

PROGRAM MIĘDZYNARODOWEJ OCENY
UMIEJĘTNOŚCI UCZNIÓW

WYNIKI BADANIA PISA 2015 W POLSCE

Warszawa, marzec 2017 r.

WYNIKI BADANIA 2015 W POLSCE

MINISTERSTWO EDUKACJI NARODOWEJ

Międzynarodowe konsorcjum realizujące badanie OECD PISA 2015:

Educational Testing Service ETS (Stany Zjednoczone)
Westat (Stany Zjednoczone)
Pearson (Wielka Brytania)
Deutsches Institut für Internationale Pädagogische Forschung (DIPF, Niemcy)
cApStAn Linguistic Quality Control (Belgia)

Rada Zarządzająca Programem PISA (PGB):

Przedstawiciel Polski – Lidia Olak, Jerzy Wiśniewski

Skład zespołu realizującego badanie PISA w Polsce:

Marcin Chrzanowski
Alina Cichomska
Dorota Cyngot
Grażyna Drążyk
Michał Federowicz
Jacek Haman
Zbigniew Marciniak
E. Barbara Ostrowska (kierownik zespołu)
Michał Sitek
Krzysztof Spalik
Agnieszka Sułowska
Grzegorz Zajączkowski
Piotr Zielonka
Piotr Walicki

Raport z badania opracowali: *Piotr Bordzoł, Krzysztof Biedrzycki, Krzysztof Bulkowski (konsultacje statystyczne), Marcin Chrzanowski (konsultacje merytoryczne), Michał Federowicz, Jacek Haman, Zbigniew Marciniak, Dominik Marszał (konsultacje merytoryczne), E. Barbara Ostrowska, Michał Sitek, Krzysztof Spalik, Agnieszka Sułowska*

Redakcja naukowa: Michał Federowicz

Redakcja językowa: Natalia Skipietrow

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

Za realizację badania w Polsce odpowiadał Instytut Badań Edukacyjnych. Badanie próbne zrealizował Zespół Współpracowników Terenowych IBE, a badanie główne GFK Polonia

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

ISBN 978-83-65115-87-4

Badanie PISA zostało sfinansowane przez Ministerstwo Edukacji Narodowej.

© Copyright by: Instytut Badań Edukacyjnych, Warszawa 2017

Spis treści

Spis treści	3
WPROWADZENIE	5
1.1. Badanie PISA. Założenia i główne wyniki	5
1.1.1. Co i jak mierzy badanie PISA?.....	6
1.1.2. Krótki i długi cykl badania PISA – kwestia porównywalności w czasie.....	7
1.1.3. Zakres wnioskowania i interpretacji wyników.....	9
1.1.4. Główne wyniki.....	10
2. INNOWACJE METODOLOGICZNE I CIĄGŁOŚĆ BADANIA	23
2.1. Pomiar i skalowanie	24
2.2. Założenia ogólne – podstawy modeli IRT	25
2.3. Skalowanie wyników PISA 2015 w porównaniu z poprzednimi edycjami badania	28
2.4. Narzędzia pomiaru: od testów papierowych do testów komputerowych	29
2.5. Podsumowanie	34
Bibliografia	35
3. POPULACJA I PRÓBA W BADANIU PISA 2015	36
3.1. Dobór próby a błędy losowe i nielosowe	39
3.2. Realizacja badania a błędy nielosowe	40
4. ROZUMOWANIE W NAUKACH PRZYRODNICZYCH	42
4.1. Założenia teoretyczne badania	43
4.1.1. Nowe założenia ramowe pomiaru.....	43
4.1.2. Kontekst zadań.....	43
4.1.3. Wiedza naukowa.....	44
4.1.4. Umiejętności.....	47
4.1.5. Postawy.....	47
4.1.6. Wymagania poznawcze.....	48
4.2. Pomiar kompetencji w naukach przyrodniczych	48
4.2.1. Charakterystyka zadań wykorzystanych w badaniu.....	48
4.2.2. Określenie poziomów umiejętności uczniów i trudności zadań.....	49
4.3. Wyniki uczniów oraz ich zmiany w latach 2006-2015	49
4.3.1. Poziomy umiejętności w pomiarze kompetencji uczniów w naukach przyrodniczych.....	54
4.4. Wyniki chłopców i dziewcząt	58
4.5. Wyniki uczniów w podziale na rodzaj umiejętności i rozwiązywalność poszczególnych zadań	58

4.6. Postawy uczniów wobec nauk przyrodniczych oraz oczekiwania względem kariery zawodowej	61
4.7. Podsumowanie	63
ANEKS. OMÓWIENIE PRZYKŁADOWYCH ZADAŃ	64
Wiązka zadań „Badanie zbocza góry”	64
Zadanie 1.	64
Zadanie 2.	65
Wiązka zadań „Bieganie podczas upału”	67
Zadanie	70
5. CZYTANIE I INTERPRETACJA	72
5.1. Założenia teoretyczne badania	72
5.2. Wyniki polskich uczniów na tle innych krajów	73
5.3. Poziomy umiejętności w dziedzinie czytania	76
5.4. Wyniki chłopców i dziewcząt	78
5.5. Pogłębiona analiza jakościowa wyników badania	80
5.5.1. Wyszukiwanie informacji.....	80
5.5.2. Interpretacja tekstu	81
5.5.3. Refleksja i argumentacja.....	83
5.6. Podsumowanie	85
6. MATEMATYKA W BADANIU PISA 2015	87
6.1. Umiejętności polskich uczniów	87
6.2. Zmiany w poziomach osiągnięć	91
6.3. Porównanie wyników chłopców i dziewcząt	94
6.4. Rozumowanie matematyczne	95
6.5. Wieloletni trend zmiany wyników	96
6.6. Podsumowanie	97
7. KORZYSTANIE Z NOWYCH TECHNOLOGII	98
7.1. Dostępność urządzeń cyfrowych w domach uczniów	98
7.2. Korzystanie z TIK w domu i szkole	102
7.3. Związek korzystania z urządzeń cyfrowych w domu i szkole z wynikami osiąganymi przez uczniów	104
7.4. Wnioski	107
Bibliografia	108
Aneks	109

WPROWADZENIE

Michał Federowicz, Jacek Haman

1.1. Badanie PISA. Założenia i główne wyniki

W marcu 2015 roku przeprowadzono w Polsce pomiar umiejętności uczniów w ramach wieloletniego międzynarodowego programu badawczego PISA (Programme for International Student Assessment – Program Międzynarodowej Oceny Umiejętności Uczniów) zorganizowanego przez OECD (Organizację Współpracy Gospodarczej i Rozwoju). Począwszy od 2000 roku, w regularnych trzyletnich odstępach, już po raz szósty zmierzono poziom umiejętności uczniów z jednego rocznika, którzy ukończyli piętnasty rok życia.

Metodologiczne rygory stosowane w badaniu PISA pozwalają na porównywanie wyników danego kraju z wynikami innych krajów i regionów uczestniczących w badaniu, dzięki czemu ogląd efektów krajowej edukacji w zakresie sprawdzanych umiejętności zyskuje znacznie szerszą płaszczyznę odniesienia. Tak pomyślane porównawcze miary osiągnięć uczniów, chociaż z natury rzeczy nie obejmują wszystkich ważnych aspektów oświaty, wnoszą istotny wkład w długofalową ocenę systemów szkolnictwa, prowokując do refleksji nad ich dopasowaniem do wyzwań współczesnego świata.

Program PISA zrodził się z konstatacji powziętej w gronie najbardziej rozwiniętych krajów świata o niewystarczającym wypełnianiu swojej roli przez współczesne systemy powszechnej oświaty. Pojawiło się pojęcie funkcjonalnego analfabetyzmu, które określa nieporadność wobec potrzeby realnego wykorzystania nabytych wcześniej elementarnych umiejętności czytania, pisania i liczenia, co w gruncie rzeczy stawia takie osoby w analogicznej sytuacji do sytuacji analfabety. W latach dziewięćdziesiątych zaczęto zdawać sobie sprawę, że ukształtowane w poprzednich pokoleniach systemy edukacyjne w coraz mniejszym stopniu przygotowują swoich absolwentów do radzenia sobie z niepewnością jaką niesie ze sobą zmieniający się świat. Skonsolidowała się też idea uczenia się przez całe życie (LLL), podkreślająca rosnące znaczenie aktywności edukacyjnej osób dorosłych. W tej perspektywie zadaniem powszechnej szkoły – inaczej niż sto, czy nawet pięćdziesiąt lat temu – jest nie tyle i nie tylko nauczenie czytania, pisania i liczenia, czy wyćwiczenia określonych schematów i algorytmów, co wyposażenie młodego człowieka w warsztat samodzielnego uczenia się. Stąd też zrodziło się pojęcie kluczowych kompetencji, z umiejętnością uczenia się jako jedną z nich, które dają podstawę dla dalszego, już samodzielnego rozwijania i poszerzania wiedzy i umiejętności, budowania własnych kompetencji. Szkoła, a dokładniej szkoła powszechna, ma te fundamenty wspólnie z uczniami uformować i w ten sposób przygotować ich do dorosłego życia.

Mała skuteczność reform edukacyjnych w większości rozwiniętych krajów świata skłaniała też do poszukiwania miar ilościowych, które by pozwoliły na usystematyzowanie poczynionych wcześniej obserwacji jakościowych i zbieranie danych przydatnych w długofalowej ocenie podejmowanych przemian. Rzecz jasna, nie da się ilościowo scharakteryzować wszystkich ważnych wymiarów edukacji, tym bardziej w sposób porównywalny międzynarodowo. Niemniej część z nich, przy dzisiejszym zaawansowaniu wiedzy psychometrycznej, można odzwierciedlić na skali probabilistycznej z wystarczająco dobrym przybliżeniem. Z kolei możliwość porównań międzynarodowych stworzona przez zastosowanie takich miar, nawet jeśli obejmują one tylko część oczekiwanych efektów uczenia się, daje nową perspektywę w refleksji nad krajowymi systemami edukacyjnymi i stanem wykształcenia społeczeństwa. Programy PISA, a potem PIAAC (skierowany do ludzi dorosłych: Programme for International Assessment of Adult Competences), powstały dla

wzmocnienia debaty nad dalszym reformowaniem systemów uczenia się, zarówno w szkole jak i poza nią.

1.1.1. Co i jak mierzy badanie PISA?

Siłą badania PISA jest staranne przygotowanie zadań sprawdzających poziom umiejętności uczniów. Są one opracowywane przez około dwa lata, zanim zostaną użyte w badaniu głównym, stanowiącym faktyczny pomiar. Podlegają wielorakim dyskusjom ekspertów i sprawdzeniom, a przede wszystkim są konfrontowane z szerokimi grupami uczniów dla empirycznego zweryfikowania ich przydatności. Początkowa liczba układanych zadań w sposób zamierzony przekracza dwu-trzykrotnie liczbę zadań potrzebnych w pomiarze głównym. W potocznym przekonaniu mocno wrośnięte jest dość stereotypowe i nadmiernie uproszczone rozumienie miar ilościowych stosowanych w edukacji zawarte w słowie „test”. Tymczasem test testowi nierówny. Można ułożyć trywialny test odpytujący z jakiegoś zakresu wiadomości, np. faktów zawartych w lekturze szkolnej, ale można też stworzyć zadania testowe, których rozwiązanie wymaga złożonych procesów myślowych, np. podjęcia własnej strategii rozwiązania danego problemu lub samodzielnego zinterpretowania przedstawionych w treści zadania faktów. Testy w badaniu PISA, czy inaczej mówiąc zestawy zadań, zawierają dużo zadań otwartych, a odpowiedzi na nie są następnie starannie kodowane z zachowaniem porównywalności procesu kodowania i sprawdzaniem jego rzetelności.

Również zadania zamknięte są tak układane, by udzielenie poprawnej odpowiedzi wymagało podjęcia pewnej drogi myślowej, a nie pamiętania wrywkowych wiadomości. Trudność zadania zamkniętego w teście PISA nigdy nie wynika ze szczegółowości pytania o konkretne wiadomości. Zróznicowanie trudności zadań wiąże się ze złożonością procesu myślowego, który trzeba podjąć dla rozwiązania zadania, zarówno otwartego jak i zamkniętego. Rozwiązanie trudniejszych zadań może wymagać odwołania się przez ucznia do jego szerszej wiedzy, ale szczegółowe wiadomości zawsze podane są w samym zadaniu.

W rozdziałach poświęconych poszczególnym dziedzinom badania zawarte są omówienia poziomów umiejętności zadań składających się na skalę PISA, które pozwalają zorientować się w sposobie różnicowania trudności poszczególnych zadań. Każde zadanie wykorzystane w teście jest precyzyjnie przypisane do określonego miejsca na skali punktowej, co wynika z empirycznie oszacowanego stopnia jego trudności. W ten sposób poszczególne poziomy trudności – odzwierciedlone wynikiem punktowym na skali – zyskują klarowną interpretację empiryczną co do zakresu umiejętności, który zademonstrowali uczniowie uzyskując wynik na danym poziomie.

W każdej dziedzinie testowej można wyodrębnić zadania, które sprawdzają dość proste umiejętności, jak np. wyszukiwanie informacji lub zastosowanie znanego algorytmu, oraz zadania które mierzą umiejętności bardziej złożone. Jest to podział o tyle owocny w analizach, że te pierwsze umiejętności często odpowiadają dość rutynowym działaniom szkoły, zaś te drugie pokazują w jakim stopniu szkoła odchodzi od rutyny i wyzwala większą aktywność umysłową uczniów. Przykłady umiejętności złożonych to: interpretacja tekstu, rozumowanie matematyczne, myślenie w kategoriach eksperymentu z zakresu nauk przyrodniczych, czy umiejętność zwięzłego przedstawienia własnej argumentacji lub odróżnienia faktów od opinii, czyli działania wymagające większej samodzielności i wyjścia poza rutynowe schematy postępowania. W każdej dziedzinie badania PISA pewna pula zadań testowych wymaga inwencji ucznia i to takie zadania budują górną część skali. Tym samym zyskujemy narzędzie pozwalające na śledzenie dynamiki efektów uczenia się w skali całego systemu edukacji, z możliwością odniesienia ich do innych systemów.

Dla porządku warto przypomnieć, że w badaniu PISA sprawdzenie umiejętności uczniów jest dokonywane równolegle w trzech dziedzinach: *reading literacy*, *mathematical literacy* oraz *scientific*

literacy. Tłumaczenie tych pojęć na inne języki napotyka na pewne trudności. Warto zwrócić uwagę, że ich zawartość zawiera także umiejętności złożone, o których jest mowa w poprzednim akapicie. Nie chodzi więc wyłącznie o elementarne umiejętności w danym obszarze, takie które przyświecały twórcom powszechnej oświaty w XIX wieku, lecz o pełne spektrum od prostych umiejętności do umiejętności złożonych związanych z daną dziedziną, które wszakże dziś stanowią troskę równie powszechnej co w poprzednich pokoleniach oświaty. Stąd też rozpiętość „w górę” dzisiejszego rozumienia przytoczonych pojęć, o czym może świadczyć fakt, że zadania testowe na elementarne umiejętności, np. czytania i wyszukania podstawowych informacji z tekstu, czy zastosowania prostego algorytmu, znajdują się na pierwszy poziomie skali, zaś wszystkich poziomów jest pięć lub sześć. W kontekście krajów wysoko rozwiniętych dysponowanie umiejętnościami wyłącznie z pierwszego poziomu uznaje się za niewystarczające, gdyż graniczy z ryzykiem wspomnianego wcześniej analfabetyzmu funkcjonalnego.

W miarę pogłębiania refleksji nad charakterem pomiaru i zakresem mierzonych umiejętności, kształtowały się polskie odpowiedniki pojęć opisujących trzy dziedziny sprawdzane w badaniu PISA. *Reading literacy to czytanie i interpretacja* (najszerzej omówione w Raporcie PISA 2009). *Mathematical literacy* to po prostu *matematyka*, niemniej warto podkreślić, że w zakresie sprawdzanych umiejętności wchodzi zarówno umiejętność stosowania algorytmów (narzędzi) jak i rozumowanie matematyczne i inne wymagania ogólne tego przedmiotu. *Scientific literacy to rozumowanie w naukach przyrodniczych*¹, którego zawartość – jako głównej dziedziny badania PISA 2015 – obszernie omówiona jest w niniejszym raporcie, wraz z podaniem przykładowych zadań.

Można powiedzieć, że zawartość głównych pojęć wyznaczających badanie PISA częściowo – w zakresie mierzonych umiejętności – odzwierciedla ewoluowanie myślenia o szkole i powszechnej oświacie. Trudno wyobrazić sobie, by miarą sukcesu powszechnej szkoły miało być, jak w XIX wieku, samo nauczanie dzieci czytania, pisanie i liczenia, w czasach gdy w naszym kręgu kulturowym coraz więcej dzieci przychodzi do szkoły już z tymi elementarnymi umiejętnościami. O sukcesie szkoły i powszechnego szkolnictwa decyduje jak duża część młodzieży jest w stanie rozwinąć umiejętności złożone, związane z najważniejszymi dziedzinami, a także jak mały odsetek młodzieży pozostaje na poziomie uznanym za minimum w danym kręgu kulturowym lub poniżej minimum. Jest sprawą poszczególnych systemów oświaty w jaki sposób modeluje sam proces uczenia się, jaką nadaje mu strukturę, jak kształtuje programy, jakie promuje metody. Niemniej w końcowej fazie powszechnej szkoły sensowne jest zestawienie ze sobą rezultatów, jakie ten czy inny system przynosi, przynajmniej w takim zakresie, w jakim daje się te rezultaty rzetelnie zmierzyć i porównać.

Przedstawione tu podejście pozwala abstrahować od programów szkolnych krajów uczestniczących w badaniu. PISA nie sprawdza uczniów czy pilnie się uczyli i wypełniali zadany w szkole program, sprawdza systemy szkolnictwa, czy wyposażają uczniów w potrzebny im kapitał założycielski najważniejszych umiejętności, który pozwoli im na pomnażanie własnych kompetencji w dalszej drodze edukacyjnej, również w dorosłym życiu. Fakt przystąpienia do badań tak wielu krajów z różnorodnych kręgów kulturowych świadczy o generalnej akceptacji dla takiego podejścia.

1.1.2. Krótki i długi cykl badania PISA – kwestia porównywalności w czasie

Wymienione trzy dziedziny: czytanie i interpretacja, matematyka oraz rozumowanie w naukach przyrodniczych obecne są w każdym badaniu głównym PISA. Ponadto na przestrzeni sześciu już

¹ Zespół PISA zachowuje pamięć i uznanie dla ś.p. Profesor Janiny Zawadowskiej, której zawdzięczamy propozycję wprowadzenia terminu rozumowania do polskiego pojęcia określającego umiejętności związane z misją szkoły w zakresie przedmiotów przyrodniczych.

pomiarów, w latach 2000-2015, występowały rozmaite dziedziny dodatkowe, które jednak na ogół były podejmowane przez znacznie mniejszą liczbę krajów, ponadto nie były powtarzane w usystematyzowany sposób. Służyły raczej sprawdzaniu możliwości dalszego rozwijania samego programu badawczego. Zatem trzy dziedziny, o których była mowa w poprzednim punkcie stanowią trzon i główną zawartość badania PISA.

Wszystkie trzy dziedziny są obecne w każdym badaniu. Cyklicznie, co trzy lata, główny pomiar PISA dostarcza danych dla oszacowania umiejętności uczniów urodzonych w tym samym roku, którzy ukończyli 15 lat i są objęci systemem szkolnym, niezależnie od tego w której klasie i w jakiego rodzaju szkole akurat się znajdują. Jednocześnie, ponieważ czas jaki podczas badania można przeznaczyć na rozwiązywanie przez uczniów zadań jest ograniczony, jedna z *tych* dziedzin jest wiodąca dla danego pomiaru, co oznacza, że liczba zadań z tej dziedziny jest większa niż w każdej z dwóch pozostałych. Obok zadań zaczerpniętych z poprzednich cykli badania w dziedzinie wiodącej wykorzystuje się także nowe zadania i na ogół poddaje się też rewizji założenia teoretyczne pomiaru. Pozostałe dwie dziedziny bazują na strukturze pomiaru z wcześniejszym cyklu, w którym każda z nich była wiodąca, z wykorzystaniem najlepszych zadań odzwierciedlających pełny zakres sprawdzanych umiejętności.

Z tego wynika, że oprócz trzyletniego cyklu badania PISA powstaje też długi, dziewięcioletni cykl, liczony pomiędzy pomiarami z tą samą dziedziną jako wiodącą. Ponieważ dziedzina wiodąca w danym badaniu jest najpełniej odzwierciedlona w zadaniach, oszacowanie jej wyników jest najdokładniejsze. Zatem porównanie wyników w długich cyklach jest obciążone najmniejszym błędem. Dla czytania i interpretacji długi cykl, czyli najlepsze porównania w czasie następują pomiędzy latami 2000 i 2009, a dla matematyki pomiędzy latami 2003 i 2012, przy czym przygotowując pomiar w 2003 roku zasadniczo rozbudowano zakres sprawdzanych umiejętności matematycznych, co oznacza, że zestawiając ze sobą wyniki dla matematyki z roku 2000 z późniejszymi należy zachować ostrożność w ich interpretowaniu. Analogicznie dla rozumowania w naukach przyrodniczych dopiero rok 2006 dostarczył pierwszego rozbudowanego pomiaru umiejętności i ściśle porównywanie wyników liczbowych z wcześniejszymi latami jest obciążone błędem systematycznym.

Ujmując to samo zagadnienie od drugiej strony, można stwierdzić, że uprawnione jest wyliczenie różnic punktowych pomiędzy kolejnymi pomiarami w następujących przedziałach czasowych: dla czytania i interpretacji od początku badania, czyli od roku 2000 do roku 2012, zaś dla matematyki od roku 2003 do 2012. Dla rozumowania w naukach przyrodniczych pomiarem „założycielskim” był pomiar w roku 2006 i do niego można porównywać wyniki późniejsze.

Do kwestii pewnego ewoluowania struktury teoretycznej pomiaru, co jest nieuniknione w miarę upływu czasu, dochodzi zmiana jaka dokonała się pomiędzy latami 2012 i 2015, tzn. przejściu w podstawowej części badania dla wszystkich trzech dziedzin z rozwiązywania zadań na arkuszu papierowym na rozwiązywanie zadań na komputerze.

Trudno oczekiwać, by długofalowy program badawczy zastygł w swoim pierwotnym kształcie i nie modyfikował podejścia metodologicznego. Międzynarodowe konsorcjum dokonało wszelkich możliwych zabiegów, by maksymalnie ograniczyć wpływ metody pomiaru na wyniki. Jednak oszacowanie „efektu narzędzia”, jakie przedstawiamy w pierwszym rozdziale niniejszego raportu nakazuje zachowanie daleko posuniętej ostrożności w wyciąganiu wniosków z zestawienia wyników liczbowych z roku 2015 z wcześniejszymi. W raporcie zestawiamy ze sobą te wyniki, jednak powstrzymujemy się przed wyciąganiem wniosków, a dostępne nam dane poddane starannej analizie sugerują, że w zasadzie rok 2015 należy traktować jako nowe otwarcie w badaniu PISA z użyciem nowego medium pomiarowego.

1.1.3. Zakres wnioskowania i interpretacji wyników

Co do porównań w czasie, powtórzmy raz jeszcze, całkowicie zasadne jest porównywanie każdego wyniku dla czytania i interpretacji w okresie od roku 2000 do 2012, dla matematyki w okresie 2003 – 2012, a dla rozumowania w naukach przyrodniczych w okresie 2006 – 2012. Pomiar w roku 2015 mimo starań konsorcjum trzeba jednak potraktować odrębnie, jako zapoczątkowanie nowej serii badań. Jak piszemy w rozdziale pierwszym, różnice pomiędzy wynikami w 2015 a wynikami z roku 2012 w każdej z trzech dziedzin w przypadku Polski najprawdopodobniej w przeważającej mierze wynikają ze zmiany medium użytego w badaniu, przy czym dokładne oszacowanie tego zjawiska, a tym bardziej ustalenie jego przyczyn wymagałoby odrębnych badań.

Oczywiście zachowuje swoją moc porównanie wyników różnych krajów i regionów w ramach badania w 2015 roku, tak samo jak uprawnione jest wnioskowanie o pozycji polskich uczniów wobec innych krajów np. w roku 2000, czy w jakimkolwiek innym roku i dotyczy to każdej z trzech dziedzin.

Skale zastosowane w badaniu PISA mają tę właściwość, że różnice punktowe odzwierciedlają konkretne różnice zakresu umiejętności (por. opisy w rozdziałach poświęconych kolejnym dziedzinom pomiaru). Przedstawione tu ograniczenia mają o tyle znaczenie, o ile interpretacje mają opierać się na wyliczeniach różnic punktowych. Jeśli jednak pozostajemy na poziomie ogólnego uchwycenia długofalowych tendencji, pozostaje ono uprawnione.

Tak jak pisaliśmy, siłą badania PISA jest staranny pomiar umiejętności, utrzymywany na najwyższym poziomie światowym. Dotyczy to zarówno sposobu tworzenia zadań testowych, który odbiega od szkolnej rutyny, jak i kodowania odpowiedzi uczniów oraz obróbki statystycznej uzyskanych danych z zastosowaniem zaawansowanych technik probabilistycznych. Jak już wspomniano, głównym celem zastosowanych modeli psychometrycznych jest minimalizacja błędu oszacowania poziomu umiejętności dla badanej populacji uczniów krajów uczestniczących, co czyni porównania międzynarodowe w danym roku badania najbardziej rzetelnymi. Perspektywa zakotwiczenia zestawu sprawdzanych umiejętności w wyobrażeniu, czego najbardziej potrzeba młodemu człowiekowi dla dalszego powodzenia w dorosłym życiu pozwala na to, by program badawczy niejako o pół kroku wyprzedzał szkołę.

Siłą rzeczy badanie stawiające sobie za cel sprawdzanie w jakim stopniu szkoła przygotowuje uczniów do radzenia sobie w dorosłym życiu nie może cały czas mierzyć idealnie tego samego. W dłuższym okresie byłoby to absurdalne. Pomiaru umiejętności nie można porównać do mierzenia jakiejś wielkości fizycznej, w której raz ustalony wzorzec zachowuje swoją przydatność na wiele pokoleń. Skoro jednak sam pomiar ewoluuje, mimo minimalizowania efektu zmian w czasie metodami psychometrycznymi, musi się to odbyć kosztem zmniejszonej porównywalności w czasie. Dla zobrazowania tego dylematu można odwołać się do badania takich wielkości jak np. inflacja, dla której wskaźnikiem jest przyrost cen ustalonego koszyka produktów. W dłuższej perspektywie, by wskaźnik taki zachował swoją adekwatność do współistniejących procesów, sama zawartość koszyka produktów musi ulegać zmianie. Struktura pomiaru umiejętności jest daleko bardziej złożona, niż przytoczony przykład i zawsze trzeba dokonywać wyborów pomiędzy jej aktualizowaniem i zachowaniem pewnej stabilności. Badanie PISA radzi sobie z tym dylematem z zachowaniem najwyższych standardów dostępnej dziś wiedzy. Przejście z rozwiązywania zadań na papierze do rozwiązywania ich na ekranie komputera kiedyś nastąpić musiało. Nie zmienia to faktu, że w zasadzie od roku 2015 jest to już nowy program badawczy, który jest wprawdzie silnie osadzony w poprzednim, ale dostarcza już nieco innych informacji. Tak jak po roku 2000 należało powstrzymać się z nadmiernymi interpretacjami wyników do uzyskania wyników w kolejnych cyklach, tak i teraz należy zachować cierpliwość. Zwłaszcza wyciąganie daleko idących wniosków na temat zmiany pomiędzy latami 2012 a 2015 byłoby przedwczesne. Niemniej dokonana zmiana pomiędzy latami 2012 a 2015 nie rzutuje na podstawowy walor badania PISA, jakim jest omawiany już staranny pomiar umiejętności w danym roku.

Walorem badania PISA jest także możliwość powiązania wyników pomiaru umiejętności z cechami statusowymi ucznia, czyli np. wykształceniem i statusem zawodowym jego rodziców, czy zasobami jego domu rodzinnego, które mogą mieć związek z nabywaniem umiejętności. W tym aspekcie także interesujące są porównania międzynarodowe. W dalszej części wprowadzenia przedstawiamy takie analizy.

Warto także zauważyć, że oprócz umiejętności uczniów badanie PISA dostarcza też wielu innych informacji, czerpanych głównie z ankiety ucznia i ankiety dyrektora szkoły. Trzeba jednak mieć na uwadze, że nie da się zrealizować takiego programu badawczego, który by obejmował wszystkie aspekty wybranego obszaru rzeczywistości społecznej jednocześnie. Każde badanie musi mieć określone priorytety badawcze i koncentrować na nich swoją metodologię i procedury. Zbieranie dodatkowych informacji jest możliwe, lecz w przeciwieństwie do danych wytwarzanych priorytetowo trzeba je traktować jako „miękkie” dane kontekstowe. Przykładowo, jeśli zadamy uczniom kilka pytań o ich relacje z nauczycielami czy inne wskaźniki klimatu szkoły, to oczywiście dostaniemy jakieś odpowiedzi. Mogą one wprawdzie pomóc w analizach głównych wyników na teście jako zmienne kontekstowe, ale już nie bardzo można traktować te odpowiedzi jako źródło wiedzy na temat samego klimatu szkoły. Trzeba pamiętać, że uczniowie wypełniający ankietę są po dwóch sześćdziesięciminutowych sesjach rozwiązywania zadań i może działać efekt zmęczenia, a jednocześnie liczba pytań dotycząca relacji z nauczycielami z konieczności musi być bardzo ograniczona. Jak wiadomo, badanie klimatu szkoły wymaga odrębnych badań, dla których stanowiłoby cel priorytetowy. Tym bardziej odpowiedzi 150 czy 200 dyrektorów szkół, pomimo że szkoły te dobrane są losowo, dostarczają jedynie orientacyjnych, miękkich informacji. Przykładowo, jeśli zapytamy dyrektorów, czy postrzegają problemy kadrowe w swojej szkole jako istotną barierę w procesie nauczania, to ciekawe będzie zestawienia takich odpowiedzi z pomierzonymi umiejętnościami uczniów, ale jeśli chcemy dowiedzieć się czegoś więcej o problemach kadrowych szkół, należy zaprojektować odrębne badanie. Dodatkowe informacje czerpane z badania PISA mają charakter „miękkich” danych, mogą oczywiście być interesujące i służyć budowaniu pewnych hipotez, które jednak następnie należałoby poddać falsyfikacji w odrębnym trybie, z metodologią adekwatną dla ich rozstrzygnięcia.

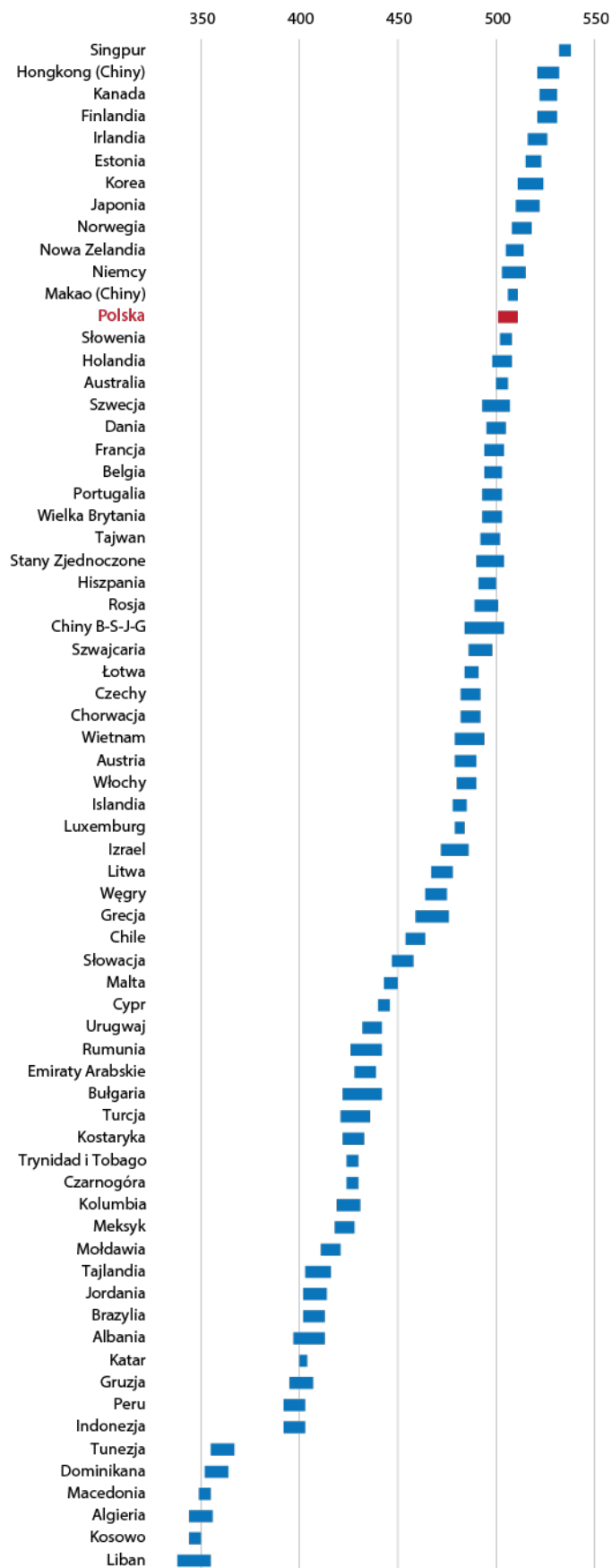
1.1.4. Główne wyniki

W niniejszym raporcie, tak jak i we wcześniejszych polskich raportach, skupiamy uwagę na priorytetowych dla badania PISA wynikach, pochodzących z pomiaru umiejętności uczniów oraz bezpośrednio z nimi związanych. W tej części Wprowadzenia przedstawiamy najpierw główne wyniki liczbowe. Ze względów wspomnianych już uprzednio na wykresach pokazujących wyniki w poszczególnych dziedzinach w całym okresie badania PISA liniami ciągłymi połączone są te pomiary, których wyniki można bezpośrednio ze sobą porównywać. Pozostałe stanowią w pewnym stopniu odrębny pomiar, który jednak mimo ograniczonej porównywalności daje szerszy ogląd sytuacji. Wyniki te są szczegółowo omówione w kolejnych rozdziałach Raportu. We Wprowadzeniu ponadto zestawiamy charakterystyki zróżnicowania wyników w powiązaniu z cechami statusowymi ucznia dla Polski i innych krajów, co także mówi o istotnych cechach systemów edukacji.

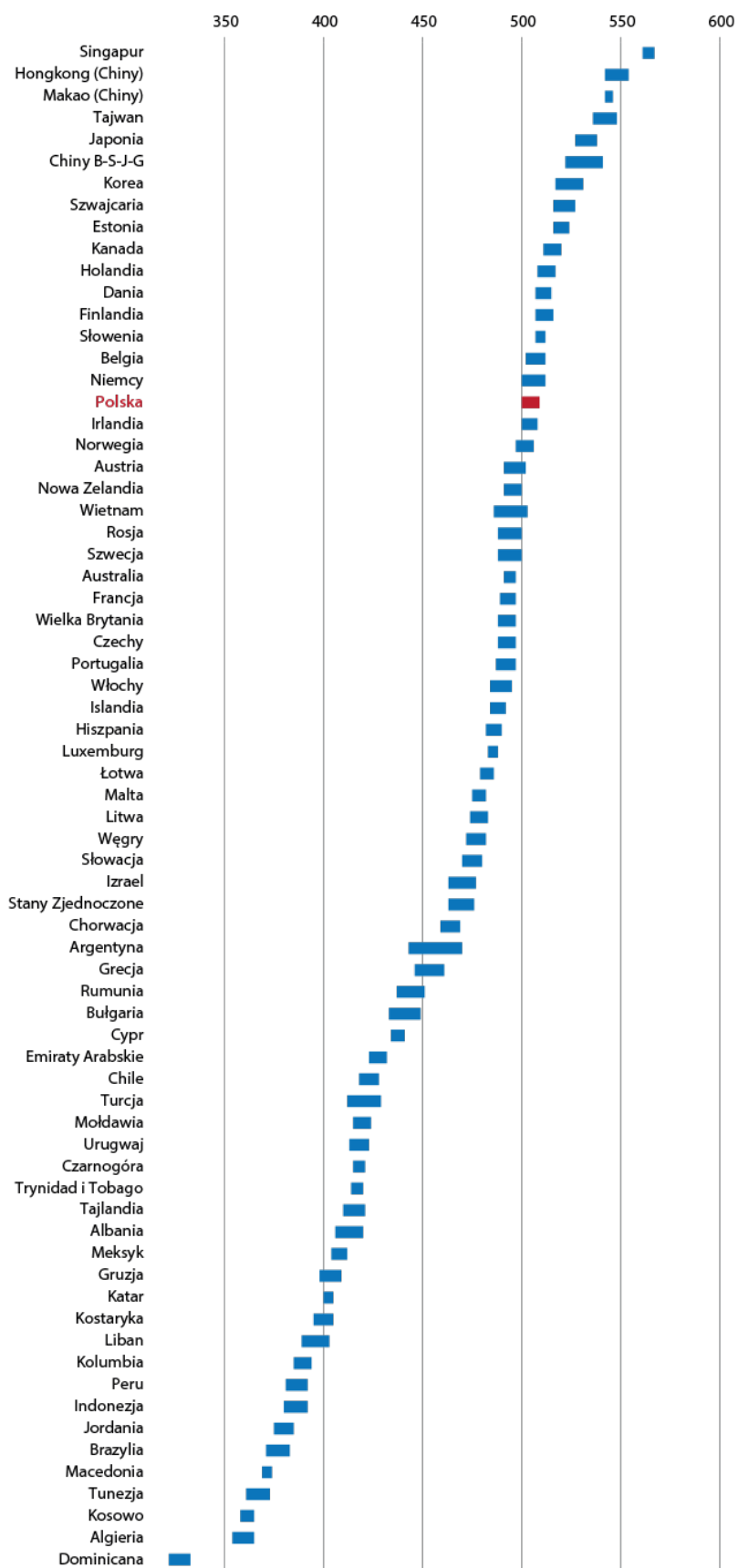
W badaniu PISA 2015 uczestniczyło ponad 500 tys. uczniów z 72 krajów i regionów. Badaniem objęto losową próbę uczniów, którzy w roku poprzedzającym badanie ukończyli lat 15; w 2015 roku byli to uczniowie urodzeni w 1999 r.

Obniżenie wyników pomiędzy latami 2012 i 2015 odnotowano w większości krajów OECD. Dotyczy to zwłaszcza krajów w wynikami powyżej średniej OECD, czyli takimi jak Polska, trudno jednak doszukać się w tej zmianie bardziej systematycznych prawidłowości. Podobnie analiza zmian na poziomie pojedynczych zadań testowych nie wykazuje istotnych tendencji.

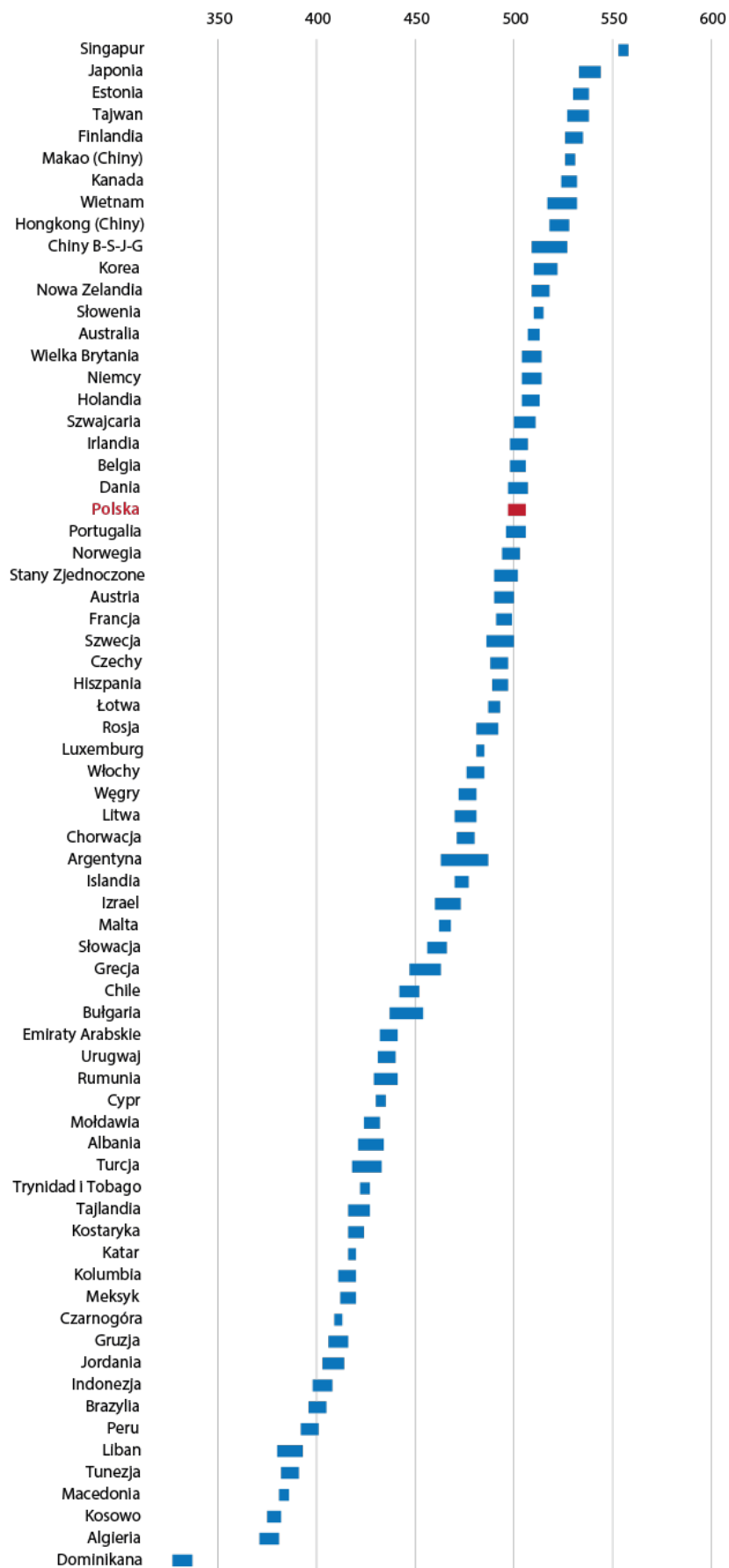
Wykres 1.1 Czytanie



Wykres 1.2 Matematyka

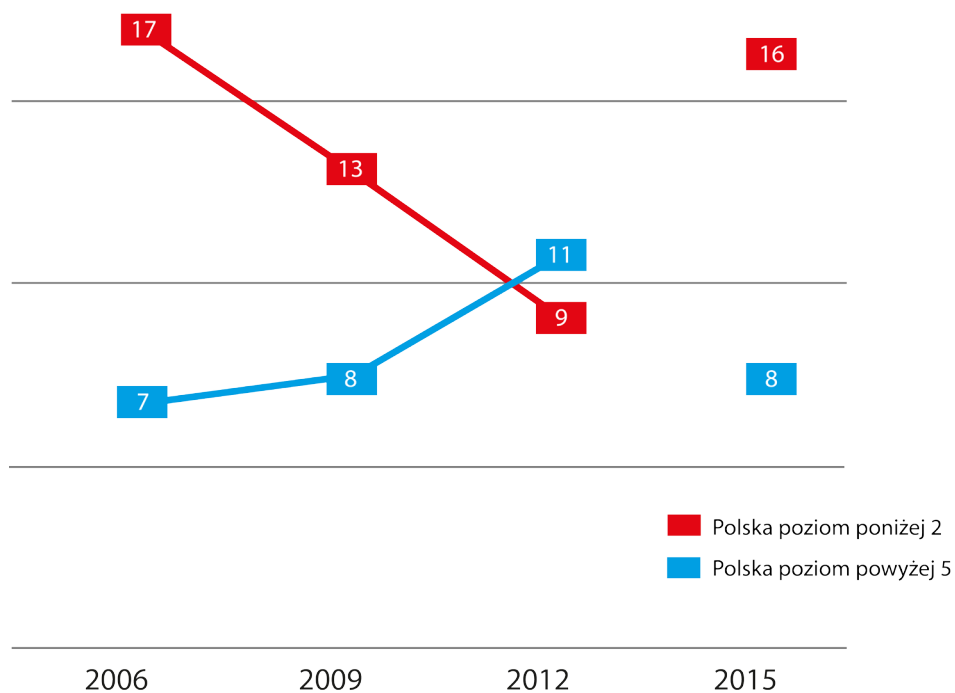


Wykres 1.3 Przyroda

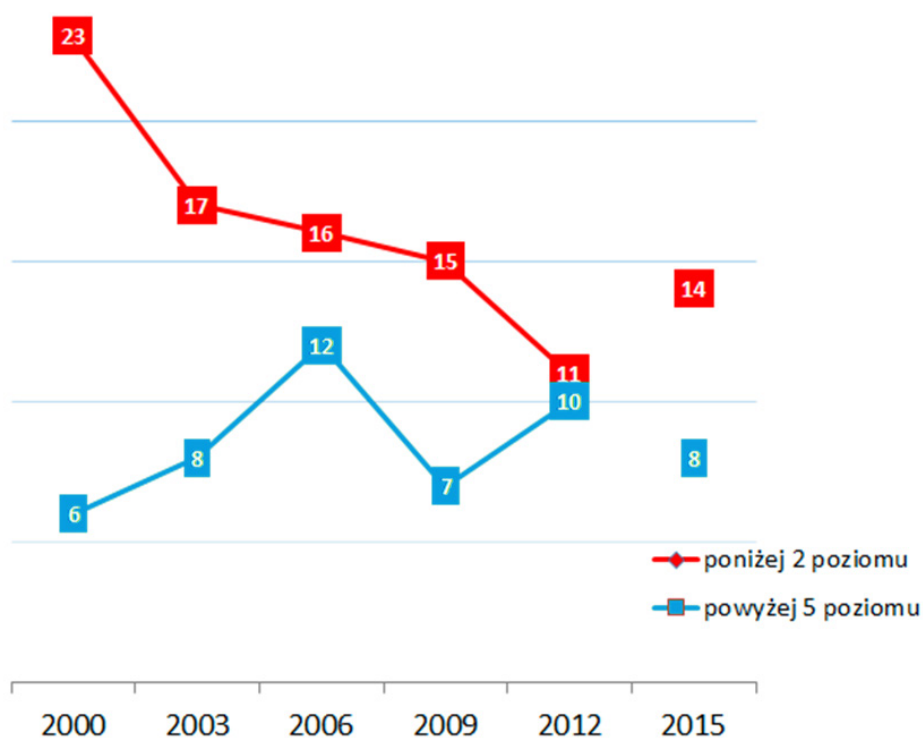


Wykres 1.4 Odsetek uczniów, którzy uzyskiwali wyniki poniżej 2 poziomu oraz powyżej 5 poziomu w Polsce w kolejnych edycjach badania (%)

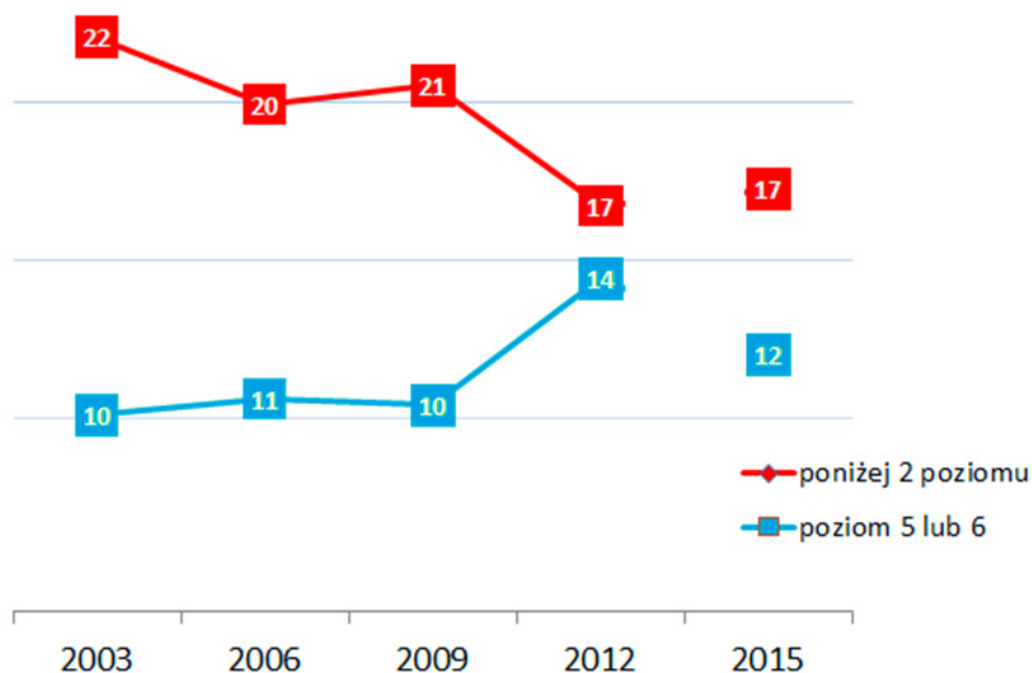
a. Przyroda



b. Czytanie i interpretacja



c. matematyka



Rozumowanie w naukach przyrodniczych: wynik polskich uczniów to 501 pkt – o 24 pkt. mniej niż w 2012 r, co jednak prawdopodobnie, jak pokazują analizy zawarte w rozdziale pierwszym, jest głównie efektem zmiany sposobu testowania. Spośród krajów Unii Europejskiej lepsze wyniki od Polski uzyskało 9 krajów (w tym, w przypadku 3, różnica nie jest istotna statystycznie). W czołówce krajów osiągających najwyższe wyniki w badaniu kompetencji przyrodniczych znajdują się kraje Dalekiego Wschodu, a także Finlandia, Estonia i Kanada. Pierwsze miejsce zajmuje Singapur (556 pkt). Z krajów europejskich przoduje Estonia z wynikiem 534 punktów, a w czołówce znajduje się też Finlandia (531 pkt).

