas ILSA, *Programme for International Student Assessment* (PISA), *Progress in International Reading Literacy Study* (PIRLS), *Trends in International Mathematics and Science Study* (TIMSS) e *International Computer and Information Literacy Study* (ICILS), as raízes assentam nas metodologias de avaliação de tendências de longo prazo na Avaliação Nacional do Progresso Educacional (NAEP) nos Estados Unidos da América (EUA). As suas metodologias foram adaptadas e ampliadas para atender aos desafios desses e de outros estudos que avaliam o desempenho educacional para além das fronteiras nacionais, garantindo, por exemplo, a comparabilidade dos itens do teste ao operar com um número crescente de países e línguas participantes (Rutkowski, Gonzalez, Joncas & von Davier, 2010).

Os esforços de reformas na educação centraram-se em preocupações relacionadas com os desafios de garantir a equidade em termos escolares. No entanto, com um maior reconhecimento dos efeitos da globalização e da competitividade económica e com uma maior preocupação com a equidade dos resultados de aprendizagem (por exemplo, em termos do que os alunos sabem e podem fazer), as ILSA são cada vez mais vistas como uma condição necessária para acompanhar e compreender os resultados dos investimentos significativos que todas as nações fazem na educação.

Ao longo dos anos, as organizações que conduzem as ILSA desenvolveram padrões técnicos com requisitos mínimos e os relatórios dos resultados para os ciclos de estudos são geralmente

acompanhados por documentação técnica abrangente, que fornecem orientações para a interpretação dos dados e a implementação de análises secundárias.

Portugal tem participado em alguns destes estudos, que são promovidos por diversas organizações internacionais: no TIMSS desde 1995, no PISA desde 2000, no PIRLS desde 2011 e no ICILS desde 2013. Para além do posicionamento de Portugal no panorama internacional, a participação nestes estudos permite obter informações sobre o sistema de ensino. Os dados recolhidos permitem depois na definição e implementação de políticas educativas.

A influência destes estudos sobre a evolução dos sistemas educativos encontra-se bem documentada (Rosa, Maia, Mascarenhas & Teodoro, 2020; Rosa, 2020). Mas não são apenas instrumentos que influenciam a decisão política na área educativa de vários países. Eles também fornecem dados únicos que incitam os investigadores do mundo inteiro a elaborar uma série de análises secundárias (Olsen & Lie, 2006).

Este artigo procura analisar, de uma forma geral, três deles, em que Portugal participa¹: o TIMSS, o PIRLS e o ICILS, procurando olhar para os seus dados gerais. Ele situa-se na sequência de outros trabalhos de investigação (Rosa, Maia, Mascarenhas & Teodoro, 2020), tendo um deles sido publicado na revista EDUSER (Rosa, 2020).

Metodologia

Sendo conduzidos pela IEA, os estudos TIMSS, PIRLS e ICILS apresentam uma metodologia semelhante e complexa. A IEA recorre a diversos testes e inquéritos por questionário aos alunos, aos diretores e coordenadores de Escola, aos docentes e aos pais/Encarregados de Educação. As avaliações são dirigidas a diferentes níveis de ensino e a diferentes literacias: o TIMSS procura avaliar os alunos do 4.º, 8.º e 12.º anos de escolaridade, analisando as literacias de matemática, ciências e física, o PIRLS avalia os alunos do 4.º ano de escolaridade, tendo presente a literacia de leitura, e o ICILS avalia as competências dos alunos do 8.º ano de escolaridade em Tecnologias da Informação e da Comunicação (TIC).

A IEA recorre a um processo de amostragem aleatória multi-etapa, garantindo que as amostras sejam representativas da população-alvo. A escala de pontuação é apresentada numa escala numérica entre 0 e 1000 pontos, com um ponto central de referência de 500 pontos e um desvio-padrão de 100 pontos. Esta pontuação traduz-se em diferentes níveis de desempenho

¹ Num artigo recentemente publicado procurámos comparar o PISA, o TIMSS e o PIRLS (Rosa, Maia, Mascarenhas & Teodoro, 2020).

dos alunos (Avançado, Elevado, Intermédio e Baixo). Os resultados são apresentados em forma de rankings dos países.

Neste estudo, recorreremos à recolha documental e à consulta das bases de dados produzidas no âmbito destes estudos, em particular os relatórios produzidos pela IEA e pelo Instituto de Avaliação Educativa, I.P. (IAVE), entidade responsável pela organização e gestão dos estudos internacionais de alunos em Portugal. Usamos uma abordagem qualitativa de carácter interpretativo.

TIMSS

O TIMSS é um estudo internacional comparativo que procura avaliar o nível de conhecimentos escolares dos alunos do 4.º e 8.º anos de escolaridade em matemática e ciências, que, na perspetiva de Marôco, Lourenço, Mendes e Gonçalves (2016a), são considerados domínios ou literacias “essenciais na formação dos alunos que optam por percursos formativos associados a áreas profissionais internacionalmente conhecidas como STEM (*Science, Technology, Engineering and Mathematics*)” (pp. 5-6). De facto, a matemática é objetiva e universal, indo para além das nações e das diferenças de culturas (Schmidt, MacKnight, Valverde, Houang & Wiley, 1997). Segundo a Etnomatemática, que surgiu na década de 1970, as ideias matemáticas, consubstanciadas nos processos de comparar, classificar, quantificar, medir, organizar, inferir e concluir, são próprias da natureza humana. Nesse sentido, a matemática é “própria do indivíduo” e é moldada pelo “meio ambiente natural, social e cultural” em que este se insere (D’Ambrósio, 2005). A ciência refere-se a qualquer conhecimento adquirido através do método científico.

