

Program Międzynarodowej Oceny Umiejętności Uczniów
PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT

PISA

Wyniki badania 2018 w Polsce



Międzynarodowe konsorcjum realizujące badanie OECD PISA 2018:

Educational Testing Servis ETS (USA)
Westat (USA)
Pearson (UK)
cApStAn Linguistic Quality Control (Belgia)

Rada Zarządzająca Programem PISA (PGB):

Przedstawiciel Polski: Jerzy Wiśniewski, Piotr Mikiewicz

Skład zespołu realizującego badanie PISA 2018 w Polsce:

Krzysztof Bulkowski
Jacek Haman
Joanna Kaźmierczak
E. Barbara Ostrowska
Michał Sitek

Eksperti współpracujący z zespołem IBE:

Mirosława Badiak
Krzysztof Biedrzycki
Jan Burski
Zbigniew Marciniak
Krzysztof Spalik
Agnieszka Sułowska

Osoby wspierające realizację badania:

Dorota Cyngot, Grażyna Drążyk, Katarzyna Jankowska, Dominik Mytkowski, Małgorzata Samsonowska-Kreczmer, Jan Urmański.

Raport z badania opracowali:

Krzysztof Biedrzycki, Krzysztof Bulkowski, Jan Burski, Wioleta Dobosz-Leszczyńska, Jacek Haman, Zbigniew Marciniak, E. Barbara Ostrowska, Michał Sitek, Krzysztof Spalik, Agnieszka Sułowska

Redakcja naukowa: Michał Sitek**Redakcja językowa:** zespół PISA

Wzór cytowania: M.Sitek (red.). *Program Międzynarodowej Oceny Umiejętności Uczniów. Wyniki badania PISA*

2018 w Polsce. Warszawa: Instytut Badań Edukacyjnych, 2019.

Zespół dziękuje Dyrektorom szkół i Kuratoriom Oświaty za życzliwą pomoc w realizacji badania. Dziękujemy także uczniom za udział w badaniu i ich rodzicom za wyrażenie na to zgody.

Za realizację badania w Polsce odpowiadał Instytut Badań Edukacyjnych.

Badanie w szkołach zrealizowała firma PBS Spółka z.o.o, ul. Junaków 2, 81-812 Sopot.

Strony internetowe badania PISA w Polsce: www.ibe.edu.pl, www.pisa.ibe.edu.pl, www.oecd.org

Badanie PISA 2018 sfinansowało Ministerstwo Edukacji Narodowej.

Spis treści

1. Populacja i próba w badaniu PISA 2018	4
O badaniu PISA	
Populacja i próba w badaniu PISA 2018 – najważniejsze informacje	
Pomiar i skalowanie	
2. Rozumienie czytanego tekstu	8
Wyniki polskich uczniów na tle innych krajów	
Zmiany wyników w latach 2000 – 2018	
Poziomy umiejętności uczniów	
Zakresy umiejętności	
Różnice między wynikami chłopców i dziewcząt	
Podsumowanie	
3. Matematyka	12
Umiejętności matematyczne polskich uczniów na tle innych krajów	
Zmiany wyników w latach 2003– 2018	
Poziomy umiejętności uczniów	
Różnice między wynikami chłopców i dziewcząt	
Podsumowanie	
4. Rozumowanie w naukach przyrodniczych	16
Wyniki polskich uczniów na tle innych krajów	
Zmiany wyników w latach 2006 – 2018	
Poziomy umiejętności uczniów	
Różnice między wynikami chłopców i dziewcząt	
Podsumowanie	

Program Międzynarodowej Oceny Umiejętności Uczniów PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT



Wyniki badania 2018 w Polsce

O badaniu PISA

Badanie PISA (Programme for International Student Assessment, Program Międzynarodowej Oceny Umiejętności Uczniów) jest organizowane przez Organizację Współpracy Gospodarczej i Rozwoju (OECD) od 2000 roku. W 2018 r. uczestniczyło w nim 79 krajów i regionów. PISA to jeden z najważniejszych edukacyjnych programów badawczych. Jego wyniki stanowią istotny punkt odniesienia dla polityki oświatowej i postrzegania jakości edukacji szkolnej w poszczególnych krajach.

Badanie jest przeprowadzane co trzy lata. Starannie dobrane zadania odnoszą się do różnych aspektów wiedzy i umiejętności z trzech dziedzin. Przedmiotem pomiaru w badaniu PISA jest: rozumienie czytanego tekstu (*reading literacy*), rozumowanie matematyczne (*mathematics literacy*) i rozumowanie w naukach przyrodniczych (*science literacy*). Każda z tych dziedzin co dziewięć lat stanowi główną dziedzinę badania, co oznacza, że uczniowie rozwiązują większą liczbę zadań z tego obszaru, a to pozwala na uzyskanie pogłębionego obrazu wiedzy i umiejętności z danej dziedziny. Część krajów uczestniczy też w dodatkowych modułach badania: od roku 2012 badanie PISA obejmuje również umiejętności finansowe (*financial literacy*) – wyniki tej części badania z 2018 roku zostaną ogłoszone w 2020 roku.

Polska uczestniczy w badaniu PISA od samego początku. Porównywanie wyników rozumienia czytanego tekstu możliwe jest od 2000 r. (gdy po raz pierwszy było ono główną dziedziną pomiaru), od roku 2003 można porównywać wyniki rozumowania matematycznego, a od 2006 – rozumowania w naukach przyrodniczych. Rozumienie czytanego tekstu ponownie było dziedziną główną w 2009 roku, a także – w najnowszej edycji badania – w 2018 roku. Rozumowanie matematyczne było dziedziną główną w latach 2003 i 2012, a rozumowanie w naukach przyrodniczych – w latach 2006 i 2015.

Badanie PISA umożliwia wnioskowanie o poziomie umiejętności piętnastolatków i jest prowadzone na reprezentatywnej próbie 15-latków, losowanych spośród uczniów różnych typów szkół. Zapewnia to porównywalność między krajami, ponieważ dobór próby nie zależy od różnic w wieku rozpoczynania nauki w szkole czy struktury szkół. W podstawowej części badania, obejmującej trzy główne dziedziny badania PISA 2018, Polskę reprezentowało 5653 uczniów z 239 szkół.

Polskie wersje zadań oraz ankiet wypełnianych przez uczniów i dyrektorów szkół, schemat losowania szkół i uczniów, ocenianie odpowiedzi i przygotowanie zbioru danych było zadaniem krajowego zespołu badania w Instytucie Badań Edukacyjnych. Było ono prowadzone

w ścisłej współpracy z międzynarodowym konsorcjum badawczym odpowiedzialnym za przygotowanie i przeprowadzenie badania. Realizacja badania w terenie odbyła się w okresie 1 marca – 13 kwietnia 2018 r. Przeprowadziła je, we współpracy z zespołem projektu PISA w IBE, agencja badawcza PBS.

Badanie w 2018 roku odbyło się w trakcie wdrażania reformy systemu szkolnictwa w Polsce. Zdecydowaną większość badanych stanowił przedostatni rocznik uczniów gimnazjum, którzy uczyli się według starej podstawy programowej (97,5% badanych było uczniami trzeciej klasy gimnazjum). Tylko nieliczni piętnastolatki uczyli się w szkołach podstawowych lub w szkołach ponadgimnazjalnych. Badanie jest więc jednym z ostatnich portretów stanu edukacji przed reformą. Efekty zmian w strukturze szkół oraz zmian programowych będą widoczne dopiero w kolejnym cyklu badania PISA, czyli w 2021 r. Badanie obejmie wtedy, podobnie jak to miało miejsce w 2000 roku, przede wszystkim uczniów pierwszych klas szkół ponadpodstawowych: liceów, techników i szkół branżowych I stopnia.

Populacja i próba w badaniu PISA 2018 – najważniejsze informacje

Zasady doboru próby do badania PISA są jednym z najtrwałszych jego elementów i pozostają praktycznie niezmiennione od pierwszej edycji badania z roku 2000. Z drugiej strony, istotny wpływ na procedury doboru próby w polskich edycjach badania PISA miały zmiany zachodzące w polskim systemie edukacyjnym w latach 2000–2018 (i zachodzące dalej – co jednak będzie miało znaczenie już dla kolejnych edycji badania). Polskie badanie PISA 2000 prowadzone było w szkołach ponadpodstawowych (liceach, technikach i zasadniczych szkołach zawodowych). W kolejnych edycjach, realizowanych w latach 2003–2015, już po wprowadzeniu gimnazjów, brali udział prawie wyłącznie gimnazjaliści („uzupełniani” przez niewielką grupę piętnastoletnich licealistów i pojedynczych uczniów szkół zawodowych). Badanie OECD PISA 2018 w Polsce realizowane było w początkowym okresie w trakcie wdrażania zmiany struktury szkół: przede wszystkim wśród gimnazjalistów z ostatnich dwóch roczników funkcjonowania tego typu szkół, z niewielkim udziałem uczniów szkół ponadgimnazjalnych i pojedynczych siódmoklasistów ze szkół podstawowych.

Badaniem OECD PISA objęci są uczniowie piętnastoletni, a dokładniej ci, którzy ukończyli 15 lat w roku poprzedzającym badanie; w przypadku PISA 2018 byli to uczniowie urodzeni w 2002 roku (znajdowali się więc wśród nich również tacy, którzy w chwili badania ukończyli już lat szesnaście). Operacyjna definicja „piętnastolatka” przyjmowana w bada-

Tabela 1. Piętnastolatki (uczniowie urodzeni w 2002 roku) w polskim systemie szkolnym (z wyłączeniem szkół specjalnych) w roku szkolnym 2017/18

Liczba piętnastoletnich gimnazjalistów ¹	319,4 tys.
Liczba piętnastoletnich uczniów liceów ogólnokształcących	2,8 tys.
Liczba piętnastoletnich uczniów średnich szkół zawodowych	1,4 tys.
Liczba piętnastoletnich uczniów szkół branżowych I stopnia	0,1 tys.
Liczba piętnastoletnich uczniów szkół podstawowych ²	4,0 tys.
Razem	327,8 tys.
Szacunkowa wielkość wyłączeń wewnątrzszkolnych	5,9 tys.
Łącznie wielkość badanej populacji	321,9 tys.
Liczba gimnazjów (bez szkół specjalnych)	6810

¹ Włączając ogólnokształcące szkoły artystyczne II stopnia.

² Do populacji badanej w PISA wśród uczniów szkół podstawowych należeli wyłącznie piętnastoletni uczniowie klas VII; brak jest danych pozwalających dokładnie określić ich liczbę. Jest ona jednak tylko nieznacznie niższa niż łączna liczba uczniów piętnastoletnich w szkołach podstawowych.

Tabela 2. Próba i jej realizacja

	Gimnazja	Licea ogólnokształcące	Technika	Szkoły branżowe I stopnia	Szkoły podstawowe	Razem
Liczba wylosowanych szkół (próba zasadnicza)	228	19	29	19	97	392
Liczba wylosowanych szkół, w których uczyli się uczniowie piętnastoletni	228	6	8	1	11	254
Liczba szkół z próby zasadniczej biorących udział w badaniu	209	6	3	0	3	221
Liczba szkół rezerwowych biorąca udział w badaniu	18	0	0	0	0	18
Łączna liczba szkół, w których zrealizowano badanie	227	6	3	0	3	239
Liczba wylosowanych uczniów	6769	13	23	1	9	6815
Liczba wylosowanych uczniów wyłączonych z próby	145	1	14	0	1	161
Próba po wyłączeniach (liczba uczniów)	6624	12	9	1	8	6654
Liczba uczniów biorąca udział w badaniu	5629	12	9	0	3	5653
Poziom realizacji próby	85,0%	100,0%	100,0%	0,0%	37,5%	85,0%

niu nie jest zatem w pełni zgodna z naturalną interpretacją słowa „piętnastolatek”). Zdecydowaną większość tej grupy stanowili w 2018 roku uczniowie III klas gimnazjum, ale należeli do niej także piętnastoletni uczniowie II klas gimnazjów, piętnastoletni uczniowie VII klas szkoły podstawowej i piętnastoletni uczniowie szkół ponadgimnazjalnych. W skład badanej populacji wchodził również, oczywiście, piętnastoletni uczniowie szkół artystycznych (przede wszystkim ogólnokształcących szkół muzycznych II stopnia). Z założenia z populacji badania wyłączeni byli natomiast uczniowie szkół specjalnych (wyłączenia na poziomie szkół) i piętnastoletni uczniowie poniżej VII klasy szkoły podstawowej, a także uczniowie, którzy nie mogli pisać testu ze względu na niepełnosprawność lub niewystarczającą znajomość języka polskiego. Dane o populacji badanej w PISA 2018 podaje tabela 1.