W czytaniu i interpretacji średni wynik Polski to 506 pkt – o 12 pkt mniej niż w 2012 r, ten spadek, analogicznie jak w rozumowaniu w naukach przyrodniczych, należy przypisać głównie zmianie medium pomiaru. Spośród krajów UE wyższą średnią osiągnęli uczniowie z Finlandii, Irlandii i Estonii oraz Niemiec (w ostatnim przypadku różnica nie jest jednak istotna statystycznie); wyższy od polskiego jest także wynik uczniów z nienależącej do UE Norwegii. Wśród krajów świata Polska znalazła się na 13 miejscu. Warto zaznaczyć, że jedynie 9 krajów uzyskało wynik statystycznie istotnie lepszy od Polski, zaś w przypadku 3 ich przewaga mieściła się w granicach błędów losowych.

Średni wynik polskich uczniów w zakresie matematyki to 504 pkt., o 14 pkt. mniej niż w 2012 r, co analogicznie, wiąże się głównie z rozwiązywaniem zadań matematycznych na komputerze, a nie na arkuszu papierowym. W Europie polscy uczniowie są na 8. pozycji, z wynikiem zbliżonym do wyników Belgii, Niemiec, Irlandii i Norwegii. W UE lepsze od Polski wyniki uzyskało 7 krajów: Estonia, Holandia, Dania, Finlandia i Słowenia oraz (na poziomie nieistotnym statystycznie) Belgia i Niemcy. Ważnym wynikiem dla matematyki jest, pomimo spadku punktowego na skali, potwierdzenie z roku 2012 generalnie lepiej rozwiniętej umiejętności rozumowania matematycznego wśród trzecioklasistów

gimnazjalnych w Polsce w porównaniu z uczniami w krajach OECD. Do kwestii tej wrócimy w końcowej części Wprowadzenia.

W pierwszym rozdziale podajemy przybliżone oszacowanie punktowe „efektu komputera” pomiędzy latami 2012 a 2015 w odniesieniu do wyników uzyskanych w Polsce. Trudno nam ocenić, jakie znaczenie mógł mieć ten efekt w przypadku innych krajów; metodologia zastosowana przez Konsorcjum przy skalowaniu wyników miała tę właściwość, że wprowadzie w skali całego badania „efekt komputera” był neutralny, to w przypadku poszczególnych krajów mógł być on zarówno pozytywny, jak i negatywny, jednakże jego ocena wymagałaby w każdym przypadku przeprowadzenia szczegółowych analiz, w tym także odwołujących się do danych, do których nie mamy dostępu. Jak już wspomniano, w zaobserwowanej zmianie trudno doszukać się prawidłowości. Przykładowo, w Anglii wyniki wszystkich trzech dziedzin pomiędzy latami 2012 i 2015 nie uległy istotnym zmianom, tymczasem w Walii i Szkocji wyraźnie obniżyły się. Trudno przypuszczać by, dajmy na to, wykorzystanie komputerów w szkole tak znacząco odbiegało od siebie różnych częściach Zjednoczonego Królestwa.

Wiele wskazuje, że wprowadzenie do pomiaru testów komputerowych miało w przypadku Polski ponadprzeciętnie duży wpływ na wyniki. Może to oznaczać, że badanie PISA 2015 sygnalizuje pewien problem występujący w polskich szkołach, ale do zdiagnozowania tego problemu należy zaprojektować odpowiednio dobraną metodologię i przeprowadzić pogłębione badanie krajowe. W raporcie, z wyjątkiem pierwszego rozdziału, nie podejmujemy szerzej tego problemu, gdyż samo badanie PISA nie dostarcza w tym celu wystarczających danych. Niemniej warto zasygnalizować, że sprawa obecności komputera w procesie dydaktycznym jest godna uwagi i pogłębionych badań. Przy czym chodzi nie tyle o to jak polskie dzieci mogą wypaść na takim czy innym teście komputerowym, co przede wszystkim o to, na ile i w jaki sposób nowe media powinny być włączone w proces dydaktyczny.

W raporcie w kilku miejscach podkreślamy potrzebę zachowania ostrożności dla uniknięcia mechanicznego wprowadzania nowych technologii. Przykładowo, nie jest oczywiste, że dzieci i młodzież rozwiązując problemy matematyczne z użyciem komputera będą lepiej rozwijały rozumowanie matematyczne, niż podejmując te problemy z ołówkiem i kartką papieru przed sobą. Tak samo nie działa żaden automat, który by sprawił, że umiejętność interpretowania tekstu rozwija się lepiej, gdy do napisania własnej interpretacji używamy klawiatury zamiast pióra. Podobnie, eksperyment w szkolnym laboratorium na przedmiotach przyrodniczych zachowuje swoją moc oddziaływania na młody umysł, niezależnie od tego na ile może być wspomagany komputerem, a nawet najlepsze symulacje komputerowe go nie zastąpią. Z drugiej jednak strony, można wyobrazić sobie zajęcia z matematyki niekiedy z wykorzystaniem komputera, pracę nad tekstem i informacjami widzianymi na ekranie, czy wykorzystanie symulacji komputerowych jako elementu wspomagającego zajęcia w laboratorium przyrodniczym.

Nie ma wątpliwości co do tego, że absolwenci obecnych szkół w swoim dorosłym życiu w coraz większym stopniu będą korzystali z rozmaitych urządzeń elektronicznych, zarówno w pracy jak i w sprawach zarówno osobistych jak i publicznych. Zatem szkoła powinna rozwijać umiejętność świadomego korzystania z tego rodzaju urządzeń. Wynik z 2015 roku pozwala na postawienie hipotezy, że polska szkoła nie robi tego w stopniu zadowalającym. Tę hipotezę trzeba starannie i wielostronnie sprawdzić, zanim podejmie się działania, by uniknąć pozornego rozwiązania problemu. Niemniej problem obecności nowych mediów w szkole i świadomego z nich korzystania istnieje.

Ważnego rodzaju wynikiem badania PISA jest informacja o odsetku uczniów najsłabszych i odsetku uczniów najlepszych w danym kraju. W dyskusji nad efektami osiąganymi przez systemy edukacyjne krajów wysokorozwiniętych przyjęto, że na pięcio- lub sześćcio-przedziałowej (ich szczegółowe omówienie zawierają rozdziały 2, 3, 4) skali osiągnięć uczniów, wynik poniżej poziomu drugiego sygnalizuje zagrożenie niepowodzeniem w dalszej karierze edukacyjnej, które może mieć negatywne

skutki w dorosłym życiu. Odsetek uczniów którzy nie wykazują umiejętności do najmniej z poziomu drugiego jest traktowany jako wskaźnik wielkości grupy ryzyka generowanej w systemie oświaty. Stąd zalecenie Unii europejskiej by do roku 2020 państwa ograniczyły ten wskaźnik poniżej 15%. Jest to granica umowna, ważne jest jednak samo postawienie takiego celu oraz zastosowanie miary porównywalnej dla wszystkich krajów. Analogicznie sprawdza się też odsetek uczniów osiągających umiejętności najwyższe, co najmniej na piątym poziomie, jako komplementarny wskaźnik charakteryzujący system edukacji.

W zakresie rozumowania w naukach przyrodniczych odsetek uczniów najslabszych, czyli poniżej drugiego poziomu umiejętności, w roku 2015 wyniósł 16,3% (w 2012 r. było to 9%, a w 2006 17%). Trudno jednoznacznie określić, na ile zmiana pomiędzy latami 2012 a 2015 wynikała z zastosowania pomiaru komputerowego. Podobne zmiany nastąpiły w pozostałych dziedzinach: w matematyce – 17,2% w 2015 r. (w 2012r. – 14,4%), a w zakresie czytania i interpretacji – 14,4% (w 2012 r. – 10,6%). Analogicznie uczniów z najlepszymi wynikami w rozumowaniu w naukach przyrodniczych było 7,3% w 2015 r. (w 2012 r. – 10,8%, a w 2006 – 6,8%). W matematyce w 2015 r. najlepszych uczniów było 12,2% (w 2012 r. – 16,7%), a w czytaniu – 8,2% (w 2012 r. – 10,1%). Ze względu na zmianę techniki pomiaru wydaje się, że ważniejsze jest odniesienie polskich wyników do wyników innych krajów. Zarówno pod względem odsetka najlepszych, jak i najslabszych uczniów wynik Polski jest lepszy niż średnio w krajach OECD, z wyjątkiem odsetka uczniów najlepszych w czytaniu i interpretacji, który jest nieco niższy i obszar ten wymaga wsparcia.

W rozumowaniu w naukach przyrodniczych w 2015 r. średni wynik chłopców w krajach OECD wyniósł 495 punktów i był o 4 pkt. wyższy od średniego wyniku dziewcząt. W Polsce średni wynik dziewcząt wyniósł 498 pkt., podczas gdy chłopcy osiągnęli 504 pkt.

Także w zakresie matematyki odnotowano przewagę chłopców: ich średni wynik to 511 punktów, a dziewcząt – 499 punktów. Po raz pierwszy od lat przewaga chłopców nad dziewczętami w matematyce jest istotna statystycznie (w latach 2009 i 2012 wynosiła tylko 4 punkty). Przewaga chłopców w tej dziedzinie dotyczy także dolnego i górnego krańca skali umiejętności – wśród najslabszych uczniów (najniższy decyl) wyniki chłopców są o 9 punktów wyższe niż dziewcząt, a wśród najlepszych aż o 19 punktów wyższe.

Natomiast w zakresie czytania i interpretacji dziewczęta są zdecydowanie lepsze: w Polsce różnica wyniosła 30 pkt na korzyść dziewcząt, jest ona zbliżona do średniej OECD (27 p.). Dziewczęta uzyskały 521 pkt, a chłopcy 491.

We wszystkich trzech dziedzinach chłopcy poprawili swoje wyniki względem dziewcząt pomiędzy latami 2012 a 2015, ale trudno określić czy i w jakim stopniu wpłynęła na to zmiana pomiaru na komputerowy.

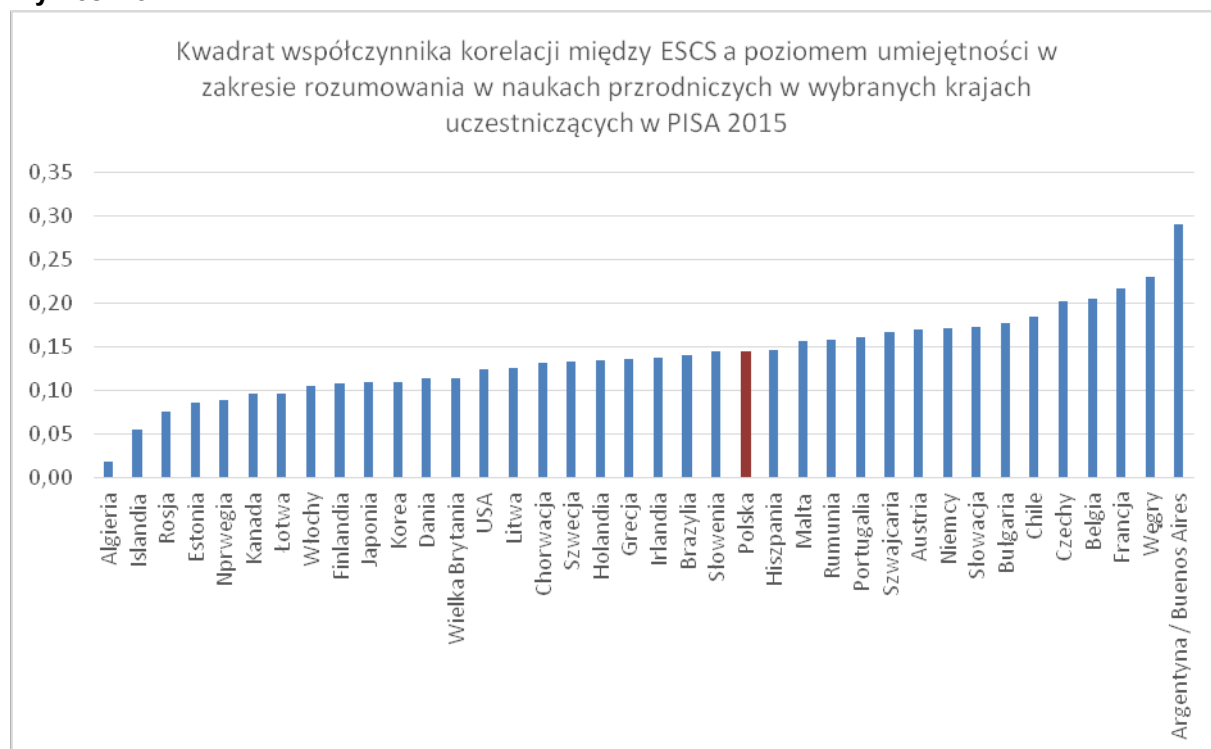
W badaniach z 2006 oraz 2015, czyli w latach z rozumowaniem w naukach przyrodniczych jako dziedziną wiodącą, pytano uczniów o różne aspekty ich postaw wobec nauk przyrodniczych, w tym, czy uczenie się przedmiotów przyrodniczych oraz poznawanie zagadnień naukowych sprawia im przyjemność i daje satysfakcję. Na podstawie tych odpowiedzi opracowano wskaźnik satysfakcji z uczenia się przedmiotów przyrodniczych (enjoyment of learning science). Krajami, w których wskaźnik ten wzrósł najbardziej między badaniami 2006 i 2015, były Irlandia i Polska. Uczniowie byli też pytani o to, jak sądzą, jaki zawód będą wykonywać w wieku 30 lat. Między 2006 a 2015 nie zmienił się odsetek uczniów wskazujących na zawody inżynierskie i bezpośrednio związane z nauką (te zawody wskazało w 2015 r. ok. 6% uczniów), zwiększył się odsetek wskazujących na zawody medyczne (wzrost z 8 do 12%), zmniejszył się natomiast odsetek uczniów wskazujących na zawody informatyczne (z 6 do 1%) oraz pozostałe zawody związane z nauką (z 6 do 1%).

O ile przypuszczenia piętnastolatków jaki zawód będą wykonywali w wieku 30 lat stanowią dość miękką informację, o tyle zależność osiągnięć uczniów od cech ich środowiska domowego jest mocno ugruntowana w analizach PISA. Jak już wspominaliśmy, walorem tych badań jest możliwość skojarzenia starannie zmierzonych umiejętności piętnastolatków z zebranymi informacjami określającymi pozycję społeczno-ekonomiczną uczniów, czyli cechami środowiska domowego w jakim wzrastają.

Podstawowym wskaźnikiem pozycji społeczno-ekonomicznej stosowanym w badaniu PISA jest indeks ESCS (indeks statusu ekonomiczno-społeczno-kulturalnego), w którego konstrukcji uwzględniane są: pozycja społeczno-ekonomiczna zawodów rodziców mierzona wskaźnikiem SEI, poziom wykształcenia rodziców, indeks zamożności rodziny, oraz indeksy zasobów edukacyjnych i kulturalnych rodziny ucznia.

Kwadrat współczynnika korelacji liniowej między wskaźnikiem ESCS a wynikami w teście PISA z rozumowania w naukach przyrodniczych wynosi w Polsce 0,14, innymi słowy, 14% wariacji umiejętności można wyjaśnić jego liniowym² związkiem ze zróżnicowaniem pozycji społeczno-ekonomicznej³. Jest to wartość podobna, jak w większości innych krajów i regionów biorących udział w badaniu PISA 2015. Średnia dla badanych społeczności wynosi 13%, przy czym omawiana zależność jest najsłabsza w takich krajach jak Algieria (2%), Islandia (6%), Rosja (8%) oraz Estonia i Norwegia (9%), zaś najsilniejsza na Węgrzech (23%), we Francji (22%), w Belgii (21%) i Czechach (20%) oraz w społeczności Benos Aires (0,29, Argentyna).

Wykres 1.5



² Założenie, że związek ten ma charakter prostej regresji liniowej jest oczywiście pewnym uproszczeniem, jednakże zastosowanie modeli krzywoliniowych prowadzi do poprawy jakości przewidywania (przyrostu R^2) o nie więcej niż 0,01.

³ Dla pozostałych dziedzin wartość tego współczynnika w Polsce również wynosi 0,14; w innych krajach różnice jego wartości dla poszczególnych dziedzin są również zwykle bardzo niewielkie.

Czynniki, które decydują o wyższej lub niższej wartości współczynnika w poszczególnych krajach mogą mieć bardzo różny charakter i bez analizowania każdego przypadku z osobna trudno się o nich wypowiadać. W każdym jednak razie, silna zależność poziomu umiejętności od charakterystyki społeczno-ekonomicznej rodziny może oznaczać ograniczenie możliwości awansu społecznego zdolnych dzieci z niższych warstw społecznych, czy utrudnienie w integracji upośledzonych grup mniejszościowych. Korelacja między pozycją społeczną rodziny a umiejętnościami ucznia nie jest w Polsce specjalnie silna; w porównaniu z innymi krajami nie można jej jednak uznać za niską.

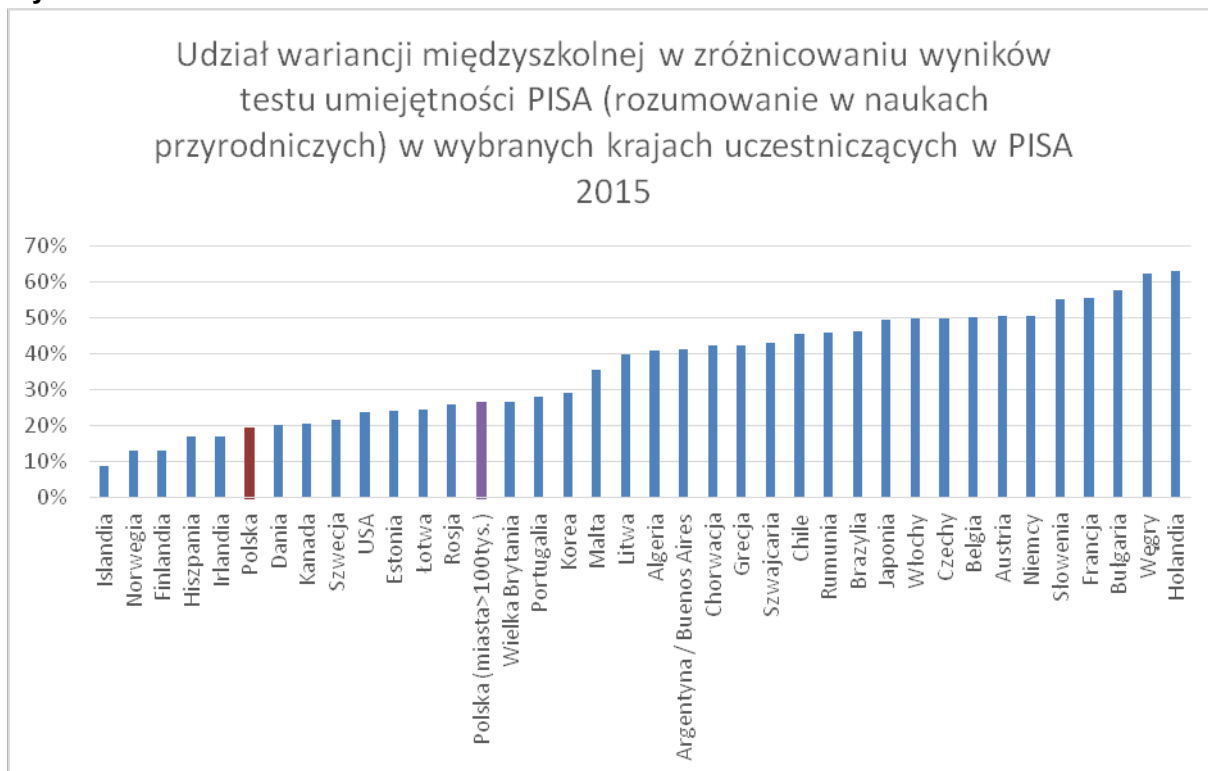
Czynnikiem, który będzie sprzyjał wyrównywaniu szans edukacyjnych, jest zapewnienie dostępu do dobrych szkół dzieciom pochodzącym z różnych warstw społecznych. „Miarą równości” może być w tym przypadku zróżnicowanie międzyszkolne, tak ze względu na poziom osiągnięć, jak i ze względu na zróżnicowanie pochodzenia społecznego uczniów. System edukacyjny, charakteryzujący się małym zróżnicowaniem szkół sprzyjać będzie wyrównywaniu nierówności oraz awansowi uczniów o niższym statusie społecznym lecz większym potencjale indywidualnym. Jednocześnie, jest to ten czynnik, w którym państwo ma do odegrania największą rolę. Z wielu względów dzieci z rodzin lepiej sytuowanych mają dostęp do szkół lepszych; wynika to chociażby z segregacji przestrzennej różnych grup społecznych (różnice między miastami a wsią, czy „lepszymi” i „gorszymi” dzielnicami w miastach) – przeciwdziałanie takim nierównościom wymaga aktywnych działań, np. w rodzaju znanego z USA „busingu”, czyli kierowania dzieci do szkół odległych od ich miejsca zamieszkania w celu zapewnienia im dostępu do lepszej szkoły i skojarzenia z dziećmi z wyżej sytuowanych rodzin. Z kolei, do separacji dzieci z różnych warstw społecznych będzie się przyczyniać działanie rodziców, dążących – w interesie swoich dzieci – do umieszczenia ich w jak najlepszych w ich mniemaniu szkołach, przy czym aktywność w tym kierunku i skuteczność takich działań jest większa w przypadku rodziców o wyższej pozycji społecznej. Choć działania takie są zrozumiałe, a państwo nie może przecież uniemożliwiać rodzicom dbania o rozwój ich dzieci, ich społeczne skutki mogą być niekorzystne. Z drugiej strony, ze względu na efektywność nauki, na jakimś jej etapie i tak następuje rozdział strumienia uczniów do szkół różniących się zarówno profilem nauki, jak i samym jej poziomem – istotne jest jednak, by podział taki nie nastąpił zbyt szybko, a także by opierał się na kryterium rzeczywistych osiągnięć uczniów, nie zaś głównie na zapobiegliwości rodziców. Badanie zróżnicowania międzyszkolnego może opierać się zarówno na danych z badań reprezentatywnych, jak i na analizie danych pochodzących z systemu egzaminacyjnego. Oba podejścia powinny mieć komplementarny charakter: dane z systemu egzaminacyjnego, jako dane wyczerpujące, nie są narażone na błędy losowe, podczas gdy dane reprezentatywne – pochodzące z relatywnie niewielkiej (ok. 150-200 w skali kraju) liczby szkół – w ograniczonym stopniu pozwalają na bardziej szczegółowe analizy (np. ze względu na zróżnicowania szkół w klasach wielkości miejscowości, co zwłaszcza w Polsce jest zjawiskiem o dużym znaczeniu). System egzaminacyjny zbiera jednak dane jedynie o poziomie osiągnięć, a poza nimi zaledwie pojedyncze charakterystyki uczniów – jak płeć, czy lokalizacja szkoły. Oparcie się na danych z badań pozwala wykorzystać znacznie pełniejszy zestaw danych kontekstowych, a oparcie się na wspólnym pomiarze kompetencji – także na czynienie porównań międzynarodowych.

Zarówno biorąc pod uwagę zróżnicowanie międzyszkolne ze względu na poziom osiągnięć uczniów (wynik w zakresie rozumowania w naukach przyrodniczych⁴), jak i na pochodzenie społeczne (mierzone wskaźnikiem ESCS) Polska należy do grupy krajów o najbardziej równościowym systemie edukacyjnym. Udział wariancji międzyszkolnej w wariancji całkowitej wyniku z zakresu rozumowania w naukach przyrodniczych (19%) daje Polsce 6. miejsce na 70 porównywalnych społeczności (zakres zmienności współczynnika od 9% w Islandii do 63% w Holandii; średnia wartość 38%). W przypadku zróżnicowania społecznego (ESCS) współczynnik ten wynosi w Polsce 22%, co daje miejsce 16.

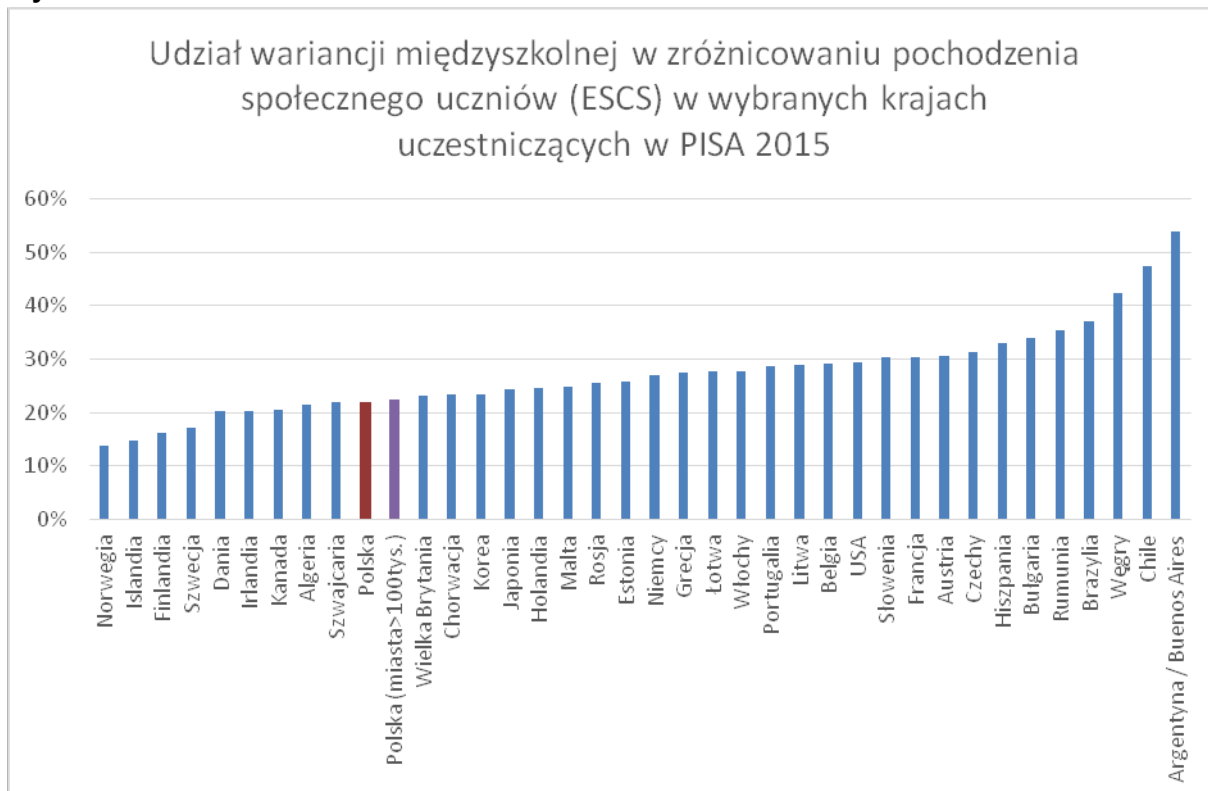
⁴ Odwołujemy się do tej dziedziny jako dziedziny wiodącej w badaniu PISA 2015, jednakże wartości współczynników byłyby prawie identyczne, a wnioski całkowicie identyczne także w przypadku odwołania się do pozostałych dziedzin.

(zakres zmienności od 14% dla Norwegii do 54% dla Buenos Aires Argentyny/Argentyna/, przy średniej 29%).

Wykres 1.6



Wykres 1.7



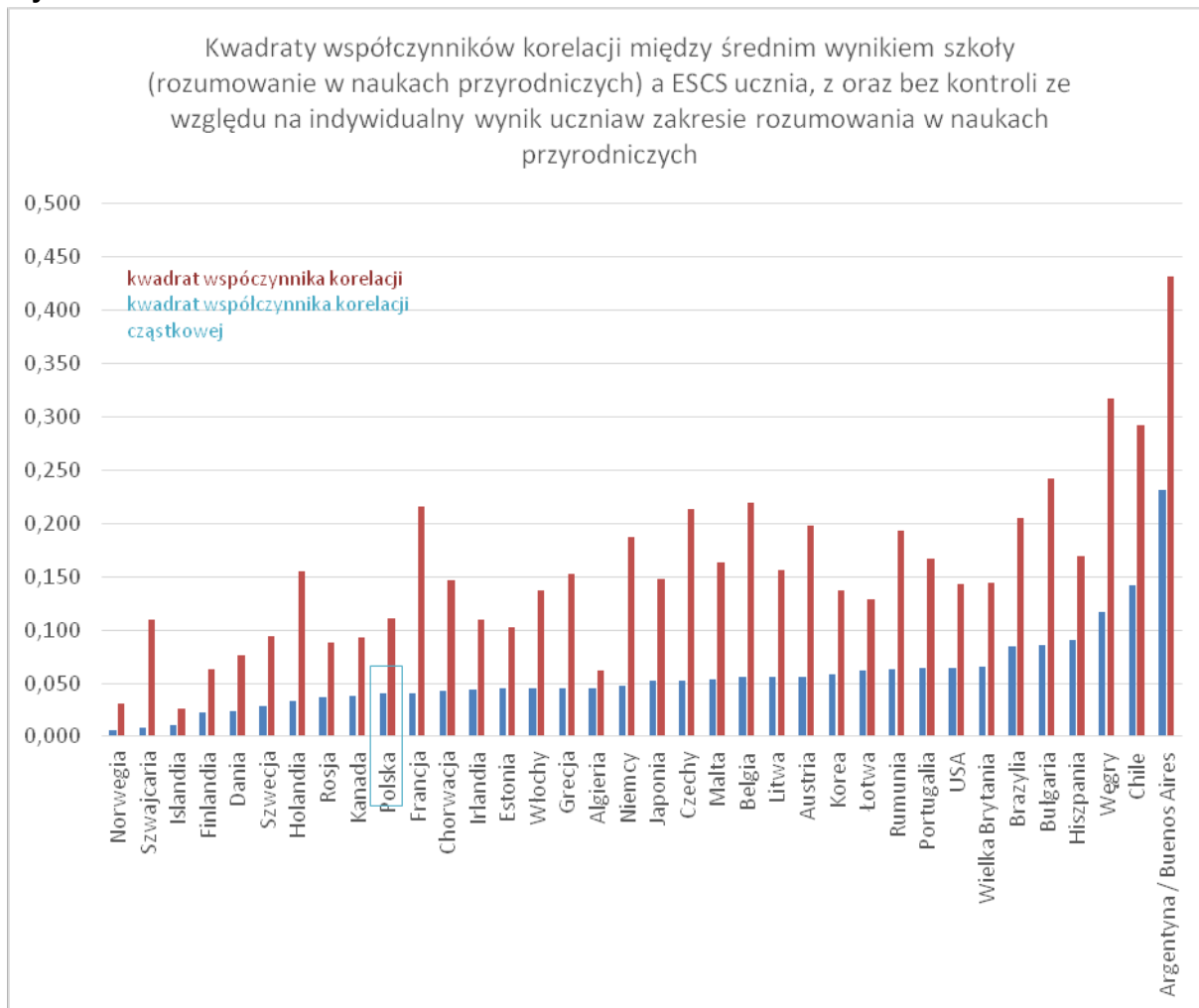
Znanym i niepokojącym zjawiskiem zachodzącym w polskich szkołach jest zwiększone zróżnicowanie szkół wielkomiejskich. Tylko tam, gdzie realnie dostępnych jest dla ucznia kilka szkół, na ich różnicowanie mogą mieć wpływ decyzje rodziców, kierujących dzieci do „lepszyc” gimnazjów, wzmacniane przez działania kierownictwa niektórych – także publicznych – szkół, dążących do „wyselekcjonowania” dla siebie najlepszych uczniów. Tam, gdzie szkoła w okolicy jest tylko jedna – a więc na wsi, czy w małym mieście, a posłanie dziecka do innej szkoły wymaga znacznie większego wysiłku związanego z dotarciem do niej, zróżnicowanie szkół z natury będzie słabsze. Wariancja międzyszkolna wyników testu PISA w dużych miastach (powyżej 100 tys. mieszkańców) Polski stanowi 26% wariancji całkowitej wyników wielkomiejskich uczniów, a więc o 7 punktów procentowych więcej, niż w skali całego kraju. Warto jednak zauważyć, że nawet tę wartość w skali całego badania można uznać za umiarkowaną (jest wciąż znacznie niższa od średniej dla wszystkich krajów i odpowiada zróżnicowaniu w kraju zajmującym 17. pozycję w pełnym rankingu międzynarodowym). Udział wariancji międzyszkolnej ESCS w dużych miastach Polski jest taki sam, jak w całym kraju (22%) – wskazuje to, że różnicowanie gimnazjów wielkomiejskich odbywa się w Polsce głównie ze względu na kryteria merytokratyczne (osiągnięcia uczniów), a nie społeczne (status rodziców).

Niskie zróżnicowanie międzyszkolne w Polsce – zwłaszcza ze względu na poziom umiejętności uczniów – wynika oczywiście w znacznej mierze z tego, że badanie PISA w latach 2003-2015 było realizowane głównie wśród uczniów gimnazjów – a więc przed zakończeniem okresu jednolitej nauki i rozdzieleniem strumieni uczniów do różnych typów szkół. Kraje o wysokim zróżnicowaniu międzyszkolnym to przede wszystkim kraje, w których podział ten następuje nieco wcześniej; nawet jednak w z tym zastrzeżeniem zróżnicowanie międzyszkolne w Polsce można uznać za niskie i to nawet w dużych miastach. Warto jednak – nie tylko z przyczyn historycznych - wrócić pamięcią do sytuacji z badania PISA 2000, realizowanego w starym systemie, w którym rozdział uczniów następował o rok wcześniej, niż obecnie. W badaniu PISA 2000 wariancja międzyszkolna umiejętności uczniów stanowiła 59% wariancji całkowitej, co w 2015 roku dałoby Polsce miejsce czwarte od końca.

Na pytanie o szanse ucznia z niższych warstw społecznych na dostanie się do szkoły o wysokim poziomie nauczania można także udzielić odpowiedzi wskazując na korelację pomiędzy wskaźnikami pochodzenia społecznego, a przeciętnymi wynikami uczniów w danej szkole. Można tu brać pod uwagę zwykły współczynnik korelacji (tzw. „rzędu zerowego”, a więc bez uwzględniania w analizie ewentualnego znaczenia wpływu innych zmiennych), tę jednak tłumaczyć może fakt, że uczniowie z rodzin o wyższej pozycji społecznej osiągają zwykle lepsze wyniki – i fakt, że są uczniami lepszych szkół wynika z ich bezpośrednio z wyższych osiągnięć, a jedynie pośrednio – z uprzywilejowanej pozycji społecznej. Z tego względu można posłużyć się również współczynnikiem korelacji cząstkowej, z wyłączeniem indywidualnego wyniku ucznia – w tym przypadku uzyskamy informację o niejako „czystej” premii za pochodzenie: w jakim stopniu uczeń z lepiej społecznie sytuowanej rodziny może liczyć na trafienie do lepszej szkoły *niezależnie* od poziomu własnych umiejętności?

Kwadrat współczynnika korelacji między przeciętnym wynikiem w szkole (rozumowanie w naukach przyrodniczych) a ESCS ucznia wynosi dla Polski 0,11 (przy zmienności od 0,02 dla Makao i 0,03 dla Norwegii do 0,43 dla Buenos Aires/Argentyna/ i średniej 0,16), jest to 20. pozycja wśród 69 społeczności dla których możliwe było obliczenie tego współczynnika. Kwadrat współczynnika korelacji cząstkowej między przeciętnym wynikiem w szkole (rozumowanie w naukach przyrodniczych) a ESCS ucznia, przy kontroli indywidualnego wyniku ucznia z rozumowania w naukach przyrodniczych wynosił w Polsce 0,040 (17. pozycja przy średniej 0,068 oraz zmienności od 0,006 dla Norwegii do 0,231 dla Buenos Aires/Argentyna/). Nie są to zatem wartości duże – Polska oświata zapewnia relatywnie równy dostęp do dobrych szkół, choć liderom w tym zakresie znacznie ustępujemy.

Wykres 1.8



Zestawienie ze sobą wyników badania PISA od roku 2000 do roku 2015 pokazuje, że w tym okresie dokonało się wiele zmian. Na samym początku XXI wieku polscy uczniowie osiągnęli wyniki znacznie poniżej średnich wyników ogółu uczniów z krajów OECD i dotyczyło to wszystkich trzech dziedzin sprawdzanych umiejętności. Następnie postępowała stopniowa poprawa, ze skokową poprawą do poziomu około średniej OECD w 2003 rok, istotnym podniesieniem szeroko rozumianej umiejętności czytania (w sensie wydobywania informacji z tekstu) i lepszymi wynikami w rozumowaniu w naukach przyrodniczych w 2006 roku, a wreszcie zasadniczą poprawą rozumowania matematycznego w 2012 roku, potwierdzona w roku 2015. Nadal jednak umiejętność interpretacji tekstu, jako pogłębiona umiejętność związana z czytaniem, pozostawia przestrzeń dla dalszej poprawy. Badanie PISA 2015 odśloniło problem sposobu wykorzystywania komputera w procesie uczenia się. Choć nie było to głównym zamierzeniem badania odkryto, że Polska jest jednym nielicznych krajów, gdzie wprowadzenie komputera do rozwiązywania zadań testowych oddziaływało najsilniej.

2. INNOWACJE METODOLOGICZNE I CIĄGŁOŚĆ BADANIA

Jacek Haman

Założeniem badania PISA, od jego pierwszej edycji w 2000 roku, jest dostarczanie szeregu porównywalnych wyników: porównywalnych międzynarodowo, ale także, w kolejnych edycjach, porównywalnych w czasie. Z drugiej strony samo badanie musi iść z duchem czasu: uwzględniać nowe technologie, a także postęp samej teorii pomiaru edukacyjnego. Zmiany w badaniu – niezbędne do utrzymania jego wysokiej jakości – stanowią wyzwanie dla utrzymania porównywalności wyników. Wiedza i doświadczenie zebrane w okresie już prawie dwudziestu lat trwania projektu OECD PISA pozwoliły na zidentyfikowanie szeregu problemów metodologicznych, wprowadzenie szeregu zmian szczegółowych w rozwiązaniach zastosowanych w kolejnych edycjach. Dzięki powszechnemu dostępowi społeczności naukowej do danych PISA i jawności metodologii badania krytyczne analizy – dotyczące w szczególności szczegółowych rozwiązań w skalowaniu danych – prowadzone są nie tylko przez badaczy związanych z konsorcjum międzynarodowym czy krajowymi zespołami realizującymi badanie, ale także przez badaczy z zewnątrz.

Postęp technologii również nie może nie mieć wpływu na badanie PISA. Gdy prowadzona była pierwsza edycja badania PISA, w roku 2000, ponad połowa (52%) polskich piętnastolatków nie miała w domu komputera, blisko jedna trzecia (29%) nigdy nie korzystała z niego w szkole; zaledwie 17% miało dostęp do Internetu w domu, a niespełna 35% kiedykolwiek korzystało z niego w szkole⁵. W roku 2015 brak dostępu do komputera i Internetu w domu zdarzał się wyjątkowo, zaś brak dostępu do Internetu w polskim gimnazjum trudno sobie nawet wyobrazić. Komputery (i inne urządzenia spełniające podobne funkcje jak smartfony czy tablety, których w roku 2000 w ogóle jeszcze nie było) odgrywają coraz większą rolę zarówno w życiu uczniów, jak i w nauczaniu, a ekran komputera czy tabletu – chcemy czy nie – staje się nie mniej naturalnym medium i narzędziem, jak wcześniej papierowy podręcznik, zeszyt, kartka papieru, ołówek. Stopniowe wprowadzanie elementów pomiaru komputerowego – ale tylko w ramach badań opcjonalnych, realizowanych obok głównych badań kompetencji – zaczęto w PISA 2006 (w polskim badaniu od PISA 2009). W badaniu 2015 narzędzia komputerowe całkowicie zastąpiły testy typu „papier i ołówek” i po raz pierwszy w ten sposób realizowany był podstawowy pomiar kompetencji uczniów. Była to – niewątpliwie – największa i najdonioślejsza zmiana i nowość w badaniu PISA 2015, ale nie jedyna.

Przedmiotem ciągłej refleksji – a zatem również ciągłych zmian – są również podstawy koncepcyjne badanych domen wiedzy. W badaniu PISA 2015 główną dziedziną po raz drugi (poprzednio w 2006) było rozumowanie w naukach przyrodniczych; w związku z tym szczególnemu przeglądowi i aktualizacji zostały poddane ramy badania tej części testu.

Czy da się pogodzić zmiany – w narzędziach, w metodologii – z ciągłością, a więc z wymogiem zapewnienia porównywalności wyników w czasie? Na tak postawione pytanie nie da się jednoznacznie odpowiedzieć, a ściślej – nigdy nie będziemy w pełni pewni, czy odpowiedź, której udzieliliśmy, jest poprawna. Najłatwiej ocenić zmiany w sposobie skalowania – można bowiem porównać wyniki wyskalowane na różne sposoby; różnice z reguły nie są wielkie, ale są zauważalne. Konsekwencje zmiany narzędzi – a więc przejścia z pomiaru „papierowego” na „komputerowy” były intensywnie analizowane przede wszystkim w ramach badania próbnego przed badaniem PISA 2015;

⁵ Na podstawie danych z zebranych w OECD PISA 2000.

jego wyniki wskazują, że porównywalność wyników została zachowana, nigdy jednak nie da się usunąć wszystkich wątpliwości.

Na pewno jednak zmiany są niezbędne: zmieniają się sposoby nauczania w szkołach, sposoby korzystania z zasobów informacyjnych przez uczniów, zmieniają się wreszcie zasoby naszej wiedzy o pomiarze dydaktycznym i rozwiązania, które były optymalne piętnaście lat temu, dziś mogą być już po części nieaktualne.

2.1. Pomiar i skalowanie

Wyniki badania PISA – a więc oszacowania kompetencji w poszczególnych obszarach badania – przypisywane są poszczególnym uczniom na podstawie procedury skalowania, uwzględniającej liczbę poprawnie rozwiązanych zadań oraz ich trudność. Zastosowanie tych procedur daje szereg istotnych korzyści, w porównaniu do oceniania kompetencji przez np. zwykle 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) powtarzanych jest w kolejnych edycjach.

Gdy w 2000 roku przeprowadzaliśmy pierwszą edycję badania OECD PISA, skalowanie wyników testów oparte na modelach teorii odpowiedzi na pytanie testowe (IRT) było w Polsce nowością. Od tego czasu IRT było wykorzystywane w szeregu badań, tak krajowych, jak i międzynarodowych, jak również było przedmiotem szeregu publikacji, zarówno o charakterze wprowadzającym, jak i bardzo zaawansowanym. Z tego względu poniższe omówienie skalowania w badaniu PISA dzieli się na dwie sekcje. Sekcja pierwsza przeznaczona jest przede wszystkim dla czytelników niemających na co dzień do czynienia z pomiarem w paradygmacie IRT – przedstawiamy w niej najważniejsze elementy tej koncepcji, niezbędne dla zrozumienia istoty pomiaru kompetencji w badaniu PISA; zainteresowanych jej głębszym poznaniem odsyłamy do dostępnej w Polsce literatury, w szczególności do książek (Jakubowski i Pokropek, 2009) i (Pokropek (red.), 2015).

W drugiej sekcji podajemy najważniejsze informacje na temat szczegółowych rozwiązań modeli IRT stosowanych w badaniu OECD PISA z podkreśleniem tych elementów, które zostały zmodyfikowane w badaniu OECD PISA 2015 względem poprzednich edycji badania.

2.2. Założenia ogólne – podstawy modeli IRT⁶

W badaniu PISA skalowanie wyników testu opiera się na teorii odpowiedzi na pytanie testowe (IRT – *Item Response Theory*), a ściślej – na uogólnionym modelu Rascha.

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 zadania⁷.

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 uwzględnianiem pewnych dodatkowych parametrów tej funkcji – obok samej trudności zadania także jego mocy dyskryminacyjnej⁸, 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 estymacja 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

⁶ Tekst zawarty w tej sekcji zawiera fragmenty powtórzone bezpośrednio lub z pewnymi modyfikacjami za raportami z badania PISA 2009 (Federowicz et al., 2010) i PISA 2012 (Federowicz et al., 2013).

⁷ 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 szeregu skal „cech kontekstowych”.

⁸ „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.

względem 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 osiągnęte 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 specyficznych 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 modelu Rascha 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.

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% lepszy 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 wiedzieć lub umieć badani powinni): skale skonstruowane są w ten sposób, by wartość 500 punktów odpowiadała średniej wyników krajów OECD w badaniu PISA 2000 oraz by jeden 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 dziedziną wiodącą).

Probabilistyczny charakter *Item Response Theory* oznacza także, że przy interpretacji wyników badania bierze się pod uwagę, że dwóch 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 o 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*). Należy jednak mieć świadomość, że to, czym różnią się modele IRT od innych podejść do testowania (jak np. klasyczna teoria testu), to nie sam fakt, że wyniki pomiaru obarczone są niepewnością – od tej bowiem nie jest wolny żaden model. Ważne jest to, że modele IRT pozwalają – w pewnym zakresie – oszacować skalę tej niepewności i zamknąć ją w modelu probabilistycznym.

Celem badania OECD PISA jest znalezienie najlepszych oszacowań przeciętnych wyników dla określonych populacji i dokonywanie porównań międzygrupowych: między uczniami z różnych krajów, szkół różnych typów, między dziewczętami a chłopcami itd. W przeciwieństwie do egzaminów nie było celem badania dostarczanie danych o wynikach konkretnych uczniów biorących w nim udział. Z tego względu zarówno organizacja testu, jak i procedury skalowania dobrane są taki sposób, aby zminimalizować błędy oszacowań średnich i wariancji poziomów umiejętności wyznaczanych dla podzbiorowości uczniów, nawet jeśli dzieje się to kosztem większych błędów oszacowań indywidualnych.

Możliwość uzyskania poprawy jakości oszacowań dla zbiorowości kosztem oszacowań indywidualnych może wydawać się paradoksalna. Przykładem takiego mechanizmu – na poziomie organizacji testu – jest wspomniana już parokrotnie zasada „rozdzielania” pełnej puli zadań między poszczególne zestawy testowe. Łatwo zauważyć, że jest to działanie korzystne z punktu widzenia oceny przeciętnego poziomu umiejętności dla zbiorowości: dzięki wykorzystaniu większej liczby zadań możemy uwzględnić więcej cząstkowych składowych umiejętności czy wiedzy wchodzącej w skład danej dziedziny, zaś z drugiej strony ewentualne specyficzne czynniki sprzyjające lub utrudniające rozwiązanie danego zadania, niepowiązane z badaną dziedziną, będą miały mniejszy systematyczny wpływ na ogólny wynik. Równie łatwo przy tym zauważyć, że zasada ta utrudnia porównywanie wyników poszczególnych uczniów, którzy rozwiązywali przecież odmienne zestawy zadań.⁹

Na poziomie skalowania, przykładanie większej wagi do poprawności oszacowań zbiorowych niż indywidualnych przejawia się, z jednej strony, w stosowaniu estymatorów *plausible values*, pozwalającymi na lepsze oszacowanie wariancji oraz błędu losowego dla średnich, kosztem dodatkowego błędu losowego na poziomie oszacowań indywidualnych, z drugiej zaś strony – w uwzględnianiu w modelach skalowania (a ściślej w modelach wyznaczania rozkładów prawdopodobieństwa, z których losowane są wartości *plausible values*) danych kontekstowych, pochodzących z kwestionariuszy ucznia. Działanie takie byłoby absolutnie nie do przyjęcia, gdyby celem badania była ocena indywidualnego poziomu umiejętności danego ucznia: trudno by zaakceptować procedurę, w której ocena ucznia zależałaby nie tylko od odpowiedzi, jakich udzieli na egzaminie, ale również od np. wykształcenia jego rodziców. W sytuacji, gdy celem badania jest ocena przeciętnych wyników dla zbiorowości, postępowanie takie pozwala jednak zmniejszyć poziom błędów losowych oszacowań. Co więcej, oszacowania poziomów poszczególnych umiejętności podawane są – na podstawie danych kontekstowych oraz wyników testów z zakresu pozostałych umiejętności – nawet w przypadku uczniów, którzy zadań z zakresu danej umiejętności w ogóle nie rozwiązywali. Na poziomie indywidualnym takie oszacowania (które w tym wypadku lepiej byłoby nazywać przewidywaniami) są, oczywiście, zwykle obciążone dużym błędem; jednakże przy szacowaniu parametrów poziomu umiejętności dla zbiorowości lepsze (dokładniejsze) oszacowania uzyskuje się, uwzględniając zarówno oszacowania pochodzące z rzeczywistego testowania danej umiejętności, jak i te wyznaczone jedynie na podstawie danych kontekstowych i z pozostałych testów.