O escopo e a complexidade do TIMSS são enormes. O sucesso deste estudo depende do esforço colaborativo entre os centros de investigação de cada país responsáveis pela gestão das tarefas, os procedimentos de amostragem das escolas e alunos e a realização das várias etapas necessárias para o processamento e análise de dados (Beaton, Martin, Mullis, Gonzalez, Smith & Kelly, 1996). Segundo Bodin e Grapin (2018), o TIMSS é um projeto de investigação que “procura esclarecer as ligações entre os programas e as instruções oficiais (curriculum desejado), as práticas de ensino (curriculum implementado) e as competências dos alunos (curriculum atingido)” (p. 71). É, na perspetiva de Fernandes (2008, p. 282), um “estudo influenciado por investigadores interessados em compreender relações entre o currículo, os contextos em que ele se desenvolve e as aprendizagens dos alunos”. Espera-se que os alunos aprendam, em cada país, de “acordo com as políticas educativas, a organização do sistema educativo e os aspetos

culturais. A organização dos planos curriculares, os conteúdos, os objetivos da aprendizagem e os processos de avaliação, enunciam os resultados esperados por cada sistema educativo” (Carvalho, Amaro, Reis e Neres, 1996, p. 2).

O quadro de referência para a matemática articula-se em torno de dois domínios: os conteúdos (números, formas geométricas e medidas, apresentação de dados) e as competências cognitivas (conhecer, aplicar, raciocinar). No TIMSS, em termos gerais, a conceção das dimensões cognitivas é semelhante às literacias de matemática e ciências, mas a identificação das competências relacionadas para cada uma é específica dos domínios avaliados. A dimensão “conhecer” refere-se aos factos, conceitos e procedimentos que o aluno deve conhecer. A dimensão “aplicar” abrange a capacidade do aluno em empregar o seu conhecimento e a compreensão dos conceitos para resolver os problemas ou na resposta a questões. A dimensão “raciocinar” diz respeito a situações complexas e problemas que requerem vários passos até que se encontre uma solução, o que envolve pensamento lógico e sistemático. A avaliação das ciências incide sobre as “ciências da vida”, as “ciências físicas” e as “ciências da terra”. Por outro lado, a avaliação é também organizada em função da tipologia dos itens e pela distribuição das áreas de conteúdo e as dimensões cognitivas, tal como se apresenta na Tabela 1.

	Itens de ciências	Itens de seleção	Itens de construção	Total	% da pontuação
Área da Conteúdo	Ciência da Vida	39	40	79	46
	Ciências Físicas	36	28	64	35
	Ciências da Terra	23	10	33	19
	Total	98	78		100
	% da pontuação	52	48	176	
Dimensão Cognitiva	Conhecer	47	25	72	41
	Aplicar	32	35	67	38
	Raciocinar	19	18	37	21
	Total	98	78		100
	% da pontuação	52	48	176	

Tabela 1 - Distribuição dos itens de ciências (área de conteúdo, dimensão cognitiva e tipo de item), TIMSS 2015

Fonte: Marôco *et al.* (2016b); Saraiva (2017).

No total de 176 itens, verifica-se que existe uma ligeira prevalência dos itens de seleção (que são de escolha múltipla) relativamente aos itens de construção (que exigem uma resposta curta), respetivamente 98 e 78. Estes dados revelam também que a área das ciências da terra assume um menor número de itens (33), correspondendo a 19% da percentagem total de pontuação. Este aspeto evidencia o menor estatuto que é atribuído a esta área do conhecimento em comparação com as áreas das ciências da vida e das ciências físicas, tal como já tinha sido

referido, em 1995, por Carvalho, Amaro, Reis e Neres (1996). Esta distribuição, na opinião de Saraiva (2017), tem correspondência nos programas oficiais em Portugal.

Da responsabilidade da IEA, o TIMSS tem por objetivo interpretar as diferenças entre os sistemas educativos, procurando melhorar o ensino e a aprendizagem dos alunos dos países participantes (Drent, Martina, Meelissen & Fabienne, 2013). Ele decorre, desde 1995, com uma periodicidade quadrienal. Portugal participou no ano inicial e depois em 2011, 2015 e 2019, sendo que neste último ano ainda não se conhecem publicamente os principais resultados.

Em complemento do TIMSS, a IEA realiza o TIMSS *Advanced*, que procura avaliar o desempenho dos alunos a frequentar o ano terminal do ensino secundário (12.º ano) em matemática avançada (A) e física. O número de países tem vindo a aumentar nos vários ciclos do TIMSS.