Badanie PISA realizowane jest na reprezentatywnej próbie losowej. Schemat doboru próby uczniów w badaniu PISA ma charakter dwustopniowego doboru warstwowego z zastosowaniem losowania systematycznego, w którym pierwszym stopniem doboru jest wybór szkoły, zaś drugim – losowanie uczniów z uprzednio wylosowanych szkół.

Do polskiej edycji badania OECD PISA 2018 wylosowano próbę 228 gimnazjów (w tej liczbie znalazła się jedna szkoła artystyczna; w jednej szkole badanie ostatecznie nie zostało zrealizowane), 19 liceów ogólnokształcących, 48 szkół zawodowych (29 techników i 19 szkół zasadniczych) oraz 97 szkół podstawowych; dodatkowo, dla każdej wylo-

sowanej szkoły, wylosowane zostały dwie szkoły rezerwowe, na wypadek gdyby wylosowana szkoła odmówiła udziału w badaniu. Spośród 67 wylosowanych szkół ponadgimnazjalnych do badania zostało zakwalifikowanych jedynie 15 szkół, w tym 6 liceów i 8 techników oraz jedna szkoła branżowa I stopnia (w pozostałych nie było żadnych uczniów piętnastoletnich); spośród wylosowanych szkół podstawowych obecność piętnastoletnich uczniów zadeklarowało 11. Ostatecznie liczba szkół podstawowych i zawodowych, których uczniowie zostali uwzględnieni w wynikach badania, była jeszcze mniejsza. Ponieważ w szkołach tych byli jedynie pojedynczy piętnastolatki, brak zgody rodziców lub po prostu nieobecność jedynego piętnastolatka w szkole powodowały, że z badania „wypadła” cała szkoła. Dane o wylosowanej i zrealizowanej próbie zawarte są w tabeli 2.

Do badania PISA (bez badania umiejętności finansowych przeprowadzonego na osobnej próbie uczniów ze szkół uczestniczących w badaniu) wylosowanych zostało pierwotnie łącznie 6815 uczniów, z czego przed badaniem wyłączonych zostało – jako nienależących do badanej populacji¹ – 161. Po tych wyłączeniach wielkość próby wynosiła 6654 uczniów, w tym 6624 uczniów gimnazjów, 12 uczniów liceów ogólnokształcących, 10 uczniów techników i 8 uczniów szkół podstawowych. Spośród nich w badaniu nie uczestniczyło – ze względu na brak zgody rodziców lub nieobecność – 1001 uczniów. Ostatecznie liczba uczniów uczestniczących w badaniu i uwzględniona w bazie danych wyniosła 5653. Oznacza to realizację wylosowanej próby na poziomie

¹ Były to przede wszystkim wyłączenia ze względu na nieznajomość języka (22 osoby), niepełnosprawność znacząco utrudniającą wypełnienie testu (109 osób), wylosowanie osób, które w chwili badania nie były już uczniami danej szkoły (20 przypadków); ponadto 10 osób było pierwotnie błędnie wpisane na listy, pomimo że nie spełniały kryterium wieku.

85%, a więc podobnym jak w poprzednich edycjach (87% w 2015 i 86% w 2012 roku) i wyraźnie powyżej poziomu wymaganego przez konsorcjum międzynarodowe.

Pomiar i skalowanie²

Wyniki badania PISA – a więc oszacowania kompetencji uczniów w poszczególnych dziedzinach – są wyliczane na podstawie złożonej procedury skalowania, uwzględniającej liczbę poprawnie rozwiązanych zadań oraz ich trudność. Zastosowanie tych procedur daje szereg istotnych korzyści, w porównaniu do oceniania kompetencji przez np. zwykłe zliczanie odsetka poprawnie rozwiązanych zadań. W szczególności możliwe jest ocenianie na tej samej skali (a więc zapewnienie porównywalności wyników) uczniów, którzy rozwiązywali częściowo różne zestawy zadań testowych (a w pewnym zakresie – nawet uczniów, którzy rozwiązywali całkowicie różne zestawy zadań), bez przyjmowania trudnego do spełnienia założenia o identycznym poziomie ich trudności. Dzięki temu w badaniu PISA możliwe jest wykorzystanie znacznie większej liczby różnorodnych zadań, niż gdyby wszyscy uczniowie mieli wykonywać te same zestawy testowe; możliwe jest również porównywanie wyników różnych edycji badania, pomimo że jedynie część zadań (tzw. zadania kotwiczące) powtarzana jest w kolejnych edycjach.

- W badaniu PISA skalowanie wyników testu opiera się na teorii odpowiedzi na pytanie testowe (IRT, *Item Response Theory*), a ściślej – na uogólnionym modelu Rascha³. Koncepcja ta odwołuje się do następujących założeń:
 - To, czy dany uczeń rozwiąże prawidłowo dane zadanie, jest zdarzeniem losowym.
 - Prawdopodobieństwo zajścia tego zdarzenia determinowane jest przez dwa czynniki:
 - poziom umiejętności ucznia,
 - poziom trudności zadania⁴.

Zakłada się przy tym określoną postać funkcji wiążącej prawdopodobieństwo rozwiązania zadania o danej trudności z poziomem umiejętności ucznia. W modelu Rascha jest to, zasadniczo, funkcja logistyczna; poszczególne warianty modeli IRT różnią się od siebie głównie uwzględnieniem pewnych dodatkowych parametrów tej funkcji – obok samej trudności zadania także jego mocy dyskryminacyjnej⁵, a w przypadku zadań zamkniętych – prawdopodobieństwo udzielenia poprawnej odpowiedzi poprzez zgadywanie. Zwyczajowo określa się poziom trudności zadania i poziom umiejętności badanego na tej samej skali, przyjmując, że badany o poziomie kompetencji k rozwiąże zadania o trudności k z prawdopodobieństwem równym $\frac{1}{2}$.

- Zarówno poziom umiejętności poszczególnych badanych, jak i poziom trudności poszczególnych zadań

(i ewentualnie ich moc dyskryminacyjna) traktowane są jako zmienne ukryte (latentne) – ich oszacowanie jest celem procesu skalowania.

W procesie skalowania jednocześnie szacowane są poziomy trudności zadań oraz kompetencje badanych – polega to na poszukiwaniu (za pomocą przede wszystkim procedur iteracyjnych) takich kombinacji ich wartości, które z największym prawdopodobieństwem prowadzą do uzyskania zaobserwowanych wyników badania (estymacja metodami „największej wiarygodności”). Drugim elementem procesu skalowania jest ocena zgodności założeń modelu z danymi obserwowanymi. Przykładowo, może okazać się, iż bardzo trudno jest utrzymać założenie, że szanse na rozwiązanie danego zadania wynikają z poziomu tej samej umiejętności, która odpowiada za pozostałe badania. W takiej sytuacji może okazać się, że trafniejsze wyniki uzyskamy, pomijając w analizie dane odnoszące się do tego zadania.

Estymacja trudności zadań może być dokonywana na całości danych z badania, możliwe jest jednak także wykorzystanie do oceny umiejętności badanych danych o poziomie trudności zadań oszacowanych uprzednio. Możliwość ta wykorzystywana jest na kilka sposobów; w szczególności:

- Użycie „zadań kotwiczących” o trudności oszacowanej już w poprzednich cyklach badania PISA pozwala zakotwiczyć skale PISA względem wcześniejszych edycji badania, a tym samym osiągnąć porównywalność i współmierność wyników kolejnych cykli PISA. Osiągnięcie tego efektu wymaga jednak wyskalowania „zadań kotwiczących” na odpowiednio bogatym materiale – z tego względu pełna porównywalność wyników kolejnych edycji badania PISA dla danej dziedziny możliwa jest jedynie od momentu, gdy dana dziedzina była głównym przedmiotem edycji (jak np. czytanie i interpretacja w edycji PISA 2000, matematyka – PISA 2003, rozumowanie w naukach przyrodniczych – PISA 2006); porównywanie wyników wcześniejszych edycji wiąże się z większym ryzykiem błędów losowych.
- Skalowanie trudności zadań odbywa się wyłącznie z użyciem wyników pochodzących z podstawowej populacji badanych, a więc populacji piętnastolatków. W badaniach uzupełniających projekt międzynarodowy – jak w prowadzonych w części poprzednich edycji polskich badaniach uczniów szkół ponadgimnazjalnych – wykorzystywane są trudności zadań oszacowane w międzynarodowej części badania. W ten sposób jednocześnie osiągnane są dwa cele: ocena umiejętności w „dodatkowych populacjach” na tych samych skalach co w przypadku piętnastolatków, a jednocześnie odseparowanie podstawowego badania międzynarodowego od dodatkowych elementów badania specyficznego dla poszczególnych krajów.

- Możliwe jest szacowanie parametrów zadań z wykorzystaniem jedynie części badanej próby, a następnie skalowanie umiejętności wszystkich badanych uczniów w oparciu o tak uzyskane parametry zadań. Metoda ta wykorzystywana była w poprzednich edycjach badania PISA ze względu na mniejsze wymagania co do mocy obliczeniowej.

Istotną korzyścią z zastosowania modeli IRT jest również możliwość oceny na tej samej skali badanych, którzy wykonywali częściowo różne zestawy zadań. W ten sposób możliwe jest wykorzystanie w badaniu znacznie większej liczby zadań, a więc zbadanie znacznie szerszego spektrum podobszarów poszczególnych umiejętności. Co więcej – możliwe jest celowe przydzielanie poszczególnym uczniom zadań o różnym, dostosowanym do ich poziomu umiejętności, poziomie trudności (testowanie adaptatywne). Wdrażanie takiego rozwiązania rozpoczęte zostało w bieżącej edycji badania.

Skale (umiejętności badanych i trudności zadań) w modelu Rascha mają charakter skal przedziałowych. Pozwalają zatem na interpretowanie i porównywanie wielkości różnic między poszczególnymi wynikami (np. między średnimi dla krajów, średnimi dla typów szkół, wynikami poszczególnych badanych). Skale te nie mają jednak obiektywnego punktu zerowego – a zatem nie jest możliwe określenie proporcji między wynikami. Tak więc, przykładowo, bezsensowne byłoby stwierdzenie, że „kraj A uzyskał wyniki o 20% lepsze od kraju B”. Jednocześnie poziom umiejętności wyrażony w punktach PISA ma charakter relatywny (i nie odnosi się do żadnych obiektywnie zdefiniowanych oczekiwań co do tego, co powinni wiedzieć lub umieć badani). Skale skonstruowane są w ten sposób, by wartość 500 punktów odpowiadała średniej wyników krajów OECD w badaniu PISA 2000 oraz by 1 punkt odpowiadał jednej setnej odchylenia standardowego wyników w populacji krajów OECD w badaniu PISA 2000 (choć ze względu na dokładności oszacowania, wsteczną porównywalność wyników PISA należy ograniczyć do edycji, w której dana dziedzina była wiodąca).

Probabilistyczny charakter Item Response Theory oznacza także, że przy interpretacji wyników badania bierze się pod uwagę, że dwoje uczniów o tym samym rzeczywistym poziomie umiejętności może uzyskać w teście różne wyniki, i *vice versa*, dwie osoby, które uzyskały taki sam wynik, mogą w rzeczywistości mieć umiejętności na różnych poziomach. Innymi słowy, probabilistyczny charakter odpowiedzi na bodziec testowy jest drugim, obok reprezentatywnego charakteru badania (błędy związane z próbą), źródłem błędów losowych w wynikach badania PISA. Sposobem uwzględniania tych błędów w analizie jest wykorzystanie do szacowania poziomów umiejętności uczniów estymatorów „wartości prawdopodobnych” (*plausible values*, PV).

Realizacja badania PISA od edycji 2015 wyłącznie za pomocą narzędzi komputerowych umożliwiła w badaniu PISA 2018 wprowadzenie kolejnej nowości metodologicznej – testowania adaptatywnego, a więc dostosowania poziomu trudności pytań do szacowanego poziomu umiejętności ucznia. Testowanie adaptatywne zostało zastosowane jedynie dla wiodącej dziedziny – rozumienia czytanego tekstu. W tym celu każdemu badanemu przypisanych było początkowo pięć podzestawów zadań: podzestaw startowy (core, 7–10 zadań), dwa warianty („łatwy” i „trudny”) podzestawu etapu I (stage I, 12–15 zadań) oraz dwa warianty („łatwy” i „trudny”) podzestawu etapu II (stage II, 12–15 zadań). Z tych podzestawów uczeń ostatecznie otrzymywał do rozwiązania trzy – startowy, jeden z podzestawów etapu I oraz jeden z etapu II. Statystyczną korzyścią z testowania adaptatywnego jest zmniejszenie błędu losowego oszacowania poziomu umiejętności ucznia w porównaniu do nieadaptatywnego testu o tej samej liczbie zadań. Test staje się także bardziej przyjazny dla ucznia – w szczególności pozwala to uniknąć sytuacji, w której uczeń musi rozwiązać wiele zadań, które są dla niego zbyt trudne. Ma to szczególne znaczenie w krajach, w których wielu uczniów osiąga bardzo słabe wyniki.