Dla celu losowania wartości *plausible values* przyjmuje się, że dla każdej kombinacji branych pod uwagę zmiennych kontekstowych¹⁰ rozkład poziomu umiejętności uczniów jest rozkładem normalnym. Model łączący wartości zmiennych kontekstowych ze średnią i wariancją tego rozkładu wyznaczany jest za pomocą modeli strukturalnych (modeli regresji zmiennych ukrytych), co z kolei, kierując się regułami rachunku prawdopodobieństwa, pozwala wyznaczyć dla każdego badanego warunkowe rozkłady prawdopodobieństwa poziomu umiejętności ze względu na jego wartości zmiennych

⁹ Warto jednak zwrócić uwagę, że właśnie dzięki zastosowaniu modeli IRT porównywanie indywidualnych wyników uczniów, którzy pisali różne zestawy zadań jest możliwe – jednakże byłoby ono obciążone większą niepewnością niż w przypadku porównywania wyników uczniów rozwiązujących dokładnie te same zadania ze względu na konieczność uwzględnienia dodatkowego składnika błędu losowego – tzw. błędu łączenia.

¹⁰ Ściślej rzecz biorąc, zmienne kontekstowe poddawane są analizie głównych składowych, po czym w modelowaniu uwzględniane są składowe odpowiedzialne łącznie za 80% ich łącznej wariancji dla danego kraju.

kontekstowych oraz wyniki testu. Z tego rozkładu losowanych jest po 10 *plausible values* dla każdej z mierzonych kompetencji.

Modele łączące wartości zmiennych kontekstowych z rozkładami umiejętności wyznaczone były dla każdego kraju biorącego udział w badaniu osobno – wynika to nie tylko z faktu, że zależności tego typu mogą być silnie warunkowane kulturowo, ale że formalnie te same cechy kontekstowe mogą w różnych krajach reprezentować niekoniecznie porównywalne stany faktyczne (przykładowo – za wykształceniem rodziców na tym samym formalnie poziomie ISCED mogą stać systemy edukacyjne bardzo silnie różniące się poziomem nauki). Oznacza to także, że ten etap procedury skalowania nie ma wpływu na oszacowania średnich osiągnięć na poziomie całego kraju.

2.3. Skalowanie wyników PISA 2015 w porównaniu z poprzednimi edycjami badania

Przy zachowaniu tych samych ogólnych założeń, szczegółowe rozwiązania przyjęte w skalowaniu wyników PISA 2015 podlegały – w porównaniu z wcześniejszymi edycjami badania – pewnym modyfikacjom:

- Zmieniony został stosowany wariant modelu IRT: o ile w cyklach PISA 2000-2012 stosowany był model jednoparametryczny (z uwzględnieniem jedynie trudności zadania), przy skalowaniu danych z PISA 2015 wykorzystano elementy¹¹ modelu dwuparametrycznego (trudność oraz moc dyskryminacji)¹².
- We wcześniejszych cyklach badania, szacowanie trudności zadań prowadzone było na podpróbie, do której włączano losowo po 500 przypadków z każdego kraju/regionu (do roku 2009 – wyłącznie z krajów OECD), jedynie z bieżącego cyklu badania. Przy skalowaniu danych z PISA 2015 szacowanie trudności zadań przeprowadzono na bazie wyników uzyskanych przez wszystkich uczniów z wszystkich dotychczasowych edycji badania.

Zastosowanie modelu dwuparametrycznego prowadzi do uzyskania lepszego dopasowania modelu do danych, a zatem trafniejszych i obciążonych mniejszymi błędami wyników; korzyści wynikające z oparcia oceny trudności zadań na pełnym zestawie danych są również oczywiste¹³.

Kolejna nowość w skalowaniu wyników PISA 2015 wiązała się ze sposobem uwzględniania różnic w trudności poszczególnych zadań w poszczególnych krajach. Co do zasady, jeśli porównania międzynarodowe mają być w ogóle możliwe, należy przyjąć, że to samo zadanie we wszystkich krajach mierzy te same umiejętności, a więc ma również taką samą trudność – jeśli jakieś zadanie okazuje się silnie zależne od czynników kulturowych, powinno zostać wyeliminowane (najlepiej na etapie badania próbnego; w razie potrzeby – także już po badaniu zasadniczym). W szczególnych przypadkach (np. specyficzne problemy z tłumaczeniem zadania) możliwe jest również pominięcie w analizie rozwiązań danego zadania pochodzących od uczniów jednego z krajów.

W badaniu PISA 2015 dopuszczono również inne rozwiązanie: wyznaczenie dla niektórych zadań poziomu trudności specyficznego dla danego kraju. W pewnym uproszczeniu – oznacza to, że dane o rozwiązaniach tego zadania nie będą miały wpływu na przeciętny poziom uzyskany przez uczniów

¹¹ Model dwuparametryczny stosowano w przypadku tych zadań, dla których uzyskiwano dzięki temu znaczącą poprawę dopasowania; jeśli wyniki uzyskane w modelu jednoparametrycznym były dopasowane w stopniu zadawalającym, pozostawano przy modelu jednoparametrycznym

¹² Ponadto, we wszystkich cyklach PISA w przypadku zamkniętych pytań testowych uwzględniane było prawdopodobieństwo przypadkowego odgadnięcia odpowiedzi.

¹³ Warto dodać, że dla danych z polskich edycji badania PISA 2000-2012 podobne analizy, uwzględniające taki sposób skalowania danych, zostały przeprowadzone wcześniej i opublikowane w pracy (Dolata, Jakubowski i Pokropek, 2013).

danego kraju na skali międzynarodowej, jednakże ich uwzględnienie pozwoli na dokładniejsze określenie poziomu umiejętności poszczególnych uczniów.

Kolejna zmiana w sposobie skalowania danych PISA 2015 dotyczy sposobu traktowania zadań, do których dany uczeń nie dotarł (zadania, na które nie udzielono żadnej odpowiedzi, znajdujące się na końcu zestawu testowego). W poprzednich cyklach zadania takie – przy szacowaniu poziomu umiejętności – były traktowane tak samo jak rozwiązane błędnie; natomiast przy szacowaniu trudności zadań – tak, jakby nie było ich w zestawie testowym. W cyklu PISA 2015 zadania takie były traktowane jako nieistniejące zarówno przy szacowaniu poziomu trudności zadania, jak i przy szacowaniu poziomu umiejętności uczniów. Głównym uzasadnieniem dla takiej zmiany jest uspoźnienie sposobu traktowania tych zadań w obu sytuacjach; trzeba jednak stwierdzić, że w przypadku tej zmiany można wskazać argumenty zarówno za jej wprowadzeniem, jak argumenty przeciwne.

Zmiany w modelu skalowania danych PISA 2015 oznaczają generalnie, że dane z obecnego cyklu są wyskalowane *lepiej* niż w cyklach poprzednich dzięki lepszemu dopasowaniu zastosowanych modeli. Oznacza to jednak, że w pewnym sensie są wyskalowane *inaczej* – a więc zmiana modelu skalowania jest czynnikiem utrudniającym możliwość analizowania zmian w czasie. Efekt ten na szczęście nie jest duży, a jego wielkość została oszacowana poprzez zastosowanie modelu skalowania danych z PISA 2015 do danych z poprzednich edycji badania. W przypadku Polski zmiana modelu skalowania ze starego na nowy oznaczałaby zmianę średniego wyniku badania PISA 2012 o –2 punkty w przypadku umiejętności matematycznych, –4 punkty w przypadku czytania i rozumienia oraz +4 punkty w przypadku rozumowania w naukach przyrodniczych.

2.4. Narzędzia pomiaru: od testów papierowych do testów komputerowych

Narzędziem pomiaru w badaniu PISA są testy. W badaniu PISA 2000-2012 miały one postać papierowych zeszytów, wypełnianych przez uczniów w czasie sesji testowej – tak samo, jak dzieje się to na np. egzaminach zewnętrznych, a często również na wewnętrznych sprawdzianach czy klasówkach. Od edycji PISA 2006 zaczęto uzupełniać badanie główne opcjonalnymi badaniami z użyciem narzędzi komputerowych; w Polsce komponent komputerowy realizowany był po raz pierwszy w PISA 2009 (badanie czytania tekstu w formie elektronicznej). W badaniu PISA 2012 moduł komputerowy był już znacznie rozbudowany i obejmował badanie kompetencji matematycznych, czytanie tekstu w formie elektronicznej, oraz zadania z „rozwiązywania problemów”, wciąż jednak był to komponent realizowany niezależnie od głównego, „tradycyjnego” badania realizowanego za pomocą testów papierowych, niejako torujący drogę do zmiany narzędzia w badaniu głównym, planowanej na edycję OECD PISA 2015. W badaniu PISA 2015 testy komputerowe stały się w większości biorących udział w badaniu krajów jedynym narzędziem badania kompetencji.

Można wskazać kilka powodów, dla których celowa była zmiana narzędzia badawczego. Realizowanie testów bezpośrednio na komputerze upraszcza bardzo złożony proces transferu materiałów – zestawów zadań do uczniów, wypełnionych testów do koderów wprowadzających do komputera odpowiedzi na pytania zamknięte i oceniających rozwiązania zadań otwartych. Komputeryzacja procesu testowania pozwala na większe zróżnicowanie zestawów testowych (a więc użycie większej liczby zadań, a także wykorzystanie większej liczby wariantów ich rotacji, co pozwala na zmniejszenie błędów związanych z czynnikami specyficznymi dla poszczególnych zadań lub dla pozycji zajmowanej w teście); dzięki temu w badaniu PISA 2015 możliwe było użycie podobnej liczby zadań z zakresu czytania i interpretacji oraz umiejętności matematycznych, co z dziedziny wiodącej – rozumowania

w naukach przyrodniczych. Korzyści na etapie sprawdzania prac (jak eliminacja ręcznego przepisywania odpowiedzi uczniów, ale również możliwości wynikające ze sprawdzania odpowiedzi na zadania otwarte w formie elektronicznej), a także późniejszego transferu danych są oczywiste.

Nie mniej poważne, ale jednocześnie – znacznie trudniejsze do jednoznacznej oceny – są kwestie o charakterze kulturowym. Komputery, tablety, smartfony w coraz większym stopniu zastępują papier i ołówek w codziennym życiu (zwłaszcza młodzieży) jako narzędzie pracy zawodowej, ale również jako narzędzie dydaktyki – zdobywania wiedzy i jej sprawdzania. Stopniowo to komputer, a nie kartka papieru staje się dla uczniów naturalnym medium. Proces ten zachodzi w różnym tempie w różnych krajach, w różnych środowiskach, również – w różnych kontekstach. Zmiana narzędzia w badaniu PISA, z założenia dotycząca wszystkich uczniów w kraju (i zdecydowaną większość uczniów uczestniczących w badaniu)¹⁴, w tym sensie nigdy nie mogła przyjść we właściwym momencie – dla części będzie to zmiana zbyt wczesna, dla części – zbyt późna.

Komputerowe badanie umiejętności matematycznych, czytania oraz „rozwiązywania problemów” realizowane w ramach PISA 2012, było prowadzone jako opcja dodatkowa, realizowana po głównym, „papierowym” badaniu PISA, i wykorzystywało nowe, specjalnie przygotowane zestawy zadań. Wyniki tego badania były zaskakujące i, zwłaszcza z polskiego punktu widzenia, niepokojące. Przypomnijmy – o ile w badaniu PISA 2012 średni wynik polskich uczniów w zakresie umiejętności matematycznych wyniósł 518 punktów, to w badaniu komputerowym był on niższy o 29 punktów (489); jeszcze gorszy był wynik czytania tekstu elektronicznego – 477 punktów, a więc o 41 mniej niż w badaniu głównym. Również wynik testu z „rozwiązywania problemów”, 481 punktów, wypadł znacząco poniżej średniej OECD. Tak silna rozbieżność między wynikiem testu papierowego a komputerowego na niekorzyść wyników badania komputerowego w badaniu OECD PISA 2012 dotknęła obok Polski tylko kilka krajów (Słowenia, Izrael). Badanie dostarczyło przy tym zbyt mało danych, by w sposób jednoznaczny można było wskazać przyczyny powstania takiej rozbieżności. Generalnie mogło być nią albo to, że testy komputerowe mierzyły w istocie inne kompetencje niż testy papierowe (takie wątpliwości zgłaszano zwłaszcza wobec zadań z matematyki), albo to, że ich wyniki były silnie zależne od lokalnych uwarunkowań (jak np. opanowanie przez dzieci w danym kraju narzędzi IT). Źródłem różnic mogą być także czynniki motywacyjne (badanie komputerowe było realizowane po głównej części badania – uczestnicy mogli być zatem zmęczeni lub zdekoncentrowani: istotne jest jednak, czy i na ile efekty z tym związane mogły występować z różnym nasileniem w różnych krajach). Zebrane dane nie pozwoliły jednak ani na pełne potwierdzenie, ani wykluczenie takich hipotez.

Aby ograniczyć ryzyka związane z planowaną zmianą narzędzi w PISA 2015, podjęto działania idące w dwóch kierunkach. Po pierwsze, zadania na testy komputerowe 2015 przygotowywano w taki sposób, by stanowiły możliwie bliskie odpowiedniki zadań w wersjach „papierowych”. Oznacza to wprawdzie, że nie wykorzystane zostały w pełni możliwości, jakie daje rozwiązywanie zadań na komputerze (innymi słowy: choć test mógłby bardziej przypominać np. grę komputerową, zdecydowano by w większym stopniu imitował test papierowy), dzięki czemu jednak zmniejszono ewentualne efekty pomiarowe związane ze zmianą narzędzia. Po drugie, w ramach badania próbnego, realizowanego (na znacznie większą skalę, niż w poprzednich edycjach PISA) na wiosnę 2014 roku testowano jednocześnie zadania w wersji papierowej i komputerowej, aby wypracować „model przejścia” umożliwiający interpretację wyników testu komputerowego na tej samej skali co

¹⁴ W OECD PISA 2015 poszczególne kraje mogły zdecydować się na pozostanie przy pomiarze za pomocą testów papierowych, jeśli ich zdaniem przeprowadzenie badań komputerowych w lokalnych warunkach byłoby zbyt skomplikowane lub zbyt ryzykowne. Zdecydowało się na to 15 z 72 biorących udział w badaniu krajów i regionów (spośród krajów UE były to jedynie Rumunia i Malta). Realizacja badania za pomocą testów papierowych oznaczała również rezygnację z wykorzystania w badaniu nowych zadań – te bowiem zostały przygotowane jedynie w wersji komputerowej.

wyników testu papierowego¹⁵. W badaniu próbnym wzięło udział 2171 polskich uczniów (w skali całego badania międzynarodowego ponad 152 tysiące uczniów), podzielonych losowo¹⁶ na trzy grupy: rozwiązujących zadania kotwiczące w wersji papierowej, rozwiązujących zadania kotwiczące w wersji komputerowej oraz rozwiązujących nowe zadania w wersji komputerowej; porównanie wyników uczniów z grupy pierwszej i drugiej służyło wypracowaniu „modelu przejścia”.

Tym samym zasady przeliczania wyników uzyskanych w testach komputerowych i papierowych zostały opracowane na mocnej podstawie empirycznej. Trzeba jednak pamiętać, że „model przejścia” jest wspólny dla wszystkich uczestniczących w badaniu krajów, a zatem nie uwzględnia np. wszystkich możliwych różnic kulturowych, w wyniku których w różnych społecznościach efekty zmiany narzędzia mogłyby być różne. Stworzenie modelu w większym stopniu uwzględniającego potencjalne różnice kulturowe wymagałoby przeprowadzenia badań na znacznie większą skalę.

Efekty związane ze zmianą narzędzia mogą mieć bowiem bardzo różne, trudne do przewidzenia z góry źródła. Mogą to być z jednej strony czynniki związane z ogólnym opanowaniem przez poszczególnych uczniów narzędzi IT (i w tym zakresie stosunkowo najłatwiej jest ewentualne efekty kontrolować), ale także ze specyfiką stosowania IT w różnych kontekstach, tak w edukacji, jak i poza edukacją. Mogą to być także czynniki związane z konkretnymi, przyjętymi w konstrukcji aplikacji testowej rozwiązaniami. Przykładowo – test papierowy może być przez ucznia rozwiązywany w dowolnej sekwencji – uczeń może dowolnie wiele razy wracać do podjętego zadania, bądź zadanie, które w pierwszej chwili uznał za zbyt trudne, może spróbować rozwiązać po rozwiązaniu łatwiejszych itp. W przypadku testu komputerowego „nawigacja” po poszczególnych zadaniach z natury rzeczy odbywa się inaczej niż w przypadku tradycyjnego testu; dodatkowo testy zastosowane w badaniu PISA 2015 uniemożliwiały „cofanie się” do zadań wcześniej porzuconych lub pominiętych (można było jedynie cofać się do wcześniejszych pytań w ramach tego samego zadania).

Jedną z wielkich zalet testów komputerowych jest możliwość sprawdzenia, jaki był rytm rozwiązywania zadań testowych przez poszczególnych uczniów (np. zapisywane są szczegółowe informacje o czasie poświęconym na rozwiązanie poszczególnych zadań czy liczba wykonywanych operacji myszką i na klawiaturze) – analiza takich danych pozwoli w przyszłości na lepsze poznanie mechanizmów określających związki między konstrukcją samej aplikacji stosowanej do testowania a wynikami uczniów; brak jest jednak tak bogatych danych w odniesieniu do testów rozwiązywanych tradycyjnie.

Wobec – z jednej strony – relatywnie słabych wyników polskich uczniów w testach komputerowych z PISA 2009 i 2012 oraz – z drugiej strony – spadku średnich wyników polskich uczniów w PISA 2015 w porównaniu z PISA 2012 nasuwa się pytanie, czy zmiana narzędzia była rzeczywiście neutralna dla polskich wyników PISA. Z przyczyn przedstawionych wyżej jakiegokolwiek odpowiedzi na to pytanie mogą mieć póki co jedynie hipotezami. Sądzymy jednak, że dysponujemy argumentami pozwalającymi traktować pewne hipotezy jako dość mocno uzasadnione. Argumenty te odnoszą się do dwóch źródeł danych: Po pierwsze, wyników badania próbnego, wskazujących, że przeciętna różnica w rozwiązywalności tych samych zadań w wersjach papierowych i komputerowych była większa

¹⁵ W istocie, wyznaczenie „modelu przejścia” polegało na wyznaczeniu wielkości korekty parametru trudności dla poszczególnych zadań w wersji komputerowej, a ściślej dla tych zadań, dla których stwierdzono znaczącą różnicę w poziomie trudności między wersją papierową a komputerową.

¹⁶ W badaniu próbnym, w przeciwieństwie do badania zasadniczego, szkoły dobierane są na zasadzie próby celowej, a nie losowej, jednakże dobór uczniów w szkole przeprowadzany jest wedle tych samych zasad co w badaniu zasadniczym, a więc losowo; to samo dotyczy również przydziału poszczególnych wersji narzędzi testowych. Tym samym na podstawie badania próbnego nie można wnioskować o przeciętnym poziomie umiejętności (nie mamy do czynienia z próbą reprezentatywną szkół, a więc i uczniów), można jednak wnioskować np. o relatywnych poziomach trudności zestawów testowych (próby, na których testowane są różne zadania, nie są wprawdzie w pełni reprezentatywne dla całej populacji, natomiast są w pełni porównywalne ze sobą).

w Polsce, w porównaniu z przeciętną z całego badania próbnego – co wskazywałoby na silniejszy, negatywny efekt narzędzia w przypadku polskich uczniów.

Tabela 2.1. podaje przeciętną rozwiązywalność zadań kotwiczących, które testowane były w badaniu próbnym w wersji komputerowej i papierowej, w próbie polskiej oraz w próbie badania międzynarodowego.

Tabela 2.1.

		Czytanie i interpretacja	Matematyka	Rozumowanie w naukach przyrodniczych
K	Przeciętna rozwiązywalność zadań w wersji komputerowej – całe badanie próbne	61,8%	48,6%	51,6%
P	Przeciętna rozwiązywalność zadań w wersji papierowej – całe badanie próbne	64,5%	51,7%	53,5%
	K-P	-2,70%	-3,10%	-1,90%
K_p	Przeciętna rozwiązywalność zadań w wersji komputerowej – badania próbne, Polska	65,7%	50,8%	56,0%
P_p	Przeciętna rozwiązywalność zadań w wersji papierowej – badania próbne, Polska	71,8%	55,5%	59,7%
	K_p-P_p	-6,10%	-4,70%	-3,70%

Same dane o przeciętnej rozwiązywalności zadań – ze względu na procedury doboru próby – nie mają bezpośredniej interpretacji (podajemy je jedynie by ułatwić czytelnikowi prześledzenie dalszego rozumowania). Uprawnione jest jednak porównywanie rozwiązywalności zadań w wersji papierowej i komputerowej. Jak widać z przytoczonych danych, w całym badaniu rozwiązywalność zadań w wersji komputerowej w zakresie czytania i interpretacji była przeciętnie o 2,7 punktu procentowego niższa, niż dla wersji papierowej; w Polsce różnica ta wynosiła 6,1 punktu procentowego; dla matematyki różnica ta wynosiła odpowiednio 3,1 i 4,7 pp, zaś dla rozumowania w naukach przyrodniczych 1,9 i 3,7 punktu procentowego. Innymi słowy, „efekt narzędzia” dla przeciętnej rozwiązywalności zadań w przypadku Polski był w przypadku matematyki o 1,6 pp, w przypadku rozumowania w naukach przyrodniczych o 1,8pp, zaś w przypadku czytania i interpretacji tekstu aż o 3,4 pp silniejszy niż przeciętnie w badaniu.

Gdyby przenieść te wyniki na wyniki badania głównego, można oszacować, że „efekt narzędzia” w przypadku matematyki odpowiadał za spadek średniego wyniku Polski o około 8 punktów, w przypadku rozumowania w naukach przyrodniczych – 9 punktów, zaś w przypadku czytania i interpretacji – aż o 19 punktów¹⁷. Biorąc pod uwagę, że w porównaniu do badania PISA 2012 średnia z matematyki w PISA 2015 spadła o 14 punktów, z rozumowania w naukach przyrodniczych o 24 punkty, zaś w czytaniu i interpretacji o 12 punktów, „efekt narzędzia” w przypadku pierwszych dwóch dziedzin mógłby odpowiadać za ok. 40-50% spadku wyników, zaś w przypadku czytania i interpretacji – za całość tego spadku, „ukrywając” wręcz niewielki wzrost przeciętnych kompetencji polskich uczniów w porównaniu do roku 2012.

¹⁷ Szacunki przyrostu przeciętnego wyniku PISA w przypadku osiągnięcia wyższego przeciętnego poziomu rozwiązywalności zadań oparte są na regresji średnich wyników punktowych krajów biorących udział w badaniu komputerowym PISA 2015 ze względu na przeciętną rozwiązywalność zadań z poszczególnych dziedzin. Wykorzystano regresję krzywoliniową (do wielomianu trzeciego stopnia), osiągając dopasowanie modelu: dla czytania i interpretacji tekstu $R^2=0,972$, matematyki $R^2=0,986$ i rozumowania w naukach przyrodniczych $R^2=0,994$.

W przypadku rozumowania w naukach przyrodniczych efekt zmian w formie zadań nie ograniczał się jednak do samego „efektu narzędzia” mierzonego w badaniu próbnym w odniesieniu do zadań stosowanych w również poprzednich edycjach badania PISA. Efektowi temu towarzyszył jeszcze drugi, silniejszy efekt związany z użyciem w badaniu baterii zadań nowego typu, znacznie różniących się od zadań stosowanych dotychczas. W związku z tym, podjęliśmy drugi typ analiz, dotyczący porównania zadań kotwiczących oraz zadań nowych w badaniu rozumowania w naukach przyrodniczych. Oba typy zadań w badaniu głównym występowały jedynie w wersji komputerowej, jednak zadania kotwiczące zasadniczo imitowały wersję papierową, podczas gdy zadania nowe wykorzystywały w większym stopniu możliwości medium komputerowego (jak możliwość „manipulowania” obiektami na ekranie komputera, czy przeprowadzania symulacji). Analizy wykazały, że (przyjmując za punkt odniesienia rozwiązywalność poszczególnych zadań w populacji uczestników całego badania, a ściślej – tych krajów, które uczestniczyły w badaniu komputerowym, gdyż tylko w tej postaci wykorzystywane były nowe zadania) polscy uczniowie z zadaniami „kotwiczącymi” radzili sobie relatywnie lepiej niż z zadaniami nowymi: przeciętna rozwiązywalność zadań kotwiczących jest wśród polskich uczniów o 8,3 punktu procentowego wyższa niż w całej próbie wszystkich krajów realizujących badanie za pomocą komputerów, podczas gdy dla zadań nowych różnica wynosi jedynie 4,9 pp na korzyść polskich uczniów. Wynik ten może wskazywać na to, że problemem polskich uczniów jest nie tylko sam fakt użycia w badaniu komputerów, ale także forma zadań, odmienna od tradycyjnej. Jest dość prawdopodobne, że gdyby w badaniu PISA 2015 w zakresie rozumowania w naukach przyrodniczych pomiar dokonywany był za pomocą tych samych narzędzi co w roku 2012, przeciętny wynik polskich uczniów również byłby bardzo zbliżony do wyniku z PISA 2012¹⁸. Oznaczałoby to, że zmiana wyniku między rokiem 2012 i 2015 nie tyle wskazuje na spadek umiejętności, co wynika z ujawnienia przez nowe zadania z roku 2015 problemów, które nie są w polskiej szkole nowe – ale wcześniej w niewielkim stopniu miały wpływ na wyniki pomiarów.

Tabela 2.2.

		Rozumowanie w naukach przyrodniczych	
		zadania kotwiczące	zadania nowe
K	Przeciętna rozwiązywalność – PISA 2015, całe badanie	47,4%	44,4%
Kp	Przeciętna rozwiązywalność – PISA 2015, Polska	55,7%	49,3%
	K-Kp	-8,3%	-4,9%

Trzeba mieć świadomość, że „efekty narzędzia” są różne w przypadku poszczególnych zadań – są zadania, których rozwiązywalność (tak wśród polskich uczniów, jak wśród ogółu badanych) w wyniku przejścia na testy komputerowe wzrosła, jak i takie, których rozwiązywalność spadła; podobnie – wśród nowych zadań z rozumowania w naukach przyrodniczych są zarówno zadania łatwe, jak i trudne. Porównując rozwiązywalność poszczególnych zadań w wersji komputerowej i papierowej w niektórych przypadkach można się domyślać, z jakiego względu dana wersja jest trudniejsza lub łatwiejsza; w niektórych przypadkach jednak trudno nawet sformułować wstępną hipotezę – tym bardziej więc trudno wskazać przyczynę, dla której w przypadku danego zadania zmiana medium ma specyficzny wpływ na łatwość rozwiązywania zadania akurat przez uczniów z Polski. Odpowiedź na te

¹⁸ Dodatkowy efekt dla średniej wyniku polskich uczniów związanych z zastosowaniem „nowych zadań” szacujemy (również na podstawie obliczeń wykorzystujących modele regresyjne) na około 12-14 punktów, co po zsumowaniu z ogólnym „efektem narzędzia” oznaczałoby łączny efekt rzędu około 21-23 punktów. Należy jednak pamiętać, że wartości te mają charakter jedynie oszacowań i mogą być obciążone różnego rodzaju błędami.

pytania mogłyby dać dopiero pogłębione badania jakościowe nad procesem rozwiązywania zadań w wersji komputerowej i papierowej¹⁹.

Musimy także powtórzyć, że skala badania próbnego - jeśli wziąć pod uwagę dane z tylko jednego kraju - była dość ograniczona²⁰. Z tego względu - trzeba to podkreślić - analizy, jakie prowadziliśmy, nie dostarczają „twardych dowodów”, dają jednak argumenty dla sformułowania hipotezy, że za przynajmniej część, i to znaczną, spadku średnich wyników polskich uczniów w badaniu PISA 2015 odpowiada przejście na testy komputerowe: innymi słowy, że gdyby badanie PISA 2015 prowadzone było za pomocą takich samych, papierowych narzędzi, jak PISA 2012, wyniki polskich uczniów byłyby znacząco wyższe. Przypuszczalnie testy komputerowe mierzyły trochę inny zestaw umiejętności niż „analogiczne” testy papierowe. Jeśli wniosek ten jest trafny, to oznacza to jednak, że polscy uczniowie (co można było już dostrzec w wynikach badań komputerowych w PISA 2009 i 2012) mają poważne problemy z wykorzystywaniem narzędzi komputerowych w innych czynnościach niż typowo młodzieżowe użytkowanie nowych mediów, kształtowane w oderwaniu od nauki w relacjach rówieśniczych. Jeśli jest to trafna hipoteza, oznaczałoby to, że przed szkołą stoi poważne zadanie wykształcenia u młodzieży bardziej świadomego korzystania z narzędzi elektronicznych, które może w przyszłości okazać się przydatne do rozwiązywania realnych problemów z użyciem nowych technologii.

2.5. Podsumowanie

Wyniki badania PISA 2015 są porównywalne (a więc wyrażone na tej samej skali) co wyniki poprzednich edycji badania. Taka „porównywalność” nigdy nie ma jednak charakteru bezwzględnego - zawsze obciążona jest pewnym marginesem niepewności, związanym z błędami o charakterze losowym lub systematycznym. Skala zmian metodologicznych - przede wszystkim zmiana narzędzia z testów papierowych na aplikację komputerową, ale także zmiany szczegółowych rozwiązań stosowanych przy skalowaniu wyników margines ten z całą pewnością poszerzają. Zmiany te jednak musiały w którymś momencie nadejść, a w każdym kolejnym badaniu realizowanym za pomocą nowoczesnych technik komputerowych niepewność, którą wprowadzają, będzie coraz mniejsza, podobnie jak w coraz większym stopniu będziemy korzystali z możliwości, jakie nowe techniki testowania dają badaczom. Trzeba jednak jeszcze raz podkreślić, że istnienie tego „marginesu niepewności” powinno skłaniać do ostrożności w formułowaniu wniosków, ale przy jej zachowaniu wnioski dotyczące porównań między krajami, jak również porównań w czasie między poszczególnymi edycjami badania OECD PISA mogą być w pełni uprawnione.

W badaniu PISA 2015 obserwujemy systematyczny spadek przeciętnych wyników w krajach, w których w poprzednich edycjach wyniki te były wysokie, i wzrost w krajach o wynikach niższych, choć spadki te i wzrosty są oczywiście różne w różnych krajach i nie da się wskazać żadnej prostej reguły je wyjaśniającej. W efekcie, pomimo że wyrażone w punktach przeciętne wyniki polskich uczniów są w 2015 roku w poszczególnych dziedzinach nieco niższe niż w roku 2012, Polska generalnie utrzymała swoją wysoką pozycję. Można postawić pytanie, na ile zmiany wyników, które obserwujemy, związane są ze zmianami w konstrukcji badania - a więc właśnie są efektami ruchów w granicach tych „marginesów niepewności”, które poprzez staranne przygotowanie badania staramy się zminimalizować, ale których nigdy nie wyeliminujemy. Znalezienie na nie pełniejszej odpowiedzi nie jest jednak łatwe i wymagać będzie jeszcze wielu analiz i dalszych badań.

¹⁹ Dodatkowym problemem jest w tym przypadku fakt, że mówimy tu o zadaniach łączących, które zgodnie z założeniami programu PISA nie mogą zostać ujawnione (będą używane w kolejnych edycjach badania); z tego względu nie możemy podać niestety konkretnych przykładów zadań łatwiejszych lub trudniejszych w wersji komputerowej.

²⁰ Osobnym problemem jest ograniczony dostęp do danych z badania próbnego; z tego względu nie mieliśmy możliwość przeprowadzenia wszystkich pożądaných analiz, w tym oceny istotności statystycznej obserwowanych różnic.

Bibliografia

- Dolata, R., Jakubowski, M., Pokropek, A. (2013). *Polska oświata w międzynarodowych badaniach umiejętności uczniów PISA OECD : wyniki, trendy, kontekst i porównywalność*. Wydawnictwa Uniwersytetu Warszawskiego.
- Federowicz, M. i in. (2013). *OECD PISA 2012. Wyniki badania 2012 w Polsce*. Ministerstwo Edukacji Narodowej.
- Federowicz, M., i in. (2010). *Program Międzynarodowej Oceny Umiejętności Uczniów OECD PISA. Wyniki badania 2009 w Polsce*. Ministerstwo Edukacji Narodowej.
- Jakubowski, M., Pokropek, A. (2009). *Badając egzaminy : podejście ilościowe w badaniach edukacyjnych*. Centralna Komisja Egzaminacyjna.
- Pokropek, A. (red.) (2015). *Modele cech ukrytych w badaniach edukacyjnych, psychologii i socjologii : teoria i zastosowania*. Instytut Badań Edukacyjnych.
- PISA 2015 Technical Report* (draft version), <http://www.oecd.org/pisa/data/2015-technical-report/> (pobrano 20 lutego 2017)

3. POPULACJA I PRÓBA W BADANIU PISA 2015

Jacek Haman

Badanie OECD PISA objęci są uczniowie piętnastoletni, a dokładniej rzecz biorąc – uczniowie, którzy ukończyli 15 lat w roku poprzedzającym badanie; w przypadku PISA 2012 odpowiadało to uczniom urodzonym w 1999 roku (znajdowali się więc wśród nich również tacy, którzy w chwili badania ukończyli już lat szesnaście; operacyjna definicja „piętnastolatka” przyjmowana w badaniu nie jest zatem w pełni zgodna z naturalną interpretacją słowa „piętnastolatek”). Zdecydowaną większość tej grupy wiekowej stanowią uczniowie gimnazjów (przede wszystkim klasy III, ale również klas II i I). Oprócz tego do badanej populacji należeli piętnastoletni uczniowie szkół artystycznych (przede wszystkim ogólnokształcących szkół muzycznych II stopnia)²¹, oraz piętnastoletni uczniowie szkół ponadgimnazjalnych (liceów lub średnich i zasadniczych szkół zawodowych). Z założenia, z populacji badania wyłączeni byli natomiast uczniowie szkół specjalnych i piętnastoletni uczniowie szkół podstawowych (wyłączenia na poziomie szkół), a także uczniowie nie mogący pisać testu ze względu na niepełnosprawność lub niewystarczającą znajomość języka polskiego (wyłączenia wewnętrzne). Do badanej populacji nie należeli także piętnastolatki, którzy z jakichkolwiek przyczyn znajdowaliby się poza polskim systemem szkolnym (w szczególności – jeśli realizują obowiązek szkolny poza Polską²²). Dane o populacji badanej w PISA 2015 podaje Tabela 3.1.

Tabela 3.1.

Piętnastolatki: rok urodzenia	1999
Liczba piętnastoletnich gimnazjalistów ²³ (bez szkół specjalnych)	352,8 tys.
Liczba piętnastoletnich uczniów Liceów Ogólnokształcących	2,6 tys.
Liczba piętnastoletnich uczniów średnich szkół zawodowych	0,6 tys.
Liczba piętnastoletnich uczniów zasadniczych szkół zawodowych	0,1 tys.
Razem	356,1 tys
Szacunkowa wielość wyłączeń wewnętrznych	3,8 tys
Łącznie wielkość badanej populacji	352,3 tys
Liczba gimnazjów (bez szkół specjalnych)	6639

Badanie PISA realizowane jest na reprezentatywnej próbie losowej. Schemat doboru próby uczniów w badaniu PISA ma charakter dwustopniowego doboru warstwowego z zastosowaniem losowania

²¹ Jeśli szkoła artystyczna prowadziła naukę na poziomie gimnazjum (np. ogólnokształcące szkoły muzyczne II stopnia), na użytek doboru próby traktowana była tak, jakby była gimnazjum; także dalej w tym tekście tam, gdzie mowa jest o gimnazjach, włącza się w to również ogólnokształcące szkoły muzyczne II stopnia.

²² W takim przypadku mogą jednak należeć do populacji badanej w innym kraju.

²³ Włączając ogólnokształcące szkoły artystyczne II stopnia.

systematycznego, w którym pierwszym stopniem doboru był wybór szkoły, zaś drugim – losowanie uczniów z uprzednio wylosowanych szkół (również w trybie systematycznego losowania warstwowego). Pierwszy etap losowania – losowanie szkół – realizowane jest przez wchodzącą w skład konsorcjum międzynarodowego amerykańską firmę badawczą WESTAT; rolą polskiego zespołu PISA jest dostarczenie pełnej bazy szkół wraz z niezbędnymi do przeprowadzenia losowania danymi. Drugi etap – losowanie uczniów – realizowany jest przez krajowy zespół PISA przy użyciu narzędzia dostarczonego przez WESTAT (co, w szczególności, oznacza, że krajowy zespół nie ma wpływu na wynik losowania, natomiast nie ma potrzeby przekazywania do WESTATu dodatkowych danych polskich uczniów).

Zachowanie ciągłości zasad doboru próby w badaniu PISA jest jednym z podstawowych środków zapewniania porównywalności wyników w poszczególnych edycjach badania – stąd procedury zastosowane w badaniu PISA 2015 były praktycznie takie same, jak w roku 2012. Drobne zmiany w porównaniu z poprzednimi edycjami badania wiązały się z faktem, że w roku 2015 w badaniu polskim nie realizowano żadnych opcji krajowych (a więc nie było ani dodatkowego badania uczniów I klas szkół ponadgimnazjalnych; ani nadreprezentacji szkół prywatnych); z drugiej strony, w porównaniu z badaniem 2015 opcjonalne międzynarodowe badanie *Financial Literacy* realizowane było na próbie głównej badania (w PISA 2012 – na dodatkowej próbie uczniów). W efekcie reguły doboru próby do polskiego badania OECD PISA 2012 były maksymalnie bliskie podstawowemu schematowi przewidzianemu dla badania międzynarodowego.

W losowaniu szkół zastosowano podział na warstwy jawne (*explicite*) wyznaczone ze względu na typ szkoły (gimnazja, licea, szkoły zawodowe); w tych trzech kategoriach losowanie było prowadzone osobno. Dzięki stosowaniu metody losowania systematycznego możliwe było wykorzystanie również warstw *implicite* (określających uporządkowanie operatu losowania) – były nimi (w hierarchii od najważniejszej do najmniej ważnej): w warstwie szkół zawodowych, podział na szkoły zasadnicze i technika; we wszystkich warstwach: publiczność (szkoły publiczne, szkoły prywatne), wielkość miejscowości (miasto na prawach powiatu, pozostałe miasta powyżej 5 tys. mieszkańców, miasta do 5 tys. mieszkańców, wieś); skład szkoły ze względu na płeć (szkoły z dominującym udziałem dziewcząt, szkoły mieszane, szkoły z dominującym udziałem chłopców); ostatnią cechą (ciągłą) porządkującą operat losowania była wielkość (liczba uczniów) szkoły.

Do polskiego badania OECD PISA 2015 wylosowano próbę 160 gimnazjów (w tej liczbie – dwie ogólnokształcące szkoły muzyczne II stopnia), 12 liceów ogólnokształcących oraz 53 szkół zawodowych (32 technika i 21 szkół zasadniczych); dodatkowo, dla każdej wylosowanej szkoły, wylosowane zostały dwie szkoły rezerwowe, na wypadek gdyby wylosowana szkoła odmówiła udziału w badaniu. Spośród 65 wylosowanych szkół ponadgimnazjalnych do badania zostało zakwalifikowanych jedynie 10: 6 liceów i 4 technika (w pozostałych – w tym we wszystkich wylosowanych szkołach zasadniczych – nie było żadnych uczniów piętnastoletnich); z czego ostatecznie uwzględniono w badaniu wyniki uczniów 9 szkół²⁴. Spośród 160 wylosowanych gimnazjów z próby zasadniczej, 17 szkół odmówiło udziału w badaniu i zostało zastąpione odpowiednimi szkołami rezerwowymi (tak więc przebadano łącznie – zgodnie z zamierzeniem – 160 gimnazjów). Dane o wylosowanej i zrealizowanej próbie zawarte są w Tabeli 3.2.

²⁴ Wynikało to z faktu, że w jednej ze szkół frekwencja wylosowanych uczniów w czasie badania wynosiła 0%; tzn. na badanie nie przyszedł jedyny uczący się w niej piętnastolatek.

Tabela 3.2.

	Gimnazja	Licea Ogólnokształcące	Technika	Zasadnicze Szkoły Zawodowe	Razem
Liczba wylosowanych szkół (próba zasadnicza)	160	12	32	21	225
Liczba wylosowanych szkół, w których uczyli się uczniowie piętnastoletni	160	6	4	0	170
Liczba szkół z próby zasadniczej biorących udział w badaniu	143	6	3	0	152
Liczba szkół rezerwowych biorąca udział w badaniu	17	0	0	0	17
Liczba wylosowanych uczniów	5182	13	5	0	5200
Liczba wylosowanych uczniów wyłączonych z próby	55	1	0	0	56
Próba po wyłączeniach	5127	12	5	0	5144
Liczba uczniów biorąca udział w badaniu	4466	8	4	0	4478
Poziom realizacji próby	87,1%	66,7%	80,0%	-	87,1%

Losowanie uczniów do badania również opierało się na schemacie losowania systematycznego z warstwami implicite (uporządkowanie operatu ze względu na płeć i klasę – a więc III, II i I-klasistów); zgodnie z tym schematem przydzielano uczniom również konkretne zestawy testowe i opcje badania. Do badania wylosowanych zostało łącznie 5200 uczniów, w tym 5182 uczniów gimnazjów, 13 uczniów liceów oraz 5 uczniów techników. Z liczby tej wyłączono 56 uczniów ze względu na niespełnianie innych kryteriów udziału w badaniu (jak nieznamość języka, niepełnosprawność uniemożliwiająca udział w teście lub ciężka dysleksja), natomiast 665 uczniów nie brało udziału z badania z powodu braku zgody rodziców bądź nieobecności. Ostateczna próba zrealizowana wyniosła 4478 uczniów, w tym 4466 uczniów gimnazjów, 8 uczniów liceów oraz 4 uczniów techników. Oznacza to realizację wylosowanej próby na poziomie 87%, a więc minimalnie lepszym niż w roku 2012 (86%).

Jak widać, badanie polskie OECD PISA 2015, choć teoretycznie obejmuje uczniów zarówno gimnazjów, jak i szkół ponadgimnazjalnych, w praktyce jest prawie wyłącznie badaniem gimnazjalistów (99,7% przebadanej próby). Wynika to, z jednej strony, z faktu, że odsetek piętnastolatków w szkołach ponadgimnazjalnych to niespełna 1%; z drugiej jednak jest efektem ścisłego stosowania procedur losowania obowiązujących w badaniu międzynarodowym, akurat w tym przypadku niedostosowanych do specyfiki polskiej sytuacji, w wyniku którego ich udział w próbie wyniósł zaledwie 0,3% (w analizie wyników wielkość ta jest następnie korygowana do prawidłowej dzięki ważeniu danych). Trzeba jednak wyraźnie stwierdzić, że nawet, gdyby próba obejmowała odpowiedni odsetek uczniów liceów i techników – a więc byłoby ich w próbie nie 13, ale około 40 – i tak nie pozwalałoby to na jakiegokolwiek uogólnianie wniosków dotyczących tej kategorii uczniów,

a w większości sytuacji postępowanie zgodnie w wymogami formułowanymi przez konsorcjum międzynarodowe przynosiło badaniu zdecydowaną korzyść, wymuszając niejako trzymanie się najwyższych standardów metodologicznych.

3.1. Dobór próby a błędy losowe i nielosowe

Próba w badaniu PISA jest **próbą losową**, co oznacza, że dla każdego ucznia z badanej populacji można wyznaczyć znane i niezerowe prawdopodobieństwo, że zostanie on do próby wylosowany (w badaniu polskim wynosiło ono przeciętnie ok. 0,01459; wartość ta mogła się wahać dla uczniów różnych szkół). Próba losowa jest z definicji **próbą reprezentatywną**, jeśli prawdopodobieństwa trafienia do niej są równe dla wszystkich członków populacji; jeśli nie są – staje się reprezentatywna po nadaniu jej elementom odpowiednich **wag**, odwrotnie proporcjonalnych do prawdopodobieństw trafienia do próby (tzw. ważenie poststratyfikacyjne). Nierówne prawdopodobieństwa trafienia do próby mogą wynikać np. z celowego nadreprezentowania jakichś podzbiorowości (takich, którym z jakichś względów chcemy poświęcić w badaniu szczególną uwagę), bądź też ich subreprezentacji (np. w celu ograniczenia kosztów związanych z badaniem tam, gdzie wiąże się ono ze szczególnie wysokimi nakładami). Takie nad- i subreprezentowanie poszczególnych części populacji nie narusza reprezentatywności próby pod warunkiem odpowiedniego jej ważenia, natomiast może być korzystne ze względów metodologicznych lub na logistykę badania. W roku 2015 nie stosowaliśmy w badaniu polskim żadnych dodatkowych procedur nad- lub subreprezentacji (w niektórych poprzednich edycjach stosowano nadreprezentację szkół prywatnych); elementem wspólnego dla wszystkich krajów uczestniczących w badaniu schematu doboru próby jest natomiast subreprezentacja szkół bardzo małych, realizowana ze względów logistycznych.

Wyniki uzyskane na próbie reprezentatywnej można uogólniać na populację, z której została ona wylosowana, jednakże są one obarczone **błędami losowymi**. Wielkość tych błędów, oczywiście, nie jest znana, jednakże, opierając się na zasadach rachunku prawdopodobieństwa, można określić wielkość, której błąd losowy nie przekroczy z określonym prawdopodobieństwem (zwykle 95%), wyznaczając w ten sposób **przedziały ufności** dla poszczególnych wyników, których szerokość (a ściślej jej połowę) traktuje się zwykle jako dokładność oszacowania (tzw. „błąd statystyczny”). Dokładność oszacowań (a zatem potencjalna wielkość błędów losowych) zależy od dwóch²⁵ czynników: wielkości próby (im większa próba, tym większa dokładność oszacowań) oraz schematu jej doboru, który może przyczynić się zarówno do zwiększenia, jak i zmniejszenia dokładności w porównaniu do tzw. „prostej próby losowej”.

Schemat doboru próby uczniów w badaniu PISA ma charakter dwustopniowego doboru warstwowego z zastosowaniem losowania systematycznego, w którym pierwszym stopniem doboru był wybór szkoły, zaś drugim – losowanie uczniów z uprzednio wylosowanych szkół. Zastosowanie doboru warstwowego umożliwia zmniejszenie skali błędów losowych, natomiast dwustopniowy charakter losowania prowadzi do ich zwiększenia w porównaniu do hipotetycznego badania realizowanego na próbie losowej prostej; wielkość błędów losowych w porównaniu do próby prostej jest wypadkową tych dwóch czynników, przy czym znacznie silniejszy jest efekt związany z doбором wielostopniowym: błędy losowe wyników uzyskiwanych na liczącej ok. 5000 elementów próbie mają skalę odpowiadającą kilkusetelementowej próbie prostej. Nie oznacza to oczywiście, że lepszą próbą byłaby właśnie próba prosta: ze względów organizacyjnych znacznie łatwiejsze i tańsze do przeprowadzenia jest badanie na próbie 5000 uczniów zgrupowanych w 180 szkołach, niż, powiedzmy, 800 uczniów, z których każdy jest uczniem innej szkoły; niezależnie od tego, badanie wielu uczniów w tej samej

²⁵ Zależy ona oczywiście również od rozkładu populacyjnego badanej cechy, ta jednak jest czynnikiem niezależnym od charakterystyki próby losowej.

szkole jest niezbędne, by móc w badaniu charakteryzować nie tylko zatomizowaną społeczność uczniów, ale także wypowiedać się o systemie edukacyjnym, w którego skład wchodzi całe społeczności szkolne.

Zastosowanie złożonego schematu doboru próby oznacza jednak, że do określania dokładności oszacowań nie można stosować standardowych technik obliczeniowych – proste wykorzystanie metod zaimplementowanych w typowych programach statystycznych prowadziłoby do niedoszacowania błędów. Z tego względu błędy losowe wyników w badaniu PISA szacowane są z wykorzystaniem metod replikacyjnych, a ściślej – techniki *balanced random replicates* w wariancie Fay'a. Metoda ta pozwala na wyznaczanie przedziałów ufności oraz weryfikację hipotez statystycznych z uwzględnieniem specyfiki przyjętego schematu doboru próby.

Mówiąc o reprezentatywności próby oraz dokładności oszacowań, należy pamiętać o kilku zasadach:

- Dokładność oszacowań zależy od liczebności próby, natomiast jest praktycznie niezależna od tego, jaką część populacji stanowi próba²⁶. W konsekwencji, przy tym samym schemacie doboru próby, próba o liczebności 5000 osób da taką samą dokładność oszacowania niezależnie od tego, czy wylosowana ona była z populacji liczącej sto tysięcy, milion czy dziesięć milionów ludzi.
- Schemat doboru próby, a także sposób jej warstwowania czy ewentualne sub- i nadreprezentacje, o ile są przeprowadzone prawidłowo i właściwie uwzględnione przy ważeniu próby, nie mają wpływu na jej reprezentatywność (choć mają wpływ na wielkość błędów losowych). W konsekwencji w pełni dopuszczalne jest porównywanie wyników uzyskanych z prób wylosowanych przy użyciu różnych schematów doboru czy wykorzystujących różne warstwowanie. Pomimo tego w kolejnych cyklach badania PISA dąży się do utrzymania stałych schematów doboru i warstwowania próby, jednakże ewentualne ich zmiany nie stanowią przeszkody w porównywaniu wyników badania PISA z różnych lat.