Este estudo internacional de larga escala recorre a um processo de amostragem aleatória multi-etapa, garantindo que a amostra seja representativa da população-alvo. Os conhecimentos e competências são estimados por modelos de teoria de resposta ao item, associando valores plausíveis numa escala de 0 a 1000 pontos, com a média de 500 pontos e o desvio-padrão de 100 pontos. Cada aluno responde a um teste, que combina diversos itens de matemática e de ciências, e tem a duração de cerca noventa minutos. Os itens são confidenciais, permitindo ao consórcio comparar os resultados ao longo das edições do estudo e identificar as tendências globais dos resultados. A IEA, a cada ciclo, divulga publicamente alguns desses itens, que deixam de integrar futuras provas, por forma a que se possa conhecer as questões que são colocadas aos alunos (IAVE, s./d., 2019).

Portugal participa no TIMSS, estando envolvido nas edições de 1995 (com os alunos do 4.º e 8.º ano de escolaridade), 2011 (com alunos do 4.º ano de escolaridade), 2015 (alunos do 4.º ano de escolaridade) e 2019 (alunos do 4.º ano de escolaridade). Em 2015, os alunos portugueses do 12.º ano participaram no TIMSS *Advanced*, com a avaliação a matemática e a física. Os resultados têm sido positivos: de 442 pontos (em 1995) passou para 525 (em 2019). Os rapazes têm vindo a obter uma pontuação média superior à das raparigas: rapazes (444 em 1995 e 533 em 2019); raparigas (440 em 1995 e 516 em 2019).

PIRLS

O PIRLS constitui uma extensa avaliação internacional dos conhecimentos e competências curriculares e escolares dos alunos de todo o mundo a frequentar este ano de escolaridade na área da literacia de leitura. A razão desta escolha, na perspetiva de Sim-Sim (2013), prende-se

com a etapa de escolaridade, em que os mecanismos básicos de aprendizagem da leitura devem estar consolidados.

O conceito de literacia surge em Portugal na década de 1990, mais concretamente com o estudo *A Literacia em Portugal*, de Benavente, Rosa, Costa e Ávila (1996). Ele tem a ver com as capacidades relativamente à utilização e interpretação de informação escrita por parte de cada indivíduo. Remete, assim, para as práticas quotidianas e para as competências, independentemente dos níveis escolares. Atualmente, a noção de literacia é utilizada de forma muito abrangente, referindo-se a diferentes domínios (literacia científica, informática, desportiva, entre outras) (Ávila, 2016). No caso do PIRLS, a literacia de leitura definida pela IEA é entendida como sendo a capacidade para compreender e utilizar as formas de linguagem escrita requeridas pela sociedade e/ou valorizadas pelo indivíduo (Campbell, Kelly, Mullis, Martin & Sainsbury, 2001; IAVE, 2017b).

O quadro de referência deste inquérito internacional centra-se na diversidade de experiências de leitura que os alunos poderão vivenciar, seja na escola, seja na sua vida quotidiana, construindo um sentido relativamente a uma ampla variedade de textos. Tal como é entendida no PIRLS, a compreensão de leitura combina dois objetivos de leitura: ler para se ter experiência literária e ler para se adquirir e utilizar a informação. Os textos literários são recitais completos acompanhados com ilustrações. Os textos devem familiarizar os alunos com a intriga, os eventos, as ações, as motivações das personagens, etc. Os textos informativos permitem tratar de aspetos do mundo real, abrangendo uma grande variedade de assuntos. São acompanhados de elementos estruturantes e ilustrativos, tais como diagramas, cartas, ilustrações, fotografias, listagens, quadros, entre outros.

No PIRLS, a população estudada não corresponde a uma geração de alunos com uma idade definida, como é o caso do Programa PISA da OCDE. O objetivo é medir as *performances* do conjunto dos alunos presentes num determinado nível de escolaridade, independentemente da sua idade, do seu percurso e da organização do sistema educativo.

Os países/regiões participantes no estudo identificam os pontos fortes e fracos nas competências de leitura dos alunos. Como referimos anteriormente, Portugal participou em 2011 e depois em 2016 (3.^a e 4.^a edições). Na edição de 2016, o estudo integrou uma nova dimensão na avaliação da literacia de leitura: *online* (ePIRLS), envolvendo 14 países, entre os quais Portugal. Nos estudos PIRLS/ePIRLS, através dos suportes papel e digital, recolhe-se informação sobre os alunos, as suas famílias, os professores e as escolas, contextualizando as oportunidades de aprendizagem de leitura, assim como a identificação dos fatores que

influenciem essas oportunidades. No âmbito do PIRLS, o protocolo da avaliação (textos e itens) está assente no cruzamento dos processos de compreensão e dos objetivos de leitura (Campbell, Kelly, Mullis, Martin & Sainsbury, 2001).