Badanie przeprowadził oraz raport przygotował zespół w składzie:

Krzysztof Biedrzycki, Krzysztof Bulkowski, Jan Burski, Wioleta Dobosz-Leszczynska, Jacek Haman, Joanna Kaźmierczak, Zbigniew Marciniak, E. Barbara Ostrowska, Michał Sitek, Krzysztof Spalik, Agnieszka Sułowska.

Więcej informacji o badaniu PISA można znaleźć na stronie internetowej www.pisa.ibe.edu.pl

Badanie zostało sfinansowane ze środków Ministerstwa Edukacji Narodowej.

² Tekst z tej sekcji zawiera fragmenty powtórzone bezpośrednio lub z pewnymi modyfikacjami za polskim raportami z badania PISA 2009, 2012 i 2015.

³ Szersze omówienie modeli IRT w języku polskim znaleźć można w książkach: Jakubowski, M. i Pokropek A. (2009) *Badając egzaminy: Podejście ilościowe w badaniach edukacyjnych*. Centralna Komisja Egzaminacyjna oraz Pokropek, A. (2015). *Modele cech ukrytych w psychologii, socjologii i badaniach edukacyjnych. Teoria i zastosowania*. Warszawa: Instytut Badań Edukacyjnych.

⁴ Określenia „poziom umiejętności” i „poziom trudności” odpowiadają podstawowemu zastosowaniu modelu Rascha – pomiaru kompetencji. W innych jego zastosowaniach mówilibyśmy raczej o „natężeniu cechy” oraz o charakterystyce konkretnego jej wskaźnika. Model Rascha w badaniu PISA używany jest – poza samym badaniem kompetencji – do konstruowania skal opracowanych z danych z ankiet wypełnianych przez uczniów.

⁵ „Moc dyskryminacyjna” zadania określa, jak silna jest zależność prawdopodobieństwa rozwiązania zadania od poziomu umiejętności ucznia: im jest ona wyższa, tym szybciej wzrasta prawdopodobieństwo rozwiązania wraz ze wzrostem umiejętności, natomiast w przypadku zadań o słabej mocy dyskryminacyjnej różnica szans na rozwiązanie zadania między uczniami o wysokich i niskich kompetencjach może być niewielka.

Program Międzynarodowej Oceny Umiejętności Uczniów PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT



Wyniki badania 2018 w Polsce

Rozumienie czytanej tekstu

Podstawową umiejętnością niezbędną w uczeniu się i poznawaniu świata jest czytanie. Zrozumienie tekstu na różnych poziomach znaczeń służy nie tylko dotarciu do poszukiwanej wiedzy, ale też jej przyswojeniu. Badanie PISA skupia się na sprawdzeniu opanowania umiejętności rozumienia czytanej tekstu. Zdefiniowano ją w następujący sposób: „rozumienie czytanej tekstu to proces obejmujący odczytanie znaczeń, wykorzystanie tekstu, ocenę, refleksję i takie zaangażowanie w lekturę, by czytelnik osiągnął postawiony przed nią cel, pogłębił wiedzę, zwiększył własny potencjał intelektualny i uczestniczył w życiu społecznym”. Badanie PISA obejmuje wszystkie te aspekty, zarówno proces, jak i efekt czytania. Umiejętności sprawdzane są w trzech zakresach: odnajdywanie informacji, rozumienie całości przekazu, ocena tekstu powiązana z refleksją, jaką on uruchamia.

Pierwszy zakres rozumienia to **odnajdywanie informacji**. Chodzi jednak nie tylko o docieranie do prostych wiadomości (co może się dokonywać nawet bez zrozumienia całej wypowiedzi), ale o ich wydobywanie z różnych warstw tekstu, a nawet (zwłaszcza w przypadku medium elektronicznego) z różnych tekstów, które czytelnik może samodzielnie łączyć w jeden hipertekst. Drugi zakres to **rozumienie całości przekazu**, zarówno jego literalnego sensu (właściwe odczytanie zdań, akapitów, a także całej wypowiedzi), jak i znaczenia, które rodzi się w procesie integracji wszystkich poziomów znaczeniowych, a następnie w powiązaniu przesłania zawartego w samym tekście z uprzednią wiedzą oraz przekonaniem czytelnika. W trzecim zakresie pojawiają się **ocena i refleksja**. Ocena tekstu dotyczy jego wiarygodności, co wiąże się z ustosunkowaniem się do źródła informacji i kompetencji autora, a także krytyczną analizą sposobu argumentowania. Refleksja obejmuje zagadnienia poruszane w tekście, ale też to, jak forma tekstu służy celowi wypowiedzi (np. perswazji, przekazywaniu informacji itd.).

W 2018 roku w badaniu wykorzystano 245 zadań, spośród których 72 były użyte też w poprzednich edycjach (tzw. zadania „kotwiczące”), dzięki czemu możliwe jest porównywanie zmian wyników w czasie. W 2018 roku, podobnie jak w 2015, papierowe zeszyty testowe zostały zastąpione systemem komputerowym. Wiązało się to nie tylko ze zmianą nośnika, ale też z wprowadzeniem wielu zadań interaktywnych i zadań trafniej odzwierciedlających czytanie i posługiwanie się tekstem w mediach elektronicznych, które współcześnie są oczywistym, a dla młodych ludzi dominującym, środowiskiem aktywności związanych z czytaniem.

Zadania różnią się stopniem i rodzajem wyzwania, które jest postawione przed uczniem. Zostały one podzielone na

sześć kategorii, obrazujących kolejne poziomy kompetencji mierzonych w badaniu PISA. Wykorzystanie wielu zadań, od najłatwiejszych do bardzo trudnych, umożliwia odtworzenie zróżnicowania umiejętności piętnastolatków w każdym z krajów biorących udział w badaniu. Pozwala także na określenie stopnia opanowania przez piętnastolatków w całym kraju umiejętności sprawdzanych w badaniu.

Wyniki uczniów są przedstawione na standaryzowanej skali, dla której punktem odniesienia jest średnia uzyskana przez uczniów z krajów należących do Organizacji Współpracy Gospodarczej i Rozwoju (OECD) w 2000 r. Przyjęto dla niej wartość 500 i odchylenie standardowe 100. Na tej samej skali określono także trudność zadań, a tym samym wymagania stawiane uczniom.

Dla dopełnienia obrazu prezentującego osiągnięcia uczniów w kraju, oprócz ogólnego wyniku dla całej populacji, ważne jest ukazanie danych dotyczących odsetka uczniów, którzy zostali zaklasyfikowani do poziomów skrajnych – najniższych i najwyższych. Umiejętności uczniów na poziomach niższych niż poziom drugi są niewystarczające do skutecznej nauki i docierania do informacji, a w konsekwencji – do sprawnego funkcjonowania w społeczeństwie. Według unijnych zaleceń państwa członkowskie powinny dokładać starań, żeby do 2020 r. odsetek uczniów uzyskujących wyniki poniżej poziomu 2 nie przekraczał 15%. Uczniowie, których umiejętności kwalifikują ich na poziomy 5 i 6, to potencjalna kadra o najwyższych kompetencjach, od której w dużej mierze będzie zależeć w przyszłości rozwój nauki, techniki, ekonomii, kultury. Odsetek uczniów na najniższych poziomach wskazuje na to, w jakim stopniu system edukacyjny pozwala na zdobycie podstawowych umiejętności niezbędnych do funkcjonowania we współczesnym świecie, natomiast odsetek uczniów na najwyższych poziomach mówi o tym, jak w systemie wykorzystuje się zdolności uczniów i ich możliwości rozwoju intelektualnego.

Wyniki polskich uczniów na tle innych krajów

W 2018 roku średni wynik polskich uczniów w dziedzinie rozumienia czytanej tekstu wyniósł 512 punktów. Był to jeden z najwyższych wyników w skali światowej. Znacząco lepsze wyniki uzyskali tylko uczniowie z krajów azjatyckich – Chin (z wybranych regionów) i Singapuru, a także uczniowie z Estonii, Kanady i Finlandii. Wynik polskich piętnastolatków był zbliżony do wyników uczniów z Irlandii, Korei Południowej, Szwecji, Nowej Zelandii i Stanów Zjednoczonych (różnice były nieistotne statystycznie).

W 2018 roku wśród krajów Unii Europejskiej lepsze wyniki uzyskali tylko uczniowie z trzech krajów – Estonii, Finlandii i Irlandii, przy czym istotnie statystycznie lepszy wynik uzyskały tylko Estonia i Finlandia.

Tabela 1. Wyniki uczniów w pomiarze rozumienia czytanej tekstu w badaniu PISA 2018.

L.p.	Kraj	Średnia (błąd standardowy)	Istotność	OECD	UE
1	Chiny B-S-J-G	555 (2,7)	↑		
2	Singapur	549 (1,6)	↑		
3	Makao (Chiny)	525 (1,2)	↑		
4	Hongkong (Chiny)	524 (2,7)	↑		
5	Estonia	523 (1,8)	↑	OECD	UE
6	Kanada	520 (1,8)	↑	OECD	
7	Finlandia	520 (2,3)	↑	OECD	UE
8	Irlandia	518 (2,2)	↑	OECD	UE
9	Korea Południowa	514 (2,9)	↑	OECD	
10	Polska	512 (2,7)	↑	OECD	UE
11	Szwecja	506 (3,0)	↑	OECD	UE
12	Nowa Zelandia	506 (2,0)	↑	OECD	
13	Stany Zjednoczone	505 (3,6)	↑	OECD	
14	Wielka Brytania	504 (2,6)	↑	OECD	UE
15	Japonia	504 (2,7)	↑	OECD	
16	Australia	503 (1,6)	↑	OECD	
17	Tajwan	503 (2,8)	↑		
18	Dania	501 (1,8)	↑	OECD	UE
19	Norwegia	499 (2,2)	↑	OECD	
20	Niemcy	498 (3,0)	↑	OECD	UE
21	Słowenia	495 (1,2)	↑	OECD	UE
22	Belgia	493 (2,3)	↑	OECD	UE
23	Francja	493 (2,3)	↑	OECD	UE
24	Portugalia	492 (2,4)	↑	OECD	UE
25	Czechy	490 (2,5)		OECD	UE
26	Holandia	485 (2,7)		OECD	UE
27	Austria	484 (2,7)		OECD	UE
28	Szwajcaria	484 (3,1)		OECD	
29	Chorwacja	479 (2,7)	↓		UE
30	Łotwa	479 (1,6)	↓	OECD	UE
31	Rosja	479 (3,1)	↓		
32	Włochy	476 (2,4)	↓	OECD	UE
33	Węgry	476 (2,3)	↓	OECD	UE
34	Litwa	476 (1,5)	↓	OECD	UE
35	Islandia	474 (1,7)	↓	OECD	
36	Białoruś	474 (2,4)	↓		
37	Izrael	470 (3,7)	↓	OECD	
38	Luksemburg	470 (1,1)	↓	OECD	UE
39	Ukraina	466 (3,5)	↓		
40	Turcja	466 (2,2)	↓	OECD	
41	Słowacja	458 (2,2)	↓	OECD	UE
42	Grecja	457 (3,6)	↓	OECD	UE
43	Chile	452 (2,6)	↓	OECD	
44	Malta	448 (1,7)	↓		UE
45	Serbia	439 (3,3)	↓		
46	ZEA	432 (2,3)	↓		
47	Rumunia	428 (5,1)	↓		UE
48	Urugwaj	427 (2,8)	↓		
49	Kostaryka	426 (3,4)	↓		
50	Cypr	424 (1,4)	↓		UE
51	Moldawia	424 (2,4)	↓		
52	Czarnogóra	421 (1,1)	↓		
53	Meksyk	420 (2,7)	↓	OECD	
54	Bułgaria	420 (3,9)	↓		UE
55	Jordania	419 (2,9)	↓		
56	Malezja	415 (2,9)	↓		
57	Brazylia	413 (2,1)	↓		
58	Kolumbia	412 (3,3)	↓		
59	Brunei	408 (0,9)	↓		
60	Katar	407 (0,8)	↓		
61	Albania	405 (1,9)	↓		
62	Bośnia i Hercegowina	403 (2,9)	↓		
63	Argentyna	402 (3,0)	↓		
64	Peru	401 (3,0)	↓		
65	Arabia Saudyjska	399 (3,0)	↓		
66	Tajlandia	393 (3,2)	↓		
67	Macedonia Północna	393 (1,1)	↓		
68	Baku (Azerbejdżan)	389 (2,5)	↓		
69	Kazachstan	387 (1,5)	↓		
70	Gruzja	380 (2,2)	↓		
71	Panama	377 (3,0)	↓		
72	Indonezja	371 (2,6)	↓		
73	Maroko	359 (3,1)	↓		
74	Liban	353 (4,3)	↓		
75	Kosowo	353 (1,1)	↓		
76	Dominikana	342 (2,9)	↓		
77	Filipiny	340 (3,3)	↓		

Kraje zaprezentowane w porządku malejącym ze względu na wynik średni. W nawiasie podano błąd standardowy. Szarym tłem wyróżnione są kraje, których średni wynik nie różni się statystycznie istotnie od średniego wyniku Polski. W odpowiednich kolumnach oznaczono kraje należące do Organizacji Współpracy Gospodarczej i Rozwoju (OECD) oraz Unii Europejskiej (UE). W powyższych wynikach nie uwzględniono Hiszpanii i Wietnamu, zgodnie z oficjalnym oświadczeniem zamieszczonym na stronie internetowej OECD.
↑ Wynik statystycznie istotnie powyżej średniej krajów OECD (kolumna Istotność) ↓ Wynik statystycznie istotnie poniżej średniej krajów OECD (kolumna Istotność)

Tabela 2. Zmiany wyników polskich uczniów w pomiarze rozumienia czytanej tekstu w badaniu PISA.