3.2. Realizacja badania a błędy nielosowe

Oprócz błędów losowych, wyniki każdego badania mogą być obciążone różnorodnymi *błędami nielosowymi (systematycznymi)*. Ich źródłem może być nie trafność narzędzi pomiarowych, błędy proceduralne w realizacji badania, a przede wszystkim – niepełna realizacja próby. W przeciwieństwie do błędów losowych, których wartości wprawdzie nie znamy, ale ich skalę możemy przewidzieć, błędy nielosowe nie poddają się opisowi probabilistycznemu i nie da się łatwo wskazać, jaką wielkość czy kierunek mogą osiągnąć. Błędów takich nie da się nigdy całkowicie uniknąć, jednak staranne przygotowanie i realizacja badania ma służyć ich minimalizacji.

Poziom realizacji próby uczniów w polskim badaniu OECD PISA 2015 wynosił 87% – był więc nieco wyższy niż w badaniu 2012 (ok. 86%)²⁷ i spełniał wymagania standardów technicznych badania PISA (min. 80%). Również odsetek szkół z próby zasadniczej uczestniczących w badaniu (89%) znacznie przekracza wymogi międzynarodowe i jest nieznacznie wyższy niż w 2012 roku (88%). Wielkości te warto porównać z poziomami realizacji prób w typowych badaniach społecznych na próbach reprezentatywnych dorosłej ludności Polski, które w przypadku najlepiej realizowanych badań

²⁶ Pomijamy tutaj sytuację, gdy próba stanowi rzeczywiście znaczącą część (np. 80%) populacji.

²⁷ Wg metodologii przyjętej przez Konsorcjum Międzynarodowe.

akademickich zwykle nie przekraczają 65%, a w przypadku badań komercyjnych (np. sondaże przedwyborcze) nie osiągają nawet 40% założonej próby.

Można jednak postawić pytanie, czy – a jeśli tak, to na ile – niepełna realizacja próby wpływa na uzyskane wyniki, w szczególności – na średnie liczby punktów uzyskanych w testach kompetencji przez polskich uczniów. Byłoby tak, gdyby przeciętny poziom umiejętności uczniów wylosowanych do próby, ale nieuczestniczących w badaniu, był inny niż uczniów, którzy udział w badaniu wzięli. Tak najprawdopodobniej rzeczywiście jest – wskazują na to np. różnice w poziomie realizacji próby wśród wylosowanych uczniów klasy III gimnazjum (88%) a klas I i II (zaledwie 66%). Pojawia się zatem pytanie, jak duży jest wpływ tego zjawiska na wynik badania oraz na możliwość porównywania wyników w czasie lub między krajami. Kwestia ta była przedmiotem szczegółowej analizy polskich wyników badań PISA 2006, 2009 i 2012, gdzie w oparciu o zbierane dane o wynikach egzaminów zewnętrznych uczniów wylosowanych do próby (zarówno biorących, jak i niebiorących udziału w badaniu) szacowano wielkość efektu niepełnej realizacji próby na około 1-3 punkty. Ponieważ parametry realizacji próby (zarówno jej ogólny poziom, jak i struktura) w roku 2015 są podobne jak w poprzednich edycjach badania, można oczekiwać, że również w badaniu PISA 2015, gdyby – hipotetycznie – uzyskano 100% realizacji próby, średnie wyników polskich uczniów okazałyby się mniejsze o również ok 1 do 3 punktów w porównaniu z ogłoszonymi wynikami badania (co jest wartością mniejszą, niż niepewność wyniku związana z błędami losowymi).

Należy jednak podkreślić, że wyniki badania PISA mają z zasady charakter relatywny. Porównując wyniki 2015 do poprzednich edycji badania, należy brać pod uwagę przede wszystkim to, czy przewidywana wielkość błędów systematycznych nie uległa zmianie. Ponieważ charakterystyka realizacji próby PISA 2015 w porównaniu z poprzednimi edycjami nie uległa zauważalnej zmianie, można przyjąć, że również ewentualne zniekształcenia wyników z nią związane są podobne – a zatem zmian średniego poziomu umiejętności między poszczególnymi edycjami badania nie można wiązać z niepełną realizacją próby, choć – zwłaszcza w bieżącej edycji badania – nie można wykluczyć, że ma ona pewien związek np. ze zmianą narzędzia badawczego z testów papierowych na komputerowe.

W kwestii porównań międzynarodowych brak jest danych dotyczących wpływu niepełnej realizacji próby dla średnich wyników uzyskanych w innych krajach niż Polska, można jednak przypuszczać, że podobne mechanizmy jak w Polsce mogą działać także w innych krajach. Ponieważ poziom realizacji próby w Polsce jest podobny jak w innych krajach uczestniczących w badaniu, można oczekiwać, że również wielkość efektów niepełnej realizacji próby będzie podobna. Tak więc niepełna realizacja próby nie stanowi istotnej przeszkody dla dokonywania międzynarodowych porównań wyników – choć, oczywiście, stanowi ona pewien dodatkowy czynnik niepewności.

4. ROZUMOWANIE W NAUKACH PRZYRODNICZYCH

Barbara Ostrowska, Krzysztof Spalik

Istotą dobrej edukacji przyrodniczej jest przede wszystkim wykształcenie u uczniów umiejętności *rozumowania naukowego*. Tę właśnie kluczową umiejętność mierzy badanie PISA w zakresie kompetencji w naukach przyrodniczych (*science literacy*). Badanie to więc służy nie tyle oszacowaniu wiedzy uczniów w zakresie nauk przyrodniczych, ile sprawdzeniu, czy potrafią oni tę wiedzę zastosować, zwłaszcza kiedy muszą twórczo rozwiązać problemy przedstawione w różnych, a nie tylko szkolnych kontekstach. Powszechna i wysokiej jakości edukacja matematyczno-przyrodnicza jest warunkiem postępu naukowego i technicznego, bez którego nie ma rozwoju cywilizacyjnego. Kanon wiedzy przyrodniczej, przede wszystkim biologicznej, zmienia się nieustannie, nowe wyzwania przynoszą także zdobyte techniki, zwłaszcza w zakresie dostępu do informacji oraz jej przetwarzania. Istnieje jednak coś, co wyróżnia dociekanie naukowe i stanowi jego fundament – to sposób weryfikacji twierdzeń nauki, bazujący na próbie ich obalenia za pomocą doświadczeń i obserwacji, czyli przez *falsyfikację*. Pewność twierdzeń nauki wynika z tego, że przetrwały one nieustanne próby wytrzymałości. Formułowanie hipotez i ich falsyfikowanie za pomocą danych empirycznych jest więc podstawą metody naukowej.

W 2015 r. rozumowanie w naukach przyrodniczych było główną dziedziną pomiaru po raz drugi w historii tego badania.

Mierzona w badaniu PISA biegłość naukowa wymaga opanowania trzech aspektów wiedzy naukowej: znajomości faktów, standardowej metodyki pozyskiwania danych (zwłaszcza prowadzenia obserwacji i doświadczeń) oraz reguł wnioskowania naukowego. Umiejętność ta odnosi się nie tylko do nauk przyrodniczych w ich aspekcie poznawczym, ale także do ich praktycznych zastosowań, np. w zakresie techniki lub medycyny. Naukowcy poszukują odpowiedzi na pytania dotyczące natury materialnego świata, podczas gdy technolodzy, inżynierowie czy lekarze szukają optymalnego rozwiązania naszych, ludzkich problemów. Oba aspekty nauki są jednak silnie ze sobą związane, a od wykształconego obywatela należy oczekiwać umiejętności uczestniczenia w dyskusji na tematy nauki i techniki, jak również podejmowania racjonalnych decyzji związanych z tymi zagadnieniami. Decyzje te dotyczą nie tylko indywidualnych wyborów, np. segregacji śmieci lub zakupu oszczędnego energetycznie sprzętu gospodarstwa domowego, ale także – przez mechanizm demokratyczny – określania kierunków przemian społeczeństwa, np. związanych z rozwojem zrównoważonym. W tym aspekcie edukacja przyrodnicza jest niezbywalnym elementem edukacji obywatelskiej.

Pomiar kompetencji piętnastolatków w zakresie rozumowania naukowego obecny był w badaniu PISA od jego pierwszej edycji w 2000 r., ale dopiero w 2006 r. był główną dziedziną pomiaru. Oznacza to, że kompetencje przyrodnicze piętnastolatków zostały wtedy wszechstronnie sprawdzone za pomocą rozbudowanego zestawu zadań, obejmujących różne aspekty wiedzy przyrodniczej oraz różne składowe umiejętności rozumowania i wnioskowania naukowego. Pozwoliło to na miarodajne wyskalowanie wyników badania. Dlatego, jak przypominano w kolejnych raportach, wyniki badań z lat 2000 i 2003 nie są w pełni porównywalne z prowadzonymi od 2006 r.

4.1. Założenia teoretyczne badania

4.1.1. Nowe założenia ramowe pomiaru

Założenia badania 2015²⁸ są podobne do tych z 2006²⁹, jednak dokonano pewnych zmian, np. określenie badanych obszarów wiedzy zostało na nowo zdefiniowane. Jednocześnie po raz pierwszy całe badanie przeprowadzone było w większości krajów z wykorzystaniem komputera, nie zaś – jak poprzednio – w formie papierowej.

W badaniu 2015 przyjęto, że na pomiar kompetencji piętnastolatków w naukach przyrodniczych składają się następujące aspekty: kontekst zadania oraz wiedza, umiejętności i postawy uczniów (Tabela 4.1.).

Tabela 4.1. Aspekty pomiaru rozumowania w naukach przyrodniczych

Kontekst	Zagadnienia osobiste, miejscowe lub narodowe oraz globalne, zarówno współczesne, jak i historyczne, które wymagają rozumienia zagadnień z zakresu nauki i techniki.
Wiedza	Znajomość najważniejszych faktów i pojęć oraz zrozumienie teorii, które tworzą podstawę wiedzy naukowej. Wiedza ta obejmuje znajomość zjawisk i procesów zachodzących w przyrodzie (<i>content knowledge</i>), znajomość procedur badawczych np. planowania doświadczeń i obserwacji (<i>procedural knowledge</i>) oraz zrozumienie podstaw wnioskowania naukowego (<i>epistemic knowledge</i>).
Umiejętności	Wyjaśnianie zjawisk przyrodniczych w sposób naukowy, planowanie i ocena poprawności procedur badawczych, interpretacja danych i dowodów naukowych.
Postawy	Zespół postaw obejmujących m.in. zainteresowanie nauką i techniką, docenianie naukowego podejścia do zdobywania wiedzy oraz świadomość wpływu człowieka na środowisko.

4.1.2. Kontekst zadań

Bardzo ważnym aspektem badania PISA jest kontekst, w jakim sprawdzane są umiejętności uczniów. Choć nawiązuje on do podstaw programowych i programów nauczania krajów biorących udział w badaniu, to zadania rozwiązywane przez uczniów nie mają typowo szkolnego charakteru, ale odnoszą się do sytuacji znanych z codziennego życia, do spraw ważnych lokalnie i globalnie, do różnorodnych zjawisk przyrodniczych. Niekiedy przywoływane są zdarzenia historyczne, aby sprawdzić, czy uczniowie rozumieją procesy czasowe oraz potrafią dostrzec znaczenie rozwoju nauki. W Tabeli 4.2. przedstawiono różnorodność zadań pod kątem ich kontekstu.

²⁸ OECD (2016). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic and Financial Literacy*. PISA, OECD Publishing, Paryż.

²⁹ Zespół Badania PISA w Polsce (2007). *Program Międzynarodowej Oceny Umiejętności Uczniów OECD PISA: Wyniki badania 2006 w Polsce*. Ministerstwo Edukacji Narodowej, Warszawa.

Tabela 4.2. Konteksty zadań w badaniu rozumowania w naukach przyrodniczych PISA 2015

	Osobisty	Lokalny/narodowy	Globalny
Zdrowie i choroby	Profilaktyka zdrowia, wypadki, odżywianie się	Zapobieganie rozprzestrzenianiu się chorób, wybór pożywienia, zdrowie publiczne	Epidemie chorób zakaźnych
Zasoby naturalne	Osobiste zużycie materiałów i energii	Potrzeby populacji ludzkich, jakość życia, bezpieczeństwo, produkcja i dystrybucja żywności, zaopatrzenie w energię	Zasoby odnawialne i nieodnawialne, wzrost ludności, zrównoważone wykorzystanie gatunków
Jakość środowiska	Działania korzystne dla środowiska, wykorzystanie i usuwanie materiałów i urządzeń	Rozmieszczenie ludności, usuwanie odpadów, wpływ różnych działań na środowisko naturalne	Różnorodność biologiczna, zrównoważone wykorzystywanie środowiska, ograniczenie zanieczyszczeń, produkcja/utrata biomasy, erozja gleby
Zagrożenia	Ocena ryzyka związanego z trybem życia	Nagłe zmiany (np. trzęsienia ziemi, gwałtowne zjawiska pogodowe), powolne, stopniowe zmiany (np. erozja brzegowa, sedymentacja), ocena ryzyka	zmiany klimatyczne, oddziaływanie współczesnej komunikacji
Nowe wyzwania nauki i techniki	Naukowe aspekty hobby, sportu, muzyki i wykorzystywania urządzeń technicznych	Nowe materiały, urządzenia i procesy, modyfikacje genetyczne, technologie medyczne, transport	Wymieranie gatunków, badania kosmosu, pochodzenie i budowa Wszechświata

4.1.3. Wiedza naukowa

Aby wykonać zadania, uczniowie powinni dysponować odpowiednią wiedzą, w której można wyróżnić:

- wiedzę o **treściach nauki** (*content knowledge*) obejmuje znajomość faktów, pojęć i teorii naukowych wyjaśniających złożoność przyrody i przebieg procesów naturalnych; do tej kategorii należy większość szkolnych treści nauczania przedmiotów przyrodniczych – biologii, chemii, fizyki i geografii;
- wiedzę o **procedurach badawczych** (*procedural knowledge*) to znajomość podstawowej metodyki badań naukowych, jak planowanie doświadczeń i obserwacji, świadomość niepewności pomiaru i wynikającej stąd konieczności powtarzania pomiarów i badań, określenie

znaczenia próby kontrolnej, znajomość sposobów analizy i prezentacji wyników;

- wiedzę o **poznaniu naukowym** (*epistemic knowledge*) to zrozumienie logicznych podstaw dociekania naukowego, na które składają się stawianie pytań badawczych i formułowanie hipotez, konstruowanie modeli zjawisk, weryfikacja hipotez przez obserwacje i doświadczenie, ale także świadomość znaczenia weryfikacji rezultatów jednych badaczy przez drugich poprzez recenzowanie prac lub powtarzanie badań, dzięki czemu twierdzenia nauki są wiarygodne.

W wyborze treści naukowych obecnych w zadaniach kierowano się następującymi przesłankami:

- treści naukowe mają powiązania z rzeczywistymi sytuacjami lub zdarzeniami;
- są to ważne koncepcje albo teorie naukowe o nieprzemijającym znaczeniu;
- treści są dostosowane do poziomu rozwojowego piętnastolatków.

Wyszczególnienie przedstawionych treści zawiera Tabela 4.3.

Tabela 4.3. Treści nauki (*content knowledge*) wymagane w zadaniach PISA 2015

Świat fizyczny
<ul style="list-style-type: none">▪ Budowa materii (model cząsteczkowy, wiązania)▪ Właściwości substancji (zmiany stanu skupienia, przewodnictwo cieplne i elektryczne)▪ Przemiany chemiczne (reakcje chemiczne, zmiany energii, kwasy i zasady)▪ Ruch i siła (prędkość, tarcie), oddziaływania fizyczne (magnetyczne, elektrostatyczne, grawitacyjne)▪ Energia i jej przemiany (zachowanie, rozpraszanie, reakcje chemiczne)▪ Zależności między energią a materią (fale świetlne, radiowe, dźwiękowe i sejsmiczne)
Świat organizmów żywych
<ul style="list-style-type: none">▪ Komórka (struktura i funkcjonowanie, DNA, komórki roślinne i zwierzęce)▪ Pojęcie organizmu (jedno- i wielokomórkowego)▪ Człowiek (zdrowie, odżywianie się, układy: pokarmowy, oddechowy, krążenia, wydalania, płciowy i zależności między nimi)▪ Populacje (gatunki, ewolucja, różnorodność biologiczna, zróżnicowanie genetyczne)▪ Ekosystemy (łańcuchy pokarmowe, przepływ energii i krążenie materii)▪ Biosfera (usługi ekosystemowe, zrównoważone wykorzystywanie)
Ziemia i kosmos
<ul style="list-style-type: none">▪ Budowa Ziemi (litosfera, atmosfera, hydrosfera)▪ Energia w układach Ziemi (źródła energii, klimat)▪ Zmiany w układach Ziemi (tektonika płyt, cykle geochemiczne, siły twórcze i niszczące)▪ Historia Ziemi (skamieniałości, powstanie i ewolucja życia)▪ Ziemia w przestrzeni kosmicznej (grawitacja, system słoneczny, galaktyki)▪ Historia i skala Wszechświata (rok świetlny, teoria wielkiego wybuchu)

Istotnym elementem kompetencji w zakresie przyrodoznawstwa jest świadomość, w jaki sposób wiedza naukowa jest pozyskiwana i weryfikowana, czyli znajomość procedur badawczych. Jest ona niezbędna nie tylko w celu prowadzenia samodzielnych obserwacji i doświadczeń, ale także aby krytycznie oceniać tezy nauki, a zwłaszcza dostrzegać, że cechują się one różnym stopniem uprawdopodobnienia. Uczeń powinien umieć wyjaśnić, czemu np. pomiar prędkości światła oraz pomiar liczebności populacji śledzi w Bałtyku zasadniczo różnią się pod względem dokładności (prędkość światła zmierzono bowiem wielokrotnie za pomocą bardzo dokładnych przyrządów, a liczebność populacji śledzi jest tylko pewnym szacunkiem, nie mamy bowiem możliwości policzenia wszystkich ryb). Zakres wiedzy o procedurach badawczych podaje Tabela 4.4.

Tabela 4.4. Wiedza o procedurach badawczych (*procedural knowledge*) wymagana w zadaniach PISA 2015

Procedury badawcze

- Pojęcie zmiennych, w tym zależnych, niezależnych i kontrolnych.
- Pojęcie pomiaru, cechy ilościowe (mieralne) i jakościowe (obserwowalne), skala, zmienne ciągłe i kategoryczne.
- Sposoby oszacowania i zmniejszenia niepewności, takie jak powtarzanie pomiaru i uśrednianie wyniku
- Sposoby oszacowania i zapewnienia powtarzalności pomiaru (zbieżność wyników wielokrotnego pomiaru) oraz jego precyzji (zbieżność między wartościami zmierzonymi a rzeczywistymi).
- Popularne metody podsumowania i prezentacji danych (tabele, schematy, wykresy) oraz ich właściwe użycie.
- Znaczenie kontroli zmiennych w planowaniu eksperymentu oraz wykorzystanie badania z randomizacją próby badanej i kontrolnej w celu uniknięcia błędu oraz określenia możliwych zależności przyczynowo-skutkowych.
- Dobór odpowiedniej procedury do określonego pytania badawczego, np. poszukiwanie prawidłowości, testowanie hipotezy przez eksperyment, obserwacja w terenie itd.

Wiedza o procedurach badawczych jest ściśle powiązana z wiedzą o poznaniu naukowym, czyli epistemiczną, której zakres w badaniu podaje Tabela 3.5. Należy podkreślić, że choć w niektórych przypadkach pojęcia i wyznaczniki nauki są w Tabeli 3.5. opisane specjalistycznym językiem (np. rodzaje rozumowania), to jednak są one dość intuicyjne i w pełni zrozumiałe dla piętnastolatków.

Tabela 4.5. Wiedza o poznaniu naukowym (*epistemic knowledge*) wymagana w zadaniach PISA 2015

Wiedza o poznaniu naukowym

Pojęcia i wyznaczniki nauki

- Natura naukowych obserwacji, faktów, hipotez, modeli i teorii.
- Nadrzędne cele nauki (wyjaśnianie świata przyrody) w odróżnieniu od techniki i technologii (optymalne rozwiązywanie potrzeb ludzi); określanie, czym są pytania i dane naukowe i techniczne.
- Wartości nauki, takie jak upowszechnienie (publikowanie) wyników, obiektywność, zapobieganie stronniczości.
- Rodzaje rozumowania wykorzystywanego w nauce, tj. dedukcyjne, indukcyjne, abdukcyjne, modelowanie, rozumowanie przez analogię.

Rola wymienionych pojęć i cech w tworzeniu wiedzy naukowej

- W jaki sposób dane i rozumowanie naukowe wspierają tezy nauki.
- Funkcja różnych form dociekania empirycznego w poszerzaniu wiedzy naukowej, ich cele (testowanie hipotez wyjaśniających albo opisywanie wzorców lub prawidłowości) oraz procedury (obserwacja, kontrolowany eksperyment, badania korelacyjne).
- W jaki sposób błąd pomiaru wpływa na stopień zaufania do wiedzy naukowej.
- Wykorzystanie i rola modeli fizycznych i abstrakcyjnych oraz ich ograniczenia.
- Rola współpracy i podejścia krytycznego oraz znaczenie recenzowania prac naukowych dla wiarygodności tez nauki.
- Rola wiedzy naukowej (wraz z innymi formami wiedzy) w określaniu i podejmowaniu problemów społecznych i technicznych.

4.1.4. Umiejętności

Na kompetencje przyrodnicze uczniów mierzone w badaniu PISA składają się trzy główne grupy umiejętności (Tabela 6.):

- wyjaśnianie zjawisk przyrodniczych w sposób naukowy,
- planowanie i ocena poprawności procedur badawczych,
- interpretacja danych i dowodów naukowych.

Umiejętności te odwołują się do opisanych powyżej kategorii wiedzy. Wyjaśnianie zjawisk przyrodniczych w sposób naukowy wymaga przede wszystkim (choć nie wyłącznie) opanowania treści nauki. Wiedza o procedurach badawczych jest niezbędna, aby uczeń mógł zaplanować eksperyment albo ocenić poprawność takiego planu, sporządzonego przez inną osobę. Interpretacja danych i dowodów naukowych wymaga wiedzy o regułach poznania naukowego.

Tabela 4.6. Umiejętności z zakresu kompetencji naukowych mierzone w badaniu PISA 2015

Wyjaśnianie zjawisk przyrodniczych w sposób naukowy
Podając, rozpoznając lub oceniając wyjaśnienia różnorodnych zjawisk z zakresu przyrody i techniki, uczeń potrafi: <ul style="list-style-type: none">▪ przywołać z pamięci i zastosować odpowiednią wiedzę naukową;▪ wskazać, wykorzystać lub stworzyć model lub inną reprezentację wyjaśniającą dane zjawisko;▪ formułować i uzasadniać odpowiednie przypuszczenia;▪ stawiać hipotezy;▪ objaśniać potencjalne następstwa wiedzy naukowej dla społeczeństwa.
Planowanie i ocena poprawności procedur badawczych
Opisując i oceniając badania naukowe i proponując sposoby odpowiedzi na pytania badawcze, uczeń potrafi: <ul style="list-style-type: none">▪ wskazać problem podejmowany w określonym badaniu naukowym;▪ odróżnić pytania, na które można odpowiedzieć w sposób naukowy, od tych, które nie mają takiego charakteru;▪ podać sposób naukowego poszukiwania odpowiedzi na określone pytanie badawcze;▪ ocenić różne sposoby naukowego poszukiwania odpowiedzi na określone pytanie badawcze;▪ 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: <ul style="list-style-type: none">▪ przetworzyć dane naukowe podane w jednej formie w inną formę;▪ analizować i interpretować dane i wyciągać odpowiednie wnioski;▪ wyodrębnić założenia, wskazać dowody i określić wnioskowanie w tekstach dotyczących nauki;▪ odróżnić argumenty bazujące na dowodach i teoriach naukowych od tych opartych na innych podstawach;▪ ocenić wiarygodność naukową tekstów z różnych źródeł (gazet, czasopism, Internetu itp.).

4.1.5. Postawy

Postawy i przekonania uczniów odgrywają niezwykle istotną rolę w ich zainteresowaniu naukami przyrodniczymi, a tym samym wpływają także na ich motywację do rozwiązywania zadań. Postawy te badano za pomocą odrębnego kwestionariusza. Rozróżniano postawę wobec nauki (*attitude towards science*) oraz postawę naukową (*scientific attitude*). Pierwsza obejmuje m.in. zainteresowanie ucznia zagadnieniami nauki i techniki, jego zaangażowanie w aktywności związane z nauką (np. obserwacje

i doświadczenia). Druga natomiast określa, w jak dużym stopniu jego światopogląd bazuje na poglądach naukowych.

W badaniu PISA 2015 skoncentrowano się na trzech obszarach:

- zainteresowaniu zagadnieniami nauki i techniki,
- świadomości środowiskowej (postawom wobec przyrody i środowiska),
- akceptacji dla naukowego sposobu poznawania świata.

4.1.6. Wymagania poznawcze

Nowością w badaniu 2015 jest określenie poziomu wymagań poznawczych (*cognitive demand*), niezbędnego dla rozwiązania poszczególnych zadań. Nie należy mylić tych wymagań z trudnością zadania. Empiryczna trudność zadania jest oszacowana na podstawie jego rozwiązywalności w badaniu. Natomiast poziom wymagań poznawczych zależy od tego, jak złożone rozumowanie trzeba przeprowadzić, by rozwiązać zadanie. Zadanie może być trudne, ponieważ wymaga odwołania się do mało znanych treści naukowych, ale może być mało wymagające pod względem rozumowania. Trudność zadania nie musi jednak polegać na znajomości treści. Zadanie może stawiać przed uczniem wysokie wymagania poznawcze, polegające na analizie, ocenie i połączeniu wielu różnorodnych elementów, które same w sobie są każdemu uczniowi znane.

Przy definiowaniu poziomów wymagań poznawczych w badaniu PISA 2015 przyjęto zmodyfikowaną kategoryzację poziomów głębokości wiedzy (*Depth of Knowledge*) autorstwa Normana L. Webba. Wyróżniono trzy poziomy:

- **niski** – rozumowanie jednostopniowe, np. przywołanie z pamięci określonych faktów, twierdzeń lub pojęć, szukanie prostej informacji w tekście, tabeli lub przedstawieniu graficznym;
- **średni** – wykorzystanie wiedzy konceptualnej do opisanie lub wyjaśnienia danego zjawiska, wybór określonej procedury obejmującej co najmniej dwa kroki, porządkowanie i prezentacja danych, interpretacja i wykorzystanie prostych zbiorów danych, schematów i wykresów;
- **wysoki** – analiza złożonych informacji lub danych, synteza i ocena dowodów naukowych, uzasadnianie, wnioskowanie na podstawie różnych źródeł informacji, planowanie działań w celu rozwiązania określonego problemu badawczego.

4.2. Pomiar kompetencji w naukach przyrodniczych

4.2.1. Charakterystyka zadań wykorzystanych w badaniu

W badaniu użyto 184 zadania; 85 z nich, tzw. zadania kotwiczące, było wykorzystywanych w poprzednich edycjach PISA od 2006 r. Nie były one publikowane i pozostały tajne, a tym samym pozwoliły na porównanie wyników między latami. W badaniu 2015 zostały one zaadaptowane do wersji komputerowej. Pozostałe zadania (99) były nowe. Należy podkreślić, że wiele zadań było interaktywnych, a zatem opracowanych specjalnie pod kątem wykorzystania komputera w badaniu. Nie mają one swoich papierowych odpowiedników. Z wszystkich 184 zadań nieco ponad połowa (98) wymagała wiedzy o treściach nauki, 60 odwoływało się do znajomości procedur badawczych, a 26 sprawdzało opanowanie rozumowania naukowego. Pod względem poszczególnych umiejętności naukowych 89 zadań sprawdzało przede wszystkim umiejętność wyjaśniania zjawisk przyrodniczych w sposób naukowy, 39 zadań dotyczyło planowania i oceny poprawności procedur badawczych, a 56 zadań wymagało wykazania się umiejętnością interpretacji danych i dowodów naukowych. 56 zadań stawiało niskie wymagania poznawcze, 15 wymagało wysokiego poziomu umiejętności rozumowania, natomiast większość (113 zadań) zostało zaklasyfikowanych do średniej poziomu wymagań.

Wykorzystano trzy sposoby konstruowania zadań.

- **Proste zadania wielokrotnego wyboru** to takie zadania, w których należało wybrać tylko jedną spośród wielu odpowiedzi albo wskazać określony element w tekście lub na obrazku. W badaniu wykorzystano 70 takich zadań.
- **Złożone zadania wielokrotnego wyboru** wymagają serii decyzji, np. określenia poprawności każdego stwierdzenia („Tak/Nie”), wyboru kilku poprawnych odpowiedzi z listy, uzupełnienia zdania przez wybór odpowiedzi z rozwijanego menu, przeciągnięcia myszką odpowiednich elementów w wybrane pola albo ustawieniu parametrów w symulacji doświadczenia. W badaniu wykorzystano 50 takich zadań.
- **Zadania otwarte** wymagały sformułowania własnej odpowiedzi, np. napisania odpowiedniego tekstu albo wykonania rysunku. Tekst mógł być krótki, jednozdaniowy, albo dłuższy, kilkuzdaniowy. Uczniowie wpisywali tekst w odpowiednie okno edytora. W badaniu wykorzystano 64 takie zadania.

4.2.2. Określenie poziomów umiejętności uczniów i trudności zadań

Sposób opracowania (skalowania) wyników przyjęty w badaniu PISA pozwala przedstawić na jednej skali poziom trudności zadania oraz wynik ucznia. Oznacza to, że jeśli np. trudność zadania wyniosła 500 pkt., to uczeń, który uzyskał podobny wynik, rozwiązuje to zadanie z prawdopodobieństwem 50%. Ponieważ każde zadanie zostało przeanalizowane pod względem wymagań poznawczych, można na tej podstawie wyróżnić i scharakteryzować poziomy umiejętności badanych piętnastolatków.

4.3. Wyniki uczniów oraz ich zmiany w latach 2006-2015

Pomiar kompetencji przyrodniczych był główną domeną badania w 2006 r., dlatego też wyniki z 2006 r. przyjęto za punkt odniesienia do wyników badania 2015 r. Średnie wyniki uzyskane przez uczniów krajów biorących udział w badaniu na ogólnej skali umiejętności przedstawiono w tabeli 5. W obu edycjach badania brała udział inna liczba krajów. W badaniu 2006 uczestniczyło 57 krajów, a średni wynik dla krajów członkowskich OECD był na poziomie 498 punktów, natomiast w roku 2015 do badania przystąpiły 72 kraje lub regiony, a średni wynik krajów OECD wyniósł 493 punkty. Należy jednak zwrócić uwagę, że pomiar umiejętności rozumowania w naukach przyrodniczych w 2006 roku był przeprowadzany za pomocą testu papierowego, a pomiar w 2015 roku był przeprowadzany na komputerze. Zmiana medium pomiaru mogła wpłynąć na obniżenie wyników.

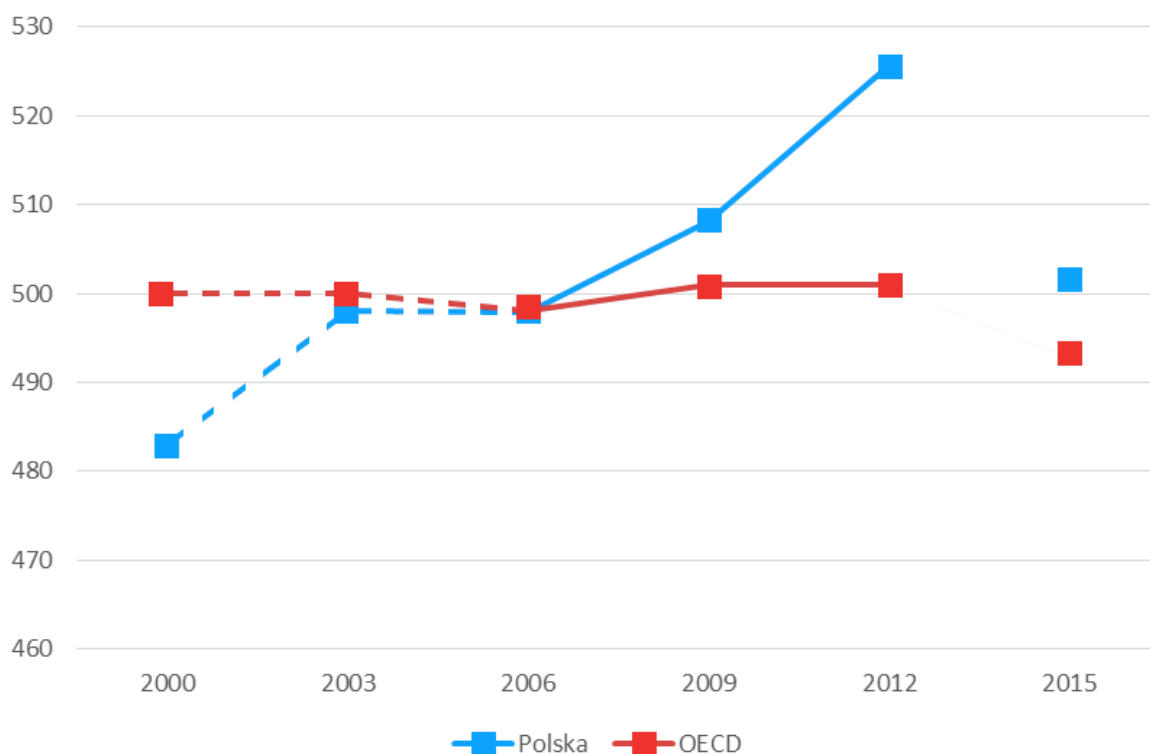
W czołówce krajów osiągających najwyższe wyniki w badaniu kompetencji przyrodniczych znajdują się kraje Dalekiego Wschodu, a także Finlandia, Estonia i Kanada. Pierwsze miejsce zajmuje Singapur (556 pkt), który w 2006 r. nie brał udziału w badaniu PISA. Uczniowie z tego kraju uzyskali średni wynik o 63 pkt wyższy niż średni wynik uczniów krajów OECD. Z krajów europejskich przoduje Estonia z wynikiem 534 punktów, a w czołówce znajduje się też Finlandia (531 pkt), która miała najwyższy wynik w 2006 r. Najniższe wyniki osiągnęli uczniowie z Dominikany (332 pkt), Algierii (376 pkt) i Kosowa (378 pkt). Polscy uczniowie uzyskali wynik 501 pkt i jest on zbliżony do wyników piętnastolatków z Irlandii, Belgii, Danii, Portugalii, Norwegii, USA, Austrii i Szwecji – różnice między Polską a tymi krajami były nieistotne statystycznie. W badaniu 2006 wyniki polskich uczniów były na poziomie średniej OECD, natomiast w 2015 roku nasi uczniowie znajdują się wśród krajów, których wynik jest istotnie wyższy od średniej OECD. W porównaniu z badaniami z lat 2009 i 2012 zarówno wynik polskich uczniów, jak i średni wynik dla krajów OECD jest niższy, co może wynikać ze zmiany medium pomiaru – z papierowego na komputerowy (Wykres 4.1., Tabela 4.7.).

Tabela 4.7. Średnie wyniki uczniów z pomiaru rozumowania w naukach przyrodniczych w badaniach z lat 2006 i 2015. Białe tło oznacza kraje, których wynik nie był istotnie różny od przeciętnego w krajach OECD, jasnym kolorem wyróżniono kraje o wynikach wyższych, a ciemnym – niższych od przeciętnego. Pogrubioną czcionką wyróżniono kraje europejskie. Dane dla Chin pochodzą tylko z czterech prowincji. Regiony Chin, które niezależnie wzięły udział w badaniu, oznaczono kursywą.

PISA 2006		PISA 2015	
Kraj lub region	Średni wynik	Kraj lub region	Średni wynik
Finlandia	563	Singapur	556
<i>Hongkong (Chiny)</i>	542	Japonia	538
Kanada	534	Estonia	534
Tajwan	532	Tajwan	532
Estonia	531	Finlandia	531
Japonia	531	<i>Makao (Chiny)</i>	529
Nowa Zelandia	530	Kanada	528
Australia	527	Wietnam	525
Holandia	525	<i>Hongkong (Chiny)</i>	523
Liechtenstein	522	Chiny B-S-J-G	518
Korea	522	Korea	516
Słowenia	519	Nowa Zelandia	513
Niemcy	516	Słowenia	513
Wielka Brytania	515	Australia	510
Czechy	513	Wielka Brytania	509
Szwajcaria	512	Niemcy	509
<i>Makao (Chiny)</i>	511	Holandia	509
Austria	511	Szwajcaria	506
Belgia	510	Irlandia	503
Irlandia	508	Belgia	502
Węgry	504	Dania	502
Szwecja	503	Polska	501
Polska	498	Portugalia	501
Dania	496	Norwegia	498
Francja	495	Stany Zjednoczone	496
Chorwacja	493	Austria	495
Islandia	491	Francja	495
Łotwa	490	Szwecja	493
Stany Zjednoczone	489	Czechy	493
Słowacja	488	Hiszpania	493
Hiszpania	488	Łotwa	490
Litwa	488	Rosja	487
Norwegia	487	Luksemburg	483
Luksemburg	486	Włochy	481
Rosja	479	Węgry	477
Włochy	475	Litwa	475
Portugalia	474	Chorwacja	475
Grecja	473	Argentyna	475
Izrael	454	Islandia	473
Chile	438	Izrael	467
Serbia	436	Malta	465
Bułgaria	434	Słowacja	461
Urugwaj	428	Grecja	455
Turcja	424	Chile	447
Jordania	422	Bułgaria	446
Tajlandia	421	Emiraty Arabskie	437
Rumunia	418	Urugwaj	435
Czarnogóra	412	Rumunia	435
Meksyk	410	Cypr	433
Indonezja	393	Moldawia	428
Argentyna	391	Albania	427
Brazylia	390	Turcja	425
Kolumbia	388	Trinidad i Tobago	425
Tunezja	386	Tajlandia	421

Azerbejdżan	382	Costa Rica	420
Katar	349	Katar	418
Kirgistan	322	Kolumbia	416
		Meksyk	416
		Czarnogóra	411
		Gruzja	411
		Jordania	409
		Indonezja	403
		Brazylia	401
		Peru	397
		Liban	386
		Tunezja	386
		Macedonia	384
		Kosowo	378
		Algieria	376
		Dominikana	332

Wykres 4.1. Zmiany wyników pomiaru rozumowania w naukach przyrodniczych uczniów w Polsce i średnio w krajach OECD w latach 2000-2015. Ponieważ nauki przyrodnicze były główną dziedziną pomiaru w 2006 r. i dopiero wtedy dokonano pełnego skalowania, wcześniejsze wyniki nie są w pełni porównywalne i dlatego zaznaczono je linią przerywaną. Z uwagi na zmianę medium pomiaru między badaniami 2012 i 2015, ich wyniki nie są też w pełni porównywalne, co zaznaczono brakiem linii łączącej.



Zestawienie wyników badania z 2015 r. z tymi z lat 2006 r. i 2012 r. dla poszczególnych krajów przedstawiono w Tabeli 3.8. Można zauważyć, że w pierwszym badaniu komputerowym, jakim było badanie 2015, średni wynik dla krajów OECD był niższy niż w badaniach przeprowadzonych z wykorzystaniem zadań drukowanych na papierze. Jednak tylko różnica między latami 2012 i 2015 była istotna statystycznie. Zaledwie 6 krajów odnotowało statystycznie istotny wzrost wyników między latami 2006 a 2012, natomiast aż 14 krajów zanotowało statystycznie istotny spadek. W tej ostatniej

grupie znalazły się również kraje-liderzy badania z 2006 r.: Finlandia, Hongkong, Nowa Zelandia i Australia. Wyniki pozostałych krajów, w tym większości krajów europejskich nie różnią się jednak istotnie między badaniami 2006 i 2015 (Estonia: 3 pkt, Słowenia: -6 pkt, Polska: 4 pkt, Irlandia: -6 pkt). Od 2006 roku do 2012 wyniki polskich uczniów z rozumowania w naukach przyrodniczych wzrastały, natomiast w roku 2015, przy zmianie medium pomiaru, nastąpił znaczny spadek wyniku.

Tabela 4.8. Różnice wyników uczniów z pomiaru rozumowania w naukach przyrodniczych między badaniem 2015 a 2006 i 2012. Kraje uporządkowano według malejącej różnicy między badaniami 2006 i 2015, w których kompetencje w naukach przyrodniczych były główną domeną pomiaru. Czerwoną czcionką zaznaczono różnice istotne statystycznie. Pogubioną czcionką wyróżniono kraje europejskie, a kursywą – regiony Chin. Gwiazdką oznaczono brak danych.

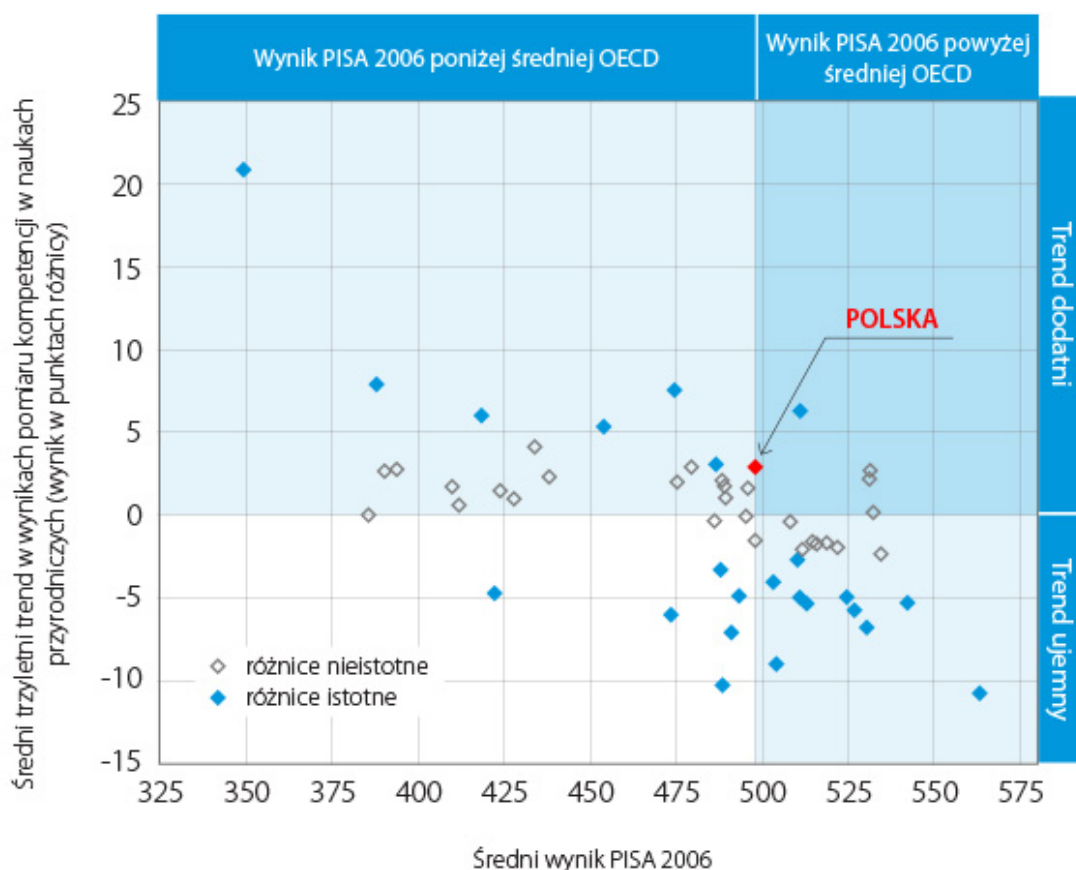
Kraj lub region	PISA 2015 a 2006		PISA 2015 a 2012	
	Wynik 2006	Zmiana od 2006	Wynik 2012	Zmiana od 2012
OECD	498	-5	501	-8
Katar	349	68	384	34
Argentyna	391	41	406	27
Kolumbia	388	28	399	17
Portugalia	474	27	489	12
<i>Makao (Chiny)</i>	511	18	521	8
Rumunia	418	16	439	-4
Izrael	454	13	470	-4
Norwegia	487	12	495	4
Bułgaria	434	12	446	-1
Indonezja	393	10	382	21
Brazylia	390	10	405	-1
Chile	438	9	445	2
Japonia	531	7	547	-8
Rosja	479	7	486	0
Urugwaj	428	7	416	20
Dania	496	6	498	3
Meksyk	410	6	415	1
Włochy	475	5	494	-13
Polska	498	4	526	-24
Hiszpania	488	4	496	-4
Estonia	531	3	541	-7
Turcja	424	2	463	-38
Łotwa	490	1	502	-12
Tunezja	386	1	398	-12
Tajwan	532	0	523	9
Francja	495	0	499	-4
Tajlandia	421	0	444	-23
Czarnogóra	412	0	410	1
Luksemburg	486	-4	491	-8
Korea	522	-6	538	-22

Słowenia	519	-6	514	-1
Wielka Brytania	515	-6	514	-5
Szwajcaria	512	-6	515	-10
Irlandia	508	-6	522	-19
Stany Zjednoczone	489	-6	497	-1
Kanada	534	-7	525	2
Niemcy	516	-7	524	-15
Belgia	510	-8	505	-3
Szwecja	503	-10	485	9
Litwa	488	-13	496	-20
Jordania	422	-13	409	-1
Holandia	525	-16	522	-13
Austria	511	-16	506	-11
Nowa Zelandia	530	-17	516	-2
Australia	527	-17	521	-12
Chorwacja	493	-18	491	-16
Islandia	491	-18	478	-5
<i>Hongkong (Chiny)</i>	542	-19	555	-32
Grecja	473	-19	467	-12
Czechy	513	-20	508	-15
Węgry	504	-24	494	-18
Słowacja	488	-28	471	-10
Finlandia	563	-33	545	-15
Liechtenstein	522	*	525	*
Singapur	*		551	4
Wietnam	*		528	-4

Znaczące różnice można zaobserwować między wynikami badań 2015 i 2012. Dla krajów OECD średni wynik spadł o 8 pkt. Osiem krajów istotnie polepszyło swoje wyniki, ale aż 24 pogorszyły. Liderem spadków była Turcja (-38 pkt), bardzo dużą zmianę odnotowano w Polsce (-24 pkt), w Irlandii (-18 pkt) i Finlandii (-15). Wysoki spadek wyniku pomiaru umiejętności uczniów stwierdzono również w Hongkongu (-32 pkt) i Korei (-22 pkt).

W ocenie systemu edukacji istotne jest nie tylko samo porównanie wyników kolejnych badań, ale także oszacowanie długoterminowych tendencji zmian. Umiejętności polskich 15-latków w naukach przyrodniczych, mimo znacznego obniżenia w latach 2012-2015, na przestrzeni pomiaru 2006-2015 wykazują wyraźną dodatnią tendencję wzrostową (Wykres 4.2). Średni trzyletni trend w wynikach pomiaru kompetencji przyrodniczych wynosi 2,9.

Wykres 4.2. Zależność między średnim trzyletnim trendem w wynikach pomiaru kompetencji w naukach przyrodniczych a średnim wynikiem uczniów w badaniu 2006.



4.3.1. Poziomy umiejętności w pomiarze kompetencji uczniów w naukach przyrodniczych

Średni wynik uczniów w PISA służy porównaniom międzynarodowym lub między cyklami badania. W tym celu przydatny jest rozkład wyników, czyli informacja o odsetku uczniów na poszczególnych poziomach umiejętności (poziomy zdefiniowano w Tabeli 4.9.). Ważne są zwłaszcza zmiany na poziomie 1 i poniżej tego poziomu oraz na poziomie 5 i powyżej. Pierwsza grupa badanych to uczniowie potencjalnego ryzyka braku umiejętności funkcjonowania w nowoczesnym społeczeństwie, w którym zagadnienia nauki i techniki odgrywają coraz większą rolę. Natomiast uczniowie na poziomach 5 i 6 to potencjalna elita intelektualna, warunkująca postęp naukowo-techniczny. Należy jednak pamiętać, że definicje poziomów umiejętności w badaniach 2006-2012 oraz 2015 nie do końca się pokrywają, ponieważ ostatnie badanie wykorzystywało, oprócz zadań kotwiczących, nową pulę zadań.