O uso da leitura apresenta duas finalidades: a experiência literária e a aquisição e utilização de informação. Ferreira e Gonçalves (2013) sublinham que

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. Assim, o quadro concetual de avaliação foi desenhado de modo a contemplar finalidades e processos de compreensão da leitura que dão significado a este conceito de literacia de leitura. (p. 12)

A avaliação do PIRLS 2016 integrou quatro tipos de processos de compreensão da leitura no âmbito de cada uma das finalidades: localizar e retirar informação explícita; fazer inferências diretas; interpretar e relacionar ideias e informação; analisar e avaliar conteúdo e elementos textuais. Os protocolos do PIRLS assentam sobre textos longos e integrais (no total de 8 no PIRLS de 2011) e sobre cada um deles existem cerca de uma dúzia de questões. Os exercícios são curtos, dado que é preciso cruzar vários fatores para obter a declinação dos itens correspondendo às características definidas (Rémond, 2006, 2007)¹. Uma metodologia complexa condiciona a construção e organização dos materiais, elaborado em inglês, para formar a “versão fonte”, servindo a base comum para as traduções noutras línguas (Araújo, Costa e Folgado, 2016; Marôco, 2018; Rémond, 2007).

Embora com objetivos diferentes, o PIRLS tem muitas similitudes, tanto na conceção como nas metodologias (Rémond, 2006), com outros estudos internacionais, nomeadamente o PISA, levado a cabo pela OCDE. Como nos outros estudos, no final existe uma escala de resultados de compreensão elaborada a partir de modelos de resposta ao item (MRI), cuja média internacional se situa nos 500 e o desvio-padrão é de 100. Desta forma, os países participantes podem ser classificados com base numa dimensão comum. Contrariamente ao PISA, no PIRLS (e no TIMSS) a avaliação das competências de leitura distribui-se por quatro níveis de desempenho: nível Baixo (de 400 a 474 pontos), nível Intermédio (de 475 a 549 pontos), nível Elevado (de 550 a 624 pontos) e nível Avançado (625 pontos ou mais). Como as percentagens

¹ A IEA divulgou publicamente 3 unidades de avaliação da leitura que fizeram parte do teste PIRLS e 2 unidades de avaliação do teste ePIRLS aplicado em 2016. Em 2011, divulgou, igualmente, um conjunto de itens libertos pelo consórcio. O IAVE compilou a informação da versão portuguesa em dois documentos (IAVE, 2018). Os questionários de contexto relativos ao PIRLS de 2011 encontram-se disponíveis no site do IAVE: <http://www.iave.pt/index.php/estudos-internacionais/pirls/instrumentos-de-avaliacao> (consultado em 09/09/2020).

são cumulativas, os alunos que alcançaram o nível Avançado, por exemplo, também alcançaram os níveis precedentes. Note-se que, com esta classificação, os alunos que obtiverem uma pontuação inferior a 400 não atingem sequer o nível Baixo.

Lafontaine (2008) aponta 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 4.º 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.

Rémond (2007) refere que os resultados do PIRLS levam a pensar que as tarefas propostas pela escola aos alunos são muito limitadas. Os alunos ficam “desconcertados” com as questões complexas e sucessivas, que é preciso responder sem a ajuda do docente (Rémond, 2007). 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.

Portugal participou neste estudo pela primeira vez em 2011 (a 3.ª edição do PIRLS) e a última participação foi em 2016. Comparando os resultados globais deste estudo nesses dois anos, podemos verificar que houve uma descida na avaliação do desempenho dos alunos portugueses do 4.º ano de escolaridade. Face a 2011, a pontuação representa uma descida de 13 pontos (de 541 passa para 528). Na escala ordenada dos resultados, entre os dois anos em causa, Portugal passou de 19.º ano para o 30.º lugar. Relativamente à distribuição dos resultados do PIRLS por género, constatamos que as raparigas têm melhores resultados do que os rapazes no âmbito da leitura em 2011 e 2016 (rapazes: 534 em 2011; 527 em 2016; raparigas: 548 em 2011; 529 em 2016).

ICILS

Com um período quinquenal, o ICILS avalia as competências dos anos do 8.º ano de escolaridade em TIC¹. Parte de uma questão: será que os alunos se encontram bem preparados para estudar, trabalhar e viver no mundo digital? O estudo foca-se em dois domínios-chave: a Literacia em Computadores e Informação (CIL) e o Pensamento Computacional (CT). Para além de influenciar a decisão política, espera-se que as conclusões possam ter impacto no trabalho realizado pela escola, melhorando, conseqüentemente, o sucesso educativo dos alunos (Fraillon *et al.*, 2020). Na perspetiva de Vanda, Nunes, Amaral, Gonçalves, Mota & Mendes (2019),

Trata-se de um estudo que avalia domínios considerados muito importantes para o desenvolvimento dos alunos, tanto do ponto de vista do contributo para o desenvolvimento curricular, dos saberes e conhecimentos disciplinares, como do desenvolvimento sociocognitivo e metacognitivo (p. 7).

Relativamente a outros estudos (PISA, TIMSS e PIRLS), o número de países/regiões participantes é menor. Em 2018, na avaliação CIL participaram 12 países (Chile, Dinamarca, Estados Unidos da América, Finlândia, França, Alemanha, Itália, Cazaquistão, República da Coreia, Luxemburgo, Portugal e Uruguai) e 2 regiões em *benchmarking* (Moscovo – Federação Russa e Renânia do Norte-Vestefália – Alemanha) e na avaliação CT participaram 8 países (Dinamarca, Finlândia, França, Alemanha, Portugal, Luxemburgo, Estados Unidos da América, República da Coreia) e 1 região (Renânia do Norte-Vestefália – Alemanha)². No total, o estudo ICILS obteve informação de 46.561 alunos do 8.º ano de escolaridade e de 26.530 professores de 2226 escolas.