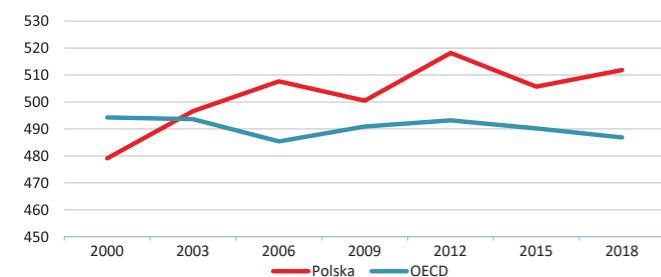
Rok badania PISA	Wynik punktowy polskich uczniów	Średni wynik krajów OECD	Liczba krajów lub regionów, które uzyskały wynik statystycznie istotnie lepszy od Polski	
			na świecie	w Europie
2000	479	494	16	10
2003	497	494	10	5
2006	508	485	6	2
2009	500	491	8	1
2012	518	493	5	0
2015	506	490	9	4
2018	512	487	7	2

Zmiany wyników w latach 2000–2018

Wynik osiągnięty przez polskich piętnastolatków w 2018 roku jest znacząco wyższy niż w poprzedniej edycji badania w 2015 roku, gdy wyniósł on 506 punktów, i niższy niż w 2012 roku (518 punktów). Jeśli jako punkt odniesienia przyjmie się wyniki od początku badania PISA w 2000 roku, a zwłaszcza wyniki

z lat, gdy rozumienie czytanej tekstu było dziedziną główną (w 2000 i 2009), zauważyć można, że utrzymywała się tendencja wzrostowa. Była ona zauważalna też na tle międzynarodowym. W roku 2000 wynik polskich uczniów był wyraźnie niższy niż średnia dla krajów OECD. W 2003 był bliski średniej, a od 2006 był powyżej średniej. Tym samym od 2006 roku Polska należała do grupy krajów o najwyższych wynikach.

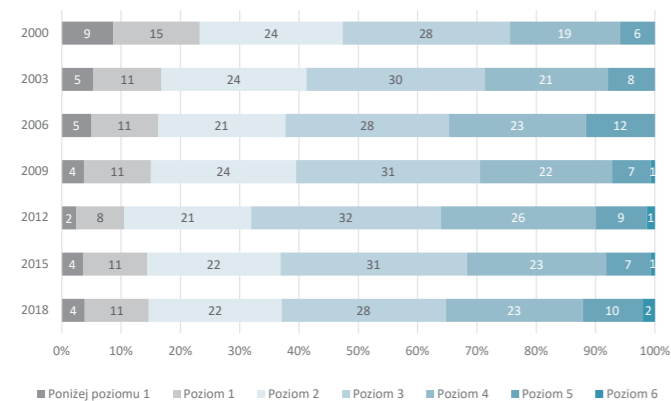
Wykres 1. Zmiany wyników pomiaru umiejętności rozumienia czytanego tekstu wśród uczniów w Polsce i średnio w krajach OECD w latach 2000–2018.



Poziomy umiejętności uczniów

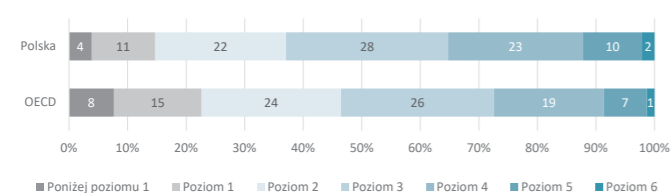
Odsetek uczniów osiągających najniższe wyniki w 2018 roku był w Polsce bardzo podobny do tego z roku 2015. Obecnie co siódmy piętnastolatek w naszym kraju nie opanował umiejętności rozumienia czytanego tekstu w stopniu wystarczającym do swobodnego korzystania ze źródeł pisanych. Warto zauważyć, że w 2000 r. w tej kategorii znalazł się niemal co 4. uczeń w Polsce (odsetek wówczas wyniósł 23%). Kolejne edycje badania pokazywały, że odsetek osób na najniższych poziomach umiejętności systematycznie spadał. Od 2009 roku wynosił on mniej niż 15%, co oznaczało, że Polska jako jeden z nielicznych krajów UE spełniła postawione wymagania. Zauważalny był również wzrost odsetka uczniów osiągających najwyższe wyniki: w roku 2000 było to tylko 6%, w kolejnych edycjach badania odsetek ten wahał się w przedziale 8–12%, w żadnym z cykli badania nie spadł poniżej tego poziomu.

Wykres 2. Odsetki uczniów na poszczególnych poziomach umiejętności rozumienia czytanego tekstu w Polsce w latach 2000–2018.



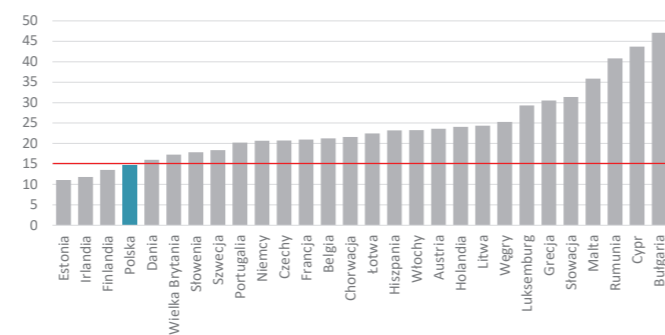
Zauważalne są różnice między Polską a krajami OECD. W naszym kraju mniej jest uczniów na najniższych poziomach umiejętności, a więcej na najwyższych. W Polsce w grupie osób o najniższych wynikach znalazło się 14,6% badanych piętnastolatek, w krajach OECD poniżej 2 poziomu usytuowało się 23% uczniów. Umiejętności na poziomach 5 i 6 w Polsce opanowało 12,2% uczniów, w krajach OECD jest to 8%.

Wykres 3. Odsetki uczniów na poszczególnych poziomach umiejętności rozumienia czytanego tekstu w Polsce i w krajach OECD w 2018 roku.



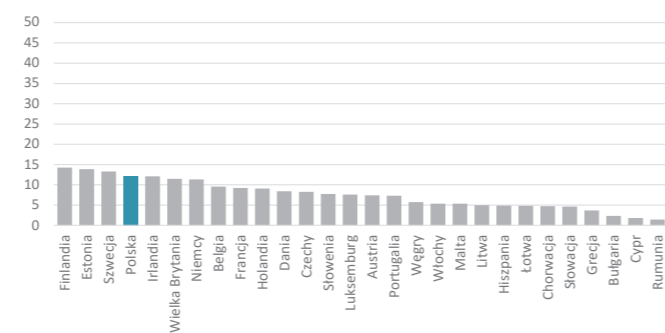
Gdy niewielki odsetek uczniów, którzy uzyskali najniższe wyniki, zestawimy z odpowiednimi danymi z pozostałych krajów Unii Europejskiej, zauważymy, że tylko w trzech krajach jest on niższy, w Estonii, Irlandii i Finlandii.

Wykres 4. Odsetek uczniów z wynikiem poniżej poziomu 2 w umiejętnościach rozumienia czytanego tekstu w krajach Unii Europejskiej.



Warto również zwrócić uwagę na porównanie odsetka polskich uczniów osiągających najwyższe wyniki z odpowiednimi wynikami z innych krajów UE. Więcej piętnastolatek, którzy sytuują się poziomach 5 i 6, można zaobserwować tylko w Finlandii, Estonii i Szwecji.

Wykres 5. Odsetek uczniów z wynikiem powyżej poziomu 4 w umiejętnościach rozumienia czytanego tekstu w krajach Unii Europejskiej.



Zakresy umiejętności

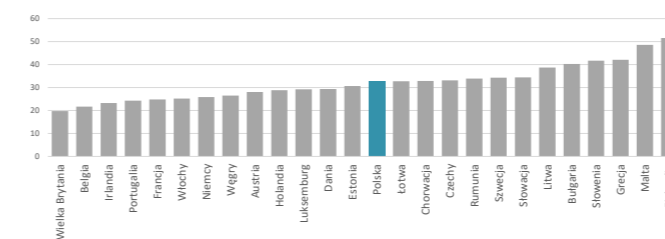
Wyniki polskich uczniów we wszystkich trzech zakresach umiejętności są wyrównane: w każdym średnio wynoszą one 514 punktów. W niewielu krajach, a zwłaszcza w niewielu krajach europejskich, można zauważyć podobne wyrównanie wyników. Oznacza to, że w polskiej dydaktyce zachowuje się równowagę w nauczaniu różnych umiejętności związanych z odbiorem tekstu.

Trzeba też zauważyć, iż w Polsce bardzo podobne są wyniki zadań badających rozumienie różnego rodzaju tekstów – pojedynczych (np. artykułów, opowiadań, instrukcji) i wielokrotnionych (np. stron internetowych, forów internetowych). Z analizy wyników można wyciągnąć ciekawy wniosek, że w wielu krajach uczniowie lepiej radzili sobie z tekstami wielokrotnionymi, co świadczy o tym, że z tego typu tekstami obcuje na co dzień i są one dla nich łatwiejsze w odbiorze niż tradycyjne teksty pojedyncze.

Różnice między wynikami chłopców i dziewcząt

Utrzymuje się duża różnica między chłopcami i dziewczętami pod względem stopnia opanowania umiejętności rozumienia czytanego tekstu. W Polsce poniżej poziomu 2 znalazł się co piąty piętnastolatek i tylko co dziesiąta piętnastolatka, z kolei na poziomach 5 i 6 usytuowało się 9,6% chłopców i 14,8% dziewcząt.

Wykres 6. Różnica średniego wyniku dziewcząt i chłopców w pomiarze rozumienia czytanego tekstu w badaniu PISA 2018.



Różnice wyników między chłopcami i dziewczętami nie są polską specyfiką i można je dostrzec też w innych krajach. Wykres 6. ilustruje to zjawisko w krajach Unii Europejskiej. Jak widać, Polska sytuuje się mniej więcej w środku skali.

Podsumowanie

W badaniu PISA w 2018 roku główną dziedziną było rozumienie czytanego tekstu. Polscy uczniowie uzyskali średni wynik wynoszący 512 punktów. Był to wynik znacząco lepszy niż w poprzedniej edycji badania w 2015 roku, gdy wyniósł on 506 punktów. Wówczas niższy wynik można było wiązać ze zmianą medium, za pomocą którego sprawdzano umiejętności. W 2015 roku po raz pierwszy badanie w Polsce (tak jak w większości krajów) w całości przeprowadzono na komputerze. Można było zatem wnioskować, że dla wielu uczniów problemem było zrozumienie tekstu wyświetlonego na ekranie a także sprawne poradzenie sobie z techniką. Wynik z roku 2018 wskazuje, że kolejna generacja piętnastolatek jest bardziej oswojona z medium elektronicznym oraz że ta forma funkcjonowania tekstu coraz częściej wykorzystywana jest na lekcjach.

Od czasu pierwszego badania PISA w 2000 roku, gdy zarazem rozumienie czytanego tekstu po raz pierwszy było dziedziną główną, nastąpiła znacząca poprawa wyników. Jest to rezultatem zmiany paradygmatu dydaktyki języka polskiego i innych przedmiotów, w której na pierwszy plan wysunięto uważną lekturę tekstu.

Badanie przeprowadził oraz raport przygotował zespół w składzie:

Krzysztof Biedrzycki, Krzysztof Bulkowski, Jan Burski, Wioleta Dobosz-Leszczynska, Jacek Haman, Joanna Kaźmierczak, Zbigniew Marciniak, E. Barbara Ostrowska, Michał Sitek, Krzysztof Spalik, Agnieszka Sułowska.

Więcej informacji o badaniu PISA można znaleźć na stronie internetowej www.pisa.ibe.edu.pl

Badanie zostało sfinansowane ze środków Ministerstwa Edukacji Narodowej.