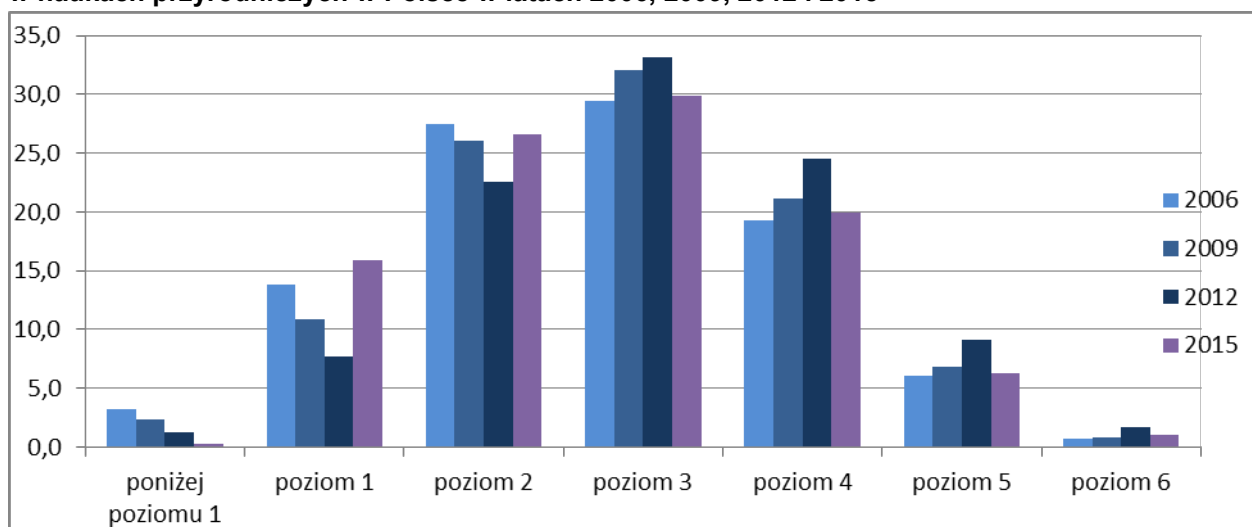
Tabela 4.9. Opis poziomów umiejętności na skali kompetencji w naukach przyrodniczych

Poziom i dolna granica przedziału	Charakterystyka umiejętności uczniów na podstawie rozwiązanych zadań
6 708 pkt.	Uczniowie wykorzystują rozległą wiedzę o treściach naukowych, procedurach badawczych i wiedzę o rozumowaniu naukowym w celu stawiania hipotez wyjaśniających nieznanne im wcześniej zjawiska, zdarzenia i procesy lub formułowania prognoz. Interpretując dane i dowody naukowe, potrafią odróżnić informacje istotne od nieistotnych oraz odwołują się do wiedzy nieobjętej szkolnym programem nauczania. Są w stanie stwierdzić, które argumenty odwołują się do teorii i faktów naukowych, a które bazują na innych rozważaniach. Analizują alternatywne

	plany badawcze złożonych eksperymentów, obserwacji terenowych i symulacji, potrafią wybrać najlepszy i uzasadnić swój wybór.
5 633 pkt.	Uczniowie potrafią wykorzystywać abstrakcyjne pojęcia lub idee naukowe, aby wyjaśnić nieznane sobie wcześniej zjawiska, zdarzenia lub procesy z wieloma zależnościami przyczynowo-skutkowymi. Wykorzystują bardziej wyrafinowaną wiedzę poznawczą, aby porównać alternatywne schematy eksperymentów i uzasadnić ich wybór. Posługują się wiedzą teoretyczną w celu interpretacji informacji oraz przewidzenia wyniku. Potrafią ocenić różne sposoby naukowego podejścia do tego samego problemu badawczego oraz określić ograniczenia związane z interpretacją danych naukowych, w tym źródła i skutki niepewności.
4 559 pkt.	Uczniowie wykorzystują bardziej złożone lub abstrakcyjne treści naukowe (podane w zadaniu lub przywołane z pamięci) w celu wyjaśniania bardziej złożonych lub mniej znanych zdarzeń i procesów. Potrafią przeprowadzić doświadczenie z dwoma lub więcej niezależnymi zmiennymi, ale w ograniczonym zakresie. Są w stanie uzasadnić plan eksperymentu, odwołując się do elementów wiedzy o procedurach i rozumowaniu naukowym. Interpretują dane pozyskane z umiarkowanie złożonego zbioru danych albo dotyczące mniej znanego kontekstu, wyciągają odpowiednie, ogólne wnioski, przedstawiają uzasadnienie swojego wyboru.
3 484 pkt.	Uczniowie wskazują lub tworzą wyjaśnienia znanych zjawisk, odwołując się do umiarkowanie złożonej wiedzy faktograficznej. W sytuacjach mniej znanych lub bardziej złożonych potrafią podać wyjaśnienie, jeśli uzyskają odpowiednią podpowiedź lub pomoc. Wykorzystując elementy wiedzy proceduralnej i poznawczej, potrafią przeprowadzić prosty eksperyment. Rozróżniają zagadnienia naukowe od nienaukowych oraz wskazują dane wspierające określoną tezę naukową.
2 410 pkt.	Odwołując się do podstawowej, codziennej wiedzy faktograficznej i z zakresu procedur badawczych, uczniowie wskazują odpowiednie wyjaśnienie naukowe, interpretują dane oraz określają pytanie badawcze w prostym doświadczeniu. Wykorzystując podstawową, codzienną wiedzę naukową, potrafią wskazać poprawny wniosek wynikający z prostego zbioru danych. Wykazują się podstawową wiedzą poznawczą, wskazując pytania, na które można odpowiedzieć w sposób naukowy.
1a 335 pkt.	Uczniowie potrafią wykorzystać podstawowe wiadomości z codziennego życia oraz znajomość procedur badawczych do rozpoznania wyjaśnienia prostego zjawiska. Z odpowiednią pomocą podejmują problemy badawcze z nie więcej niż dwoma zmiennymi. Potrafią zidentyfikować proste zależności przyczynowo-skutkowe i korelacje oraz interpretować graficzne przedstawienia danych, które nie stawiają wysokich wymagań poznawczych. Potrafią dobrać najlepsze wyjaśnienie naukowe dla określonych danych w znanym już kontekście osobistym, lokalnym lub globalnym.
1b 261 pkt.	Uczniowie dysponują podstawową, potoczną wiedzą i wykorzystują ją jedynie do rozpoznania znanych sobie lub prostych zjawisk przyrodniczych. Potrafią dostrzec proste wzorce w danych, znają podstawowe terminy naukowe, przeprowadzają procedury naukowe na podstawie jednoznacznych poleceń.

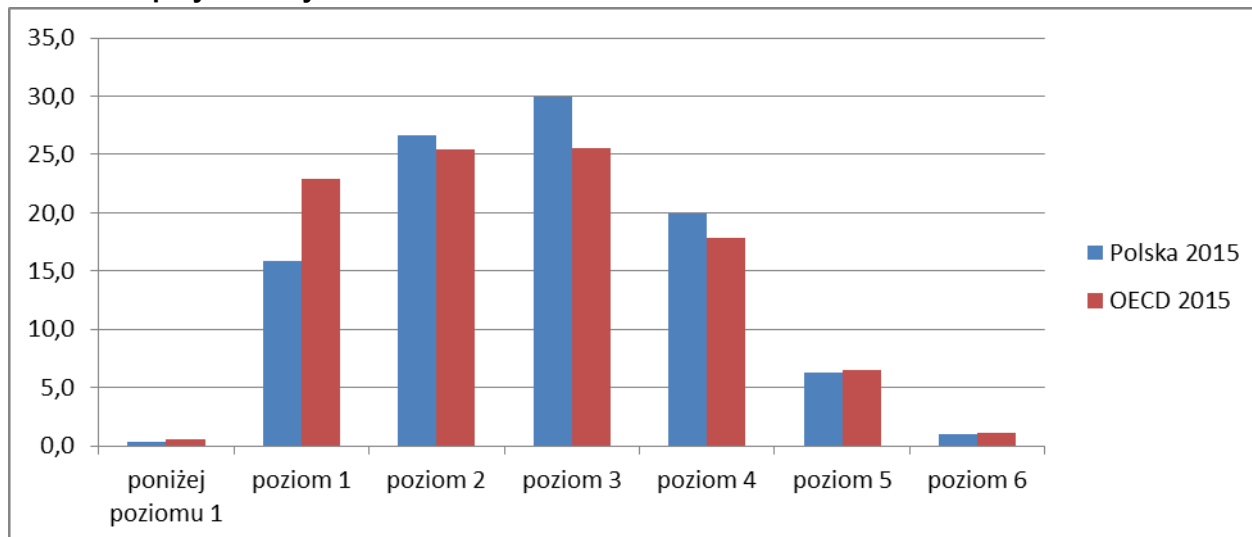
Źródło: Raport międzynarodowy: PISA 2015 Results Excellence and equity in education, Volume I, OECD 2016

Wykres 4.3. Odsetki uczniów na poszczególnych poziomach umiejętności rozumowania w naukach przyrodniczych w Polsce w latach 2006, 2009, 2012 i 2015



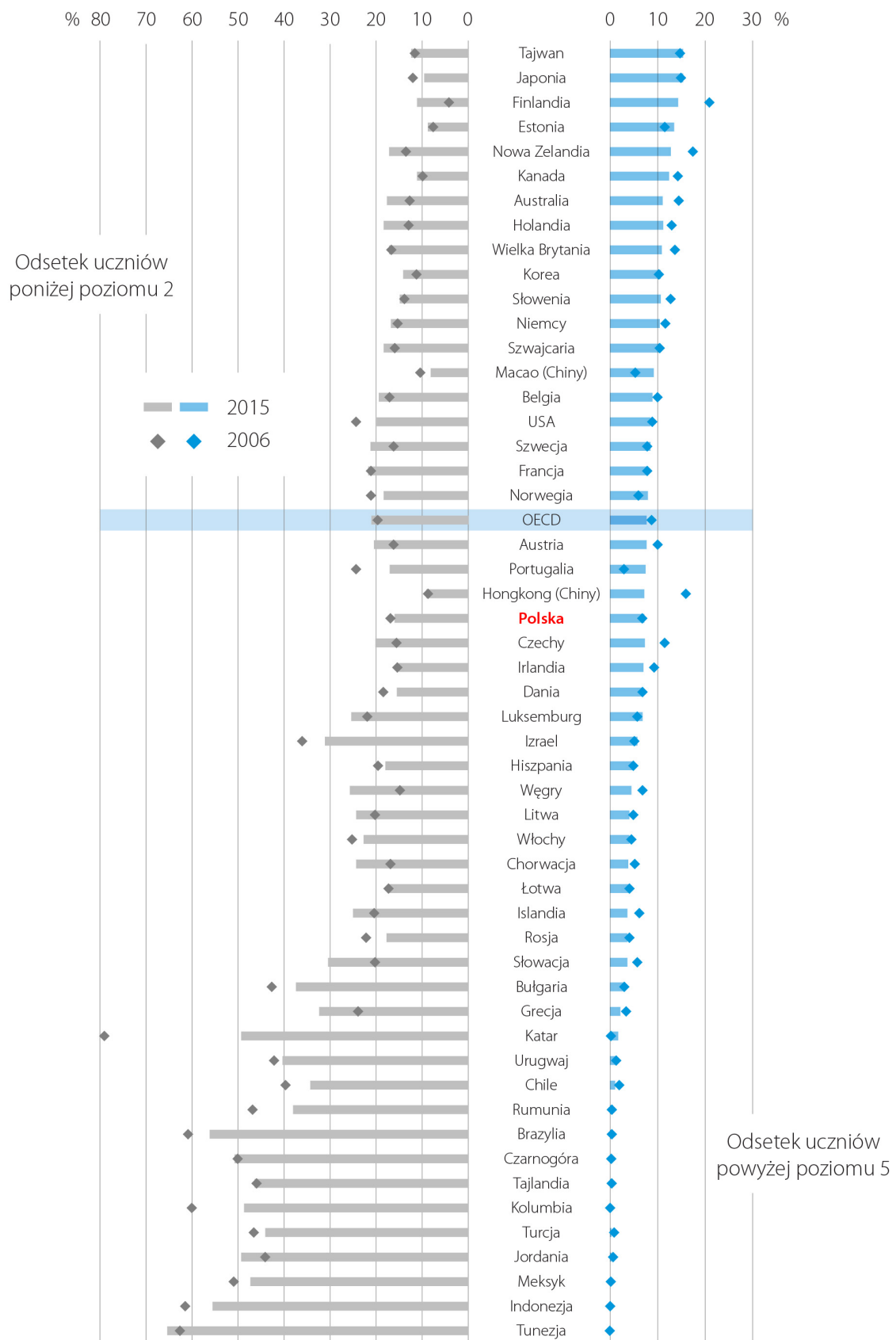
W 2015 roku odsetek uczniów w Polsce na poziomie 1 i 2 wzrósł, a na poziomach od 3-6 spadł w odniesieniu do 2012 roku (Wykres 4.3.). Ogółem, rozkład wyników z 2015 r. jest zbliżony do wyników z 2006 r. Porównując wyniki dla Polski i OECD (Wykres 4.4.), można odnotować, że w Polsce jest wyższy odsetek uczniów na poziomach 2-4, a niższy na pozostałych. Choć zatem cieszy niższy odsetek uczniów najsłabszych, to martwi niski odsetek uczniów na najwyższych poziomach umiejętności.

Wykres 4.4. Odsetki uczniów na poszczególnych poziomach umiejętności rozumowania w naukach przyrodniczych w Polsce i OECD w 2015 r.



Należy jednak zauważyć, że Polska należała do nielicznych krajów o niskich różnicach w odsetku uczniów na najniższych i najwyższych poziomach umiejętności między badaniami 2006 i 2015 (Wykres 4.5.).

Wykres 4.5. Odsetki uczniów na poziomach 1 i poniżej oraz 5 i powyżej w latach 2006 i 2015 w pomiarze umiejętności rozumowania w naukach przyrodniczych PISA.

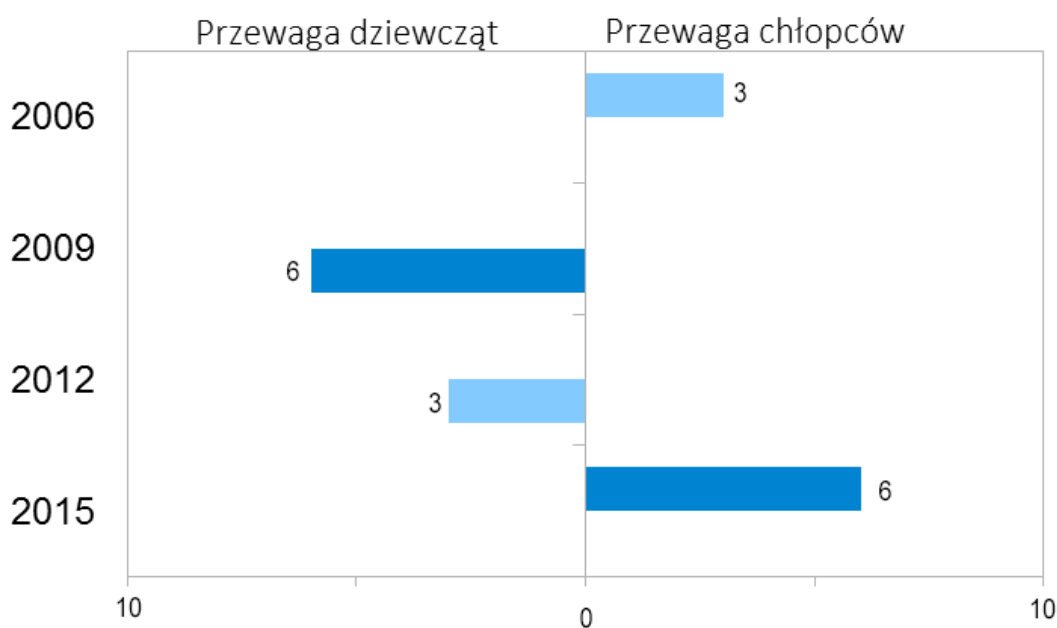


4.4. Wyniki chłopców i dziewcząt

W 2015 r. średni wynik chłopców w krajach OECD wyniósł 495 punktów i był o 4 pkt. wyższy od średniego wyniku dziewcząt. W Polsce średni wynik dziewcząt wyniósł 498 pkt., podczas gdy chłopcy osiągnęli 504 pkt. Zarówno w wynikach OECD, jak i Polski, różnica między chłopcami a dziewczętami była istotna statystycznie. Odsetek chłopców i dziewcząt na poziomie poniżej 2 był podobny i wynosił około 16%; był przy tym niższy od średniej OECD (około 22%). Odsetek chłopców na poziomie powyżej 4 był natomiast istotnie wyższy niż odsetek dziewcząt na tym poziomie, zarówno w Polsce, jak i średnio w krajach OECD.

Na Wykresie 4.6. przedstawiono różnicę między wynikami chłopców i dziewcząt w kolejnych badaniach PISA. W 2006 r. wyższy wynik osiągnęli chłopcy, ale różnica była nieistotna statystycznie. W 2009 r. przewaga dziewcząt była wyraźna (i istotna statystycznie), ale spadła w badaniu 2012 r., a różnica między płciami nie była istotna. W 2015 r., w badaniu umiejętności uczniów na komputerze, różnica stała się ponownie istotna, tym razem na korzyść chłopców.

Wykres 4.6. Różnice w wynikach dziewcząt i chłopców w pomiarze rozumowania w naukach przyrodniczych w latach 2006-2015. Statystycznie istotne różnice zaznaczone są ciemniejszą barwą.



4.5. Wyniki uczniów w podziale na rodzaj umiejętności i rozwiązywalność poszczególnych zadań

Jak opisano powyżej, kompetencje w zakresie nauk przyrodniczych obejmują trzy umiejętności: (1) wyjaśnianie zjawisk w sposób naukowy, (2) planowanie i ocena poprawności procedur badawczych, (3) interpretacja danych i dowodów naukowych. W kształtowaniu tych umiejętności niezbędne są trzy kategorie wiedzy: o treściach nauki, o procedurach badawczych (o metodyce badawczej) i o podstawach poznania naukowego. Dla wyjaśniania zjawisk przyrodniczych w sposób naukowy niezwykle istotna jest wiedza o treściach naukowych, przy planowaniu i ocenie procedur badawczych ważna jest znajomość tych procedur, natomiast wiedza o poznaniu naukowym jest niezbędna w interpretacji danych i dowodów naukowych. Wszystkie zadania zostały podzielone na trzy grupy

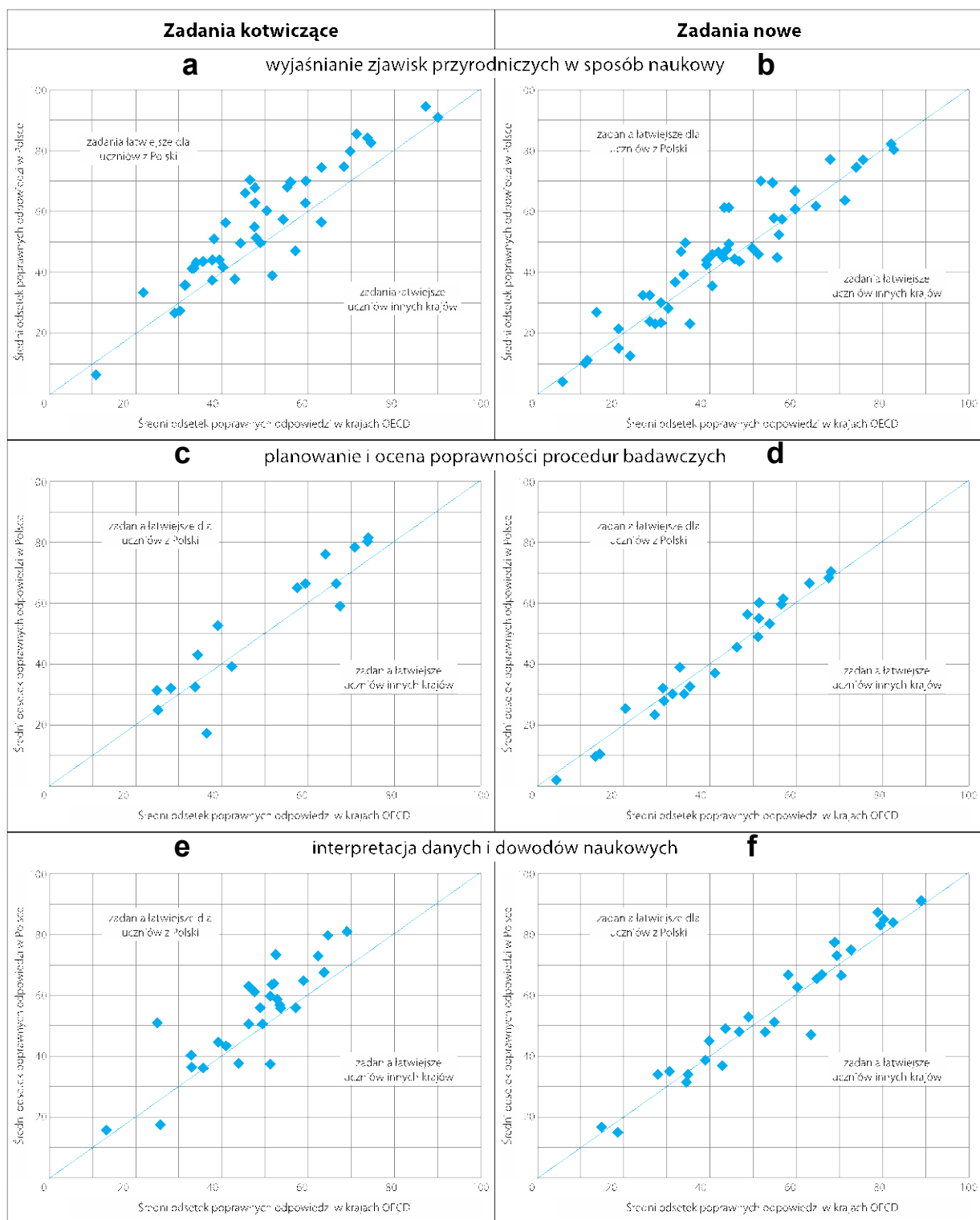
umiejętności i dla każdej z tych umiejętności opracowano oddzielną skalę. Warto zauważyć, że wyniki polskich piętnastolatków są podobne na wszystkich trzech skalach umiejętności, co wskazuje na zrównoważone kształtowanie tych umiejętności w szkole (Tabela 4.10). Również na trzech podskalach opisujących treści nauki w 2015 roku uczniowie uzyskali podobne wyniki. W badaniu w 2006 r. natomiast umiejętności odwołujące się do wiedzy faktograficznej przeważały nad umiejętnościami wymagającymi rozumowania naukowego.

Tabela 4.10. Wyniki uczniów z Polski oraz z OECD na skalach umiejętności przyrodniczych oraz pod względem grup treści naukowych w latach 2006 i 2015.

	OECD		Polska	
	2006	2015	2006	2015
Umiejętności				
Wyjaśnianie zjawisk przyrodniczych w sposób naukowy	500	493	506	501
Planowanie i ocena poprawność i procedur badawczych	499	493	483	502
Interpretacja danych i dowodów naukowych	499	493	494	501
Wiadomości – treści nauki				
Świat fizyczny		493	497	503
Świat organizmów żywych		492	501	501
Ziemia i kosmos		494	509	501

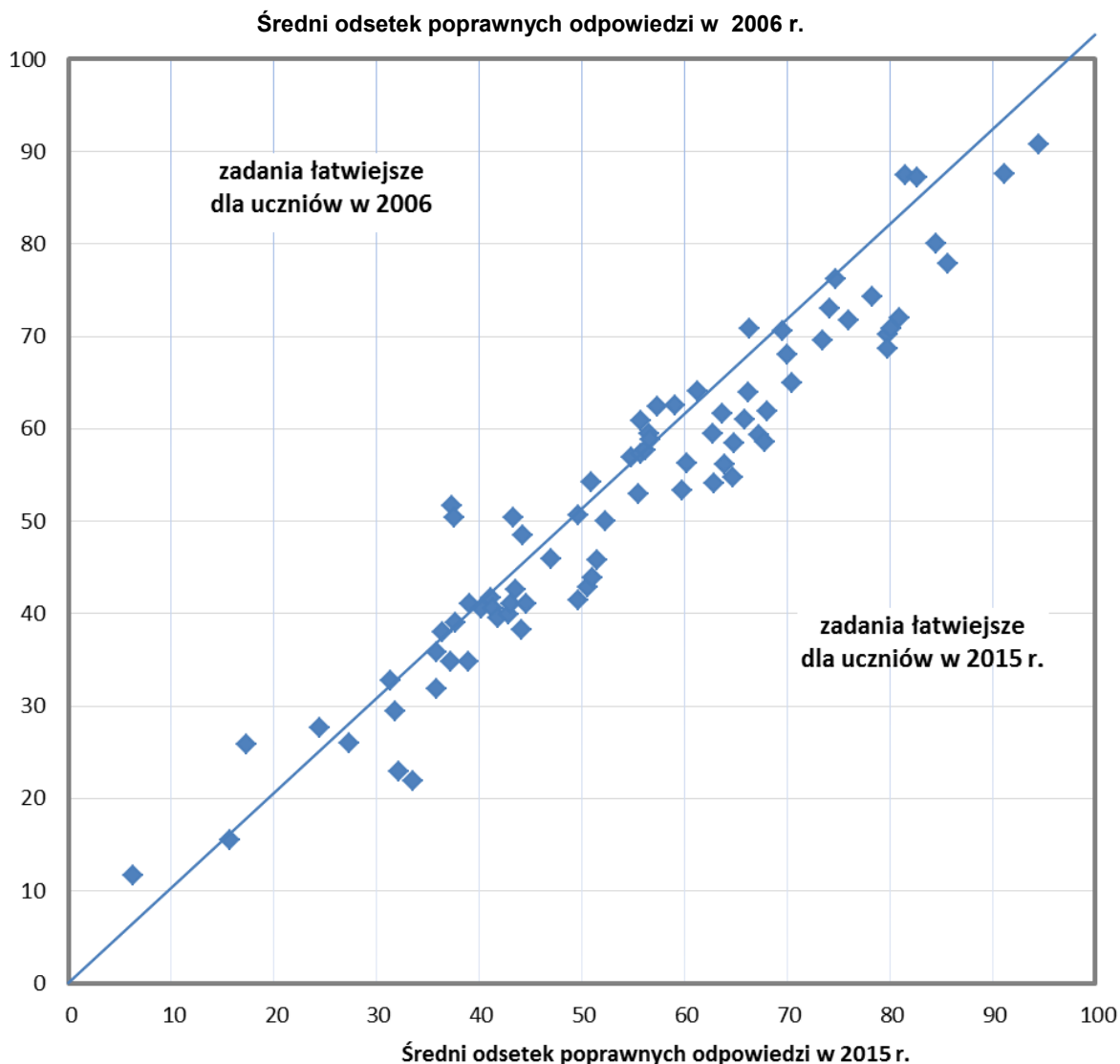
Wyniki polskich uczniów w rozwiązywaniu poszczególnych zadań porównano z wynikami, jakie osiągnęli uczniowie z krajów OECD (Wykres 4.7). Polscy uczniowie rozwiązywali większość zadań lepiej od rówieśników z OECD, a różnica nie miała zwykle związku z trudnością zadania ani z kluczową umiejętnością, jaką sprawdzało dane zadanie. Jedynie w wypadku nowych zadań mierzących umiejętności z zakresu planowania i oceny poprawności procedur badawczych (Wykres 4.7 d) zaobserwowano, że różnica między wynikami OECD a Polski zwiększa się wraz z trudnością zadania (korelacja wyniosła 0,42 i była istotna statystycznie). Znaczy to, że nasi uczniowie mieli większy kłopot z rozwiązywaniem zadań przygotowanych na komputer o wysokiej trudności w porównaniu z ich rówieśnikami z OECD.

Wykres 4.7. Porównanie rozwiązywalności zadań kotwiczących oraz nowych w Polsce oraz w krajach OECD.



W badaniu PISA 2015 wykorzystano 79 zadań z 2006 roku. W Polsce w 2015 roku w porównaniu z rokiem 2006 procent poprawnych odpowiedzi zwiększył się dla 50 zadań, a zmniejszył dla 29 (Wykres 4.8.).

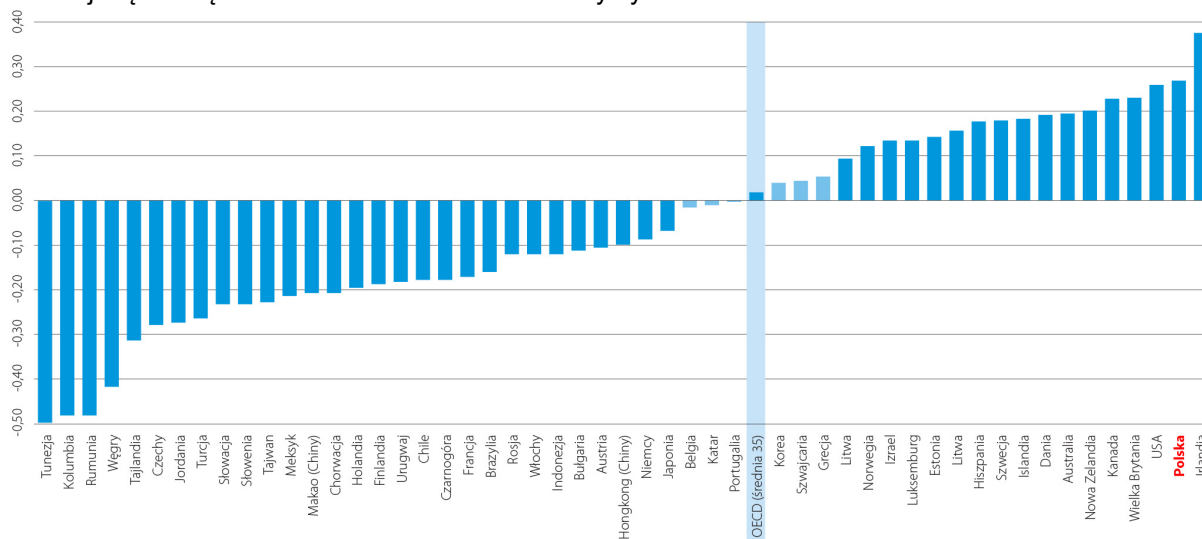
Wykres 4.8. Wyniki polskich uczniów w poszczególnych zadaniach z przedmiotów przyrodniczych w latach 2006 i 2015.



4.6. Postawy uczniów wobec nauk przyrodniczych oraz oczekiwania względem kariery zawodowej

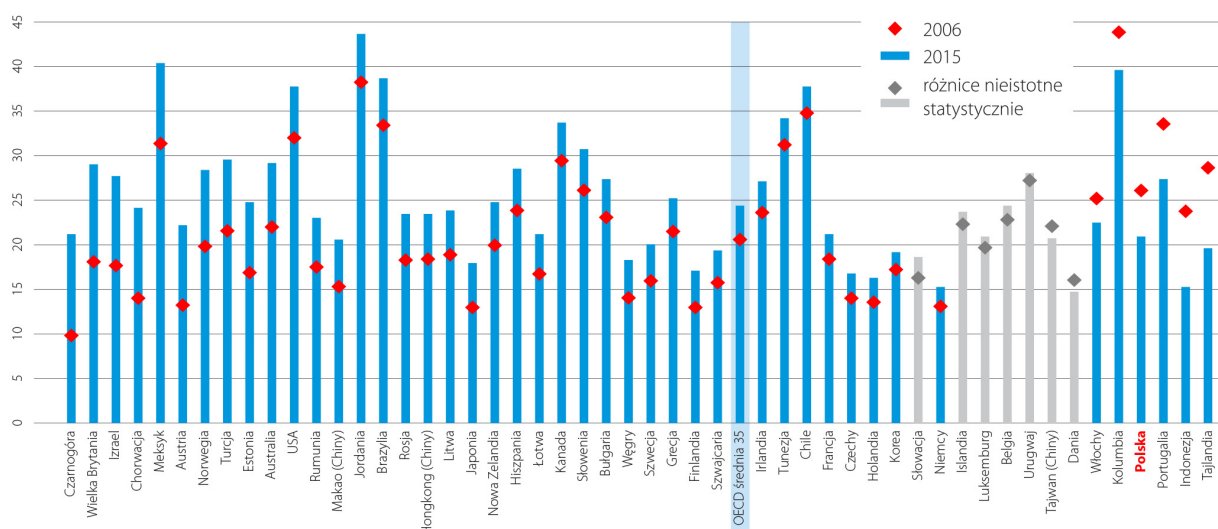
W badaniu 2006 oraz 2015 pytano uczniów o różne aspekty ich postaw wobec nauk przyrodniczych, w tym, czy uczenie się przedmiotów przyrodniczych oraz poznawanie zagadnień naukowych sprawia im przyjemność i daje satysfakcję. Na podstawie tych odpowiedzi opracowano wskaźnik satysfakcji z uczenia się przedmiotów przyrodniczych (*enjoyment of learning science*). Krajami, w których wskaźnik ten wzrósł najbardziej między badaniami 2006 i 2015, były Irlandia i Polska (Wykres 4.9).

Wykres 4.9. Zmiana wskaźnika satysfakcji z uczenia się o przyrodzie między badaniami PISA 2006 i 2015. Kraje uszeregowano od najmniejszej do największej zmiany tego wskaźnika. Ciemniejszą barwą zaznaczono różnice istotne statystycznie



Uczniowie byli też pytani o to, jak sądzą, jaki zawód będą wykonywać w wieku 30 lat. Odsetek uczniów wskazujących na zawody związane z nauką zmniejszył się, w porównaniu z 2006 r. z ok. 26 do 21 procent (Wykres 4.10.). Między 2006 a 2015 nie zmienił się odsetek uczniów wskazujących na zawody inżynierskie i bezpośrednio związane z nauką (te zawody wskazało w 2015 r. ok. 6% uczniów), zwiększył się odsetek wskazujących na zawody medyczne (wzrost z 8 do 12%), zmniejszył się natomiast odsetek uczniów wskazujących na zawody informatyczne (z 6 do 1%) oraz pozostałe zawody związane z nauką (z 6 do 1%).

Wykres 4.10. Odsetek uczniów w badaniach PISA 2006 i 2015, którzy oczekują, że w wieku 30 lat będą pracować w zawodach związanych z naukami przyrodniczymi. Kraje uszeregowano od najwyższej do najniższej wartości różnicy tego odsetka między latami 2015 i 2006. Szarym kolorem oznaczono różnice nieistotne statystycznie.



4.7. Podsumowanie

W badaniu 2015 polscy uczniowie osiągnęli wynik podobny do uzyskanego w 2006 r., kiedy to kompetencje przyrodnicze były także główną dziedziną pomiaru PISA, ale tym razem był to wynik istotnie powyżej średniej dla OECD. Choć jest to dobra wiadomość, to jednak jest to rezultat znacznie poniżej oczekiwań, do jakich skłaniał bardzo wysoki wynik polskich piętnastolatków w badaniu 2012. Rozważając tę rozbieżność, należy pamiętać, że w latach 2006-2012 wykorzystywano papierowe zeszyty z zadaniami, natomiast w 2015 r. badanie zrealizowane było za pomocą komputerów. We wszystkich badaniach w latach 2006-2015 wykorzystano te same zadania kotwiczące, ale w 2015 r. zostały one zaadaptowane do rozwiązywania na komputerze, natomiast nowe zadania nie miały już odpowiednika w formie papierowej. Sugeruje to, że głównym problemem polskich uczniów było właśnie rozwiązywanie zadań z wykorzystaniem komputera. Do takiego wyjaśnienia skłaniają też wyniki badania PISA 2012 dotyczącego rozwiązywania problemów (*Creative Problem Solving*) z wykorzystaniem komputera. Zostało ono przeprowadzone równoległe do badania głównego w 2012 r. W badaniu głównym w zakresie umiejętności rozumowania w naukach przyrodniczych polscy uczniowie uzyskali średni wynik 526 pkt, plasujący ich w czołówce światowej. Natomiast w badaniu rozwiązywania problemów z wykorzystaniem komputera osiągnęli wynik zaledwie 481 pkt (przy średniej OECD 500 pkt), czyli znacznie poniżej oczekiwań.

Niepokój budzi także analiza rozkładu poziomów umiejętności uczniów. W latach 2006-2012 stopniowo malał odsetek uczniów na najniższych poziomach, a rósł na najwyższych. Natomiast w pierwszym badaniu komputerowym rozkład ten jest podobny, jak w 2006 r. W porównaniu z OECD może nas cieszyć niższy odsetek uczniów najsłabszych, jednak niepokoi niewysoki odsetek uczniów na najwyższych poziomach umiejętności, stanowiących przyszłą elitę naukowo-techniczną.

Przejście z badania papierowego na komputerowe ujawniło problem różnic między płciami – w 2015 r. chłopcy byli o 6 pkt istotnie lepsi od dziewcząt, podczas gdy w ostatnim badaniu papierowym (2012 r.) różnica między płciami była nieistotna statystycznie, a średni wynik dziewcząt był o 3 pkt wyższy od wyniku chłopców.

Pozytywnym wynikiem badania 2015 jest to, że Polska jest w grupie krajów o dodatniej średniej długoterminowej tendencji zmian. Pokazuje to stałą poprawę poziomu umiejętności polskich 15-latków. Dalsze badania pokażą, czy jest to tendencja trwała.

Bardzo dobrym rezultatem są także wyniki pomiaru składowych umiejętności w zakresie kompetencji przyrodniczych. W 2006 r. badania pokazały, że polscy uczniowie cechują się ponadprzeciętną wiedzą o treściach naukowych, natomiast słabiej wypadali w zakresie umiejętności wykorzystywania tej wiedzy do wnioskowania, planowania doświadczeń czy w interpretacji faktów naukowych. Natomiast w badaniu 2015 wszystkie komponenty biegłości naukowej były bardziej zrównoważone.

Niepokoi spadek oczekiwań względem kariery zawodowej w nauce – najwyraźniej uczniowie w coraz mniejszym stopniu uznają taki rozwój zawodowy za atrakcyjny. Bardzo dobrą wiadomością płynącą z badań jest znaczący wzrost satysfakcji uczniów z nauki przedmiotów przyrodniczych, co jest niezwykle ważne w świetle coraz częstszej i obecnej nawet w mediach publicznych postawy antynaukowej, nieopartej na faktach, ale bazującej na irracjonalnych uprzedzeniach, np. sprzeciw wobec szczepionek i organizmów genetycznie zmodyfikowanych albo zaprzeczanie, że ocieplanie się klimatu stanowi problem globalny.

ANEKS. OMÓWIENIE PRZYKŁADOWYCH ZADAŃ

Wiązka zadań „Badanie zbocza góry”

Prezentowana wiązka zadań dotyczy problemu konstrukcji układu badawczego oraz interpretacji otrzymanych danych. We wstępie nakreślono plan eksperymentu oraz postawiono problem badawczy. Wiązka ta ma charakter tradycyjny i równie dobrze mogłaby być wykorzystana w papierowej wersji badania.

PISA 2015

Badanie zbocza wzgórza
Wstęp

Przeczytaj wstęp. Następnie kliknij na strzałkę DALEJ.

BADANIE ZBOCZA WZGÓRZA

Grupa uczniów zauważyła znaczną różnicę w stanie roślinności na dwóch zboczach tworzących dolinę: roślinność jest znacznie bardziej zielona i obfita na zboczu A niż na zboczu B. Tę różnicę pokazano na ilustracji po prawej stronie.

Uczniowie badają, dlaczego roślinność na tych dwóch zboczach tak bardzo się od siebie różni. W ramach tego badania uczniowie przez określony czas mierzą trzy czynniki środowiskowe:

- **Promieniowanie słoneczne:** ile światła słonecznego pada na dane miejsce
- **Wilgotność gleby:** jaki jest poziom wilgotności gleby w danym miejscu
- **Opady deszczu:** ile deszczu spada w danym miejscu



Zadanie 1.

Pierwsze zadanie wymaga sformułowania własnej odpowiedzi i sprawdza umiejętność planowania doświadczenia. Wymaga przede wszystkim wiedzy o poznaniu naukowym, a dotyczy bardzo ważnego zagadnienia źródeł niepewności pomiaru w badaniach.

PISA 2015

Badanie zbocza wzgórza
Pytanie 1 / 2

Zapoznaj się z tekstem "Zbieranie danych" po prawej stronie. Wpisz odpowiedź na pytanie.

Dlaczego, badając różnicę w stanie roślinności między dwoma zboczami tworzącymi dolinę, uczniowie umieszczają na każdym zboczach dwie sztuki każdego urządzenia?

BADANIE ZBOCZA WZGÓRZA
Zbieranie danych

Na każdym ze zboczycy uczniowie umieszczają dwie sztuki każdego z trzech urządzeń, jak pokazano poniżej.

- Czujnik promieniowania słonecznego:** mierzy ilość światła słonecznego, w megadžulach na metr kwadratowy (MJ/m²)
- Czujnik wilgotności gleby:** mierzy ilość wody wyrażonej jako procent masy gleby
- Deszczomierz:** mierzy ilość opadów deszczu, w milimetrach (mm)



Aby prawidłowo rozwiązać to zadanie, uczeń powinien rozumieć, w jakim celu w badaniach naukowych dokonuje się co najmniej dwóch niezależnych pomiarów określonej wielkości. Za poprawne odpowiedzi uznawano takie, w których uczniowie wykazali znaczenie takiego podejścia. Poniżej podano przykładowe poprawne odpowiedzi, zachowując oryginalne sformułowania uczniów, opatrzone je także komentarzem. Warto zaznaczyć, że oceniana była jedynie poprawność rozumowania, nie zaś forma czy ewentualne błędy językowe.

- „Żeby można było określić, czy różnica między zboczami jest istotna”. Uczeń rozumie, że pojedynczy pomiar może być – z różnych względów – niewiarygodny.
- „Ponieważ mogą występować różnice na tym samym zboczach”. Uczeń zauważa, że mierzona wielkość fizyczna może cechować się zmiennością.
- „Żeby zwiększyć precyzję pomiaru na każdym ze zboczycy”. Uczeń wskazuje, że dla poprawnego wnioskowania należy oszacować zmienność cechy.
- „Na wypadek awarii jednego z dwóch”. Uczeń dostrzega, że urządzenia pomiarowe mogą być zawodne.
- „Żeby porównać różne ilości słońca na zboczach”. Jak powyżej – uczeń odwołuje się do lokalnej zmienności mierzonej wielkości.

Zadaniu przypisano 3. poziom umiejętności na skali kompetencji w naukach przyrodniczych z wartością 517 pkt. Rozwiązało je prawidłowo 60,7% polskich uczniów, podczas gdy średnia dla OECD wyniosła 51,3%.

Zadanie 2.

Drugie zadanie z prezentowanej wiązki dotyczy problemu zrozumienia wpływu niepewności pomiarowej na możliwość interpretacji otrzymanych danych. Wymagało przede wszystkim umiejętności interpretacji danych oraz odwoływało się do wiedzy o poznaniu naukowym. Rozwiązując zadanie, uczeń musiał wybrać jedną z dwóch odpowiedzi, a następnie wyjaśnić, dlaczego wybrał tę odpowiedź.

PISA 2015

Badanie zbocza wzgórza
Pytanie 2 / 2

Zapoznaj się z tekstem "Analiza danych" po prawej stronie. Kliknij, aby wybrać odpowiedź, a następnie wpisz wyjaśnienie, uzasadniając swój wybór.

Dwoje uczniów nie zgadza się co do przyczyn różnicy w poziomie wilgotności gleby między oboma zboczami.

- Uczeń nr 1 uważa, że różnica w poziomie wilgotności gleby wynika z różnicy w promieniowaniu słonecznym na obu zboczach.
- Uczeń nr 2 uważa, że różnica w poziomie wilgotności gleby wynika z różnicy w opadach deszczu na obu zboczach.


Na podstawie danych powiedz, który z uczniów ma rację?

Uczeń nr 1
 Uczeń nr 2

Wyjaśnij swoją odpowiedź.

BADANIE ZBOCZA WZGÓRZA
Analiza danych

Uczniowie obliczają średnie wielkości pomiarów zebranych w danym okresie z każdej pary urządzeń umieszczonych na każdym ze zboczy i obliczają niepewność tych średnich. Uzyskane przez nich wyniki zapisano w poniższej tabeli. Niepewność została zapisana po znaku "±".



	Średni poziom promieniowania słonecznego	Średni poziom wilgotności gleby	Średni poziom opadów deszczu
Zbocze A	3 800 ± 300 MJ/m ²	28 ± 2%	450 ± 40 mm
Zbocze B	7 200 ± 400 MJ/m ²	18 ± 3%	440 ± 50 mm

Aby wynik pomiaru mógł służyć za podstawę do dalszego wnioskowania, niezbędna jest znajomość dokładności, z jaką został on wykonany. Analiza i interpretacja wyników to jeden z kluczowych etapów badania naukowego. W przedstawionym zadaniu uczniowie mają przeanalizować dane dotyczące trzech parametrów charakteryzujących badane zbocza: średniego poziomu promieniowania słonecznego, średniej wilgotności gleby oraz średniego poziomu opadów deszczu. Wartości mierzonych parametrów zostały zestawione w tabeli razem z niepewnościami pomiarowymi średnich. Aby wybrać i uzasadnić prawidłową odpowiedź, niezbędne jest zrozumienie, że pomiar jest nierozzerwalnie związany z niepewnością pomiarową.

Rozwiązując zadanie, badany piętnastolatek powinien dostrzec, że rację ma „uczeń nr 1”, ponieważ błąd pomiaru poziomu promieniowania słonecznego jest niski w porównaniu z różnicą między średnimi, co pozwala wnioskować, że różnica jest rzeczywista. Natomiast różnica między średnim poziomem opadów deszczu na obu zboczach jest wielokrotnie mniejsza niż błąd pomiaru – czego prawdopodobnie nie zauważył uczeń nr 2. Za odpowiedzi poprawne uznawano zatem takie, w których badani uczniowie wybrali odpowiedź „Uczeń 1”, a w uzasadnieniu napisali np.:

- „Na zbocze B dociera o wiele więcej promieniowania słonecznego niż na zbocze A, za to ilość deszczu jest ta sama.”
- „Nie ma różnicy w ilości deszczu pomiędzy oboma zboczami.”
- „Jest duża różnica między ilością światła słonecznego docierającego na zbocze A w porównaniu do zbocza B.” Zadanie to prawidłowo rozwiązało 31,5 % polskich uczniów, nieco mniej niż średnio w OECD (34, 7%). Na skali trudności zadań uzyskało ono 589 pkt. (poziom 4).

Wiązka zadań „Bieganie podczas upału”

Omawiana wiązka zadań jest jednym z narzędzi interaktywnych, które stanowią nową formułę pomiaru kompetencji w naukach przyrodniczych w badaniu PISA. Wiązka *Bieganie podczas upału* dotyczy zagadnienia dociekania naukowego związanego z problemem termoregulacji u biegaczy długodystansowych, biegających w zmiennych warunkach pogodowych. Wyniki symulacji przedstawione w wiązce oparte są o uproszczony model matematyczny funkcjonowania organizmu konkretnej osoby po biegu trwającym jedną godzinę w różnych warunkach. Ta wiązka pytań występowała tylko w badaniu pilotażowym.

PISA 2015

Bieganie podczas upału
Wstęp

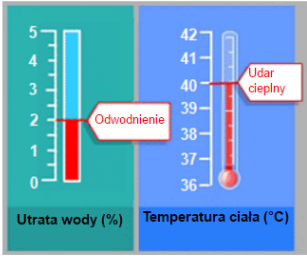
Przeczytaj wstęp. Następnie kliknij na strzałkę DALEJ.

BIEGANIE PODCZAS UPALU

W czasie biegów na długie dystanse temperatura ciała podnosi się i występuje pocenie.

U biegaczy, którzy nie piją dostatecznych ilości wody, aby uzupełnić wodę utraconą przez pocenie się, może wystąpić odwodnienie. Uznaje się, że stan odwodnienia występuje przy utracie wody stanowiącej 2% lub więcej masy ciała. Ta wielkość procentowa została oznaczona na pokazanej niżej podziałce wskazującej utratę wody.

Jeżeli temperatura ciała podniesie się do 40°C lub wyżej, u biegaczy może wystąpić stan powodujący zagrożenie życia, zwany udarem cieplnym. Temperaturę tę oznaczono na pokazanym niżej termometrze temperatury ciała.



Utrata wody (%)	Temperatura ciała (°C)
0	36
1	37
2	38
3	39
4	40
5	41

Po przeczytaniu wstępu uczniowie wykonują wstępną symulację, zapoznając się z działaniem modelu i ze sposobem zbierania danych.

PISA 2015

Bieganie podczas upału

Wstęp

Symulacja opiera się na modelu, w którym oblicza się objętość potu, utratę wody oraz temperaturę ciała biegacza po biegu trwającym jedną godzinę.

Aby zobaczyć, jak działają poszczególne ustawienia w tej symulacji, wykonaj poniższe kroki:

1. Przesuń suwak **temperatury powietrza**.
2. Przesuń suwak **wilgotności powietrza**.
3. Przy **picu wody** kliknij na odpowiedź Tak lub na odpowiedź Nie.
4. Kliknij na przycisk "Start", aby zobaczyć wyniki. Zauważ, że utrata wody stanowiąca 2% lub więcej powoduje odwodnienie, a temperatura ciała wynosząca 40°C lub więcej powoduje udar cieplny. Wyniki zostaną również wyświetlone w tabeli.

Uwaga: Wyniki symulacji są oparte na uproszczonym modelu matematycznym funkcjonowania organizmu konkretnej osoby po biegu trwającym jedną godzinę w różnych warunkach.

Temperatura powietrza (°C) 20 25 30 35 40

Wilgotność powietrza (%) 20 40 60

Picie wody Tak Nie

Temperatura powietrza (°C)	Wilgotność powietrza (%)	Picie wody	Objętość potu (litry)	Utrata wody (%)	Temperatura ciała (°C)

W przypadku prezentowanego zadania uczniowie mogą modelować funkcjonowanie organizmu człowieka po biegu trwającym jedną godzinę w różnych warunkach – przy zmieniającej się temperaturze i wilgotności powietrza oraz w zależności od tego, czy biegacz pije w czasie biegu wodę czy też nie. Po wyborze parametrów przez ucznia i naciśnięciu przycisku „Start”, w tabeli pojawiają się wartości wybranych parametrów i odpowiadające im wyniki modelowania zachowania się organizmu biegacza: objętość potu, utrata wody oraz temperatura ciała. Wyniki bieżącej symulacji są dodatkowo prezentowane graficznie na górnym schemacie. Jeśli wybrane warunki wywołają stan odwodnienia lub udaru cieplnego, zagrożenie zdrowia lub życia biegacza jest sygnalizowane czerwonymi agami. Po wykonaniu wstępnej symulacji uczniowie otrzymują informację, czy dobrze ją przeprowadzili.

Bieganie podczas upału

Wstęp

Symulacja opiera się na modelu, w którym oblicza się objętość potu, utratę wody oraz temperaturę ciała biegacza po biegu trwającym jedną godzinę.