A IEA segue a mesma estrutura do TIMSS e do PIRLS relativamente à escala numérica, que varia entre os 0 e os 1000 pontos e com um ponto central fixo em 500 pontos (desempenho médio). O desvio-padrão é de 100 pontos.

O ICILS é estruturado num quadro conceptual de referência, onde são definidas as dimensões de análise e as áreas de conteúdo avaliadas nos dois domínios em apreço (CIL e CT). O teste³ é composto por diferentes níveis de dificuldade das tarefas, assim como por níveis de proficiência de desempenho. O nível de proficiência 1 situa-se entre os 407 e 491 pontos, o

¹ As TIC, em Portugal, são uma disciplina obrigatória para os alunos do 5.º ao 9.º ano de escolaridade. As competências curriculares estão organizadas em quatro domínios: 1) cidadania digital; 2) investigar e pesquisar; 3) comunicar e colaborar; 4) criar e inovar. A Direção-Geral da Educação (DGE) tem vindo a promover em vários anos letivos (2015-2017) iniciativas de iniciação à programação, dirigida a alunos do 3.º e 4.º ano de escolaridade. Em 2017/2018 passou a designar-se de “Probótica”.

² No ICILS 2018, o CT foi considerado como domínio optativo.

³ O teste é composto por 5 módulos (com questões e tarefas) e tem a duração total de 30 minutos cada. O teste CIL tem uma duração de 60 minutos. A avaliação CT foi organizada em 2 módulos de 25 minutos cada.

nível de proficiência 2 situa-se entre os 492 e 576 pontos, o nível de proficiência 3 situa-se entre 577 e os 661 pontos e o nível de proficiência 4 é mais de 661 pontos¹.

De acordo com a IEA, a CIL

Refere-se à capacidade de um indivíduo utilizar computadores para investigar, criar e comunicar, de modo a participar ativamente nas sociedades contemporâneas, seja em casa, na escola, no local de trabalho e nos contextos comunitário e educativo (IAVE, 2019, p. 23).

No que diz respeito ao CT, a IEA define-o da seguinte forma:

Refere-se à capacidade de um indivíduo reconhecer aspetos e problemas do mundo real passíveis de uma formulação computacional, bem como a sua capacidade de avaliar e de desenvolver soluções algorítmicas para esses problemas, que possam ser operacionalizadas em computador (IAVE, 2019, p. 25).

As Tabelas 2 e 3 apresentam a informação mais detalhada sobre as dimensões e as áreas de conteúdo (CIL e CT).

Dimensões CIL	%	Áreas de conteúdo CIL	%
Compreender a utilização do computador	14	Conhecer os fundamentos da utilização do computador	2
		Conhecer as convenções na utilização do computador	12
Recolher informação	25	Aceder e avaliar a informação	15
		Gerir informação	10
Produzir informação	50	Transformar informação	20
		Criar informação	30
Comunicar digitalmente	11	Partilhar informação	8
		Utilizar informação responsável e segura	3

Tabela 2 - Dimensões e áreas de conteúdo (CIL) e percentagens

Fonte: IAVE (2019)

Dimensões CT	%	Áreas de conteúdo CT	%
Conceptualizar problemas	41	Conhecer e compreender sistemas digitais	18
		Formular e analisar problemas	10
		Recolher e representar dados relevantes	13
Operacionalizar soluções	59	Planear e avaliar soluções	33
		Desenvolver algorítmicos, programas e interfaces	26

Tabela 3 - Dimensões e áreas de conteúdo (CT) e percentagens

Fonte: IAVE (2019)

Existem várias definições das TIC. Ricoy e Couto (2012) destacam que a denominação TIC surge nos finais da década de 1990 e são constituídas por meios técnicos para manipular

¹ Para uma informação mais detalhada sobre os níveis de proficiência, intervalos da escala, características do nível e exemplos, ver o relatório do IAVE (2019, pp. 27-30).

informação e promover a comunicação, incluindo o *hardware* e o *software*, com associação às redes computacionais. As TIC vinculam-se também às telecomunicações, como um meio de difusão da comunicação. Para Blurton (1999), as TIC são constituídas por uma diversidade de ferramentas e recursos tecnológicos que se utilizam para comunicar, criar, disseminar e obter informação. Spanhel (2008), por sua vez, esclarece que as TIC são meios tecnológicos ou eletrónicos, baseados nos princípios da digitalização e da conexão em rede. No setor educativo, quando se faz referência a estes dispositivos, refere-se a novas técnicas de informação e comunicação.

No âmbito do ICILS Portugal¹ registou uma pontuação média de 516 pontos na CIL, colocando-o acima da média internacional do ICILS 2018 (496 pontos). No caso do CT, a nível nacional, obteve-se 482 pontos, ficando a 18 pontos abaixo da média internacional. Tal como no TIMSS e PIRLS, existe uma diferenciação de género no ICILS. Em todos os países, as raparigas apresentaram melhores resultados médios do que os rapazes em CIL. Portugal segue a tendência internacional, tendo as raparigas obtido 522 pontos e os rapazes 511 pontos. Na avaliação CT alteraram-se os resultados na tendência internacional. São os rapazes a obter os resultados médios mais elevados. A nível nacional, verifica-se que os rapazes alcançaram uma pontuação média de 490 pontos, o que equivale a 16 pontos acima da pontuação obtida pelas raparigas.