Program Międzynarodowej Oceny Umiejętności Uczniów
PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT



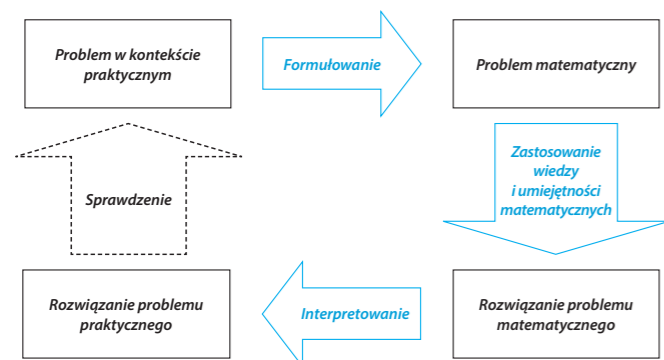
Wyniki badania 2018 w Polsce

Matematyka

Celem badania PISA w zakresie matematyki jest określenie, w jakim stopniu uczniowie potrafią wykorzystać swoją wiedzę i umiejętności matematyczne, gdy stają przed koniecznością rozwiązywania problemów otaczającego świata. Dlatego dobry wynik w badaniu PISA uzyskują uczniowie, którzy umieją rozumować matematycznie, a także skutecznie wykorzystywać pojęcia i narzędzia matematyczne do opisu, analizy i prognozowania różnych zjawisk. Nawet bardzo dobra, ale tylko teoretyczna znajomość narzędzi matematyki jest w tym badaniu sprawą drugorzędną; przede wszystkim liczy się umiejętność zastosowania tych narzędzi w kontekście praktycznym.

Na schemacie 1 przedstawiono cykl rozwiązywania problemów w kontekście praktycznym (*modelling cycle*), tak jak go zdefiniowano w dokumencie ramowym opisującym założenia teoretyczne badania PISA 2018.

Schemat 1. Cykl rozwiązywania problemów w kontekście praktycznym.



Formułowanie polega na matematyzacji problemu, czyli na wyborze lub skonstruowaniu modelu matematycznego adekwatnego dla danej sytuacji praktycznej. **Zastosowanie wiedzy i umiejętności matematycznych** polega na rozwiązaniu problemu matematycznego za pomocą narzędzi i metod matematyki. **Interpretowanie** polega na krytycznym odniesieniu uzyskanego wyniku do praktycznego kontekstu, w którym problem powstał.

W założeniach teoretycznych, które są podstawą mierzenia osiągnięć uczniów w badaniu, podkreślono, że podstawowymi celami kształcenia matematycznego jest rozumowanie i argumentacja, myślenie strategiczne oraz modelowanie matematyczne. Te same cele, jako tzw. wymagania ogólne, wskazuje od 2008 roku polska podstawa programowa dla matematyki.

Część matematyczna badania PISA 2018 zawierała 70 zadań o różnym stopniu trudności. Były to dokładnie te same zadania, które rozwiązywali uczniowie w badaniu PISA 2015. Zadania miały zróżnicowany format: w 26 z nich nale-

żało wybrać poprawną odpowiedź spośród kilku podanych lub ocenić, czy podane zdania są prawdziwe, czy fałszywe; w 44 pozostałych należało samodzielnie sformułować odpowiedź i jej uzasadnienie. W badaniach PISA 2015 oraz PISA 2018 uczniowie rozwiązywali zadania na ekranach komputerów; do roku 2012 wpisywali rozwiązania w zeszytach testowych.

Po przeprowadzeniu badania w roku 2003, w którym matematyka po raz pierwszy była dziedziną główną, została określona skala, na której prezentowane są wyniki z matematyki uzyskane w kolejnych edycjach badania. Parametry tej skali zostały dobrane tak, by średni wynik uczniów z krajów należących do OECD dla 2003 roku był równy 500, a odchylenie standardowe 100. Na tej samej skali określono wyniki uczniów, a także trudność zadań. Wyniki badania są również prezentowane w postaci rozkładów procentowych uczniów w poszczególnych krajach pomiędzy 6 poziomów umiejętności. Pokazują one zróżnicowanie umiejętności uczniów, w tym odsetek uczniów uzyskujących słabe wyniki oraz tych uzyskujących wyniki bardzo dobre.

Umiejętności matematyczne polskich uczniów na tle innych krajów

W badaniu PISA 2018 polscy uczniowie uzyskali wynik 516 punktów, o 27 punktów więcej niż średnia dla krajów OECD (równa 489 punktów). Wśród 79 krajów lub regionów biorących udział w badaniu PISA 2018, najlepsze wyniki z matematyki uzyskali uczniowie z krajów lub regionów azjatyckich, w tym uczniowie z Chin, reprezentowanych w badaniu przez cztery miasta i prowincje (Pekin, Szanghaj, Jiangu, Guangdong). Oprócz 7 krajów lub regionów azjatyckich wynik istotnie wyższy od Polski uzyskał tylko jeden kraj europejski: Estonia. Wyniki nieodróżnialne statystycznie od wyniku polskich uczniów uzyskali uczniowie z Holandii, Szwajcarii i Kanady. Pozostałe 66 krajów lub regionów biorących udział w badaniu uzyskało wyniki niższe niż Polska.

Zmiany wyników w latach 2003–2018

Matematyka była główną dziedziną badania PISA po raz pierwszy w 2003 r. i począwszy od tej edycji wyniki z matematyki są porównywalne w czasie. Średni wynik uzyskany przez polskich uczniów w 2018 r. jest statystycznie nieodróżnialny do wyniku z 2012 r., który był najwyższy w historii polskich edycji badania PISA. Kiedy porównuje się wyniki uzyskiwane przez polskich uczniów w kolejnych latach, widoczna jest bardzo duża poprawa: różnica między dwiema ostatnimi edycjami badania wynosi 12 punktów, a między pierwszym badaniem w 2003 r. a ostatnim w 2018 r. wynosi aż 26 punktów.

Tabela 1. Wyniki uczniów z matematyki w badaniu PISA 2018.

L.p.	Kraj	Srednia (błąd standardowy)	Istotność	OECD	UE
1	Chiny B-S-J-G	591 (2.5)	↑		
2	Singapur	569 (1.6)	↑		
3	Makao (Chiny)	558 (1.5)	↑		
4	Hongkong (Chiny)	551 (3.0)	↑		
5	Tajwan	531 (2.9)	↑		
6	Japonia	527 (2.5)	↑	OECD	
7	Korea Południowa	526 (3.1)	↑	OECD	
8	Estonia	523 (1.7)	↑	OECD	UE
9	Holandia	519 (2.6)	↑	OECD	UE
10	Polska	516 (2.6)	↑	OECD	UE
11	Szwajcaria	515 (2.9)	↑	OECD	
12	Kanada	512 (2.4)	↑	OECD	
13	Dania	509 (1.7)	↑	OECD	UE
14	Słowenia	509 (1.4)	↑	OECD	UE
15	Belgia	508 (2.3)	↑	OECD	UE
16	Finlandia	507 (2.0)	↑	OECD	UE
17	Szwecja	502 (2.7)	↑	OECD	UE
18	Wielka Brytania	502 (2.6)	↑	OECD	UE
19	Norwegia	501 (2.2)	↑	OECD	
20	Niemcy	500 (2.6)	↑	OECD	UE
21	Irlandia	500 (2.2)	↑	OECD	UE
22	Czechy	499 (2.5)	↑	OECD	UE
23	Austria	499 (3.0)	↑	OECD	UE
24	Łotwa	496 (2.0)	↑	OECD	UE
25	Francja	495 (2.3)	↑	OECD	UE
26	Islandia	495 (2.0)	↑	OECD	
27	Nowa Zelandia	494 (1.7)	↑	OECD	
28	Portugalia	492 (2.7)	↑	OECD	UE
29	Australia	491 (1.9)	↑	OECD	
30	Rosja	488 (3.0)			
31	Włochy	487 (2.8)		OECD	UE
32	Słowacja	486 (2.6)	↓	OECD	UE
33	Luksemburg	483 (1.1)	↓	OECD	UE
34	Hiszpania	481 (1.5)	↓	OECD	UE
35	Litwa	481 (2.0)	↓	OECD	UE
36	Węgry	481 (2.3)	↓	OECD	UE
37	Stany Zjednoczone	478 (3.2)	↓	OECD	
38	Białoruś	472 (2.7)	↓		
39	Malta	472 (1.9)	↓		UE

Kraje zaprezentowane w porządku malejącym ze względu na wynik średni. W nawiasie podano błąd standardowy. Szarym tłem wyróżnione są kraje, których średni wynik nie różni się statystycznie istotnie od średniego wyniku Polski. W odpowiednich kolumnach oznaczono kraje należące do Organizacji Współpracy Gospodarczej i Rozwoju (OECD) oraz Unii Europejskiej (UE). W powyższych wynikach nie uwzględniono Wietnamu, zgodnie z oficjalnym oświadczeniem zamieszczonym na stronie internetowej OECD. ↑ Wynik statystycznie istotnie powyżej średniej krajów OECD (kolumna Istotność) ↓ Wynik statystycznie istotnie poniżej średniej krajów OECD (kolumna Istotność)

Poprawa wyników polskich uczniów w tym okresie jest również widoczna, kiedy porównamy średnie wyniki Polski z wynikami innych krajów: w pierwszych edycjach badania wyniki wyższe niż Polska osiągało około 20 krajów świata, a w 2018 r. takich krajów lub regionów świata jest tylko 8. Jeszcze bardziej imponująco wypada porównanie z krajami należącymi do Unii Europejskiej: w pierwszych edycjach badania wyniki wyższe niż Polska osiągało około 10 krajów Unii, a tymczasem w 2012 r. polscy gimnazjaliści uzyskali

Tabela 2. Umiejętności polskich uczniów w zakresie matematyki w kolejnych edycjach badania PISA.

Rok badania PISA	Wynik punktowy polskich uczniów	Średni wynik krajów OECD	Liczba krajów lub regionów, które uzyskały wynik statystycznie istotnie lepszy od Polski	
			na świecie	w Unii Europejskiej
2003	490	499	21	12
2006	495	490	20	10
2009	495	492	20	8
2012	518	490	9	0
2015	504	487	14	5
2018	516	489	8	1

L.p.	Kraj	Srednia (błąd standardowy)	Istotność	OECD	UE
40	Chorwacja	464 (2.5)	↓		UE
41	Izrael	463 (3.5)	↓	OECD	
42	Turcja	454 (2.3)	↓	OECD	
43	Ukraina	453 (3.6)	↓		
44	Grecja	451 (3.1)	↓	OECD	UE
45	Cypr	451 (1.4)	↓		UE
46	Serbia	448 (3.2)	↓		
47	Malezja	440 (2.9)	↓		
48	Albania	437 (2.4)	↓		
49	Bułgaria	436 (3.8)	↓		UE
50	ZEA	435 (2.1)	↓		
51	Brunei	430 (1.2)	↓		
52	Rumunia	430 (4.9)	↓		UE
53	Czarnogóra	430 (1.2)	↓		
54	Kazachstan	423 (1.9)	↓		
55	Moldawia	421 (2.4)	↓		
56	Baku (Azerbejdżan)	420 (2.8)	↓		
57	Tajlandia	419 (3.4)	↓		
58	Urugwaj	418 (2.6)	↓		
59	Chile	417 (2.4)	↓	OECD	
60	Katar	414 (1.2)	↓		
61	Meksyk	409 (2.5)	↓	OECD	
62	Bośnia i Hercegowina	406 (3.1)	↓		
63	Kostaryka	402 (3.3)	↓		
64	Peru	400 (2.6)	↓		
65	Jordania	400 (3.3)	↓		
66	Gruzja	398 (2.6)	↓		
67	Macedonia Północna	394 (1.6)	↓		
68	Liban	393 (4.0)	↓		
69	Kolumbia	391 (3.0)	↓		
70	Brazylia	384 (2.0)	↓		
71	Argentyna	379 (2.8)	↓		
72	Indonezja	379 (3.1)	↓		
73	Arabia Saudyjska	373 (3.0)	↓		
74	Maroko	368 (3.3)	↓		
75	Kosowo	366 (1.5)	↓		
76	Panama	353 (2.7)	↓		
77	Filipiny	353 (3.5)	↓		
78	Dominikana	325 (2.6)	↓		

najwyższy wynik z matematyki spośród wszystkich krajów Unii Europejskiej. Także w 2018 r. wynik polskich gimnazjalistów jest znakomity – Estonia jest jedynym krajem Unii i Europy z wynikiem wyższym niż Polska.