Aby zobaczyć, jak działają poszczególne ustawienia w tej symulacji, wykonaj poniższe kroki:

1. Przesuń suwak **temperatury powietrza**.
2. Przesuń suwak **wilgotności powietrza**.
3. Przy **picciu wody** kliknij na odpowiedź Tak lub na odpowiedź Nie.
4. Kliknij na przycisk "Start", aby zobaczyć wyniki. Zauważ, że utrata wody stanowiąca 2% lub więcej powoduje odwodnienie, a temperatura ciała wynosząca 40°C lub więcej powoduje udar cieplny. Wyniki zostaną również wyświetlone w tabeli.

Uwaga: Wyniki symulacji są oparte na uproszczonym modelu matematycznym funkcjonowania organizmu konkretnej osoby po biegu trwającym jedną godzinę w różnych warunkach.



Temperatura powietrza (°C) 20 25 30 35 40
 Wilgotność powietrza (%) 20 40 60
 Picie wody Tak Nie

Start

Temperatura powietrza (°C)	Wilgotność powietrza (%)	Picie wody	Objętość potu (litry)	Utrata wody (%)	Temperatura ciała (°C)
30	40	Tak	1,2	0,0	39,3

Dobrze!
Aby kontynuować, kliknij na strzałkę DALEJ.

Bieganie podczas upału

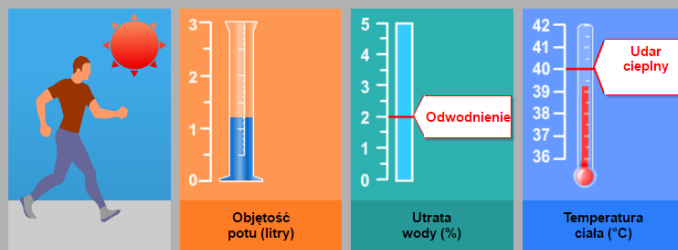
Wstęp

Symulacja opiera się na modelu, w którym oblicza się objętość potu, utratę wody oraz temperaturę ciała biegacza po biegu trwającym jedną godzinę.

Aby zobaczyć, jak działają poszczególne ustawienia w tej symulacji, wykonaj poniższe kroki:

1. Przesuń suwak **temperatury powietrza**.
2. Przesuń suwak **wilgotności powietrza**.
3. Przy **picciu wody** kliknij na odpowiedź Tak lub na odpowiedź Nie.
4. Kliknij na przycisk "Start", aby zobaczyć wyniki. Zauważ, że utrata wody stanowiąca 2% lub więcej powoduje odwodnienie, a temperatura ciała wynosząca 40°C lub więcej powoduje udar cieplny. Wyniki zostaną również wyświetlone w tabeli.

Uwaga: Wyniki symulacji są oparte na uproszczonym modelu matematycznym funkcjonowania organizmu konkretnej osoby po biegu trwającym jedną godzinę w różnych warunkach.



Temperatura powietrza (°C) 20 25 30 35 40
 Wilgotność powietrza (%) 20 40 60
 Picie wody Tak Nie

Start

Temperatura powietrza (°C)	Wilgotność powietrza (%)	Picie wody	Objętość potu (litry)	Utrata wody (%)	Temperatura ciała (°C)
30	40	Tak	1,2	0,0	39,3

Tak wygląda symulacja, jeśli temperaturę powietrza ustawiono na 30, wilgotność powietrza ustawiono na 40, w przypadku picia wody wybrano ustawienie "Tak", a następnie kliknięto "Start".

Aby kontynuować, kliknij na strzałkę DALEJ.

Zadanie

Zadaniem ucznia jest wskazanie maksymalnej temperatury, w której człowiek może biec przez godzinę przy wilgotności powietrza 40%, nie dostając udaru ciepłego. Uczniowie muszą przeprowadzić symulację dla podanych w pytaniu wartości temperatury powietrza. Udzielają odpowiedzi nie tylko wybierając odpowiednią temperaturę, ale także wskazując dwa wiersze z wynikami, które uzasadniają ten wybór. Na koniec wyjaśniają, w jaki sposób wybrane wiersze z wynikami modelowania uzasadniają wskazaną temperaturę.

PISA 2015

Bieganie podczas upału
Pytanie 4 / 5

Jak przeprowadzić symulację

Przeprowadź symulację, aby zebrać dane, opierając się na informacjach podanych poniżej. Kliknij, aby wybrać odpowiedź, zaznacz dane w tabeli, a następnie wpisz wyjaśnienie, uzasadniając swój wybór.

Odpowiedz na podstawie symulacji: Jeżeli wilgotność powietrza wynosi 40%, to jaka jest najwyższa temperatura powietrza, przy której człowiek może biec przez jedną godzinę, nie doznając udaru ciepłego?

20°C
 25°C
 30°C
 35°C
 40°C

★ Zaznacz w tabeli dwa wiersze z danymi, które uzasadniają wybraną przez Ciebie odpowiedź.

Wyjaśnij, w jaki sposób dane te uzasadniają Twoją odpowiedź.

Temperatura powietrza (°C)	Wilgotność powietrza (%)	Picie wody	Objętość potu (litry)	Utrata wody (%)	Temperatura ciała (°C)

Omawiane zadanie sprawdza umiejętność planowania doświadczenia i wymaga wiedzy o procedurach badawczych. Warto zwrócić uwagę na kontekst zadania, jakim jest zdrowie i choroby człowieka w aspekcie osobistym.

W ocenie wykonania tego zadania uwzględniono zarówno odpowiedzi całkowicie poprawne, jak i częściowo poprawne. Na w pełni poprawną odpowiedź składały się: wybór temperatury 35 °C, wskazanie wierszy z symulacjami dla wilgotności powietrza 40% i temperatur 35 °C oraz 40 °C oraz wyjaśnienie, które wskazuje lub sugeruje, że najwyższa temperatura powietrza niegrożąca udarem ciepłym wynosi 35 °C, ponieważ temperatura powietrza 40 °C taki udar już wywołuje. Oto przykłady takich wyjaśnień:

- „Kiedy temperatura na zewnątrz rośnie od 35 do 40 °C, temperatura ciała wzrasta powyżej 40 °C, co wywołuje u biegacza udar ciepły.”
- „Przy 40% wilgotności bieganie w temperaturze 40 °C powoduje udar ciepły, ale przy 35 stopniach temperatura ciała biegacza jest poniżej poziomu wystąpienia udaru ciepłego.”
- „Kiedy zwiększa się temperaturę powietrza, przy 40 stopniach biegacz doznaje udaru ciepłego.”

- „Kiedy poziom wilgotności wynosi 40%, u biegacza udar występuje dopiero przy 40 °C. 35 stopni to druga najwyższa temperatura.”
- „40 °C udar ciepły, nie 35.”

Za częściowo poprawne odpowiedzi uznano takie, w niepoprawny był jeden element: wyjaśnienie było niejasne bądź błędne albo nie były wybrane prawidłowe wiersze, albo wskazana była temperatura 40 stopni przy poprawnych pozostałych elementach. W badaniach pilotażowych zadanie uzyskało 592 punkty na skali trudności, co plasuje je na 4. poziomie.

5. CZYTANIE I INTERPRETACJA

Krzysztof Biedrzycki, Piotr Bordzół

Czytanie należy do podstawowych umiejętności, w które młodego człowieka wyposaża szkoła. Bez tej umiejętności niemożliwy jest jakikolwiek postęp w pozyskiwaniu wiedzy. Gdy jednak mówimy o czytaniu, możemy mieć na myśli różne sprawności. W sensie podstawowym czytanie oznacza proste dekodowanie tekstu, czyli zrozumienie zawartego w nim przekazu. W szerokim znaczeniu czytanie obejmuje jednak szereg operacji intelektualnych, które pozwalają dotrzeć do głębokich warstw tekstu. Najczęściej operacje te określane są jako interpretacja. Dlatego, żeby uniknąć nieporozumienia, w niniejszym raporcie badana umiejętność określana jest jako czytanie i interpretacja

W latach 2000 oraz 2009 czytanie i interpretacja stanowiły w badaniu PISA główną dziedzinę. W 2000 r. użyto 140 zadań, a w 2009 – 99 zadań. W latach 2003 oraz 2006 do pomiaru tej umiejętności wykorzystano 28 zadań, z kolei w 2012 – 45 zadań. W 2015 – wykorzystano 96 zadań z poprzednich edycji. W latach 2000, 2003, 2006, 2009 oraz 2012 uczniowie rozwiązywali zadania w tradycyjnej formie na papierowych arkuszach – umiejętności czytania tekstów w formie elektronicznej poświęcono osobne, uzupełniające części komputerowe. W 2015 roku w Polsce i w większości innych krajów do pomiaru umiejętności czytania i interpretacji wykorzystane zostały wyłącznie komputery.

5.1. Założenia teoretyczne badania

W badaniu PISA umiejętność czytania oznacza zrozumienie pisanych tekstów (w różnej formie i na różnych nośnikach), poddanie refleksji, wykorzystanie do różnych celów (np. pogłębienia wiedzy czy działania w życiu społecznym) oraz zaangażowanie czytelnika w niesione przez nie treści. Zadania sprawdzają umiejętność czytania w trzech obszarach: wyszukiwania informacji, interpretacji tekstu oraz refleksji i oceny. W badaniu z roku 2015 wyszukiwaniu informacji poświęcone było 21 zadań, interpretacji 53 zadania, refleksji i ocenie 22 zadania).

Dla opisu umiejętności uczniów ważne jest też przyjrzenie się zróżnicowaniu wyników – odsetkom uczniów na poszczególnych poziomach. Wykorzystane w badaniu PISA metody skalowania pozwala ją umieścić poziom umiejętności uczniów i trudność zadań na jednej skali. Aby ułatwić interpretację wyników i zobrazować rodzaj i trudność zadań oraz związane z nimi umiejętności, skalę podzielono na sześć poziomów, a pierwszy, najniższy poziom podzielono dodatkowo na trzy części.

Tabela 5.1. Opis poziomów umiejętności na skali osiągnięć czytania i interpretacji

Poziom	Umiejętności typowe dla danego poziomu
Poniżej poziomu 1b (poniżej 262,04 pkt.)	Uczeń nie radzi sobie z tekstem, nawet na podstawowym poziomie.
Poziom 1b (262,04–334,75 pkt.)	Uczeń wydobywa z tekstu informacje wyrażone wprost, radzi sobie z tekstem łatwym, o znanej mu formie, dostrzega proste związki między tekstem a osobistym doświadczeniem
Poziom 1a (334,75–407,47pkt.)	Uczeń wyszukuje jasno wskazane informacje w tekście o znanej mu formie, określa temat tekstu, dostrzega związki między informacjami zawartymi w tekście a wiedzą potoczną.

Poziom 2 (407,47–480,18 pkt.)	Uczeń wydobywa z tekstu jedną lub kilka informacji określonych w poleceniu, dostrzega główną ideę tekstu, dostrzega związki tekstu z wiedzą posiadaną z innych źródeł.
Poziom 3 (480,18–552,89 pkt.)	Uczeń wyszukuje informacje, wydobywa z tekstu główną ideę, potrafi dokonywać połączeń napotkanych w tekście elementów.
Poziom 4 (552,89–625,61 pkt.)	Uczeń wydobywa z tekstu kilka zawartych w nim informacji, potrafi ogarnąć sens całości tekstu, wykorzystuje posiadaną wiedzę do dokonywania ocen, które potrafi właściwie uzasadnić.
Poziom 5 (625,61–698,32 pkt.)	Uczeń potrafi wydobyć szereg informacji, które są zawarte głęboko w tekście, interpretuje szczegóły tekstu i tekst w całości, dokonuje krytycznej oceny tekstu.
Poziom 6 (powyżej 698,32 pkt.)	Uczeń bezbłędnie dociera do wszystkich informacji, dokonuje głębokiej interpretacji tekstu, stawia hipotezę dotyczącą całościowej oceny tekstu.

5.2. Wyniki polskich uczniów na tle innych krajów

W roku 2015 polscy uczniowie w badaniu PISA w dziedzinie czytania i interpretacji uzyskali średnio 506 punktów. Ich wynik był wyższy od średniego wyniku krajów OECD o 13 punktów (w 2012 roku był wyższy o 22 punkty). W porównaniu do innych krajów świata Polska znalazła się na wysokim, trzynastym miejscu; w Europie, jeśli uwzględnić błąd statystyczny, lepszy średni wyniki osiągnęli tylko uczniowie z pięciu krajów. Wyniki polskich uczniów porównywalne są z wynikami uczniów z następujących krajów: Nowa Zelandia, Niemcy, Makao, Słowenia, Holandia, Australia, Szwecja, Dania, Francja, Belgia. W 2015 roku Polska była więc nadal w gronie krajów o najwyższym wyniku, mimo że jej średni wynik, w porównaniu do roku 2012, był niższy.

Tabela 5.2. Średnie wyniki w zakresie czytania i interpretacji w krajach objętych badaniem. Polska na tle świata w 2015 r.; pogrubioną czcionką zaznaczono kraje europejskie. Białym tłem oznaczone są kraje, których wynik nie jest istotnie różny od średniego wyniku w krajach OECD, jasnym kolorem wyróżniono kraje o wynikach lepszych, a ciemnym – gorszych od średniego wyniku OECD.

Kraj/region	Średni wynik
Singapur	535
Hongkong (Chiny)	527
Kanada	527
Finlandia	526
Irlandia	521
Estonia	519
Korea	517
Japonia	516
Norwegia	513
Nowa Zelandia	509
Niemcy	509
Makao	509
Polska	506
Słowenia	505
Holandia	503
Australia	503
Szwecja	500
Dania	500
Francja	499

Belgia	499
Portugalia	498
Wielka Brytania	498
Tajpej (Republika Chińska)	497
Stany Zjednoczone	497
Hiszpania	496
Rosja	495
B-S-J-G (Chiny)	494
Szwajcaria	492
Łotwa	488
Czechy	487
Chorwacja	487
Wietnam	487
Austria	485
Włochy	485
Islandia	482
Luksemburg	481
Izrael	479
Autonomiczne Miasto Buenos Aires	475
Litwa	472
Węgry	470
Grecja	467
Chile	459
Słowacja	453
Malta	447
Cypr	443
Urugwaj	437
Rumunia	434
Zjednoczone Emiraty Arabskie	434
Bułgaria	432
Turcja	428
Kostaryka	427
Trinidad i Tobago	427
Czarnogóra	427
Kolumbia	425
Meksyk	423
Mołdawia	416
Tajlandia	409
Jordania	408
Brazylia	407
Albania	405
Katar	402
Gruzja	401
Peru	398
Indonezja	397
Tunezja	361
Dominikana	358
Macedonia	352
Algieria	350
Kosowo	347
Liban	347

Warto spojrzeć na wyniki osiągnięte przez polskich uczniów w dłuższej perspektywie. Od roku 2000 do 2006 polscy uczniowie poprawiali swoją pozycję. W 2000 ich wynik wyniósł 479 punktów, w 2003 – 497, w 2006 – 508. Od tamtej pory wyniki ulegają wahaniom, ale pozostają porównywalne z wynikami krajów znacznie przewyższających średnią OECD (2009 – 500, 2012 – 518, 2015 – 506 pkt). Między 2000 a 2015 rokiem w Polsce średni wynik z czytania i interpretacji wzrósł o 27 pkt. Dla porównania w Tabeli 5.3. zamieszczono wyniki z badań z poprzednich edycji w krajach Unii Europejskiej. Warto zauważyć, że średni wynik krajów OECD między rokiem 2000 a 2015 zmniejszył się, o ile w pierwszym z tych lat wyniósł on 500 punktów, o tyle w ostatnim badaniu spadł do 493 punktów.

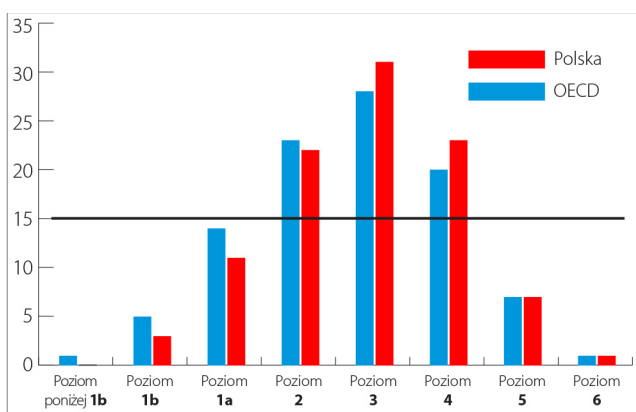
Tabela 5.3. Średnie wyniki uczniów z pomiaru umiejętności czytania i interpretacji z lat 2000, 2003, 2006, 2009 i 2012 i 2015. Polska, kraje Unii Europejskiej biorące udział w badaniu. Białe tło oznacza kraje, których wynik nie był istotnie różny od przeciętnego w krajach OECD, jasnym kolorem wyróżniono kraje o wynikach lepszych, a ciemnym – gorszych od średniego wyniku.

PISA 2000		PISA 2003		PISA 2006		PISA 2009		PISA 2012		PISA 2015	
Kraj	Średni wynik	Kraj	Średni wynik	Kraj	Średni wynik	Kraj	Średni wynik	Kraj	Średni wynik	Kraj	Średni wynik
Finlandia	546	Finlandia	543	Finlandia	547	Finlandia	536	Finlandia	524	Finlandia	526
Irlandia	527	Irlandia	515	Irlandia	517	Holandia	508	Irlandia	523	Irlandia	521
Szwecja	516	Szwecja	514	Polska	508	Belgia	506	Polska	518	Estonia	519
Austria	507	Holandia	513	Holandia	507	Estonia	501	Estonia	516	Niemcy	509
Belgia	507	Belgia	507	Szwecja	507	Polska	500	Holandia	11	Polska	506
Francja	505	Polska	497	Estonia	501	Szwecja	497	Belgia	509	Słowenia	505
Dania	497	Francja	496	Belgia	501	Niemcy	497	Niemcy	508	Holandia	503
Hiszpania	493	Dania	492	Niemcy	495	Irlandia	496	Francja	505	Szwecja	500
Czechy	492	Niemcy	491	Wielka Brytania	495	Francja	496	Wielka Brytania	499	Dania	500
Włochy	487	Austria	491	Dania	494	Dania	495	Dania	496	Francja	499
Niemcy	484	Łotwa	491	Słowenia	494	Wielka Brytania	494	Czechy	493	Belgia	499
Węgry	480	Czechy	489	Austria	490	Węgry	494	Włochy	490	Portugalia	498
Polska	479	Węgry	482	Francja	488	Portugalia	489	Austria	490	Wielka Brytania	498
Grecja	474	Hiszpania	481	Czechy	483	Włochy	486	Łotwa	489	Hiszpania	496
Portugalia	470	Luksemburg	479	Węgry	482	Łotwa	484	Węgry	488	Łotwa	488
Łotwa	458	Portugalia	478	Łotwa	479	Słowenia	483	Hiszpania	488	Czechy	487
Bułgaria	430	Włochy	476	Luksemburg	479	Grecja	483	Luksemburg	488	Chorwacja	487
Rumunia	428	Grecja	472	Chorwacja	477	Hiszpania	481	Portugalia	488	Austria	485
		Słowenia	469	Portugalia	472	Czechy	478	Chorwacja	485	Włochy	485
				Litwa	470	Słowacja	477	Szwecja	483	Luksemburg	481
				Włochy	469	Chorwacja	476	Słowenia	481	Litwa	472
				Słowacja	466	Luksemburg	472	Litwa	477	Węgry	470
				Hiszpania	461	Litwa	468	Grecja	477	Grecja	467
				Grecja	460	Bułgaria	429	Słowacja	463	Słowacja	453
				Bułgaria	402	Rumunia	424	Rumunia	438	Malta	447
				Rumunia	396			Bułgaria	436	Cypr	443
										Rumunia	434
										Bułgaria	432

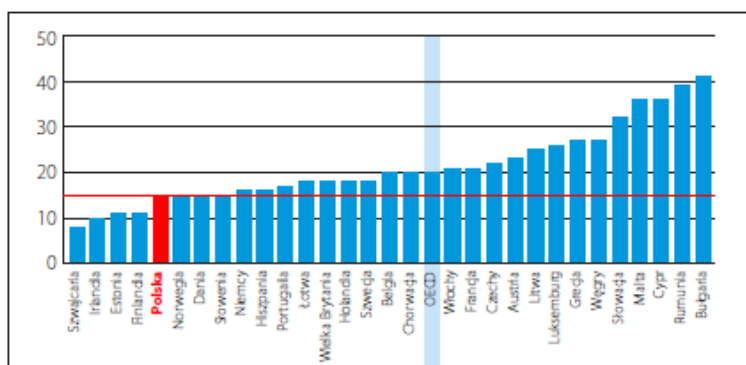
5.3. Poziomy umiejętności w dziedzinie czytania

Wiele informacji przenosi analiza osiągnięć uczniów na poszczególnych poziomach umiejętności. Poziom 2 można przyjąć za minimalny poziom kompetencji, poniżej którego rośnie ryzyko nieradzenia sobie w dalszej edukacji i w dorosłym życiu. Takie kryterium uzgodniły kraje Unii Europejskiej jako ważny wskaźnik osiągania celów polityki edukacyjnej – wykorzystywany jako wskaźnik zagrożenia wykluczeniem społecznym. Ostatnia edycja badania potwierdza, że wyniki polskich uczniów należą do mało zróżnicowanych. Stosunkowo nieliczne są grupy skrajne: uczniów poniżej poziomu 2, a także na poziomach 5 i 6 (Wykresy 4.1, 4.2 i 4.3). Trzeba podkreślić, że w Polsce między rokiem 2000 a 2015 zdecydowanie spadła liczba uczniów uzyskujących wyniki poniżej 2 poziomu. Obecnie mniejszy niż w Polsce odsetek tych uczniów odnotowuje się w niewielu krajach i regionach: Hongkongu (9%), Irlandii (10%), Estonii, Kanadzie, Finlandii, Singapurze (po 11%), Makao (12%) oraz Japonii (13%). Warto zwrócić uwagę, że w Europie są tylko trzy kraje, w których odsetek uczniów uzyskujących najniższe wyniki jest mniejszy niż w Polsce. Mniej zadowolający jest efekt pracy z uczniami najzdolniejszymi. Obecnie odsetek polskich uczniów, którzy w dziedzinie czytania i interpretacji uzyskali wyniki na 5 i 6 poziomie (8,2%), jest zbliżony do średniej OECD (8,4%). W Europie w jedenastu krajach pod tym względem jest lepiej.

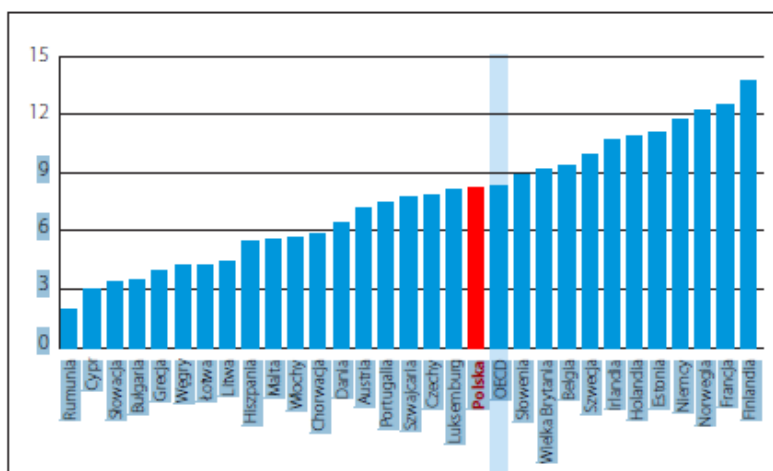
Wykres 5.1. Odsetek uczniów, którzy uzyskali wyniki na poszczególnych poziomach w 2015 r. (Polska i średnia OECD, %)



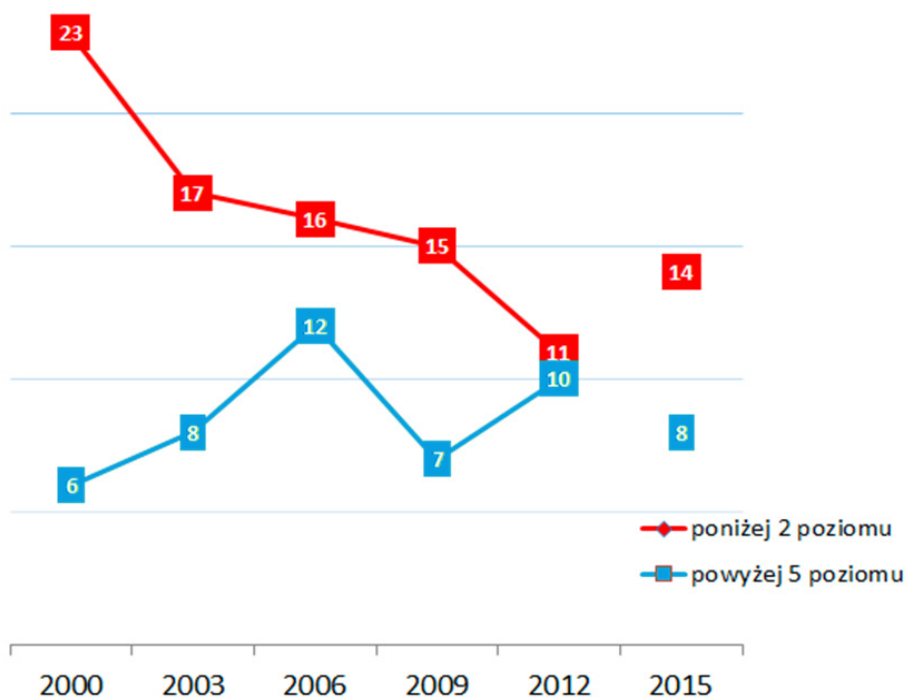
Wykres 5.2. Odsetek uczniów, którzy uzyskali wyniki poniżej 2 poziomu w krajach Unii Europejskiej i związanych z Unią Europejską w 2015 r. (%). Kraje uporządkowano według wzrastającego odsetka uczniów.



Wykres 5.3. Odsetek uczniów, którzy uzyskali wyniki powyżej poziomu 5 w krajach Unii Europejskiej i związanych z Unią Europejską w 2015 r. (%). Kraje uporządkowano według wzrastającego odsetka.



Wykres 5.4. Odsetek uczniów, którzy uzyskiwali wyniki poniżej 2 poziomu oraz powyżej 5 poziomu w Polsce w kolejnych edycjach badania (%)

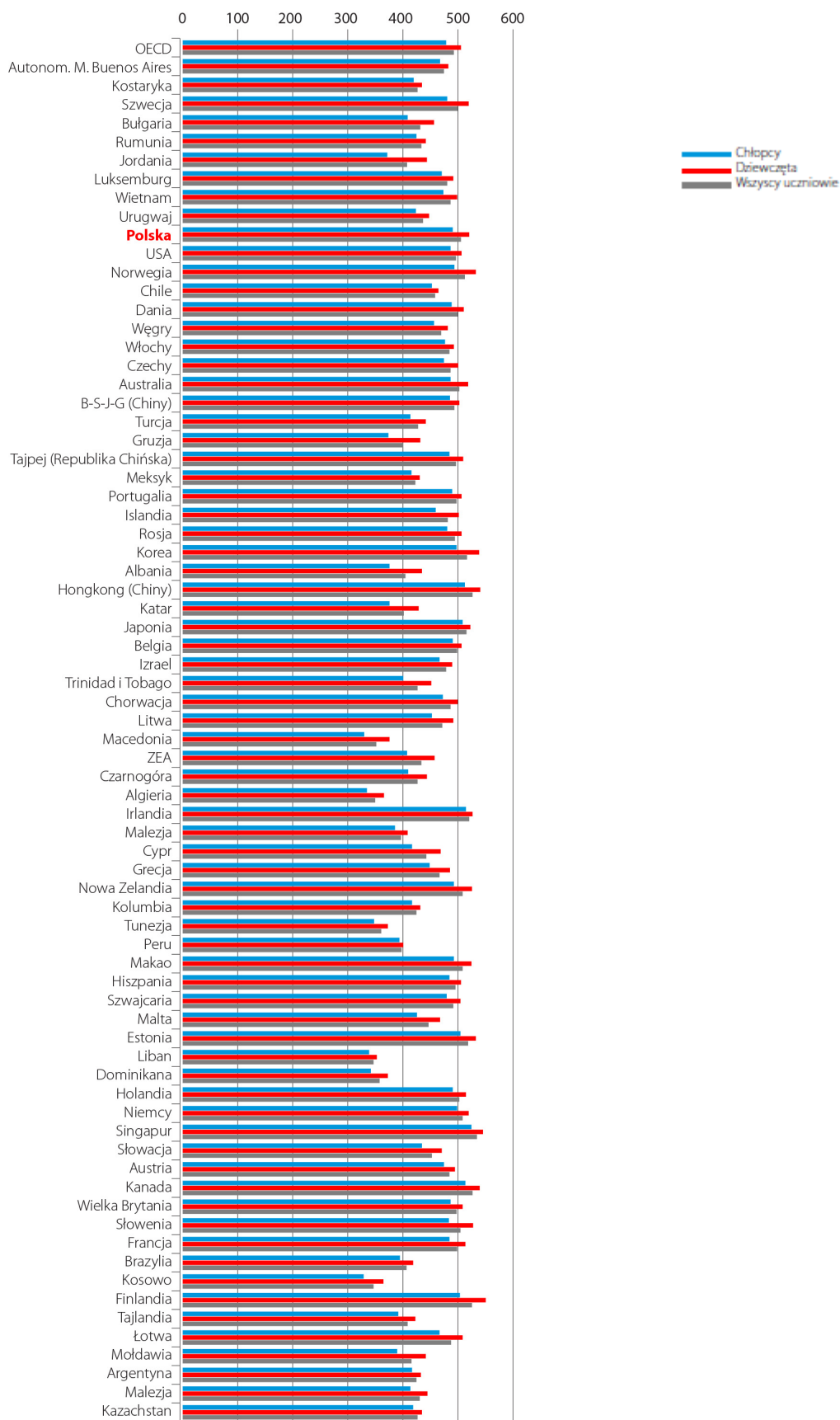


5.4. Wyniki chłopców i dziewcząt

W badaniu umiejętności czytania i interpretacji zauważalna jest różnica między wynikami chłopców i dziewcząt, na korzyść dziewcząt. W 2015 w OECD średni wynik chłopców wyniósł 479, a dziewcząt 506. W Polsce wyniki wyglądały następująco: chłopcy 491 pkt., dziewczęta 521 pkt. Zatem różnica wyniosła 30 pkt. na korzyść dziewcząt, jest ona zbliżona do średniej OECD (27 pkt., Wykres 5.5). Między rokiem 2000 a 2015 nieznacznie zmniejszyła się ta różnica: w 2000 roku chłopcy w Polsce osiągnęli średni wynik 461 pkt., a dziewczęta 497 pkt., toteż różnica wtedy wynosiła 36 pkt. W 2012 roku różnica między chłopcami a dziewczętami wynosiła 42 pkt.

Interesujące jest porównanie różnic w rozbiciu na poziomy umiejętności. Od 2009 roku odsetek chłopców uzyskujących najniższe wyniki utrzymuje się na względnie stabilnym poziomie. Warto w tym miejscu zauważyć, że tylko w pięciu krajach UE odsetek chłopców uzyskujących najniższe wyniki był w 2015 roku mniejszy niż w Polsce. Badanie z 2012 pokazało znaczne zmniejszenie się odsetka dziewcząt uzyskujących wyniki poniżej drugiego poziomu. W 2015 r. odsetek uczennic z takimi wynikami nieco wzrósł, mimo to wyniki te należą nadal do najkorzystniejszych w UE – podobnie jak w przypadku chłopców tylko w pięciu krajach (w tym w pozostającej formalnie poza UE Norwegii) jest mniejszy odsetek dziewcząt uzyskujących najsłabsze wyniki. Jeśli chodzi o chłopców uzyskujących najwyższe wyniki, to ich odsetek w Polsce jest zbliżony do przeciętnego w UE. W długiej perspektywie czasowej jest on dosyć stabilny, jakkolwiek daje się zauważyć wyraźny wzrost liczby uczniów uzyskujących najwyższe wyniki w stosunku do badania z 2009 roku (niemal podwojenie). Odsetek dziewcząt uzyskujących najwyższe wyniki również jest zbliżony do średniej dla UE. W ujęciu czasowym można zauważyć względną stabilizację od 2009 roku.

Wykres 5.5. Średnie wyniki dziewcząt i chłopców w krajach i regionach uczestniczących w badaniu PISA 2015.



5.5. Pogłębiona analiza jakościowa wyników badania

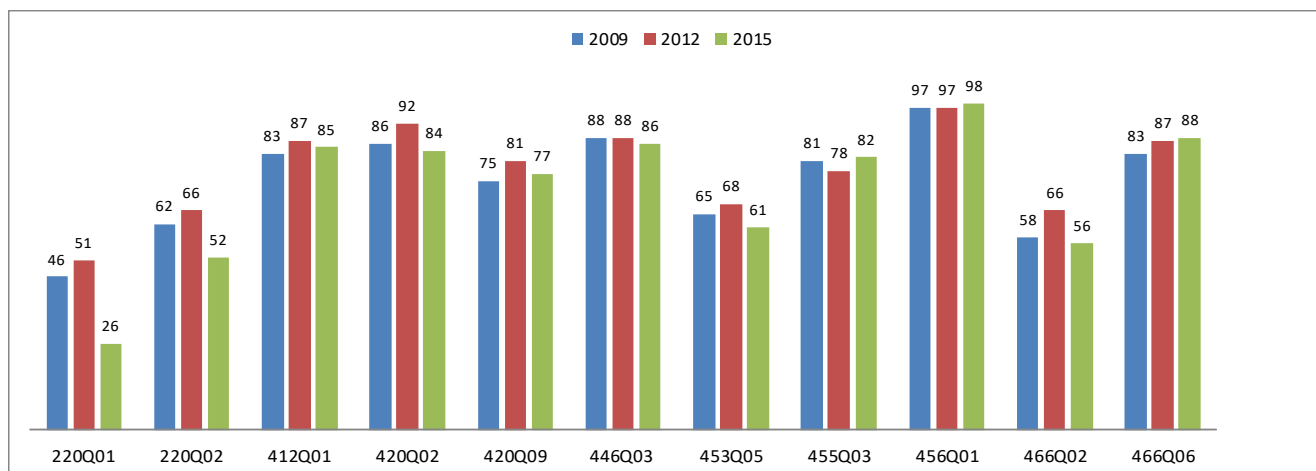
Pogłębiona analiza wyników pozwala na sformułowanie dalszych wniosków dotyczących umiejętności polskich uczniów na tle osiągnięć młodzieży w innych krajach oraz polskich piętnastolatków w poprzednich edycjach badania.

Zadania sprawdzają umiejętności w trzech głównych obszarach: wyszukiwania informacji, interpretacji tekstu oraz refleksji i argumentacji. Jeśli chodzi o wyszukiwanie informacji, to uczeń ma przede wszystkim zlokalizować je w tekście, niekiedy jednak musi dokonać hierarchizacji różnych wiadomości, żeby wskazać tę, o którą chodzi w pytaniu. Interpretacja polega na tym, żeby wydobyc z wypowiedzi sens, który nie został wyrażony wprost, dostrzec główną myśl tekstu, ogarnąć go w całości, zauważyć związki między jego poszczególnymi częściami. Umiejętność refleksji i argumentacji sprawdzana jest w ten sposób, że uczeń ma skonfrontować informacje zaczerpnięte z tekstu z informacjami pochodzącymi z innych źródeł, sformułować własny sąd i uzasadnić go, wśród zadań z tego obszaru pojawiają się takie, w których wymaga się dostrzeżenia celowości rozwiązań konstrukcyjnych zastosowanych w tekście. Każdy obszar obejmuje szereg konkretnych, szczegółowo opisywanych umiejętności. Poniżej zostaną przedstawione wnioski, które można wyciągnąć z analizy wyników zadań.

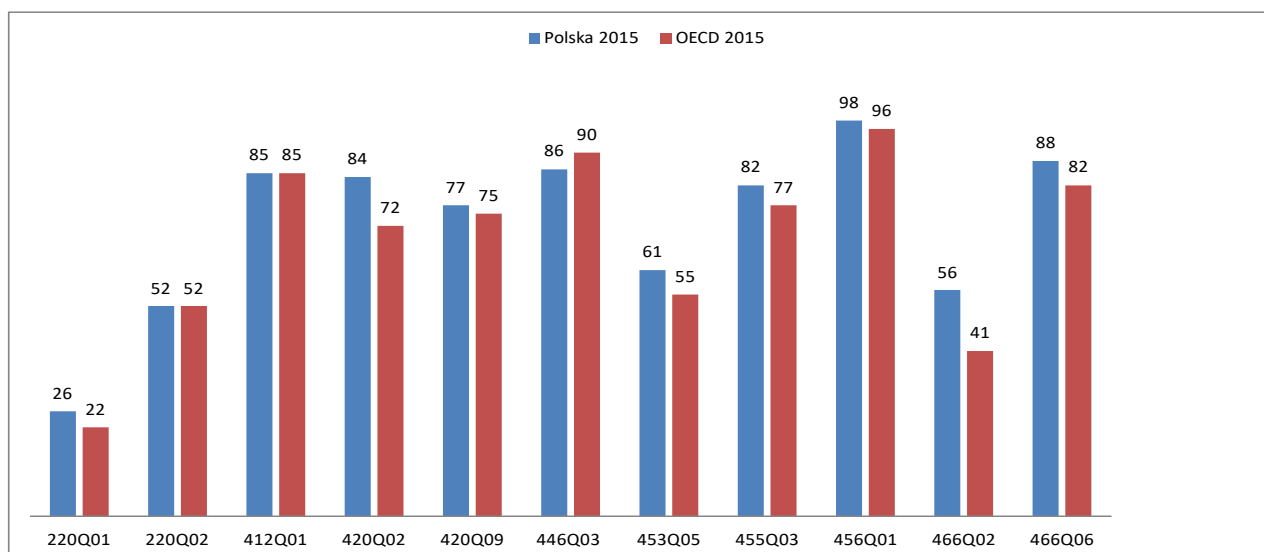
5.5.1. Wyszukiwanie informacji

W przypadku większości zadań z tego obszaru wyniki uzyskiwane w ciągu lat na ogół są stabilne, podobne jak w edycjach badania z 2009 i 2012 roku są to wyniki wysokie (między 80% a 90% prawidłowych odpowiedzi), wyższe niż średnie wyniki OECD.

Wykres 5.6. Średnie wyniki polskich uczniów w latach 2009, 2012, 2015. Zadania z obszaru *Wyszukiwanie informacji*. Ujęte zostały tylko zadania powtarzające się w kolejnych edycjach badania (tzw. zadania kotwiczące)



Wykres 5.7. Średnie wyniki polskich uczniów oraz w krajach OECD w 2015 r. Zadania z obszaru Wyszukiwanie informacji (zadania kotwiczące).



Uczniowie bardzo dobrze potrafili wydobyć informacje tak z tekstu ciągłego, jak z wykresu (niekiedy dokonując ich konfrontacji). Rzecz ciekawa, polscy uczniowie znacznie lepiej niż ich koledzy z Finlandii, Estonii czy Słowenii (krajów zajmujących wyższe lokaty w rankingu na podstawie średniej liczby punktów) potrafili odpowiedzieć na pytanie dotyczące konkretnej informacji z tekstu, które jednak wymagało odnalezienia cytatu i zinterpretowania go. W ciągu kilku lat (między badaniami 2009 i 2015) nastąpiła wyraźna poprawa w zrozumieniu treści zawartej na ulotce reklamowej, pod tym względem piętnastolatki w naszym kraju radzą sobie tak jak uczniowie w najlepszych krajach europejskich, warto zauważyć, że niektóre zadania sprawdzające tę umiejętność należą do względnie trudnych, polski wynik jest wyraźnie wyższy niż średnie w krajach OECD.

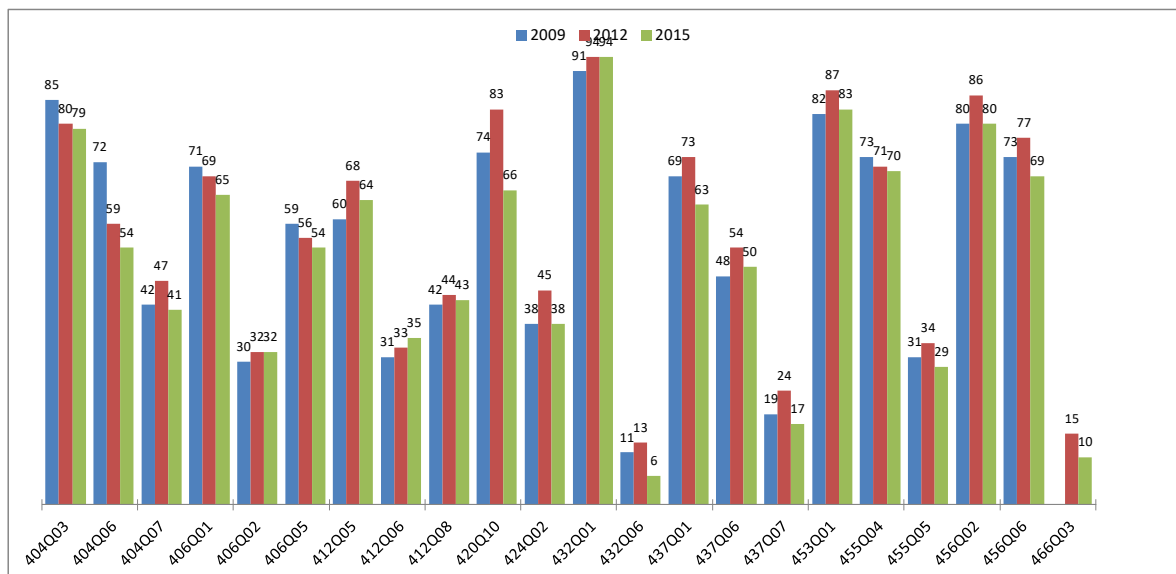
Bardzo znaczny spadek wyniku nastąpił w przypadku zadań, w których trzeba było połączyć informacje pochodzące z kilku (więcej niż dwóch) tekstów o różnym charakterze – tekst ciągły, mapa, wykres (zadania 220Q01 i 220Q02). Ów spadek nastąpił jednak w większości krajów, co można wiązać ze zmianą medium. Po pierwsze, nie jest wykluczone, że trudność stanowiło porównanie tekstów, które trudno było przedstawić równocześnie na ekranie (na papierze można było je zestawić obok siebie). Po drugie, uczniowie mieli wykonać rysunek na mapie, a więc powinni umieć poprawnie posługiwać się kursorem. Główną barierą mógł być zatem brak biegłości w pracy na komputerze, tym bardziej że inne zadania polegające na konfrontowaniu informacji zwartych w różnego typu tekstach wykonywane były bez problemu.

5.5.2. Interpretacja tekstu

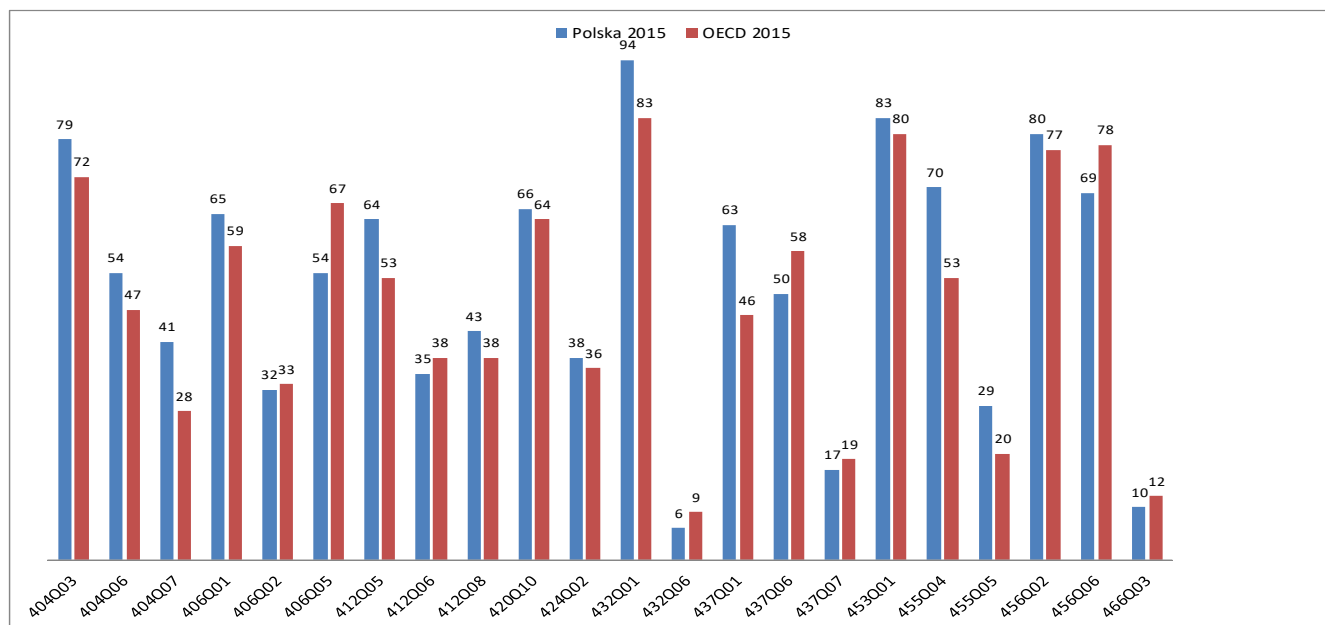
Najwięcej zadań należało do obszaru interpretacji. Są to zadania znacznie trudniejsze niż w obszarze wyszukiwania informacji, wobec czego wyniki w oczywisty sposób okazują się słabsze. Istotne jest jednak ich porównanie z wynikami innych krajów i z wynikami z poprzednich edycji badania PISA w Polsce. Otóż w przypadku większości zadań z tego obszaru polskie rezultaty są lepsze niż przeciętne w OECD, ale przeważnie gorsze niż w krajach osiągających najwyższe lokaty (jak Finlandia, Irlandia czy Estonia). W tym obszarze nastąpił też najbardziej widoczny spadek wyników. Ponieważ zadania te najczęściej wymagają wykazania się umiejętnościami złożonymi, można postawić tezę, że pomimo poprawy w długiej perspektywie (od 2000 roku), nadal w polskiej szkole mamy spore rezerwy, jeśli chodzi

o kształcenie samodzielności myślenia, wnikliwego interpretowania tekstu, odczytywania myśli niewyrażonych wprost

Wykres 5.8. Średnie wyniki polskich uczniów w latach 2009, 2012, 2015. Zadania z obszaru *Interpretacja tekstu*. Ujęte zostały tylko zadania powtarzające się w kolejnych edycjach badania (tzw. zadania kotwiczące). Zadanie 466Q03 w 2009 nie było uwzględnione w wynikach, tutaj zaprezentowane jest jako przykład zadania wyjątkowo trudnego.



Wykres 5.9. Średnie wyniki polskich uczniów oraz w krajach OECD w 2015 r. (%). Zadania z obszaru *Interpretacja tekstu* (zadania kotwiczące).



Są zadania, z którymi polscy uczniowie poradzili sobie znacznie lepiej niż ich koledzy nawet w krajach o bardzo dobrych wynikach. Zwłaszcza gdy trzeba było z tekstów wydobyć informacje niewyrażone wprost, lecz zawarte w narracji, i połączyć je. To umiejętność pokrewna umiejętnościom z poprzedniego obszaru, jednak wymagająca więcej sprawności polegającej na wnikliwej lekturze tekstu, zrozumieniu jego sensu, wskazaniu odpowiedniego fragmentu,

hierarchizacji i konfrontacji informacji, wreszcie eliminacji treści w tym kontekście nieistotnych. Nie było istotne, czy wypowiedź miała charakter literacki, czy popularnonaukowy, polscy uczniowie radzili sobie z nimi równie dobrze. Dotyczyło to niekiedy zadań dosyć trudnych. W kilku zadaniach zauważono wyraźny wzrost wyników, przede wszystkim tam, gdzie chodziło o wnikliwe czytanie tekstu. Niekiedy były to zadania bardzo trudne (rozwiązywalność w krajach OECD na poziomie 30%), wymagające zrozumienia sensu fragmentu opowiadania lub głównej myśli w tekście wyrażonych językiem naukowym.