Fatores explicativos dos resultados do TIMSS, PIRLS e ICIL em Portugal

Um estudo do Conselho Nacional de Educação (CNE), órgão independente e que tem por competências emitir opiniões, pareceres e recomendações sobre a educação em Portugal, procura identificar alguns dos fatores explicativos do desempenho dos alunos no âmbito das três principais literacias (leitura, matemática e ciências), com referência ao TIMSS 2015 e ao PIRLS 2016. Os autores desse estudo (Félix, Perdigão & Lourenço, 2020) investigam como é que esses fatores poderão promover a igualdade de oportunidades no acesso à educação e como explicam a diferenciação no desempenho no PIRLS dos alunos e, consequentemente, no seu desempenho escolar. Para o efeito, comparam os resultados de vários países europeus (Finlândia, Noruega, Holanda, Polónia, Alemanha, Eslováquia, Espanha, Itália, França, Irlanda e Portugal). Quando comparado com outros países europeus, Portugal insere-se no grupo de

¹ Para mais informações e aprofundamento do ICILS em Portugal, confronte Rosa (2020).

países com uma escolaridade obrigatória mais longa¹ (European Commission/EACEA/Eurydice, 2020). Alguns dos fatores explicativos são os seguintes:

- 1) Os alunos com origem em famílias com elevado capital familiar (este indicador integra o nível de escolaridade, a qualificação profissional dos encarregados de educação, os livros disponíveis em casa, com destaque para os livros infantis, os materiais de apoio ao estudo)² apresentam melhores desempenhos do que os alunos com origem em famílias com menos recursos socioeconómicos. Esta questão não se coloca apenas no TIMSS e no PIRLS. O PISA 2018 e o TIMSS 2015 também revelaram que o estatuto socioeconómico é um forte preditor da *performance* dos alunos nacionais.
- 2) Quanto maior for o domínio das questões básicas de literacia (e de numeracia) antes da escolaridade, maior será a probabilidade de os alunos terem bons desempenhos no quarto ano de escolaridade.
- 3) A frequência continuada em programas de educação e os cuidados para o desenvolvimento da primeira infância revela-se importante para os alunos de famílias com menores recursos socioeconómicos. No caso de Portugal, os autores referem que “a frequência de três ou mais anos representa um aumento significativo no desempenho em Leitura para os alunos com ‘Poucos ou alguns recursos’, mas não tem um resultado estatisticamente significativo para o grupo com ‘Muitos recursos’” (p. 11).
- 4) Os alunos que depositam mais confiança nas suas competências são os que alcançam melhores resultados nos três domínios principais da literacia.
- 5) Dos vários países europeus, Portugal apresenta a percentagem mais elevada de alunos provenientes de escolas de meios desfavorecidos, conseguindo alcançar, em todos os domínios, pontuações acima da média internacional. Os resultados dos alunos portugueses, quando comparados com os outros países, sugerem uma boa capacidade do sistema de ensino para reduzir as diferenças decorrentes de contextos socioeconómicos distintos.
- 6) As escolas mais orientadas para o sucesso escolar permitem aos seus alunos a obtenção de melhores desempenhos. De facto, o clima escolar revela-se um preditor importante.

¹ Em Portugal, a Lei n.º 85/2009, de 27 de agosto, estabeleceu o alargamento da idade de cumprimento da escolaridade obrigatória até aos 18 anos e consagrou a universalidade da educação pré-escolar para as crianças a partir dos 5 anos de idade.

² De referir que “capital”, na teoria de Bourdieu (1986), é sinónimo de poder, consistindo em ativos económicos, culturais e sociais que se reproduzem e promovem a mobilidade social.

- 7) Os alunos que frequentam escolas ditas “muito seguras e organizadas” estão mais representados em meios socioeconómicos favorecidos. Neste aspeto, o indicador “Problemas de Disciplina” é um bom preditor dos desempenhos em leitura.

Os estudos ICILS colocam em evidência que nascer num mundo digital não significa necessariamente que alguém seja digitalmente competente. Ao contrário da visão comum, de que a geração jovem de hoje é uma geração de ‘nativos digitais’, as descobertas dos primeiros dois ciclos do ICILS indicam que os jovens não desenvolvem habilidades digitais sofisticadas. Cresce apenas o uso digital de dispositivos. Por outro lado, existe uma grande variação entre os países na conquista da alfabetização informacional. O foco não deve estar somente nos jovens com baixos recursos socioeconómicos, mas nos que têm níveis de proficiência superior em competência digital. Há também diferenciação de género na utilização das TIC. As raparigas têm desempenhos superiores do que os rapazes na CIL, mas essa diferenciação é menos evidente na avaliação CT. Os resultados do ICILS sugerem, igualmente, a necessidade de uma abordagem holística no uso pedagógico das TIC nas escolas. Fornecer aos alunos e aos professores equipamentos de TIC não é suficiente para melhorar as suas habilidades digitais. Eles devem ser encorajados e apoiados no uso de ferramentas digitais.