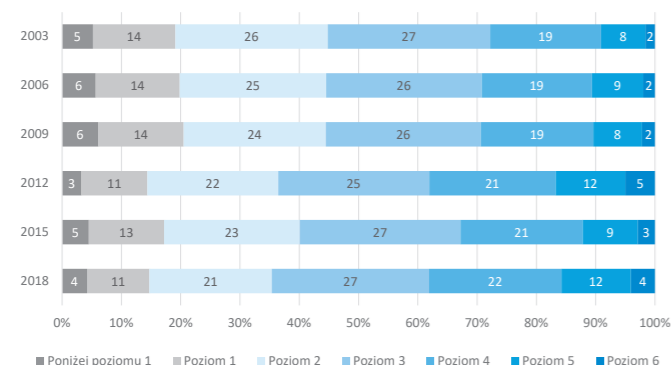
Poprawa wyników polskich uczniów jest tym bardziej godna podkreślenia, że w okresie 2003–2018 w wielu krajach świata i Europy średnie wyniki obniżyły się, co znajduje odzwierciedlenie w malejącym średnim wyniku dla krajów OECD.

Poziomy umiejętności uczniów

Dla zobrazowania zróżnicowania wyników uczniów skalę matematyczną badania PISA podzielono na sześć poziomów. Do każdego z poziomów przyporządkowano zadania oraz odsetek uczniów, których umiejętności matematyczne są na danym poziomie. Szczególnie pomocne w interpretacji zróżnicowania wyników w poszczególnych krajach i różnic między krajami są odsetki uczniów uzyskujących najłabsze wyniki (poniżej 2 poziomu) i uzyskujących najlepsze wyniki (na 5 i 6 poziomie).

Uczniowie z poziomu 1 radzą sobie z typowymi zadaniami, w których wszystkie dane są bezpośrednio podane, a zadane pytania są proste. Potrafią wykonać czynności rutynowe, postępując zgodnie z podanym prostym przepisem i podejmują działania oczywiste, wynikające wprost z treści zadania. Uczniowie osiągający poziom 5 potrafią modelować złożone sytuacje, identyfikując ograniczenia i precyzując zastrzeżenia. Umieją porównywać, oceniać i wybierać odpowiednie strategie rozwiązania problemu. Ich rozwiązania zadań pokazują, że potrafią krytycznie ocenić swoje działania, przedstawić swoją interpretację i sposób rozumowania, używając do tego odpowiednich reprezentacji, w tym symbolicznych i formalnych.

Wykres 1. Odsetki uczniów w Polsce na poszczególnych poziomach umiejętności matematycznych w latach 2003–2018.

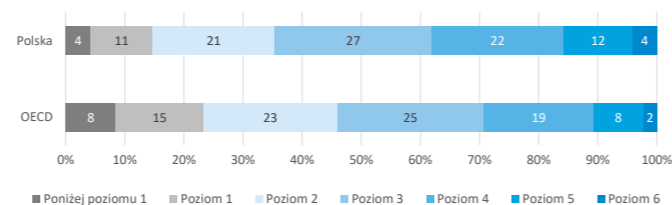


Porównanie odsetków uczniów na poszczególnych poziomach umiejętności w latach 2003–2018 pokazuje, że poprawa średnich wyników została osiągnięta zarówno przez zmniejszenie odsetka uczniów najłabszych, jak i zwiększenie odsetka uczniów najlepszych. W 2018 r. na dwóch najniższych poziomach znajdowało się łącznie tylko 14,7% polskich uczniów (w 2003 r. aż 22%), a odsetek uczniów na dwóch najwyższych poziomach wynosił łącznie 15,8% (w 2003 r. tylko 10,1%).

Poza Polską są jeszcze tylko 3 kraje na świecie, którym w latach 2003–2018 udało się zarówno obniżyć odsetek uczniów o najniższych umiejętnościach, jak i zwiększyć odsetek uczniów o najwyższych umiejętnościach. Są to: Makao (Chiny), Portugalia i Włochy. Spośród nich tylko Makao ma średni wynik z matematyki wyższy niż średnia OECD i wyższy niż Polska. Natomiast wyniki Portugalii i Włoch są na poziomie średniej krajów OECD, czyli znacznie niższe niż Polska. Porównanie odsetków uczniów z Polski i z krajów OECD na poszczególnych poziomach umiejętności matematycznych przedstawiono na wykresie 2. Pokazuje

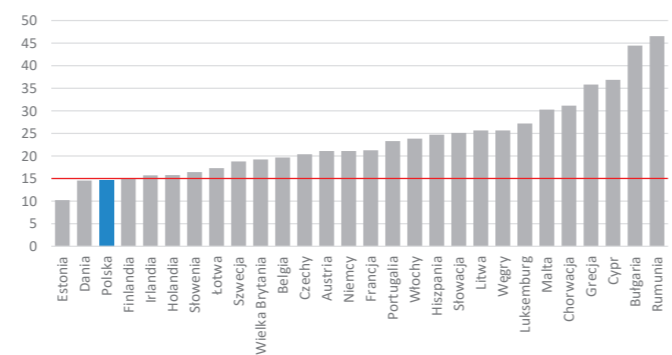
on, że wśród polskich gimnazjalistów było znacząco mniej uczniów uzyskujących niskie wyniki i znacznie więcej uczniów osiągających wysokie wyniki niż średnio w krajach OECD. Na dwóch najniższych poziomach umiejętności (poziom 1 i poniżej) znajdowało się 14,7% polskich uczniów (w OECD średnio 24%). Na dwóch najwyższych poziomach (poziom 5 i 6) było łącznie 15,8% polskich uczniów (w OECD średnio 10,9%).

Wykres 2. Odsetki uczniów z Polski i z krajów OECD na poszczególnych poziomach umiejętności matematycznych w 2018 roku.

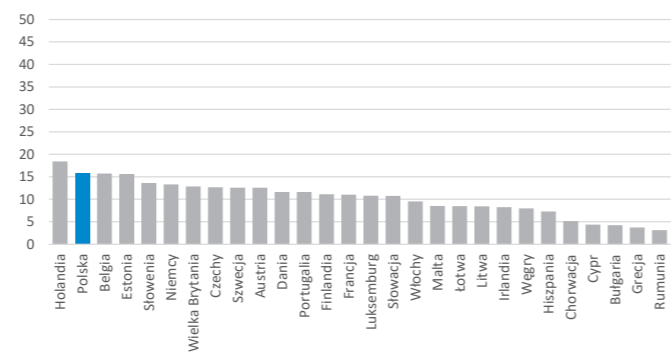


Odsetki uczniów osiągających poszczególne poziomy umiejętności bardzo różnią się między krajami. Porównanie odsetków uczniów uzyskujących najłabsze i najlepsze wyniki w poszczególnych krajach Unii Europejskiej przedstawione jest na wykresach 3 i 4.

Wykres 3. Odsetek uczniów z wynikiem poniżej poziomu 2 z matematyki w krajach Unii Europejskiej.



Wykres 4. Odsetek uczniów z wynikiem powyżej poziomu 4 z matematyki w krajach Unii Europejskiej.



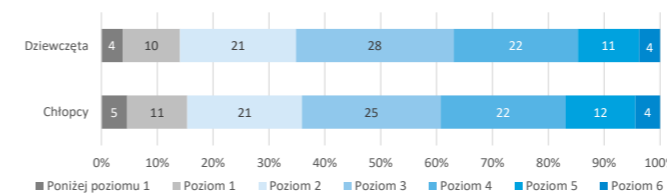
Wykresy te pokazują znakomite osiągnięcia kształcenia matematyki w Polsce – polskie szkoły na tle systemów edukacji krajów Unii Europejskiej charakteryzowały się bardzo małym odsetkiem uczniów o najniższych umiejętnościach (niższy odsetek był jedynie w Estonii, a podobny jak w Polsce – w Danii i Finlandii), oraz bardzo wysokim odsetkiem uczniów o najwyższych umiejętnościach (wyższy odsetek był tylko w Holandii, a podobny jak w Polsce – w Belgii i Estonii).

Różnice między wynikami chłopców i dziewcząt

W badaniu PISA 2018 średni wynik z matematyki chłopców w krajach OECD wyniósł 492 punkty, a średni wynik dziewcząt był o 5 punktów niższy. W Polsce, w każdym z dotychczas przeprowadzonych badań PISA, z wyjątkiem edycji z 2015 roku, wyniki dziewcząt i chłopców były podobne. Również w 2018 r. średnie wyniki dziewcząt i chłopców praktycznie się nie różniły (odpowiednio 515 i 516 punktów).

Natomiast wyniki chłopców są bardziej zróżnicowane niż dziewcząt: więcej chłopców znajduje się grupach uczniów o najniższych i najwyższych umiejętnościach, natomiast dziewcząt jest więcej wśród uczniów o średnim poziomie umiejętności.

Wykres 5. Odsetki chłopców i dziewcząt na poszczególnych poziomach umiejętności matematycznych w badaniu PISA 2018.



Podsumowanie

W badaniu PISA 2018 r. polscy uczniowie osiągnęli wynik z matematyki równy 516 punktów. Jest on statystycznie istotnie lepszy od średniej dla krajów OECD. Wynik Polski należy do najwyższych w Unii Europejskiej: spośród krajów europejskich istotnie wyższym wynikiem z matematyki może się pochwalić jedynie Estonia.

Polska zdecydowanie wyróżnia się też na tle innych krajów świata biorących udział w badaniu: wyższą niż Polska średnią uzyskały, poza Estonią, jedynie kraje azjatyckie oraz miasta i prowincje reprezentujące Chiny.

Wynik polskich gimnazjalistów z matematyki osiągnięty w 2018 r. jest o 12 punktów wyższy od wyniku z poprzednie-

go badania w 2015 r. Jest on podobny do wyniku z 2012 r., który był najwyższy w Unii Europejskiej i najwyższy w historii polskich edycji badania PISA.

Analizując umiejętności matematyczne polskich gimnazjalistów, należy podkreślić bardzo mały odsetek uczniów uzyskujących najniższe wyniki i równocześnie bardzo wysoki odsetek uczniów o najwyższych wynikach. Taki rozkład umiejętności matematycznych uczniów wskazuje na wysoką skuteczność nauczania matematyki w polskich szkołach, w tym gimnazjach, na tle systemów edukacji krajów Unii Europejskiej.

Polski system edukacji pozytywnie wyróżniał się także z powodu niedużej i nieistotnej różnicy między wynikami chłopców i dziewcząt.

Część zadań w badaniu PISA wymaga od ucznia przeprowadzenia samodzielnego rozumowania matematycznego. Zadania tego typu stanowiły w roku 2003 dla polskich uczniów duży problem: znacznie lepiej wypadali oni w zadaniach odtwórczych, wymagających zastosowania znanego sposobu postępowania, wyćwiczonego w szkole. Po zmianie podstawy programowej z matematyki w 2008 roku, rozumowanie matematyczne stało się jednym z najważniejszych celów kształcenia matematycznego w polskich szkołach. System egzaminacyjny konsekwentnie egzekwował to wymaganie, także na egzaminie gimnazjalnym. W badaniu PISA 2012 po raz pierwszy wyniki polskich uczniów w zadaniach wymagających rozumowania były takie same lub lepsze od średnich wyników w krajach OECD. Podobnie rzecz się miała w roku 2015, a badanie PISA 2018 potwierdziło trwałość tego zjawiska wśród uczniów gimnazjów.

Kolejne edycje badania PISA przeprowadzane w latach 2003–2018 pokazały systematyczny wzrost umiejętności matematycznych polskich uczniów kształconych w systemie edukacji powszechnej funkcjonującym od 1999 do 2019 r. Następne badania PISA będą sprawdzeniem skuteczności nowego systemu szkolnego.

Badanie przeprowadził oraz raport przygotował zespół w składzie:

Krzysztof Biedrzycki, Krzysztof Bulkowski, Jan Burski, Wioleta Dobosz-Leszczynska, Jacek Haman, Joanna Kaźmierczak, Zbigniew Marciniak, E. Barbara Ostrowska, Michał Sitek, Krzysztof Spalik, Agnieszka Sułowska.

Więcej informacji o badaniu PISA można znaleźć na stronie internetowej www.pisa.ibe.edu.pl

Badanie zostało sfinansowane ze środków Ministerstwa Edukacji Narodowej.