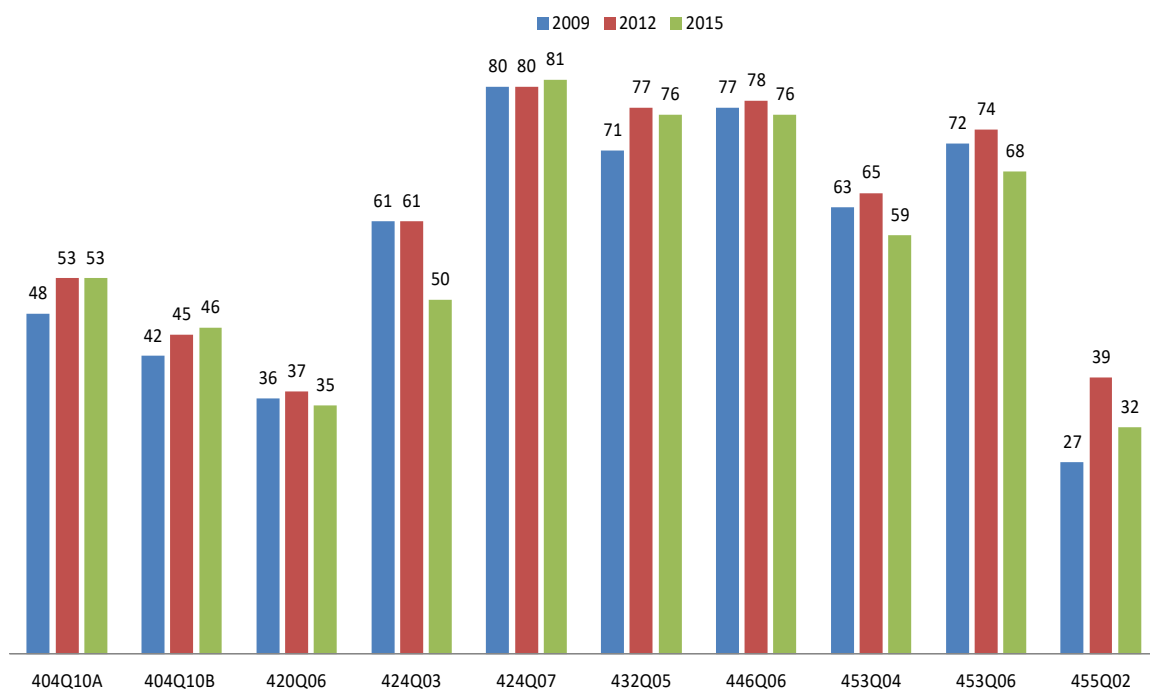
Największą trudność dla polskich uczniów stanowią zadania, w których trzeba się wykazać wnikliwością interpretacyjną, samodzielnością myślenia i umiejętnością zwięzłego sformułowania myśli (są to często zadania otwarte). Znacznie gorzej niż ich koledzy z krajów OECD (nie mówiąc o takich krajach jak Finlandia czy Estonia) radzą sobie z interpretacją zachowań i wypowiedzi bohaterów, ze zrozumieniem znaczenia nieoczekiwanych zwrotów akcji czy z wnioskowaniem na podstawie przesłanek zawartych w tekście. Widoczny jest lęk przed niekonwencjonalnym rozwiązywaniem problemów intelektualnych związanych z lekturą. Niekiedy bardzo dobre zrozumienie fragmentu wypowiedzi paradoksalnie utrudnia dostrzeżenie sensu całości, co widać po analizie rozwiązań różnych zadań dotyczących tego samego tekstu.

W przypadku prawie połowy zadań z obszaru interpretacji nastąpił spadek wyników. Można wśród nich wyodrębnić dwa typy zadań. W pierwszym typie spotykamy zadania wymagające porównania trzech tekstów różnego rodzaju (np. tekst ciągły, wykres, tabela, rysunek). Można przypuszczać, że podobnie jak w obszarze wyszukiwania informacji problemem była nieumiejętność porównania sobie z tekstami, które trzeba by ustawić obok siebie na ekranie. Drugi typ dotyczy zadań obejmujących umiejętność samodzielnej interpretacji tekstu i logicznego wnioskowania. Widać to szczególnie w przypadku bardzo trudnych zadań (około 10% rozwiązywalności w krajach OECD), w których na podstawie informacji zawartych w otrzymanych tekstach (przesłanek) trzeba było zrekonstruować informacje pochodne (wnioski). W tych przypadkach nastąpiło wyraźne pogorszenie. Podobnie sytuacja wyglądała z zadaniami wymagającymi wnikliwszej interpretacji intencji i wypowiedzi postaci w opowiadaniach czy zrozumienia przesłania utworu.

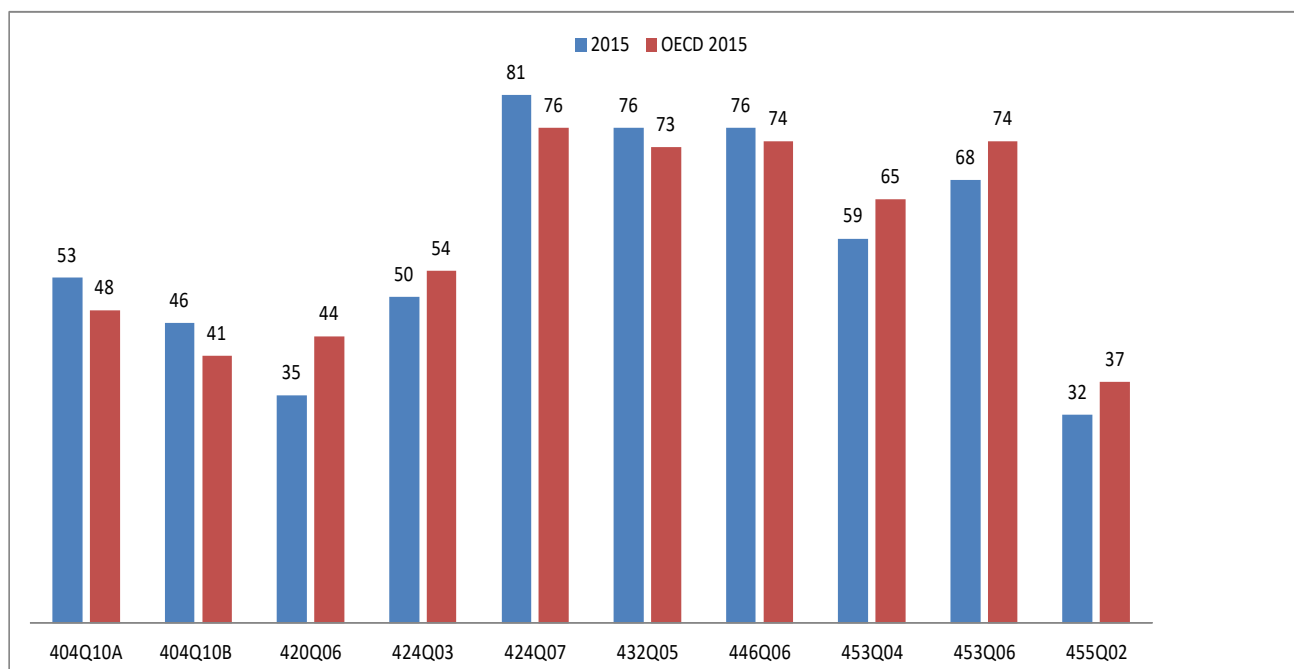
5.5.3. Refleksja i argumentacja

Trzeci obszar obejmował zadania wymagające od ucznia podjęcia wysiłku oceny tekstu, zrozumienia celowości zastosowanych w nim rozwiązań formalnych, połączenia zawartych w nim informacji z wiedzą uzyskaną wcześniej, uzasadnienia sformułowanych sądów.

Wykres 5.10. Średnie wyniki polskich uczniów z lat 2009, 2012, 2015. Zadania z obszaru Refleksja i argumentacja. Ujęte zostały tylko zadania powtarzające się w kolejnych edycjach badania (tzw. zadania kotwiczące).



Wykres 5.11. Średnie wyniki polskich uczniów oraz w krajach OECD w 2015 r. Zadania z obszaru Refleksja i argumentacja (zadania kotwiczące).



Polscy uczniowie bardzo dobrze, znacząco lepiej niż przeciętnie w OECD, porównywalnie z najlepszymi krajami, potrafią podejmować refleksję na temat celowości zastosowanych

w tekście rozwiązań, potrafią także włączać pozyskane informacje w inne konteksty. Na przykład odczytują sens wprowadzenia konkretnej opowieści do artykułu publicystycznego. Na podstawie pozyskanej informacji na temat jakiegoś zjawiska w jednym kraju, dosyć dobrze potrafią ocenić możliwość wystąpienia tego samego zjawiska w innym kraju. Ciekawie wygląda umiejętność oszacowania funkcji perswazyjnej tekstu, gdyż w pokrewnych zadaniach dotyczących tej kwestii wyniki były wyraźnie rozbieżne. Zapewne decydujący wpływ na wykonanie poleceń miały inne czynniki niż sama umiejętność dostrzeżenia perswazji (być może to, że łatwiej wskazać funkcję fragmentu w jednolitym tekście niż funkcję wykresu wobec tekstu ciągłego).

Piętnastolatki w Polsce mają jednak poważny kłopot z uzasadnianiem formułowanych sądów i podawaniem argumentów. Wyraźnie słabiej wypadają zadania otwarte, niejednokrotnie wynik jest niższy niż przeciętnie w OECD, a dramatycznie niższy niż w krajach z czołówki europejskiej. O tym, że problemem jest argumentacja, a nie samo przemyślenie zagadnienia, świadczy fakt, że inne zadania wskazują na duże umiejętności analizy tekstu i zjawisk, których on dotyczy.

5.6. Podsumowanie

Wyniki badania PISA z 2015 roku pokazują, że polscy uczniowie dobrze opanowali umiejętności z dziedziny czytania i interpretacji. Spadek średniej liczby punktów (dający się również zaobserwować dla średniej liczby punktów w OECD) mieści się w zakresie niedającym powodów do niepokoju, zwłaszcza że porównanie wyników z kolejnych edycji badania wskazuje na względną stabilizację, która została osiągnięta po reformie przeprowadzonej na przełomie stuleci (między wynikami z 2000 i 2015 roku odnotowuje się zauważalny wzrost średniej liczby punktów). W okresie 2000–2015 zauważalny jest także spadek odsetka uczniów, którzy uzyskują wyniki poniżej 2 poziomu oraz nieznaczny wzrost odsetka tych najlepszych. Trudno jednoznacznie stwierdzić, co jest przyczyną słabszego wyniku w roku 2015 w porównaniu z rokiem 2012.

Polscy uczniowie bardzo dobrze radzą sobie z lekturą tekstu, której celem jest uzyskanie informacji. Nie tylko z lekturą prostą, bo również taką, w której informacje muszą być konfrontowane, hierarchizowane, a także wyszukiwane w strukturze tekstu, jego narracji. Coraz lepiej piętnastolatki radzą sobie w wnikliwym i uważnym czytaniu. Odczytują sens rozwiązań formalnych stosowanych w tekstach. Potrafią pozyskane informacje łączyć ze swoją wiedzą.

Jak w poprzednich edycjach badania można dostrzec słabość w opanowaniu wielu umiejętności złożonych. Kłopot stanowi konfrontacja tekstów różnego rodzaju. Trudność sprawia samodzielność myślenia – stawianie niestandardowych tez, interpretowanie nieoczywistych znaczeń zawartych w wypowiedziach, odczytywanie sensów zachowań postaci literackich, wnioskowanie na podstawie zawartych w tekstach przesłanek. W pracy dydaktycznej powinno się więcej uwagi zwracać właśnie na umiejętności złożone, zwłaszcza na samodzielność rozwiązywania problemów intelektualnych.

Polscy uczniowie bądź unikają zadań otwartych, bądź są w nich dosyć bezradni. Słabo uzasadniają swoje stanowiska. To może się łączyć z tym, że mają za mało praktyki w zwięzłym argumentowaniu. Zauważalne są niedostatki w umiejętności logicznego dowodzenia stawianych tez, w tym tez interpretacyjnych.

W edycji badania PISA 2015 pojawił się dodatkowy kłopot związany z nowym medium. Jego wpływ jest dostrzegalny w przypadku zadań, w których uczniowie mieli konfrontować więcej niż dwa teksty o różnym charakterze. Podobne zadania w poprzednich edycjach, nawet jeśli sprawiały trudności, wykonywane były z lepszymi rezultatami. Słabiej wypadły również zadania otwarte, co także może się wiązać z tym, że niektórzy uczniowie mogli sobie nie radzić z pisaniem na komputerze. Niekiedy, na co wskazują analizy, problemy związane z samą umiejętnością czytania były pogłębiane przez opór, jaki części uczniów stawiało medium.

Badanie PISA z 2015 roku potwierdza obserwacje z poprzednich edycji, że polski system kształcenia umiejętności czytania i interpretacji jest skuteczny, jakkolwiek konieczne jest zwrócenie większej uwagi na rozszerzenie grupy uczniów biegłych w wykonywaniu operacji najtrudniejszych, wymagających znacznej samodzielności myślenia, którzy byliby zdolni do uzyskiwania wyników na najwyższych poziomach umiejętności.

6. MATEMATYKA W BADANIU PISA 2015

Zbigniew Marciniak, Agnieszka Sułowska

Pomiar umiejętności matematycznych w badaniu PISA koncentruje się na tym, w jakim stopniu 15-letni uczniowie są w stanie wykorzystać swoją wiedzę i umiejętności matematyczne, gdy stają przed koniecznością rozwiązania problemów praktycznych, jakich dostarcza im otaczający świat. Nawet bardzo dobra, ale tylko teoretyczna znajomość narzędzi matematyki nie wystarczy. Dobry wynik uzyskują uczniowie, którzy umieją rozumować matematycznie, a także potrafią skutecznie wykorzystywać pojęcia i narzędzia matematyczne do opisu, analizy i prognozowania różnych zjawisk. Są to wymagania znacznie wykraczające poza umiejętność mechanicznego odtworzenia wyuczonych pojęć i procedur szkolnych. Aby im sprostać potrzebne są solidne umiejętności i dobre rozumienie matematyki na poziomie gimnazjalnym.

W 2015 roku w badaniu PISA wzięli udział uczniowie z aż 72 krajów lub regionów świata. Daje nam to, po raz kolejny, możliwość przyjrzenia się, jak z określonym powyżej niełatwym wyzwaniem radzą sobie polscy uczniowie w porównaniu z rówieśnikami z innych krajów.

Część matematyczna badania zawierała 70 zadań o różnym stopniu trudności. W 26 z nich można było wybrać odpowiedź z zaproponowanej listy; jednak aż w 44 pozostałych należało przedstawić samodzielne opracowane rozwiązanie.

6.1. Umiejętności polskich uczniów

Matematyka była główną dziedziną badania PISA w latach 2003 oraz 2012. Po badaniu w roku 2003 została zdefiniowana skala matematyczna, do której odnoszone są wyniki uczniów uzyskane w kolejnych cyklach. Skala ta została skalibrowana tak, by średni wynik krajów OECD (w roku 2003) był równy 500, a odchylenie standardowe 100.

Zastosowanie uogólnionego modelu Rascha daje możliwość umieszczenia na tej skali zarówno wszystkich badanych uczniów, jak i wszystkich zadań. Zadania, które na tej skali znajdują się wysoko, mierzą umiejętności złożone. Uczniowie uzyskujący wysoki wynik radzą sobie z takimi zadaniami lepiej od innych.

Dla lepszego objaśnienia tych różnic oraz uzyskanych wyników na skali wyróżniono sześć poziomów oraz podano typowe umiejętności uczniów, plasujących się na każdym z nich (tabela 6.1.).

Tabela 6.1. Opis poziomów umiejętności na skali osiągnięć matematycznych

Poziom (dolna granica)	Umiejętności typowe dla danego poziomu umiejętności matematycznych
Poziom 6 (669 pkt.)	Uczeń potrafi analizować i uogólniać informacje zgromadzone w wyniku zbadania samodzielnie zbudowanego modelu złożonej sytuacji problemowej. Umie połączyć informacje pochodzące z różnych źródeł i swobodnie przemieszczać się między nimi. Potrafi wykonywać zaawansowane rozumowania i umie wnioskować matematycznie. Umie połączyć rozumowanie z biegłością w wykonywaniu operacji symbolicznych i formalnych podczas twórczej pracy nad nowym dla siebie kontekstem. Potrafi precyzyjnie formułować komunikat o swoim rozumowaniu, uzasadniając podjęte działania.
Poziom 5 (607 pkt.)	Uczeń umie modelować złożone sytuacje, identyfikując ograniczenia i precyzując zastrzeżenia. Potrafi porównywać, oceniać i wybierać odpowiednie strategie rozwiązywania problemów związanych ze zbudowanym modelem. Wykorzystuje dobrze rozwinięte umiejętności matematyczne, z użyciem odpowiednich reprezentacji, w tym symbolicznych i formalnych. Potrafi krytycznie ocenić swoje działania, zakomunikować swoją interpretację oraz sposób rozumowania.
Poziom 4 (545 pkt.)	Uczeń umie efektywnie pracować z podanymi wprost modelami złożonych sytuacji realnych, identyfikując ograniczenia i czyniąc niezbędne założenia. Potrafi wybierać oraz łączyć informacje pochodzące z różnych źródeł, wiążąc je bezpośrednio z kontekstem realnym. Umie w takich kontekstach stosować ze zrozumieniem dobrze wyuczone techniki. Potrafi konstruować komunikaty opisujące swoje interpretacje, argumenty i działania.
Poziom 3 (482 pkt.)	Uczeń umie wykonać jasno opisany algorytm, także wymagający sekwencyjnego podejmowania decyzji. Potrafi wybierać i stosować proste strategie rozwiązywania problemów. Potrafi interpretować i wyciągać bezpośrednie wnioski z danych pochodzących z kilku źródeł. Umie przedstawić wyniki nieskomplikowanych interpretacji i rozważań.
Poziom 2 (420 pkt.)	Uczeń umie rozpoznać i zinterpretować sytuację wymagającą tylko prostego kojarzenia. Potrafi wydobyć istotną informację z pojedynczego źródła i użyć na raz jednej formy reprezentacji danych. Umie zastosować prosty wzór lub przepis postępowania. Potrafi wyciągnąć bezpośrednie wnioski i dosłownie zinterpretować wyniki.
Poziom 1 (358 pkt.)	Uczeń umie rozwiązywać typowe zadania, w których wszystkie dane są bezpośrednio podane, a zadane pytania są proste. Potrafi wykonywać czynności rutynowe, postępując zgodnie z podanym prostym przepisem. Podejmuje działania oczywiste, wynikające wprost z treści zadania.
Poniżej poziomu 1	Uczeń wykazuje brak umiejętności nawet na poziomie 1.

Źródło: Raport międzynarodowy: PISA 2015 Results Excellence and equity in education, Volume I, OECD 2016

Tabela 6.2. Średnie wyniki uczniów z matematyki w badaniu PISA 2015.

Pogrubioną czcionką wyróżniono kraje europejskie. Kolor tła wyróżnia kraje, które uzyskały wyniki (odpowiednio, od góry) statystycznie istotnie wyższe, statystycznie nieodróżnialne oraz statystycznie istotnie niższe niż średnia krajów OECD.

Kraj/region	Średni wynik 2015	Kraj/region	Średni wynik 2015
Singapur	564	Węgry	477
Hongkong	548	Słowacja	475
Makao	544	Israel	470
Tajwan	542	USA	470
Japonia	532	Chorwacja	464
B-S-J-G (Chiny)	531	Kazachstan	460
Korea	524	Grecja	454
Szwajcaria	521	Malezja	446
Estonia	520	Rumunia	444
Kanada	516	Bułgaria	441
Holandia	512	Cypr	437
Dania	511	Emiraty	427
Finlandia	511	Chile	423
Słowenia	510	Turcja	420
Belgia	507	Moldawia	420
Niemcy	506	Urugwaj	418
Polska	504	Czarnogóra	418
Irlandia	504	Trynidad	417
Norwegia	502	Tajlandia	415
Austria	497	Albania	413
Nowa Zelandia	495	Argentyna	409
Wietnam	495	Meksyk	408
Rosja	494	Gruzja	404
Szwecja	494	Katar	402
Australia	494	Kostaryka	400
Francja	493	Liban	396
Portugalia	492	Kolumbia	390
Wielka Brytania	492	Peru	387
Czechy	492	Indonezja	386
Włochy	490	Jordania	380
Islandia	488	Brazylia	377
Hiszpania	486	Macedonia	371
Luksemburg	486	Tunezja	367
Łotwa	482	Kosowo	362
Malta	479	Algieria	360
Litwa	478	Dominikana	328

Źródło: Raport międzynarodowy PISA 2015 Results Excellence and equity in education, Volume I, OECD 2016, tabela L502

W badaniu PISA 2015 średni wynik krajów OECD z matematyki wyniósł 490 punktów.

O wyniku kraju decyduje średnia wyników wszystkich uczniów z tego kraju uczestniczących w badaniu. Wyniki te obarczone są błędem statystycznym, który można oszacować wykorzystując odpowiednie techniki statystyczne. Dzięki temu wiemy, że dwa kraje, które mają różne wyniki punktowe, mogą mimo to być statystycznie nieodróżnialne. Dlatego należy zachować dużą ostrożność przy interpretowaniu listy krajów uszeregowanych według średniego wyniku punktowego.

Dla każdego kraju można wskazać grupę krajów, których wyniki nie różnią się od jego wyniku w sposób statystycznie istotny. Dla Polski tymi krajami są Belgia, Niemcy, Irlandia i Norwegia. Wynik istotnie wyższy od Polski uzyskało 14 krajów i regionów, w tym 6 krajów europejskich: Szwajcaria, Estonia, Holandia, Dania, Finlandia i Słowenia.

Na wyniki badania PISA 2015 oczekiwano w wielu krajach z dużym zainteresowaniem także ze względu na to, że towarzyszyła mu zmiana, która mogła istotnie te wyniki zaburzyć: w 59 krajach (w tym w Polsce) przeprowadzono badanie nie jak dotychczas w wersji papierowej, lecz za pomocą komputerów. Uczniowie rozwiązywali z matematyki dokładnie te same zadania co w 2012 roku, ale nie otrzymali ich w formie książeczki, lecz czytali je odpowiednio nawigując na ekranach, a swoje rozwiązania wpisywali z klawiatury komputera lub zaznaczali myszką. Nie było jasne, jak ta zmiana wpłynie na wyniki badania.

Okazało się, że w zdecydowanej większości krajów, które w 2012 roku miały wyniki zbliżone do średniej OECD lub wyższe, w roku 2015 nastąpił spadek wyniku o kilka, kilkanaście lub nawet kilkadziesiąt punktów. Tylko trzy kraje (Słowenia, Dania i Norwegia) na 29 krajów z tej grupy zanotowały istotny statystycznie wzrost wyniku (tabela 5.3.).

Jednak mimo tych zmian, lista krajów, które w badaniu PISA 2015 osiągnęły najlepsze wyniki z matematyki uległa tylko niewielkim zmianom w stosunku do badania PISA 2012. Można zatem stwierdzić, że w zakresie umiejętności matematycznych polscy gimnazjaliści utrzymali swoją wysoką pozycję.

Tabela 6.3. Średnie wyniki uczniów z matematyki w krajach o najwyższych wynikach w badaniu PISA 2012 oraz zmiana wyniku między rokiem 2012 a 2015 (w punktach).

Kolorem czerwonym zaznaczono obniżenie wyniku punktowego, a czarnym poprawę wyniku. Zmiany istotne statystycznie zaznaczono pogrubioną czcionką. Gwiazdka oznacza, że kraj lub region nie brał udziału w badaniu PISA 2015.

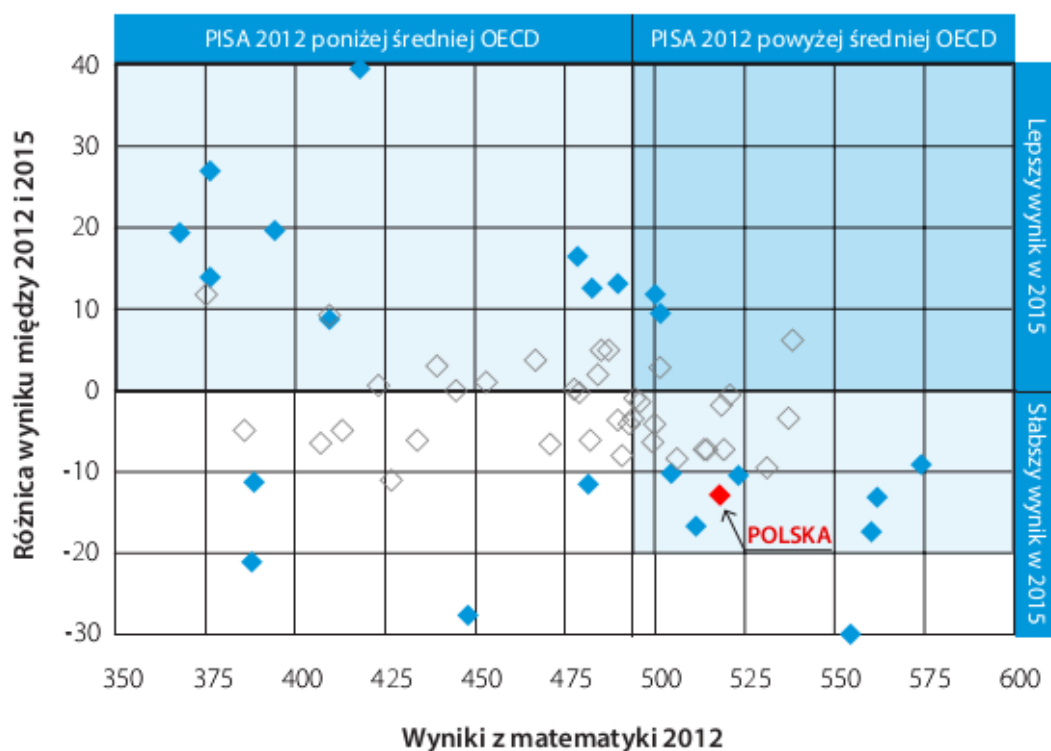
Kraj/region	Średni wynik w 2012	Zmiana od 2012
Szanghaj (Chiny)	613	*
Singapur	573	-9
Hongkong	561	-13
Tajwan	560	-18
Korea	554	-30
Macao (Chiny)	538	+6
Japonia	536	-4
Liechtenstein	535	*
Szwajcaria	531	-10
Holandia	523	-11
Estonia	521	-1
Finlandia	519	-8
Kanada	518	-2
Polska	518	-14
Belgia	515	-8
Niemcy	514	-8
Wietnam	511	-16
Austria	506	-9
Australia	504	-10
Irlandia	501	+3
Słowenia	501	+9
Dania	500	+11
Nowa Zelandia	500	-5
Czechy	499	-7
Francja	495	-2
Wielka Brytania	494	-2
Islandia	493	-5

Łotwa	491	-9
Luksemburg	490	-4
Norwegia	489	+13
Portugalia	487	+5
Hiszpania	487	-1
Włochy	485	-5
Rosja	482	+12
Słowacja	482	-7
USA	481	-11

Źródło: Raport międzynarodowy: PISA 2015 Results Excellence and equity in education, Volume I, OECD 2016, tabela I.504

To samo zjawisko ilustruje wykres 6.1. Pokazuje on także, że znaczna większość krajów, których wyniki z matematyki są istotnie statystycznie lepsze w 2015 roku niż w 2012, to kraje o bardzo niskich wynikach – poniżej 425 punktów.

Wykres 6.1. Różnica wyniku z matematyki między badaniami PISA 2012 i PISA 2015 w poszczególnych krajach, w zależności od wyniku z matematyki w badaniu PISA 2012 w tych krajach. Niebieskim rombem zaznaczono zmiany istotnie statystycznie, przezroczystym rombem zmiany nieistotnie statystycznie.



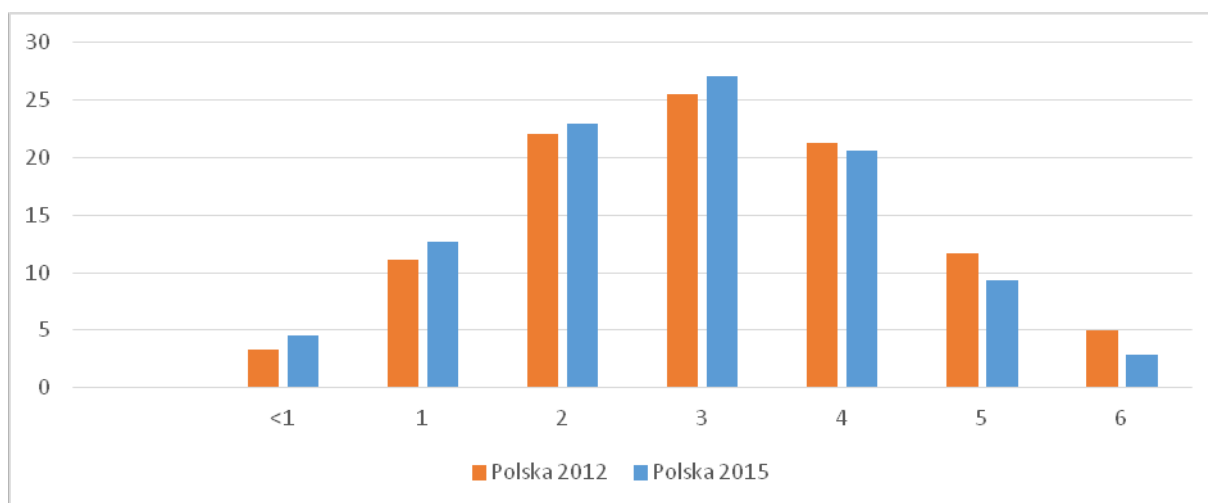
Źródło: Raport międzynarodowy: PISA 2015 Results Excellence and equity in education, Volume I, OECD 2016, diagram I.505

6.2. Zmiany w poziomach osiągnięć

Słabszemu wynikowi kraju towarzyszy z reguły zwiększenie odsetka uczniów na najniższych poziomach umiejętności, a także obniżenie odsetka uczniów na najwyższych poziomach.

Podobnie jak w wielu najlepszych krajach świata, także w Polsce nastąpiło takie przesunięcie (wykres 6.2.).

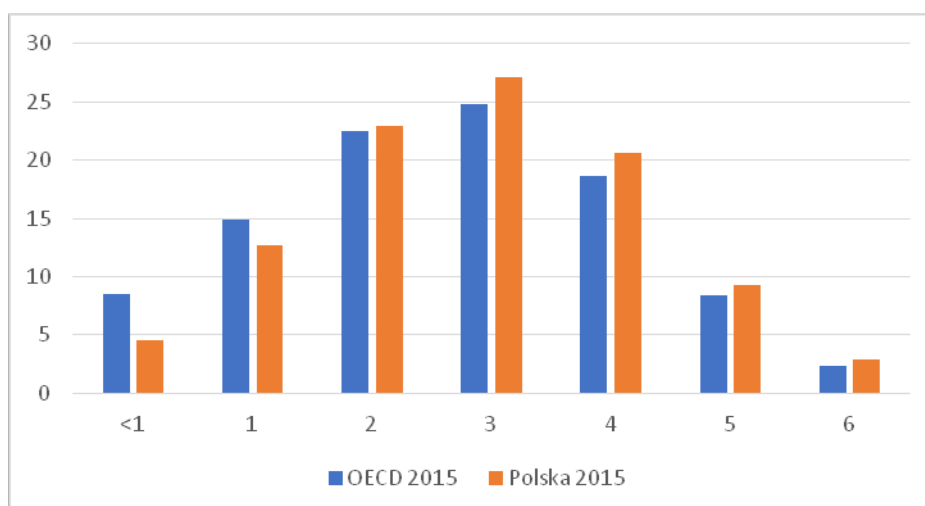
Wykres 6.2. Odsetki uczniów w Polsce na poszczególnych poziomach umiejętności w badaniu PISA 2012 i PISA 2015.



W badaniu PISA 2015 odsetek uczniów o umiejętnościach poniżej poziomu 2 wzrósł w Polsce z 14,4% w 2012 roku do 17,2% w roku 2015. Jednocześnie odsetek uczniów na dwóch najwyższych poziomach umiejętności zmniejszył się z 16,7% do 12,2%.

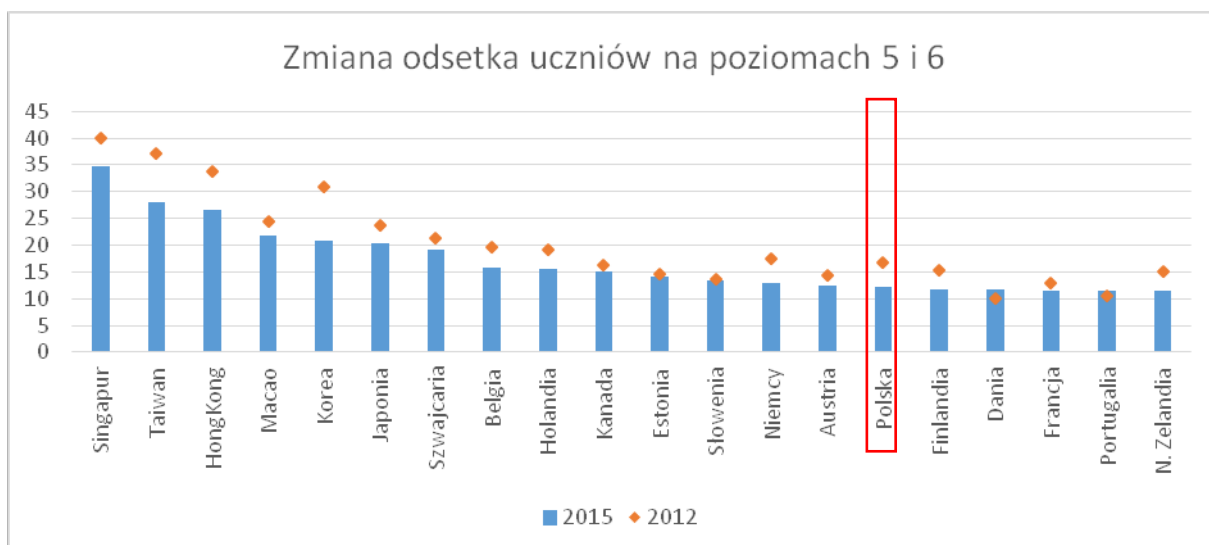
Należy przy tym podkreślić, że zarówno pod względem odsetka najlepszych, jak i najslabszych uczniów wynik Polski jest wciąż bardzo dobry na tle innych krajów świata oraz krajów OECD. Ilustrują to wykresy 6.3., 6.4. i 6.5.

Wykres 6.3. Odsetki uczniów w Polsce i w krajach OECD na poszczególnych poziomach umiejętności w badaniu PISA 2015.



Wykres 6.3. wyraźnie pokazuje, że w Polsce w porównaniu z krajami OECD jest znacznie mniej uczniów osiągających najniższe wyniki – odsetek uczniów o umiejętnościach poniżej poziomu 2 w OECD jest równy 23,4%, a w Polsce 17,2%.

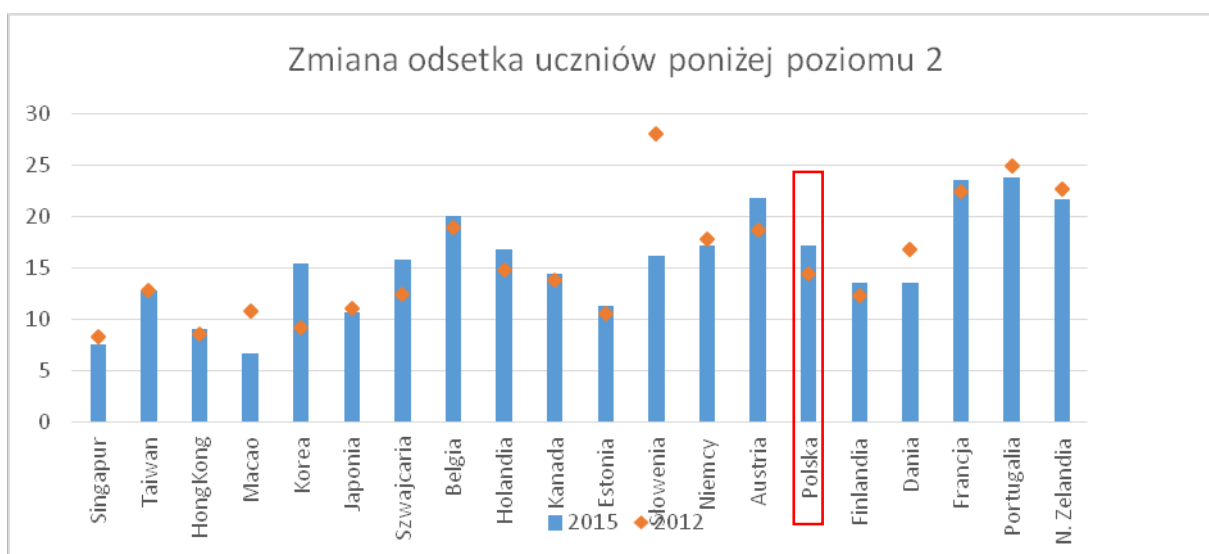
Wykres 6.4. Odsetki uczniów na poziomach 5 lub 6 w poszczególnych krajach w latach 2012 i 2015.



Na wykresie uwzględniono kraje, które w badaniu PISA 2015 mają najwyższe odsetki uczniów na tych poziomach umiejętności. Kraje uporządkowano według malejącego odsetka uczniów na tych poziomach w 2015 roku.

Wykres 6.5. Odsetki uczniów na 1. poziomie lub niżej w poszczególnych krajach w latach 2012 i 2015.

Na wykresie uwzględniono te same kraje, co na wykresie 6.3 i uporządkowano je w takiej samej kolejności jak na wykresie 6.4.

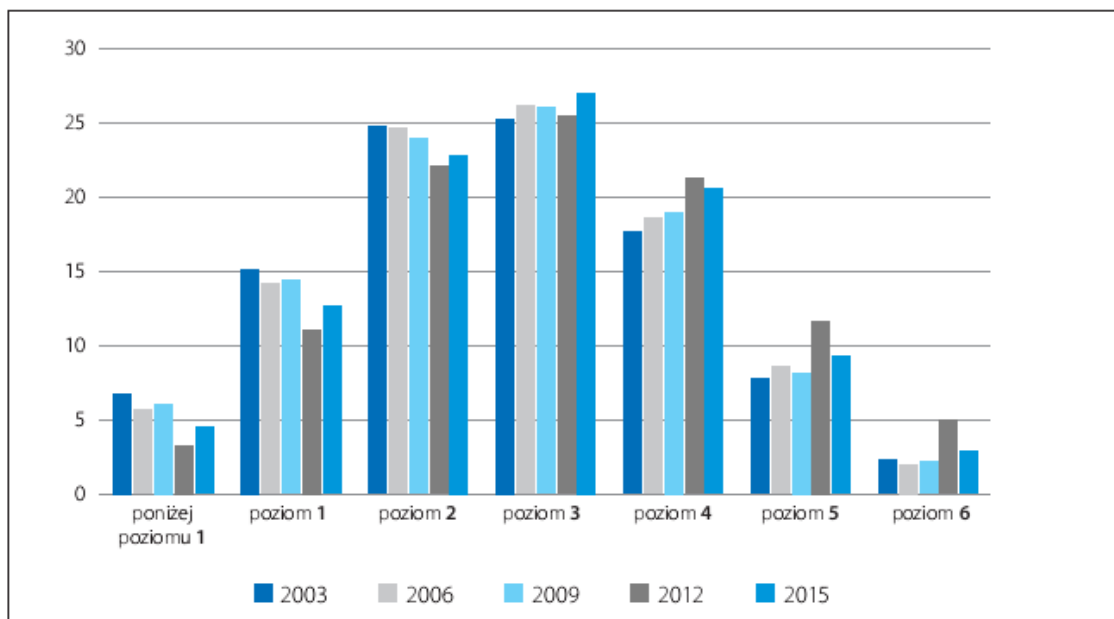


Opracowano na podstawie raportu międzynarodowego: *PISA 2015 Results Excellence and equity in education, Volume I, OECD 2016, diagram I.509*

Warto spojrzeć na tę kwestię z perspektywy historycznej. Na wykresie 6.6. przedstawiona jest zmiana rozkładu uczniów pomiędzy poszczególne poziomy umiejętności matematycznych na przestrzeni pięciu ostatnich edycji badania PISA. Pokazuje on, że w 2015 roku uczniów o umiejętnościami poniżej poziomu 2 było procentowo mniej niż w latach 2003-2009. Podobnie

odsetki uczniów na poziomach 5 i 6 są wyższe w roku 2015 niż w latach 2003-2009, w których były one zbliżone do średniego odsetka w krajach OECD.

Wykres 6.6. Odsetki uczniów w Polsce na poszczególnych poziomach umiejętności w badaniu PISA 2015.

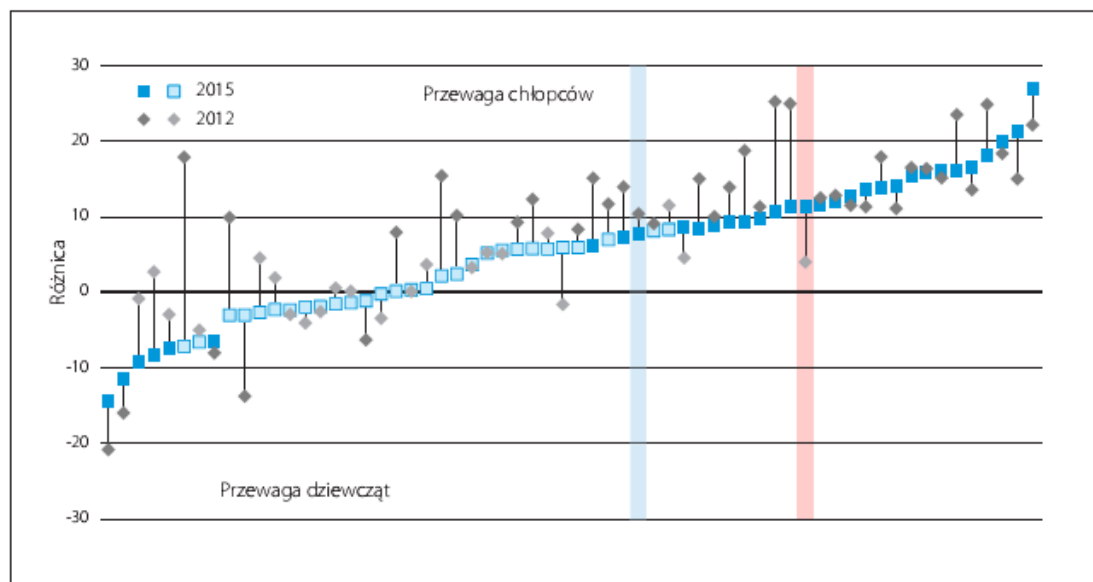


6.3. Porównanie wyników chłopców i dziewcząt

W badaniu PISA 2015 w wielu krajach nastąpiła znaczna zmiana zróżnicowania wyników dziewcząt i chłopców w stosunku do notowanych wcześniej. Efekt ten przedstawiony jest na wykresie 6.7.

Wykres 6.7. Różnica między wynikami z matematyki osiąganymi przez dziewczęta i chłopców w latach 2012 i 2015 (wynik chłopców minus wynik dziewcząt w punktach).

Dane uporządkowano począwszy od kraju o największej różnicy na korzyść dziewcząt w 2015 roku, do największej różnicy na korzyść chłopców w 2015 roku. Niebieskim kwadratem i ciemnym rombem zaznaczono różnice wyniku istotne statystycznie, natomiast błękitnym kwadratem i jasnym rombem różnice nieistotne statystycznie. Błękitny pasek wskazuje średnie różnice dla krajów OECD, a czerwony pasek dla Polski.



Źródło: Raport międzynarodowy: PISA 2015 Results Excellence and equity in education, Volume I, OECD 2016, diagram I.511

W przeciwieństwie do lat 2009 i 2012 w badaniu PISA 2015 chłopcy w Polsce uzyskali istotną statystycznie przewagę nad dziewczętami w zakresie umiejętności matematycznych – nieistotna statystycznie czteropunktowa różnica z roku 2012 zwiększyła się w 2015 roku do 12 punktów. Średni wynik chłopców wyniósł 511 punktów, zaś dziewcząt 499 punktów. To zróżnicowanie wyników występuje także na dolnym i górnym krańcu skali umiejętności. W dolnym (najniższym) decylny wyniki chłopców i dziewcząt to odpowiednio 396 i 387 punktów (9 punktów różnicy na korzyść chłopców), zaś w ostatnim (najwyższym) decylny te wyniki to odpowiednio 627 oraz 608 punktów (aż 19 punktów różnicy na korzyść chłopców). Pojawienie się tak dużych różnic między dziewczętami i chłopcami przypuszczalnie może mieć związek z użyciem komputerów w trakcie badania.

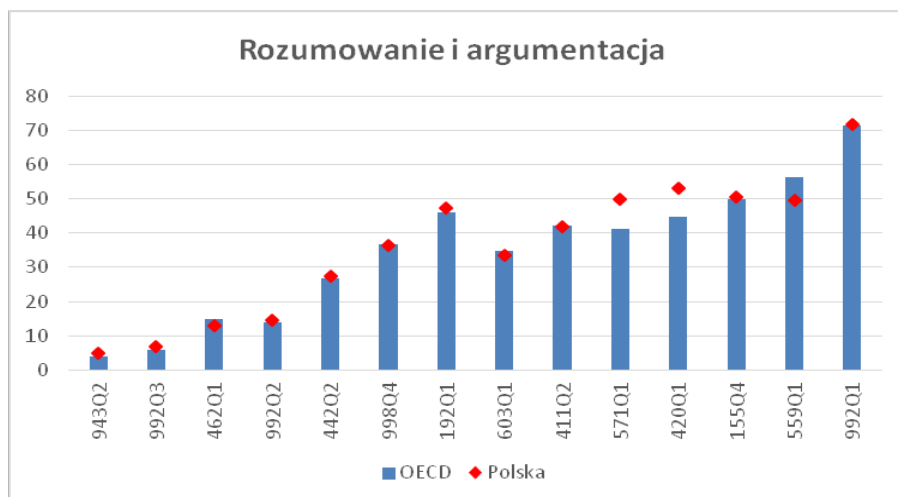
6.4. Rozumowanie matematyczne

Polski zespół ekspertów matematycznych, analizujący wyniki badania PISA, od roku 2003 przeprowadza szczegółowe analizy wyników uzyskanych przez polskich uczniów na poziomie poszczególnych zadań, a także grup zadań sprawdzających podobne umiejętności.

Jednym z wyznaczników sukcesu polskich gimnazjalistów w 2012 roku była istotna poprawa wyników w zadaniach wymagających rozumowania matematycznego (mierzących V. wymaganie ogólne podstawy programowej). Wtedy, po raz pierwszy od początku badania PISA, polskim uczniom udało się uzyskać dla prawie każdego z tych trudnych zadań wynik lepszy od średniej OECD.

Badanie PISA 2015 potwierdza trwałość tego zjawiska. Wykres 6.8. pokazuje porównanie wyników polskich uczniów ze średnim wynikiem OECD dla wszystkich zadań, które wymagały rozumowania matematycznego.

Wykres 6.8. Wyniki uczniów w Polsce i w krajach OECD w badaniu PISA 2015 w zadaniach wymagających rozumowania matematycznego (w procentach).



Po badaniu PISA 2015 przeprowadzono także wiele analiz mających na celu identyfikację powodów związanych z umiejętnościami uczniów, które mogłyby – choćby częściowo – uzasadnić obniżenie wyniku z matematyki z 518 punktów w roku 2012 do 504 punktów w roku 2015. W szczególności porównano, zadanie po zadaniu, wyniki uzyskane przez uczniów w obu edycjach badania. Niestety, mimo licznych prób, analiza treści zadań oraz ich wyników nie pozwoliła na postawienie wiarygodnej hipotezy dotyczącej umiejętności matematycznych, która znalazłaby systematyczne potwierdzenie w danych – nawet w grupach zadań mierzących podobne umiejętności niektóre zadania okazały się w roku 2015 łatwiejsze, a inne trudniejsze niż w roku 2012.

Przeprowadzono także analizę porównawczą, której celem było wyodrębnienie tych zadań, które w sposób istotny zmieniły swoją funkcję pomiarową w obrębie obu badań (przypomnijmy, że w części matematycznej badania uczniowie rozwiązywali te same zadania, co w roku 2012). W przekonujący sposób udało się wyodrębnić tylko jedno takie zadanie. Nie zdradzając jego treści można sformułować dość wiarygodną hipotezę tłumaczącą zmianę jego funkcjonowania – rozwiązując to zadanie w wersji papierowej uczeń mógł zmierzyć linijką pewne odległości, co mogło mu pomóc w rozwiązaniu. W wersji komputerowej taki sposób rozwiązania zadania był niemożliwy.

6.5. Wieloletni trend zmiany wyników

Raport międzynarodowy komentuje trend poziomu umiejętności matematycznych obserwowany w kolejnych cyklach badania od 2003 roku. Stwierdzono w nim, że średni wynik pomiaru umiejętności matematycznych w krajach OECD systematycznie obniża się, średnio o 1,7 punktu na każdy cykl badania PISA. W tym kontekście odnotowano, że polscy gimnazjaliści zachowują trend wznoszący, poprawiając swój wynik średnio o 5 punktów z cyklu na cykl. Wraz z Polską w tym kontekście wymieniona jest jeszcze tylko Brazylia oraz Tunezja, poprawiające swoje wyniki średnio odpowiednio o 6,2 oraz 3,8 punktu na cykl. Oba te kraje,

w przeciwieństwie do Polski, pozostają jednak ciągle w strefie wyników słabszych od średniej OECD (odpowiednio 377 oraz 367 punktów w 2015 roku).

6.6. Podsumowanie

W 2015 roku uczniowie rozwiązywali z matematyki te same zadania, co w roku 2012, ale po raz pierwszy pomiar umiejętności odbywał się przy pomocy komputerów. Być może w związku z tą zmianą w prawie wszystkich krajach, które w 2012 roku miały wyniki wyższe od średniej OECD, w roku 2015 nastąpił spadek średniego wyniku. Polska, która znajduje się w tej grupie krajów, także uzyskała niższy wynik – 504 punkty. Jednak pomimo mniejszej liczby uzyskanych punktów, wynik polskich gimnazjalistów jest wciąż bardzo dobry na tle innych krajów, także pod względem odsetka najlepszych i najgłabszych uczniów.