Conclusão

Nas últimas décadas a avaliação em larga escala adquiriu uma grande importância no âmbito educacional. Os principais objetivos das ILSA, especialmente aqueles empreendidos no contexto escolar, são o de procurar melhorar a qualidade e a equidade da educação, bem como atender à crescente procura mundial relativamente aos investimentos feitos na oferta educativa. Em geral, os ILSA compartilham objetivos comuns que explicitam ou implicitamente incluem um ou mais dos seguintes elementos:

- Fornecimento de dados de alta qualidade para melhorar a compreensão dos formuladores de políticas sobre os principais fatores escolares e não escolares que influenciam o ensino e a aprendizagem;
- Fornecimento de dados de alta qualidade como um recurso para identificar áreas de preocupação e ação e para preparar e avaliar reformas educativas;
- Desenvolvimento e melhoria da capacidade dos sistemas educacionais de se comprometerem em estratégias nacionais de monitorização e melhoria educacional.

Os ILSA podem ser divididos em duas categorias: as avaliações baseadas nos programas escolares e as avaliações baseadas nas competências. As primeiras avaliam em que medida os

alunos apreenderam o programa de ensino e as segundas avaliam em que medida os alunos os conseguem colocar em prática na vida real (Addey & Sellar, 2019). Vários estudos são implementados, levando a adesão de um número cada vez maior de países e regiões/economias. As avaliações em larga escala procuram contribuir para a melhoria da qualidade da educação, procurando aferir as competências e habilidades, mas também são ferramentas contínuas de trabalho, onde se identificam condições problemáticas. Identificados os constrangimentos, propõem-se novas possibilidades pedagógicas na escola. Os resultados dos testes indicam possíveis deficiências no processo que resultam no não domínio de certas competências e habilidades que devem ser desenvolvidas no âmbito da escola. Várias revistas e obras especializadas fazem eco destes estudos, insistindo sobre os seus objetivos, os métodos utilizados e os seus resultados. Os meios de comunicação divulgam os resultados através de *rankings*, mas oferece muito poucos esclarecimentos sobre os processos desses estudos.

Portugal tem vindo a participar em vários estudos internacionais de larga escala, permitindo obter informações sobre o sistema educativo e o contexto socioeconómico das famílias e pessoais dos alunos. Esses dados permitem influenciar na definição e implementação de políticas educacionais. Foi possível verificar que os alunos portugueses têm vindo a apresentar resultados positivos, apesar das oscilações consoante os ciclos das avaliações internacionais, nesses estudos.

Agradecimentos

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

Referências

- Addey, C., & Sellar S. (2019). Cela en vaut-il la peine ? Raisons de participation (ou non) aux évaluations internationales à grande échelle des apprentissages. *Recherche et Prospective en Éducation*, n° 24. UNESCO. <https://fr.unesco.org/>
- Araújo, L., Costa, P., & Folgado, C. (2016). Avaliação da leitura: PIRLS 2011. In Fernando Azevedo e Ângela Balça (Coord.) (2016), *Leitura e educação literária* (pp. 15-30). Pactor – Edições de Ciências Sociais, Forenses e da Educação.
- Ávila, P. (2005). *A literacia de adultos: competências-chave na sociedade do conhecimento*. Tese de doutoramento em Sociologia. ISCTE.

- Beaton, A., Martin, M., Mullis, I., Gonzalez, E., Smith, T., & Kelly, D. (1996). *Science achievement in the middle school years*. Boston College Chestnut Hill, MA, USA and International Association for the Evaluation of Educational Achievement (IEA).
- Benavente, A. (Coord.), Rosa, A., Costa, A., & Ávila, P. (1996). *A literacia em Portugal – resultados de uma pesquisa extensiva e monográfica*. Fundação Calouste Gulbenkian – Serviço de Educação/Conselho Nacional de Educação.
- Blurton, C. (1999). New directions in education. In M. Tawfik (Org.) (1999). *The world communication and information* (pp. 46-61). UNESCO.
- Bodin, A., & Grapin, N. (2018). Un regard didactique sur les évaluations du PISA et de la TIMSS : mieux les comprendre pour mieux les exploiter. *Mesure et évaluation en éducation*, 41(1), 67-96. <https://doi.org/10.7202/1055897ar>
- Bourdieu, P. (1986). The forms of capital. In Richardson, J. (1986). *Handbook of theory and research for the sociology of education* (pp. 241-258). Westport, CT: Greenwood.
- Campbell, J., Kelly, D., Mullis, I., Martin, M., & Sainsbury, M. (2001). *Framework and specifications for PIRLS Assessment 2001: Progress in international reading literacy study*. Chestnut Hill: Boston College: IEA – International Association for the Evaluation of Educational Achievement.
- Carvalho, J., Amaro, G., Reis, P., & Neres, F. (1996). *Terceiro estudo internacional em matemática e ciências (TIMSS): semelhanças num contexto de diferenças*. Instituto de Inovação Educacional.
- D'Ambrósio, U. (2005). Sociedade, cultura, matemática e seu ensino. *Revista Educação e Pesquisa*, 31, 99-120.
- Drent, M., Martina, M., Meelissen, & Fabienne, K. (2013). The contribution of TIMSS to the link between school and classroom factors and student achievement. *Journal of Curriculum Studies*, 45(2), 198-224.
- European Commission/EACEA/Eurydice (2020). *Compulsory Education in Europe – 2020/21. Eurydice Facts and Figures*. Publications Office of the European Union.
- Félix, P., Perdigão, R., & Lourenço, V. (2020). *Desempenho e equidade: uma análise comparada a partir dos estudos internacionais TIMSS e PIRLS*. Conselho Nacional de Educação (CNE).
- Ferreira, A., & Gonçalves, C. (2013). *TIMSS & PIRLS 2011 – Relações entre os desempenhos em leitura, matemática e ciências, 4.º ano*. ProjAVI Grupo de Projeto para a Avaliação Internacional de Alunos.
- Fraillon, J., Ainley, J., Wolfram, S., Friedman, T., & Duckworth, D. (2020). *Preparing for life in a digital world: IEA International Computer and Information Literacy Study 2018 International Report*. Switzerland: Springer.