Program Międzynarodowej Oceny Umiejętności Uczniów PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT



Wyniki badania 2018 w Polsce

Rozumowanie w naukach przyrodniczych

Umiejętności uczniów w zakresie nauk przyrodniczych mierzone w badaniu PISA, określane zbiorczo jako *science literacy*, obejmują zastosowanie wiedzy przyrodniczej do rozwiązywania zadań odnoszących się do problemów życia codziennego, funkcjonowania społeczeństwa oraz globalnych wyzwań przyrodniczych i środowiskowych. Wymaga to nie tylko opanowania wiadomości z zakresu biologii, chemii, fizyki i geografii, ale przede wszystkim rozumienia podstaw funkcjonowania nauki. Zadania wykorzystane w badaniu PISA mierzą więc zarówno wiedzę o faktach naukowych, jak i wiedzę o samej nauce – o drodze dochodzenia do prawdy naukowej. Kluczowym obszarem wiedzy o nauce jest metoda naukowa: stawianie pytań badawczych i hipotez, a następnie sprawdzanie tych hipotez za pomocą obserwacji i doświadczeń. Hipotezę uznajemy za potwierdzoną, jeśli nie udaje się jej empirycznie odrzucić. Teorię uznajemy za prawdziwą, jeśli mimo usilnych starań nie udaje się jej empirycznie obalić. Niezwykle ważny jest także krytycyzm i samokrytycyzm. Pozwala on na odrzucenie wyników błędnie przeprowadzonych doświadczeń i obserwacji, powstrzymuje przed formułowaniem pochopnych lub nieuprawnionych wniosków, pozwala demaskować oszustwa naukowe i odróżnić twierdzenia naukowe od pseudonauki – a tym samym zapewnia nauce wiarygodność.

Rozumowanie w naukach przyrodniczych obejmuje takie zagadnienia, jak zdrowie, zasoby naturalne, zagrożenia oraz nowe wyzwania nauki i techniki. Odnosi się do wiedzy o treściach nauki, procedurach badawczych i poznaniu naukowym. Na kompetencje przyrodnicze mierzone w badaniu złożyły się trzy główne umiejętności, opisane poniżej.

■ **Wyjaśnianie zjawisk przyrodniczych w sposób naukowy.** Podając, rozpoznając lub oceniając wyjaśnienia różnorodnych zjawisk z zakresu przyrody i techniki, uczeń potrafi: (1) przywołać z pamięci i zastosować odpowiednią wiedzę naukową, (2) wskazać, wykorzystać lub stworzyć model lub inną reprezentację wyjaśniającą dane zjawisko, (3) formułować i uzasadniać odpowiednie przypuszczenia, (4) stawiać hipotezy, (5) objaśniać potencjalne następstwa wiedzy naukowej dla społeczeństwa.

■ **Planowanie i ocena poprawności procedur badawczych.** Opisując i oceniając badania naukowe oraz proponując sposoby odpowiedzi na pytania badawcze, uczeń potrafi: (1) wskazać problem podejmowany w określonym badaniu naukowym, (2) odróżnić pytania, na które można odpowiedzieć w sposób naukowy od tych, które nie mają takiego charakteru, (3) podać sposób naukowego poszukiwania odpowiedzi na określone pytanie badawcze, (4)

ocenić różne sposoby naukowego poszukiwania odpowiedzi na określone pytanie badawcze, (5) opisać i ocenić, w jaki sposób naukowcy starają się zagwarantować rzetelność danych oraz obiektywizm i uniwersalność wniosków.

■ **Interpretacja danych i dowodów naukowych.** Analizując i oceniając dane naukowe, tezy i argumenty podane w różnej formie, a także wyciągając odpowiednie wnioski, uczeń potrafi: (1) przetworzyć dane naukowe podane w jednej formie w inną formę, (2) analizować i interpretować dane i wyciągać odpowiednie wnioski, (3) wyodrębnić założenia, wskazać dowody i określić wnioskowanie w tekstach dotyczących nauki, (4) odróżnić argumenty bazujące na dowodach i teoriach naukowych od tych opartych na innych podstawach, (5) ocenić wiarygodność naukową tekstów z różnych źródeł (gazet, czasopism, internetu itp.).

W badaniu użyto 115 zadań. Wszystkie były wykorzystane w badaniu PISA 2015, kiedy to rozumowanie w naukach przyrodniczych było główną dziedziną pomiaru. Większość zadań (76) zostało opracowanych do komputerowej wersji testu – pozostałe były wykorzystane przed 2015 r. w tradycyjnych, papierowych zeszytach testowych. Wyjaśnianie zjawisk przyrodniczych w sposób naukowy sprawdzano w 49 zadaniach, planowanie i ocenę poprawności procedur badawczych w 30 zadaniach, a interpretację danych i dowodów naukowych – w 36 zadaniach.

Wyniki uczniów są przedstawione na standaryzowanej skali, dla której punktem odniesienia jest średnia uzyskana przez uczniów z krajów należących do Organizacji Współpracy Gospodarczej i Rozwoju (OECD) w 2006 r. Przyjęto dla niej wartość 500 i odchylenie standardowe 100. Na tej samej skali określono także trudność zadań, a tym samym wymagania stawiane uczniom. Pozwoliło to na wyróżnienie sześciu poziomów umiejętności. Uczniowie z wynikami poniżej poziomu 2 mogą w przyszłości mieć problemy z funkcjonowaniem w nowoczesnym społeczeństwie, w którym znajomość zagadnień nauki i techniki jest coraz bardziej istotna. Uczniowie z najwyższymi wynikami – na poziomach 5 i 6 – to potencjalne przyszłe kadry naukowe i techniczne. Dlatego odsetki uczniów z wynikami poniżej poziomu 2 oraz powyżej poziomu 4 są ważnymi miarami efektywności systemu edukacji.

Wyniki polskich uczniów na tle innych krajów

Polscy uczniowie uzyskali w 2018 r. średni wynik 511 pkt, co plasuje ich na 11. miejscu wśród wszystkich 79 krajów biorących udział w badaniu, na 6. miejscu wśród krajów OECD, a na 3. miejscu wśród krajów Unii Europejskiej (tabela 1). Wynik ten był o 22 punkty wyższy od średniej dla

Tabela 1. Wyniki uczniów w pomiarze rozumowania w naukach przyrodniczych w badaniu PISA 2018.

L.p.	Kraj	Średnia (błąd standardowy)	Istotność	OECD	UE
1	Chiny B-S-J-G	590 (2,7)	↑		
2	Singapur	551 (1,5)	↑		
3	Makao (Chiny)	544 (1,5)	↑		
4	Estonia	530 (1,9)	↑	OECD	UE
5	Japonia	529 (2,6)	↑	OECD	
6	Finlandia	522 (2,5)	↑	OECD	UE
7	Korea Południowa	519 (2,8)	↑	OECD	
8	Kanada	518 (2,2)	↑	OECD	
9	Hongkong (Chiny)	517 (2,5)	↑		
10	Tajwan	516 (2,9)	↑		
11	Polska	511 (2,6)	↑	OECD	UE
12	Nowa Zelandia	508 (2,1)	↑	OECD	
13	Słowenia	507 (1,3)	↑	OECD	UE
14	Wielka Brytania	505 (2,6)	↑	OECD	UE
15	Holandia	503 (2,8)	↑	OECD	UE
16	Niemcy	503 (2,9)	↑	OECD	UE
17	Australia	503 (1,8)	↑	OECD	
18	Stany Zjednoczone	502 (3,3)	↑	OECD	
19	Szwecja	499 (3,1)	↑	OECD	UE
20	Belgia	499 (2,2)	↑	OECD	UE
21	Czechy	497 (2,5)	↑	OECD	UE
22	Irlandia	496 (2,2)	↑	OECD	UE
23	Szwajcaria	495 (3,0)	↑	OECD	
24	Francja	493 (2,2)	↑	OECD	UE
25	Dania	493 (1,9)	↑	OECD	UE
26	Portugalia	492 (2,8)	↑	OECD	UE
27	Norwegia	490 (2,3)	↑	OECD	
28	Austria	490 (2,8)	↑	OECD	UE
29	Łotwa	487 (1,8)	↑	OECD	UE
30	Hiszpania	483 (1,6)	↓	OECD	UE
31	Litwa	482 (1,6)	↓	OECD	UE
32	Węgry	481 (2,3)	↓	OECD	UE
33	Rosja	478 (2,9)	↓		
34	Luksemburg	477 (1,2)	↓	OECD	UE
35	Islandia	475 (1,8)	↓	OECD	
36	Chorwacja	472 (2,8)	↓		UE
37	Białoruś	471 (2,4)	↓		
38	Ukraina	469 (3,3)	↓		
39	Turcja	468 (2,0)	↓	OECD	
40	Włochy	468 (2,4)	↓	OECD	UE
41	Słowacja	464 (2,3)	↓	OECD	UE
42	Izrael	462 (3,6)	↓	OECD	
43	Malta	457 (1,9)	↓		UE
44	Grecja	452 (3,1)	↓	OECD	UE
45	Chile	444 (2,4)	↓	OECD	
46	Serbia	440 (3,0)	↓		
47	Cypr	439 (1,4)	↓		UE
48	Malezja	438 (2,7)	↓		
49	ZEA	434 (2,0)	↓		
50	Brunei	431 (1,2)	↓		
51	Jordania	429 (2,9)	↓		
52	Mołdawia	428 (2,3)	↓		
53	Tajlandia	426 (3,2)	↓		
54	Urugwaj	426 (2,5)	↓		
55	Rumunia	426 (4,6)	↓		UE
56	Bułgaria	424 (3,6)	↓		UE
57	Meksyk	419 (2,6)	↓	OECD	
58	Katar	419 (0,9)	↓		
59	Albania	417 (2,0)	↓		
60	Kostaryka	416 (3,3)	↓		
61	Czarnogóra	415 (1,3)	↓		
62	Kolumbia	413 (3,1)	↓		
63	Macedonia Północna	413 (1,4)	↓		
64	Peru	404 (2,7)	↓		
65	Argentyna	404 (2,9)	↓		
66	Brazylia	404 (2,1)	↓		
67	Bośnia i Hercegowina	398 (2,7)	↓		
68	Baku (Azerbejdżan)	398 (2,4)	↓		
69	Kazachstan	397 (1,7)	↓		
70	Indonezja	396 (2,4)	↓		
71	Arabia Saudyjska	386 (2,8)	↓		
72	Liban	384 (3,5)	↓		
73	Gruzja	383 (2,3)	↓		
74	Maroko	377 (3,0)	↓		
75	Kosowo	365 (1,2)	↓		
76	Panama	365 (2,9)	↓		
77	Filipiny	357 (3,2)	↓		
78	Dominikana	336 (2,5)	↓		

Kraje zaprezentowane w porządku malejącym ze względu na wynik średni. W nawiasie podano błąd standardowy. Szarym tłem wyróżnione są kraje, których średni wynik nie różni się statystycznie istotnie od średniego wyniku Polski. W odpowiednich kolumnach oznaczono kraje należące do Organizacji Współpracy Gospodarczej i Rozwoju (OECD) oraz Unii Europejskiej (UE). W powyższych wynikach nie uwzględniono Wietnamu, zgodnie z oficjalnym oświadczeniem zamieszczonym na stronie internetowej OECD.

↑ Wynik statystycznie istotnie powyżej średniej krajów OECD (kolumna Istotność) ↓ Wynik statystycznie istotnie poniżej średniej krajów OECD (kolumna Istotność)

OECD, która wyniosła 489 punktów (różnica była istotna statystycznie). Był zbliżony do wyników piętnastolatków z Hongkongu, Tajwanu, Nowej Zelandii, Słowenii i Wielkiej Brytanii – różnice między Polską a tymi krajami były nieistotne statystycznie.

Na świecie pod względem kompetencji przyrodniczych przodują piętnastolatkowie z krajów Dalekiego Wschodu – Chin (z czterech wybranych regionów), Tajwanu, Singapuru, Japonii, Korei – a także Estonii, Finlandii i Kanady. Pierwsze miejsce zajmują Chiny (Pekin, Szanghaj, Jiangsu i Guangdong, 590 pkt) z wynikiem o 101 pkt wyższym od średniej OECD. Z krajów europejskich najlepsza jest Estonia z wynikiem 530 punktów, a za nią Finlandia (522 pkt).

Zmiany wyników w latach 2006–2018

Rozumowanie w naukach przyrodniczych było głównym pomiarem badania PISA w 2006 r. i dopiero wtedy po

raz pierwszy wszechstronnie sprawdzono kompetencje uczniów oraz opisano je na skali umiejętności. Natomiast w 2015 r. papierowe zeszyty testowe zastąpił test komputerowy, w którym wykorzystano, oprócz tradycyjnych zadań wyświetlanych na ekranie, nowe, interaktywne zadania np. obejmujące symulacje określonych procesów. W ostatnim badaniu, w porównaniu z badaniem 2015, średni wynik polskich uczniów poprawił się o 10 punktów, co było jednym z najwyższych przyrostów, a drugim pod względem wielkości wśród krajów o wynikach powyżej średniej dla OECD (tabela 2). Należy jednak pamiętać, że wcześniej zanotowano duży spadek – w pierwszym badaniu w wersji komputerowej średnia była aż o 25 punktów niższa niż w ostatnim badaniu z wykorzystaniem wersji papierowej. Pomimo tego spadku, w ostatnich badaniach wynik polskich piętnastolatków był statystycznie istotnie wyższy od średniej dla krajów OECD. Był także statystycznie istotnie wyższy o 13 pkt od wyniku badania z 2006 r., w którym po raz pierwszy określono ramy ewaluacji kompetencji przyrodniczych.

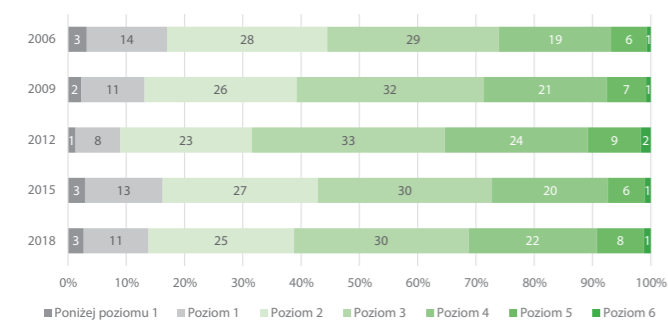
Tabela 2. Zmiany wyników polskich uczniów w pomiarze rozumowania w naukach przyrodniczych w badaniu PISA.

Rok badania	Wynik punktowy polskich uczniów	Średni wynik krajów OECD	Liczba krajów lub regionów, które uzyskały wynik statystycznie istotnie lepszy od Polski	
			na świecie	w Unii Europejskiej
2006	498	495	20	10
2009	508	498	15	4
2012	526	498	7	2
2015	501	491	18	6
2018	511	489	8	2

Poziomy umiejętności uczniów

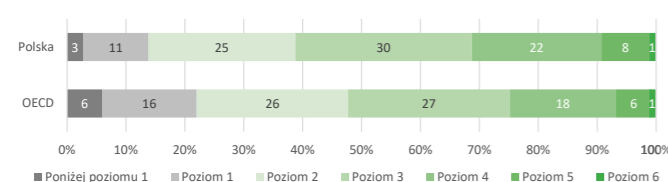
W Polsce w badaniu 2018 udział uczniów, którzy osiągnęli wynik poniżej 420,07 pkt (dolna granica poziomu 2), wyniósł 13,8%, co oznacza poprawę w porównaniu z rokiem 2015, kiedy miał on wartość 16,2% (wykres 1). Jednak wciąż jest wyższy niż w ostatnim badaniu z wykorzystaniem testu w wersji papierowej (2012 r.), kiedy to wyniósł 9%. Odsetek uczniów na poziomach 5 i 6 wyniósł w 2018 r. łącznie 9,3%, czyli więcej niż w 2015 r. (7,3%), choć mniej niż w 2012 r. (10,8%).

Wykres 1. Odsetki uczniów na poszczególnych poziomach umiejętności rozumowania w naukach przyrodniczych w Polsce w latach 2006–2018.



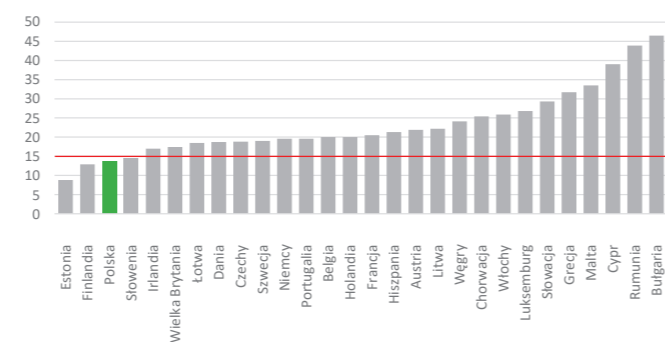
Porównanie rozkładów procentowych wyników według poziomów umiejętności uczniów z Polski z ich rówieśnikami z krajów OECD wypada korzystnie dla odchodzących już w przeszłość gimnazjów (wykres 2): mniej jest uczniów z najniższymi wynikami (drugiego poziomu nie osiągnęło 13,8% uczniów w Polsce, a w OECD – 22%), a więcej uczniów z najwyższymi wynikami (łącznie poziom 5 i 6 osiągnęło odpowiednio 9,3% i 6,8% uczniów).

Wykres 2. Odsetki uczniów na poszczególnych poziomach umiejętności rozumowania w naukach przyrodniczych w Polsce i w krajach OECD w 2018 roku.



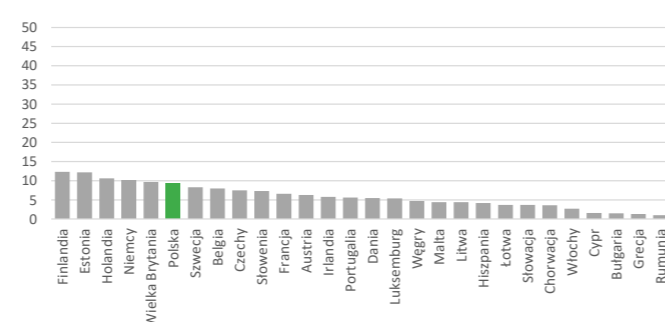
Jednym z celów współpracy europejskiej w zakresie edukacji jest podejmowanie działań na rzecz zmniejszenia do 2020 r. odsetka uczniów osiągających najłabsze wyniki (poniżej poziomu 2) do wartości poniżej 15%. Jedynie cztery kraje osiągnęły ten cel (wykres 3), a wśród nich Polska.

Wykres 3. Odsetek uczniów z wynikiem poniżej poziomu 2 w umiejętnościach rozumowania w naukach przyrodniczych w krajach Unii Europejskiej.



Pod względem odsetka uczniów z bardzo wysokimi wynikami Polska także uzyskała wysoką pozycję. Uplasowała się na 6. miejscu (wykres 4), a oprócz Estonii i Finlandii wyprzedziły nas także Holandia, Niemcy i Wielka Brytania.

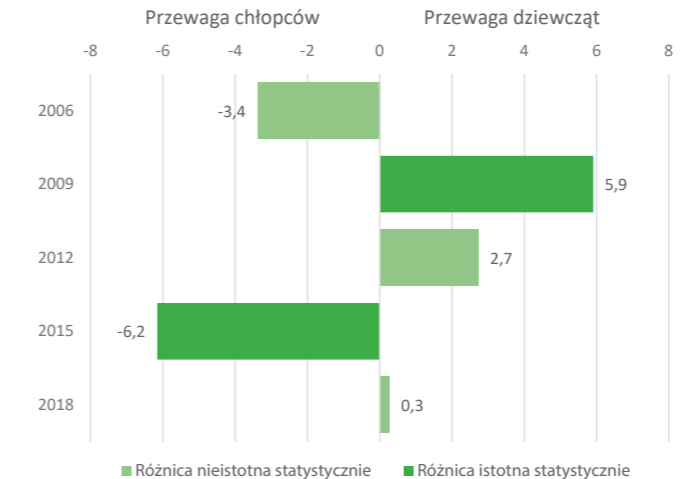
Wykres 4. Odsetek uczniów z wynikiem powyżej poziomu 4 w umiejętnościach rozumowania w naukach przyrodniczych w krajach Unii Europejskiej.



Różnice między wynikami chłopców i dziewcząt

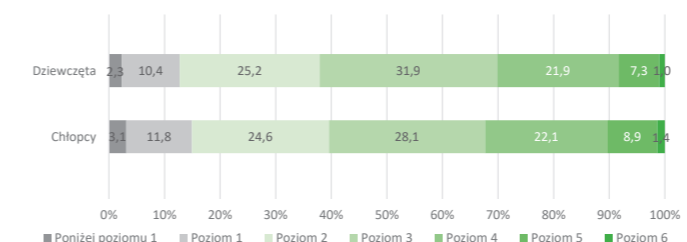
W 2018 r. średni wynik dziewcząt w krajach OECD wyniósł 490 i był o 2 punkty wyższy od wyniku chłopców, co jest znaczącą zmianą w porównaniu z badaniem w 2015 r., w którym chłopcy mieli przeciętny wynik o 4 punkty wyższy od dziewcząt. W większości krajów uczestniczących w badaniu lepsze wyniki osiągnęły dziewczęta – różnica na korzyść dziewcząt była statystycznie istotna w 34 krajach, podczas gdy chłopcy mieli statystycznie istotnie wyższy średni wynik jedynie w sześciu krajach. W Polsce średnie wyniki dziewcząt i chłopców były zbliżone (511 pkt), podczas gdy w badaniu z 2015 r. średni wynik chłopców był wyższy o 6 pkt, a w 2009 r. wyższy wynik osiągnęły dziewczęta (wykres 5).

Wykres 5. Różnice między średnimi wynikami dziewcząt i chłopców w pomiarze rozumowania w naukach przyrodniczych w Polsce w latach 2006–2018.



Mimo że średnie wyniki chłopców i dziewcząt są podobne, to można zauważyć znaczące różnice w odsetkach chłopców i dziewcząt z wynikami poniżej poziomu 2 i powyżej poziomu 4 (wykres 6). Odsetek chłopców, którzy osiągnęli niepokojąco niskie wyniki, wyniósł 15% i był statystycznie istotnie wyższy od analogicznego odsetka dziewcząt. Jednak odsetek chłopców z wynikami na najwyższych poziomach umiejętności był również istotnie wyższy.

Wykres 6. Odsetki chłopców i dziewcząt na poszczególnych poziomach umiejętności w pomiarze rozumowania w naukach przyrodniczych w Polsce.



Podsumowanie

W badaniu 2018 r. polscy uczniowie osiągnęli wynik 511 pkt, który jest statystycznie istotnie lepszy od uzyskanego w 2015 r. oraz od średniego wyniku dla krajów OECD. Jest to wciąż wynik niższy niż w rekordowym 2012 r.

Badanie przeprowadził oraz raport przygotował zespół w składzie:

Krzysztof Biedrzycki, Krzysztof Bulkowski, Jan Burski, Wioleta Dobosz-Leszczynska, Jacek Haman, Joanna Kaźmierczak, Zbigniew Marciniak, E.Barbara Ostrowska, Michał Sitek, Krzysztof Spalik, Agnieszka Sułowska.

Więcej informacji o badaniu PISA można znaleźć na stronie internetowej www.pisa.ibe.edu.pl

Badanie zostało sfinansowane ze środków Ministerstwa Edukacji Narodowej.



W raporcie z badania 2015 r. sugerowano, że głównym problemem polskich uczniów było rozwiązywanie zadań interaktywnych z wykorzystaniem komputera. Do takiej interpretacji skłaniały nie tylko wyniki z 2015 r., ale także z 2012 r. – w głównym badaniu umiejętności rozumowania w naukach przyrodniczych polscy uczniowie uzyskali wtedy średnio 526 pkt, co zapewniło im miejsce w czołówce światowej, natomiast w dodatkowym, komputerowym badaniu z zakresu rozwiązywania problemów osiągnęli wynik zaledwie 481 pkt przy średniej OECD 500 pkt, czyli poniżej oczekiwań. Obecnie wydaje się, że rozwiązywanie zadań na komputerze, a nie na papierze, nie stanowi dla uczniów problemu.

Z satysfakcją należy odnotować spadek odsetka uczniów o wynikach poniżej poziomu 2, a zatem tych zagrożonych trudnościami z radzeniem sobie z praktycznymi problemami, w których niezbędne są umiejętności rozumowania przyrodniczego, do wartości poniżej 15%. Wzrósł odsetek uczniów na najwyższych poziomach umiejętności, choć jest on wciąż stosunkowo niski (poniżej 10%). Z perspektywy rozwoju konkurencyjności i innowacyjności gospodarki należy dokładać większych starań, aby zwiększyć tę grupę uczniów. W pierwszym badaniu z wykorzystaniem komputerów (2015 r.) niepokojąca była statystycznie istotna różnica w wynikach na korzyść chłopców, podczas gdy w ostatnim badaniu papierowym (2012 r.) istotnej różnicy statystycznej między płciami nie odnotowano, chociaż dziewczęta osiągnęły lepsze wyniki od chłopców. Ta różnica zdawała się potwierdzać tezę o sprawniejszych technicznie chłopcach i sprawniejszych werbalnie dziewczętach – która jednak, w świetle najnowszego badania, okazała się jedynie stereotypem. Na uwagę zasługuje jednak silniejsze zróżnicowanie umiejętności chłopców niż dziewcząt, co pokazują różnice w odsetkach chłopców i dziewcząt na najniższych i najwyższych poziomach umiejętności. Różnice te były obserwowane także w poprzednich edycjach badania PISA.

PISA 2018 jest ostatnim badaniem, które diagnozuje system edukacji powszechnej wprowadzony w 1999 r. i zreformowany w 2017 r. Diagnoza ta wypada pozytywnie. Wzrost kompetencji przyrodniczych polskich uczniów w latach 2006–2018 świadczy o tym, że w szkołach, w tym w gimnazjach, coraz większy nacisk kładziono na kształtowanie umiejętności, a nie jedynie na przekazywanie wiadomości. Należy mieć nadzieję, że ta tendencja się utrzyma.