- Instituto de Avaliação Educativa (IAVE) (2017b). *Resultados Globais PIRLS 2016 - ePIRLS 2016 – Portugal. Literacia de leitura & literacia de leitura online*. Autor.
- Instituto de Avaliação Educativa (IAVE) (2018). *PIRLS 2016 - ePIRLS2016 – Literacia de leitura & literacia de leitura online. Unidades de avaliação*. Autor.
- Instituto de Avaliação Educativa (IAVE) (s./d.). *TIMSS 2011 – Trends in International Mathematics and Science Study – itens de matemática – 4.º ano, disponibilizados ao público*. Autor.
- Lafontaine, A. (2008). *PIRLS 2006. Progress in reading literacy study. Note de synthèse*. Université de Liège.
- Lietz, P., Cresswell, J. C., Rust, K. F., & Adams, R. J. (2017). Implementation of large-scale education assessments. In P. Lietz, J. C. Cresswell, K. F. Rust, & R. J. Adams (Eds.), *Wiley series in survey methodology. Implementation of large-scale education assessments* (pp. 1-25). John Wiley & Sons.
- 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.13768>
- Marôco, J. (Coord.), Lourenço, V., Mendes, R., & Gonçalves, C. (2016a). *TIMSS Advanced 2015 Portugal - Volume I, Desempenhos em matemática e em física*. IAVE.
- Marôco, J. (Coord.), Lourenço, V., Mendes, R., & Gonçalves, C. (2016b). *TIMSS 2015 – Portugal – Volume I, desempenhos em matemática e em ciências*. IAVE.
- Olsen, R., & Lie, S. (2006). Les évaluations internationales et la recherche en éducation : principaux objectifs et perspectives. *Revue française de pédagogie*, 157, 11-26. <https://doi.org/10.4000/rfp.393>
- Rémond, M. (2006). Éclairages des évaluations internationales PIRLS et PISA sur les élèves français. *Revue française de pédagogie*, 157, 71-84. <https://doi.org/10.4000/rfp.433>
- Rémond, M. (2007). Que nous apprend PIRLS sur la compréhension des élèves français de 10 ans ? *Repères, recherches en didactique du français langue maternelle*, 35, 53-72.
- Ricoy, M., & Couto, M. (2012). Os recursos educativos e a utilização das TIC no ensino secundário na matemática. *Revista de Educação Portuguesa*, 25(2), 241-262.
- Rosa, V., Maia, J., Daniela, M., & Teodoro, A. (2020). PISA, TIMSS e PIRLS em Portugal: uma análise comparativa. *Revista Portuguesa de Educação*, 33(1), 94-120. <https://doi.org/10.21814/rpe.18380>
- Rosa, V. (2020). A participação de Portugal no estudo ICILS. *EDUSER: revista de educação*, 12(2), 1-16.

- Rosine, L., & Postlethwaite, N. (1994). Les études internationales de l'IEA. *Revue internationale d'éducation de Sèvres*. <https://doi.org/10.4000/ries.4294>
- Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. 2010. International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142-145. <https://doi.org/10.3102/0013189X10363170>
- Schmidt, W., MacKnight, C., Valverde, G., Houang, R., & Wiley, D. (Eds.) (1997). *Many visions, many aims: a cross-national investigation of curricular intentions in school mathematics, volume 1*. Kluwer Academic Publishers.
- Spanhel, D. (2008). La importancia de las nuevas tecnologías en el sector educativo. In M. L. Sevillano (Coord.) (2008). *Nuevas tecnologías en Educación Social* (pp. 29-52). McGraw-Hill: Madrid.
- Vanda, L., Nunes, A., Amaral, A., Gonçalves, C., Mota, M., & Mendes, R. (2019). *ICILS 2018 – PORTUGAL. Literacia em Tecnologias da Informação e da Comunicação*. IAVE.
- Wagemaker, H. (2014). International large-scale assessments: From research to policy. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Statistics in the social and behavioral sciences series. Handbook of international large-scale assessment. Background, technical issues, and methods of data analysis* (pp. 11-36). Boca Raton: CRC Press.