Zmiany, które zaszły w badaniu PISA 2015 w stosunku do poprzednich cykli badania, wprowadziły zapewne wiele czynników pozamatematycznych, a nawet pozainformatycznych, mogących mieć wpływ na wyniki uczniów. Dlatego należy zachować dużą rozwagę w wyciąganiu z uzyskanych wyników daleko idących wniosków na temat nauczania matematyki. W szczególności, zalecana jest duża ostrożność w wyciąganiu wniosku, że intensywna komputeryzacja nauczania matematyki jest prostą drogą do przyszłych sukcesów.

7. KORZYSTANIE Z NOWYCH TECHNOLOGII

Michał Sitek

Badanie PISA daje okazję przyjrzenia się roli nowych technologii w życiu nastolatków. Badanie PISA daje też unikalną możliwość zestawienia informacji o dostępie do nowych technologii i sposobach korzystania z nich z wynikami osiąganymi przez uczniów w testach PISA. W tej części raportu wykorzystamy też dane z badania kompetencji komputerowych i informacyjnych ICILS 2013, przeprowadzonego 2 lata wcześniej niż badanie PISA 2015 i obejmującego uczniów drugiej klasy gimnazjum, a więc populacji bardzo zbliżonej do tej, którą badano w PISA.

Wdrażanie nowych technologii w szkołach jest jednym z ważniejszych trendów w polityce edukacyjnej ostatnich dziesięcioleci. Brakuje jednak przekonujących dowodów naukowych na pozytywny wpływ zwiększania obecności nowych technologii w szkole na wyniki uczniów – badania wskazują raczej na złożoność tej relacji: aby technologie mogły być wykorzystywane z pożytkiem dla uczniów, spełnionych musi być szereg warunków: wiążących się między innymi z postawami i umiejętnościami nauczycieli, jak i dobrą jakością zasobów edukacyjnych (Henessy i London, 2013). Badania PISA 2012 wyraźnie pokazały, że wyposażanie szkół w komputery i inne urządzenia cyfrowe i częstsze korzystanie z nowych technologii na lekcji nie przynosi pozytywnych efektów, co dobitnie pokazuje opublikowany przez OECD raport dotyczący roli komputerów w szkole (OECD 2015). Podkreśla się też, że w wielu krajach i regionach osiągających dobre wyniki tj. Korea, Japonia, Finlandia, Polska czy Szanghaj komputery są wykorzystywane bardzo rzadko na lekcjach. W Polsce ważnym doświadczeniem był program „Cyfrowa szkoła”, który, jak pokazują dotychczasowe analizy, nie przełożył się znaczącą zmianę sposobu nauczania i nie przyczynił się też do poprawy umiejętności uczniów mierzonych na sprawdzianie po szkole podstawowej (Penszko 2013; Penszko i Zielonka 2015).

Do dokładnego zmierzenia efektów wprowadzania nowych technologii potrzebne byłyby dobrze zaprojektowane badania eksperymentalne i pogłębione badania jakościowe, które pomogłyby lepiej zrozumieć specyfikę korzystania z nowych technologii przez nastolatków i skuteczności korzystania z mediów elektronicznych w nauczaniu. Zaletą wyników międzynarodowych badań edukacyjnych jest możliwość pokazania ogólnego obrazu korzystania z nowych technologii i związków między korzystaniem z nich a wynikami osiąganymi przez uczniów. Dane też są pomocne w ocenie, w jakim stopniu polscy uczniowie różnią się od swoich rówieśników z innych krajów i jakie zmiany zaszły w tym zakresie w ostatnich latach.

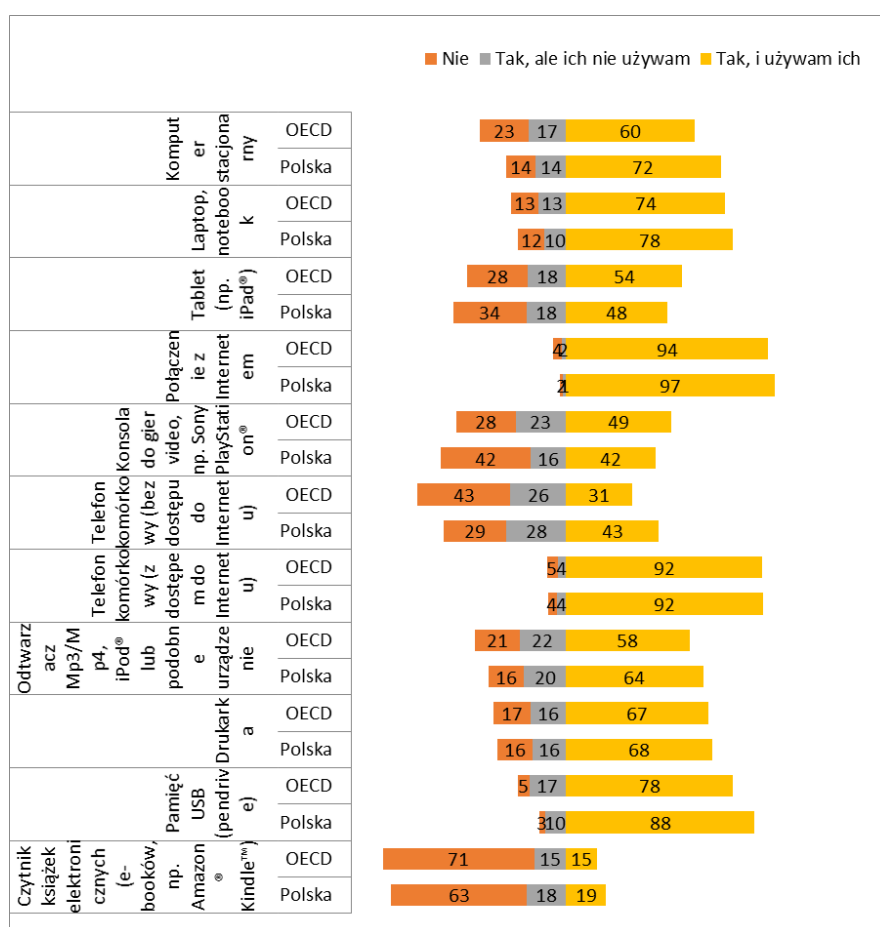
7.1. Dostępność urządzeń cyfrowych w domach uczniów

W ostatniej dekadzie Polska i inne kraje regionu nadrobiły dystans, jaki dzielił je od bogatszych krajów. Zdecydowana większość polskich nastolatków nie ma problemów z dostępem do mediów elektronicznych. Brak dostępu do Internetu w domu deklaruje zaledwie 1,6% 15-latków w porównaniu z ok. 5% w 2012 r. i 15% w 2009 r. Z innych badań wiemy, że pod względem

dostępu do mediów elektronicznych gospodarstwa domowe z dziećmi wyróżniają się na tle ogółu polskiego społeczeństwa, w którym dostęp do nowych technologii jest znacznie gorszy³⁰.

Nastolatki korzystają w domu z różnorodnych urządzeń cyfrowych. Między 2012 a 2015 zmniejszył się odsetek korzystających z komputerów stacjonarnych (z 79 do 72%) na rzecz komputerów przenośnych (wzrost odsetka korzystających z 63 do 78%), tabletów (z 13 do 48%) i telefonów z dostępem do Internetu (z 69 do 92%). Zazwyczaj uczniowie mają dostęp do kilku urządzeń: jedynie 4% piętnastolatków w Polsce deklaruje, że nie korzysta w domu z komputera stacjonarnego lub przenośnego. W porównaniu ze średnią OECD sytuacja w Polsce wypada korzystnie³¹. Polscy piętnastolatki mają nieznacznie lepszy dostęp do komputerów, smartfonów, Internetu czy czytników książek. Nieznacznie rzadziej od swoich rówieśników z innych krajów korzystają jedynie z tabletów i konsol do gier.

Wykres 7.1. Czy możesz korzystać w domu z wymienionych poniżej urządzeń?



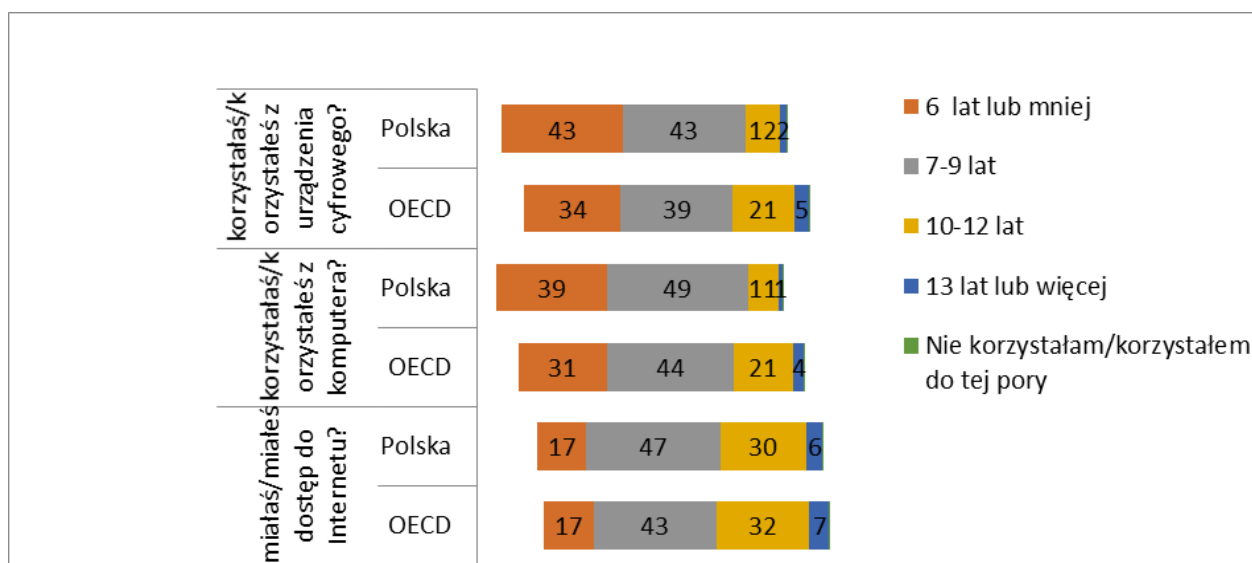
³⁰ Z corocznych badań Eurostatu wynika, że w 2015 r. wskaźnik dostępu do Internetu (isoc_cj_in_h) w gospodarstwach domowych z dziećmi wyniósł w Polsce 2015 r. 95%, podczas gdy średnia dla wszystkich gospodarstw wyniosła zaledwie 75%.

³¹ Część ankiety ucznia dotycząca korzystania z urządzeń cyfrowych nie była przeprowadzana we wszystkich krajach. Podawana w tym rozdziale średnia OECD opiera się na danych z 31 krajów i nie obejmuje uczniów z Kanady, Norwegii, Turcji i Stanów Zjednoczonych.

* Średnia OECD nie obejmuje danych z Kanady, Norwegii, Turcji i Stanów Zjednoczonych, w których nie przeprowadzono części ankiety dotyczącej korzystania z nowych technologii.

Polscy piętnastolatki mają też dłuższe doświadczenie w korzystaniu z urządzeń cyfrowych³² od swoich rówieśników z krajów OECD – mniej jest też uczniów, którzy zaczęli używać ich pod koniec szkoły podstawowej lub po rozpoczęciu nauki w gimnazjum. Jedynie w przypadku Internetu odpowiedzi polskich uczniów są zbliżone do średniej OECD. Pod względem „stażu” korzystania z urządzeń cyfrowych polscy uczniowie są w czołówce, tuż za krajami skandynawskimi i Estonią. Ma to znaczenie, ponieważ doświadczenie komputerowe przekłada się na umiejętności cyfrowe i informacyjne (Sijko 2014).

Wykres 7.2. Ile miałś/miałaś lat, kiedy po raz pierwszy...



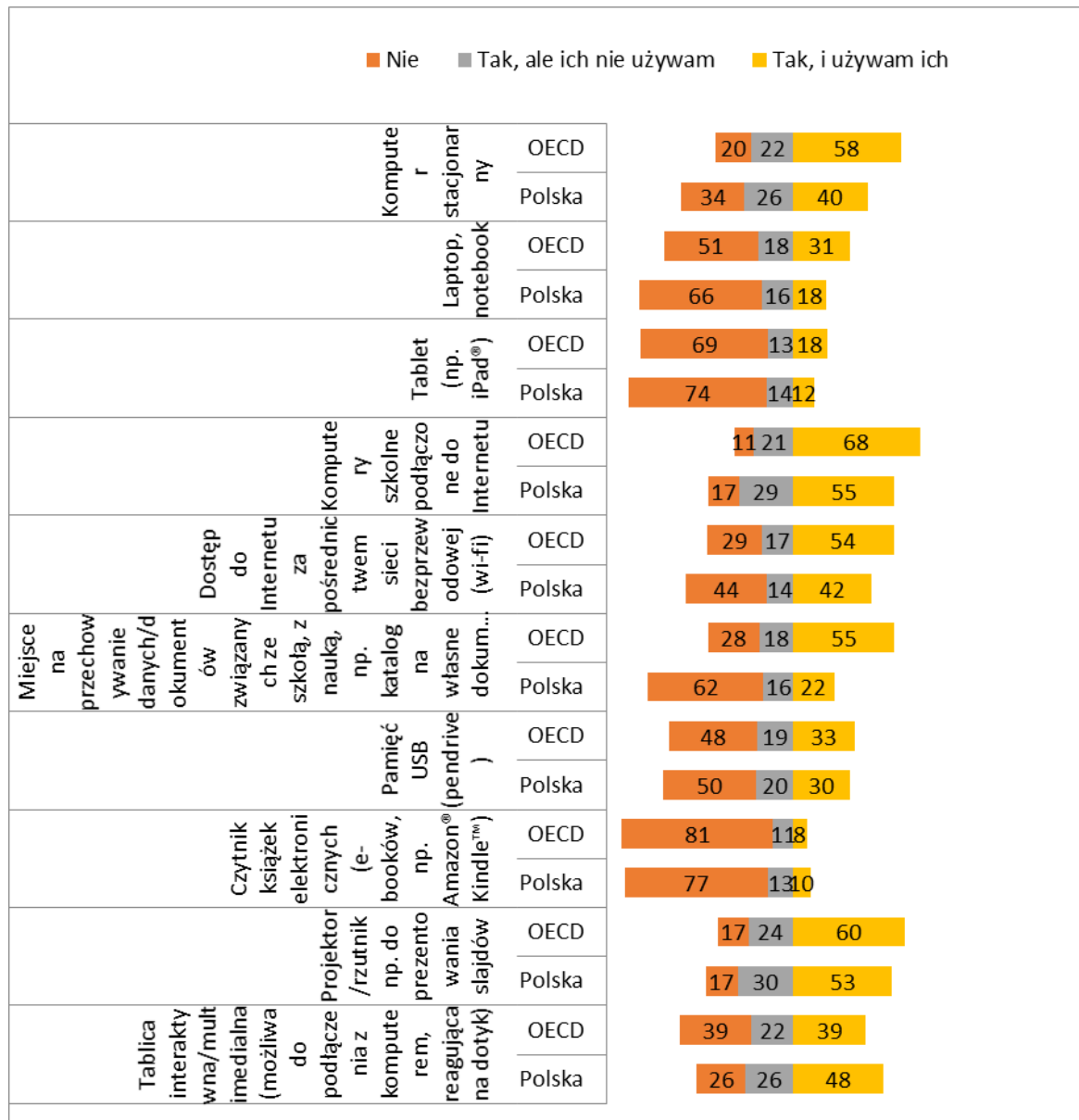
Nieźle wyposażenie w sprzęt komputerowy domów rodzinnych uczniów kontrastuje ze słabym wyposażeniem szkół. Nie jest to zaskoczeniem. Z prowadzonych w poprzednich latach badań wiemy, że polskie szkoły są gorzej wyposażone w urządzenia cyfrowe od szkół z innych krajów. Potwierdziło to także badanie PISA 2015. Według deklaracji dyrektorów szkół biorących udział w badaniu PISA na 1 gimnazjalistę w trzeciej klasie przypadało 0,46 komputera w porównaniu ze średnią OECD 0,76. W Czechach wskaźnik ten wyniósł 1,02, Słowacji 0,91, Węgrzech 0,75 (zob. OECD 2016, tabela II.6.4). W 2012 r. wskaźnik ten wyniósł dla Polski 0,36, a dla OECD 0,68 (OECD 2013, tabela iv.3.18).

W ankiecie wypełnianej przez uczniów pytano o dostęp do wybranych urządzeń i korzystanie z nich w szkole. W porównaniu ze średnią OECD mniejszy odsetek uczniów deklaruje, że korzysta w szkole z komputerów, gorszy jest też deklarowany dostęp do sieci Wi-Fi czy komputerów podłączonych do Internetu. Jedynie 22% polskich piętnastolatków ma dostęp do szkolnego dysku sieciowego, na którym mogliby przechowywać swoje dokumenty (w porównaniu ze średnią 55% dla OECD). Jedynymi wyjątkami są tablice interaktywne, z których polscy uczniowie korzystają częściej niż w wielu innych krajach, oraz czytniki książek, z których w szkole korzysta podobny odsetek uczniów co w OECD. Odpowiedzi uczniów

³² W ankiecie urządzenia cyfrowe zdefiniowano jako „komputery stacjonarne, laptopy, notebooki, smartfony, tablety, telefony komórkowe bez dostępu do Internetu, konsole do gier i telewizory z dostępem do Internetu”.

świadczą nie tylko o infrastrukturze szkół, ale też o polityce szkół w zakresie nowych technologii, np. możliwości korzystania ze szkolnego konta e-mail czy szkolnego serwera, umożliwianiu korzystania przez uczniów ze szkolnego wifi na własnych urządzeniach czy dostępności infrastruktury IT dla uczniów poza lekcjami (np. dostępności pracowni komputerowych, komputerów w szkolnych bibliotekach). Wiąże się to zapewne z tym, że polskie szkoły są zazwyczaj mniejsze od szkół w innych krajach i często nie mają własnej, rozbudowanej infrastruktury sieciowej, ani informatyków specjalizujących się w zarządzaniu siecią szkolną. Jak pokazało badanie ICILS, komputery w polskich szkołach są dostępne głównie w pracowniach komputerowych i bibliotekach, a nie w salach lekcyjnych (Sijko 2014).

Wykres 7.3. Czy możesz korzystać w szkole z wymienionych poniżej urządzeń?

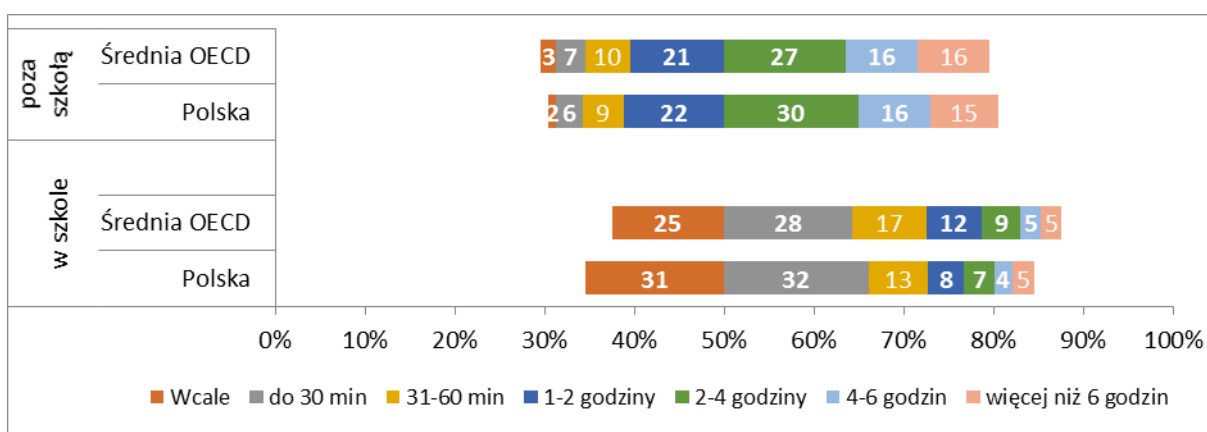


7.2. Korzystanie z TIK w domu i szkole

Młodzi Polacy są online przez znaczącą część dnia. W badaniu PISA pytano osobno o czas przebywania w Internecie w typowym dniu tygodnia w szkole i poza szkołą. Poza szkołą niemal co trzeci piętnastolatek jest online więcej niż 4 godziny, a do rzadkości należą osoby deklarujące, że nie korzystają z Internetu wcale lub korzystają krócej niż pół godziny. Relatywnie dobry dostęp do Internetu w domach i umiarkowane korzystanie z niego przez uczniów jest niewątpliwie pozytywnym wynikiem. W Polsce narasta natomiast problem nadużywania mediów elektronicznych. Odsetek nastolatków deklarujących korzystanie z Internetu poza szkołą w zwykły dzień tygodnia przez więcej niż 6 godzin wzrósł z 7,5% w 2012 r. do 15% w 2015 r. W czasie weekendu przebywanie w Internecie więcej niż 6 godzin deklaruje aż 23% uczniów (w 2012 r. 16%) Wzrost ten jest zgodny z trendami w innych krajach i stanowi ważne wyzwanie dla rodziców i nauczycieli, a także szkół i poradni psychologiczno-pedagogicznych, które muszą zmierzyć się z potrzebą diagnozy e-uzależnień i zwracać większą uwagę na kwestie edukacji medialnej.

Łączny bilans czasu korzystania z Internetu przez polskich nastolatków poprawia szkoła, co jest szczególnie widoczne, jeśli porównujemy polskie dane z danymi z innych krajów. Polscy uczniowie krócej od swoich rówieśników korzystają z Internetu w szkole. Aż 31% z nich twierdzi, że nie korzysta z Internetu w szkole wcale, a drugie tyle twierdzi, że w szkole korzysta z Internetu krócej niż pół godziny (w 2012 r. było to, odpowiednio, 50 i 30% uczniów). Nie widać przy tym różnic ze względu na wielkość miejscowości: odsetek uczniów deklarujących, że nie korzystają z Internetu w szkołach zlokalizowanych na wsi i małych miastach (31%) jest podobny do odsetka uczniów korzystających z Internetu w szkołach z miast liczących powyżej 100 tys. mieszkańców (28,5%). Korzystanie z Internetu jest częstsze w szkołach dających dostęp do szkolnego wifi, ale nie jest ważną barierą, ponieważ zdecydowana większość nastolatków ma możliwość korzystania z Internetu w smartfonie.

Wykres 7.4. Częstość korzystania z Internetu w szkole i poza szkołą (Jak długo korzystasz z Internetu w szkole/poza szkołą podczas typowego dnia tygodnia?)



Uczniów poproszono o wskazanie częstotliwości wykonywania wybranych czynności w domu. Wśród nich znalazły się: korzystanie z nowych technologii do komunikacji i kontaktów społecznych (Facebook, czat, poczta elektroniczna), korzystania z gier, celów użytkowych

(czytanie wiadomości, wyszukiwanie praktycznych informacji) oraz udostępniania treści (brzmienie pytań i szczegółowe rozkłady odpowiedzi przedstawiono w załączniku). Podobnie jak w 2012 r. najczęściej wykonywaną czynnością okazało się korzystanie z portali społecznościowych oraz oglądanie filmików w Internecie. Relatywnie często czytane są też informacje w Internecie oraz ściągane są filmy i muzyka. 15-latki rzadko grają w gry i korzystają z poczty elektronicznej. Niezbyt częste jest też tworzenie i udostępnianie treści. Widać też wyraźne różnice ze względu na płeć: chłopcy znacznie częściej grają w gry i deklarują częstsze wykonywanie większości czynności. Jedynie pod względem częstości korzystania z portali społecznościowych widać przewagę dziewcząt.

Uczniów zapytano też o to, jak często korzystają z urządzeń cyfrowych w związku z nauką w szkole. Pytano o częstotliwość komunikowania się z innymi uczniami i nauczycielami, przeszukiwanie Internetu w związku z nauką oraz odrabianie prac domowych na komputerze lub urządzeniu mobilnym. Polscy piętnastolatki dość często odrabiają też prace domowe na komputerze lub urządzeniu mobilnym. Rzadko komunikują się z nauczycielami i innymi uczniami za pośrednictwem poczty elektronicznej – częściej wykorzystują do tego portale społecznościowe. Według deklaracji uczniów tematy związane ze szkołą są często poruszane w komunikacji elektronicznej z rówieśnikami. Polscy uczniowie wyróżniają się na tle swoich rówieśników w OECD pod względem deklarowanej częstości korzystania z nowych technologii w związku ze szkołą. Częściej niż w Polsce korzysta się z nich tylko w Danii i Estonii. Kraje, w których uczniowie najrzadziej korzystają z nowych technologii w związku ze szkołą, to Japonia i Finlandia.

Według deklaracji uczniów w szkole urządzenia cyfrowe i media elektroniczne wykorzystywane jest bardzo rzadko – co nie powinno dziwić, biorąc od uwagę słabość szkolnej infrastruktury informatycznej. Poza tym pytania ankiety odnosiły się niemal wyłącznie do czynności, które uczniowie mogą wykonywać poza lekcjami, takich jak korzystanie z czatu, wysyłanie poczty elektronicznej czy odrabianie prac domowych. Odrabianie prac domowych na szkolnym komputerze deklaruje ok. 26% uczniów, a poszukiwanie w Internecie informacji potrzebnych do nauki ok. 2/3 uczniów. Zauważyć można tylko nieznaczny wzrost częstości wykonywania poszczególnych czynności w porównaniu z 2012. Polska, razem z Japonią, Koreą, Niemcami i Irlandią, należy do grupy krajów OECD, w których uczniowie korzystali w szkole z urządzeń cyfrowych najrzadziej. Kraje, w których uczniowie deklarują najczęstsze korzystanie z nowych technologii, to Dania, Australia, Szwecja i Holandia.

Odpowiedzi uzyskane w badaniu PISA 2015 warto zestawić z wynikami międzynarodowego badania kompetencji komputerowych i informacyjnych ICILS 2013. Badanie ICILS przeprowadzono w drugiej klasie gimnazjów i wprowadziło ono mniejszą liczbę krajów, to bardziej dokładnie prześledzono sposoby korzystania z nowych technologii przez uczniów (zob. Sijko 2013; Fraillon i in. 2014). Polscy gimnazjaliści wyróżniali się na tle uczniów innych krajów pod względem korzystania z aplikacji użytkowych takich jak edytory tekstu, oprogramowanie edukacyjne czy proste programy graficzne. Częściej od uczniów z innych krajów używali też komputerów w związku z nauką, korzystając z różnych źródeł informacji w Internecie, w tym Wikipedii, używając forów internetowych i komunikując się z innymi uczniami. Rzadziej natomiast korzystali w domu z testów komputerowych i współpracowali z innymi uczniami. Polscy gimnazjaliści rzadko korzystali z nowych technologii w szkole: przygotowując prezentacje, współpracując z innymi uczniami ze swojej szkoły lub z innych szkół czy rozwiązując zadania. W ankiecie ICILS pytano także o korzystanie z nowych technologii na lekcjach. Tu pozytywnie wyróżniały się jedynie lekcje informatyki – na lekcjach z polskiego,

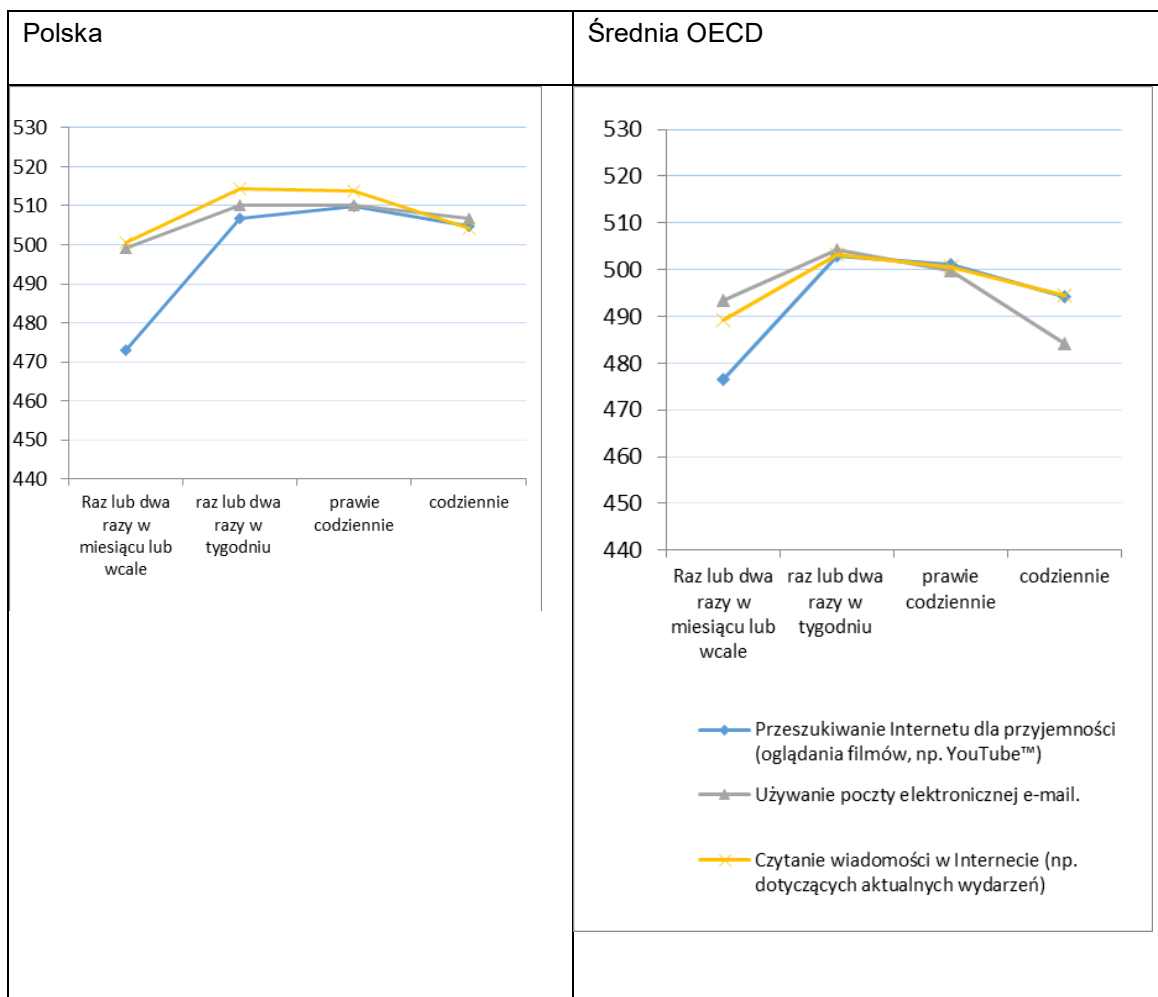
języka obcego, matematyki, historii, przedmiotów przyrodniczych czy zajęć artystycznych polscy gimnazjaliści rzadziej korzystali z nowych technologii od uczniów z innych krajów (słabiej wykorzystanie TIK w szkole prezentowało się jedynie w Chorwacji i Niemczech).

7.3. Związek korzystania z urządzeń cyfrowych w domu i szkole z wynikami osiąganymi przez uczniów

Podobnie jak to miało miejsce w 2009 i 2012 r. relacja między czasem spędzonym w Internecie a wynikami uczniów w testach PISA jest nieliniowa. Zarówno w rozumowaniu w naukach przyrodniczych, jak i w czytaniu i matematyce najslabsze wyniki charakteryzują uczniów, którzy spędzają w Internecie więcej niż 6 godzin, słabsze są także wyniki uczniów, którzy nie korzystają z Internetu lub korzystają z Internetu sporadycznie. W niewielkim stopniu te różnice można to wyjaśnić charakterystykami uczniów. Rzadkie korzystanie z Internetu w domu (do 30 minut dziennie) jest częstsze wśród dziewcząt, uczniów gimnazjów zlokalizowanych na wsi lub w małych miejscowościach, uczniów deklarujących słabsze wyposażenie domu rodzinnego w urządzenia cyfrowe. Pod tym względem wyróżniają się także dzieci osób z wyższym wykształceniem. Uczniowie deklarujący korzystanie z Internetu w domu przez 6 godzin i więcej mają zazwyczaj rodziców o niższym poziomie wykształcenia, mieszkają w dużym mieście i deklarują lepsze wyposażenie w urządzenia cyfrowe.

Zależność między częstością korzystania z urządzeń do konkretnych celów a wynikami PISA warto prześledzić na przykładzie konkretnych pytań. W przypadku korzystania z mediów elektronicznych dla rozrywki najlepsze wyniki osiągają zazwyczaj uczniowie deklarujący wykonywanie poszczególnych czynności od czasu do czasu (kilka razy w miesiącu lub kilka razy w tygodniu), najslabsze ci, którzy deklarują, że nie wykonują tych czynności wcale lub bardzo rzadko. Nie ma pod tym względem dużych różnic między czytaniem, matematyką i rozumowaniem w naukach przyrodniczych. Wyniki dla Polski nie różnią się znacząco od średniej OECD.

Wykres 7.5. Wyniki uczniów z rozumowania w naukach przyrodniczych w zależności od deklarowanej częstości wykonywania wybranych czynności dla przyjemności.



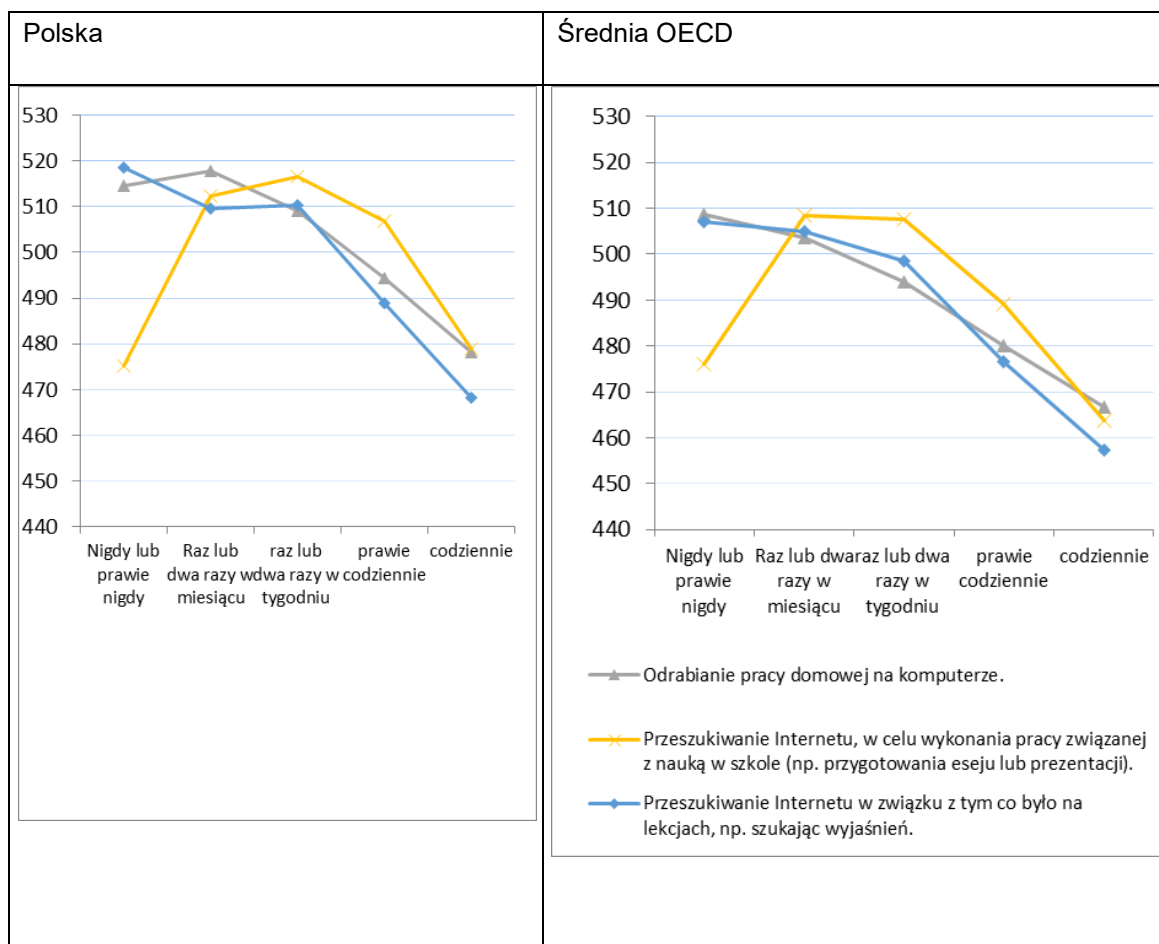
* Ze względu na niewielki odsetek uczniów wybierających odpowiedź „Wcale” połączono tą kategorię z kategorią („1-2 razy w miesiącu”). Oszacowania z modeli regresji, w których kontrolowano płeć i status społeczno-ekonomiczny uczniów (wskaźnik ESCS).

W 2012 r. dla niektórych czynności obserwowaliśmy wyraźną pozytywną zależność. Przykładem może być czytanie wiadomości w Internecie. W 2015 r. zależność nie jest już liniowa – uczniowie deklarujący, że czytają wiadomości codziennie mają nieco niższy poziom umiejętności niż czytający je od czasu do czasu. Inną obserwacją są różnice między płaciami. W przypadku grania w gry i czytania wiadomości w Internecie związek między częstością wykonywania danej czynności a wynikami wygląda nieco inaczej dla dziewcząt i chłopców, co może wiązać się z odmiennymi wyborami dotyczącymi gier i rodzaju czytanych informacji w Internecie.

Čyęstsze korzystanie z nowych technologii w związku ze szkołą zazwyczaj wiąże się z niższymi wynikami uczniów. Można było przypuszczać np., że częstsze przeglądanie Internetu w związku z tym, co było na lekcji, będzie wiązać się z lepszymi wynikami. Tak jednak nie jest: im więcej czasu uczniowie poświęcają na tę czynność, tym słabsze są ich wyniki. Ujemna korelacja między czasem spędzonym na naukę przy komputerze a wynikami uczniów widoczna jest nie tylko w danych ogólnopolskich, ale też dla większości szkół uczestniczących w badaniu.

Wy tłumaczeniem mogłaby być specyfika uczniów deklarujących częstsze wykonywanie tych czynności. Faktycznie, okazuje się, że są to nie tylko uczniowie osiągający słabsze wyniki, ale też mający rodziców z niższym poziomem wykształcenia i posiadający lepszy dostęp do TIK w domu. Jednak uwzględnienie różnic w statusie społeczno-ekonomicznym uczniów nie zmienia kształtu tej zależności (wykres xx).

Wykres 7.6. Wyniki uczniów z rozumowania w naukach przyrodniczych w zależności od deklarowanej częstości wykonywania wybranych czynności dla przyjemności.



* Oszacowania z modeli regresji, w których kontrolowano płeć i status społeczno-ekonomiczny uczniów (wskaźnik ESCS).

Do podobnych wniosków prowadzi analiza związku między umiejętnościami uczniów a korzystaniem z nowych technologii w szkole. Jeśli porównać uczniów deklarujących, że nie wykonują danych czynności w szkole wcale, z uczniami, którym się to zdarza, to w przypadku wszystkich czynności uczniowie korzystający z nowych technologii w szkole mieli niższe wyniki od uczniów niekorzystających i dotyczyło to wszystkich rodzajów działań wymienionych w ankiecie. Część czynności, takich np. jak umieszczanie własnych materiałów na stronie szkoły, uczenie się i utrwalanie wiedzy z wykorzystaniem szkolnych urządzeń, odrabianie prac domowych na szkolnym komputerze czy używanie szkolnych komputerów do porozumiewania się było częstsze w gimnazjach zlokalizowanych na wsi. Ale uwzględnienie różnic między uczniami ze względu na płeć, lokalizację szkoły czy wykształcenia rodziców pokazuje, że uczniowie deklarujących korzystanie z urządzeń cyfrowych w szkole uzyskują gorsze wyniki niż uczniowie niewykonyjący czynności, o które pytano w ankiecie.

Można twierdzić, że, korzystając z nowych technologii, uczniowie rozwijają inne umiejętności niż te, które mierzy badanie PISA. Ale analizy wyników badania ICILS pokazują, że częstotliwość, z jaką jej uczniowie korzystają z komputerów, nie wykazuje prostego liniowego związku z umiejętnościami informacyjnymi i komputerowymi: w Polsce jest on wyraźny jedynie w przypadku korzystania z nowych technologii dla rozrywki i korzystania z mediów społecznościowych, a w przypadku korzystania z komputerów związanego ze szkołą najlepsze kompetencje komputerowe i informacyjne mają uczniowie korzystający z komputerów od czasu do czasu. Wyniki te sugerują, że zależność korzystania z nowych technologii a umiejętnościami jest dość uniwersalna, bez względu na rodzaj mierzonej umiejętności. Pokazuje to, że w nabywaniu kompetencji ważna jest nie tyle ilość czasu korzystania z urządzeń cyfrowych i mediów elektronicznych. Tym, czego nie uwzględniają pytania zadane w badaniu, są też różnice, które mogą występować w ramach poszczególnych czynności. Dobrym przykładem są gry komputerowe, które są bardzo zróżnicowane i korzystanie z nich może być w różny sposób powiązane z umiejętnościami uczniów.

7.4. Wnioski

Polscy uczniowie mają dobry dostęp do urządzeń cyfrowych w domu i często z nich korzystają. W praktyce edukacyjnej warto zwrócić większą uwagę na dwie grupy uczniów. Pierwszą grupę stanowią uczniowie, którzy mają słabszy dostęp do urządzeń cyfrowych i korzystają z nich bardzo rzadko. W Polsce jest to nieliczna grupa, co oznacza, że stosunkowo łatwo można byłoby zaadresować do niej odpowiednie wsparcie w poprawie dostępu do urządzeń cyfrowych. Drugą, i jak się wydaje bardziej wymagającą wsparcia grupą, są uczniowie korzystający z urządzeń cyfrowych bardzo często. W tej grupie są uczniowie osiągający zazwyczaj słabsze wyniki, mający już wcześniej problemy w szkole. Są to zazwyczaj uczniowie mających gorzej wykształconych rodziców i mieszkający w dużych miastach. Warto uczyć rodziców na czas i sposób spędzany przez dzieci przed komputerem i rozwijać różne formy pomocy dla uczniów z ryzykiem e-uzależnień.

Polska szkoła nie stwarza wielu możliwości korzystania z nowych technologii i nie ułatwia uczniom poruszania się w cyfrowym świecie. Tym, co odróżnia doświadczenie polskich uczniów od uczniów z innych krajów, jest relatywnie częste wykorzystywanie przez uczniów nowych technologii w domu i bardzo rzadkie korzystanie z nich w szkole. Wydawać by się mogło, że jest to idealny schemat „odwróconej klasy”. Niestety, ten brak równowagi zdaje się nie sprzyjać rozwojowi umiejętności cyfrowych uczniów. Wydaje się jednak, że aktywności uczniów są mało ukierunkowane, dostępne dla uczniów zasoby są słabej jakości, a nauka w szkole nie uczy efektywnego korzystania z mediów. Posiadanie dostępu do nowych technologii i często jej wykorzystanie przez młodzież nie jest równoznaczne z posiadaniem przez nich kompetencji komputerowych i informacyjnych. Badanie ICILS pokazało, że spora część uczniów posiada słabe kompetencje komputerowe i informacyjne. Badanie PISA 2012 pokazało z kolei, że polscy uczniowie gorzej od uczniów z innych krajów radzą sobie z zadaniami mierzącymi umiejętność czytania tekstów elektronicznych, w mniejszym stopniu wykorzystują też w zadaniach umiejętność nawigacyjne. Kształcenie tego rodzaju umiejętności jest przede wszystkim wyzwaniem dla edukacji polonistycznej.

Badania PISA i wiele innych pokazują, że nowe technologie nie są prostą receptą na poprawianie osiągnięć uczniów. Ważne jest oczywiście zapewnienie podstawowych warunków wygodnego wykorzystywania ich w szkołach, takich np. jak dostęp do szerokopasmowego

Internetu czy komputerów i mobilnych urządzeń, które można wykorzystywać nie tylko na lekcjach informatyki, ale też na lekcjach z innych przedmiotów. Dużo większe znaczenie ma poprawa umiejętności cyfrowych i edukacja medialna, która pomogłaby uczniom bardziej produktywnie i świadomie korzystać z nowych technologii.

Ujemna korelacja między częstością korzystania z nowych technologii w związku z nauką nie oznacza oczywiście, że w szkole powinno się unikać korzystania z komputerów, tablic interaktywnych czy smartfonów uczniów. Umiejętności cyfrowe uczniów mają znaczenie dla przyszłego funkcjonowania absolwentów szkół na studiach, rynku pracy czy w życiu społecznym. Komputery i technologie informacyjno-komunikacyjne radykalnie zmieniają zawartość pracy w większości zawodów, a możliwości, jakie stwarzają, są podstawą nowych zawodów, w których zapewne pracować będzie w przyszłości znacząca część absolwentów polskich szkół. Brak podstawowych umiejętności cyfrowych jest i będzie w coraz większym stopniu przeszkodą w dostępie do informacji, funkcjonowaniu w kulturze i komunikowaniu się z innymi, a nawet korzystaniu z usług publicznych. Nowe technologie odgrywają też ważną rolę w codziennym życiu nastolatków, co stwarza zarówno szanse, jak i zagrożenia dla rozwoju ich kompetencji. Jednak sposób wprowadzania nowych technologii do szkół nie może ulegać automatyzmowi. Niejednoznaczne wyniki badań PISA i ICILS dotyczące wpływu korzystania z mediów elektronicznych w szkole i w związku ze szkołą na wyniki uczniów wskazują na to, że ważniejsze od czasu przeznaczanego na korzystanie z tych technologii jest jakość zasobów edukacyjnych oraz umiejętne ich wykorzystanie.

Bibliografia

Frailon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age*. Springer-Verlag GmbH.

Henessy, S., London, L. (2013). Learning from International Experiences with Interactive Whiteboards: The Role of Professional Development in Integrating the Technology. OECD Education Working Papers, No. 89 (polski przekład opublikowany przez Instytut Badań Edukacyjnych: <http://eduentuzjasci.pl/images/stories/publikacje/ibe-oecd-wnioski-z-miedzynarodowych-doswiadczen-tablice-interaktywne.pdf>)

OECD (2015). Students, computers and learning. Making the connection. Paris: Organisation of Economic Co-operation and Development.

OECD(2016). PISA 2015 Results. Policies and Practices for Successful Schools. Tom II. Paris: Organisation of Economic Co-operation and Development.

Penszko P. (red.) (2013). Ewaluacja ex-post rządowego programu rozwijania kompetencji uczniów i nauczycieli w zakresie stosowania technologii informacyjno-komunikacyjnych – „Cyfrowa szkoła”. Instytut Badań Edukacyjnych.

Penszko, P., Zielonka P. (2015). Analiza wpływu programu “cyfrowa szkoła” na wyniki sprawdzianu szóstoklasisty. Analizy IBE 3/2015.

Sijko, K. (red.) (2014). Kompetencje komputerowe i informacyjne młodzieży w Polsce. Raport z międzynarodowego badania kompetencji komputerowych i informacyjnych ICILS 2013, Warszawa: Instytut Badań Edukacyjnych.

Aneks

Tabela 7.1. Częstość wykonywania czynności dla rozrywki (Jak często korzystasz z urządzeń cyfrowych do wykonywania następujących czynności poza szkołą?)

	Nigdy lub prawie nigdy	Raz lub dwa razy w miesiącu	raz lub dwa razy w tygodniu	prawie codziennie	codziennie	Razem
Korzystanie z portali społecznościowych (np. Facebook).	5.6	3.6	7.2	20.2	63.5	100
Przeszukiwanie Internetu dla przyjemności (oglądania filmów, np. YouTube™)	3.4	5.1	14.1	27.9	49.4	100
Czytanie wiadomości w Internecie (np. dotyczących aktualnych wydarzeń).	10	15.8	24.4	25.4	24.3	100
Czatowanie (np. Gadu-Gadu).	36.6	10.7	12.2	16.8	23.7	100
Uzyskiwanie praktycznych informacji z Internetu (np. miejsca i daty wydarzeń).	9.9	15.1	27.7	26.1	21.1	100
Ściąganie muzyki, filmów, gier czy programów z Internetu.	9.7	20.5	26.9	22.7	20.3	100
Granie on-line w gry z udziałem kilku graczy.	44.4	12.2	12.3	14.4	16.6	100
Pobieranie na urządzenie mobilne nowych aplikacji.	15.4	33.8	24	13.7	13.1	100
Używanie poczty elektronicznej e-mail.	15.8	30.7	27.9	14	11.6	100
Granie w jednoosobowe gry.	37.6	18.2	19.5	15.1	9.6	100
Dodawanie i udostępnianie stworzonych przez siebie treści (np. muzyki, wierszy, filmów, programów komputerowych).	50.3	19.1	12.7	9.2	8.6	100
Granie w gry online w portalach społecznościowych	63.2	12.9	9.2	6.7	8	100

Tabela 7.2. Częstość wykonywania w domu czynności związanych ze szkołą (Jak często korzystasz z urządzeń cyfrowych do wykonywania następujących czynności poza szkołą?)

	Nigdy lub prawie nigdy	Raz lub dwa razy w miesiącu	raz lub dwa razy w tygodniu	prawie codziennie	codziennie	Razem
Korzystanie z portali społecznościowych do porozumiewania się z innymi uczniami na temat nauki szkolnej (np. Facebook).	9.8	12.7	23.1	25.7	28.7	100
Odrabianie pracy domowej na komputerze.	16.8	22	28.7	20.4	12.1	100
Sprawdzanie ogłoszeń na stronie internetowej szkoły np. nieobecności nauczycieli.	37.9	17.7	18.4	14.8	11.3	100
Odrabianie pracy domowej na urządzeniu mobilnym.	28.3	19.5	23.5	17.6	11.3	100
Przeszukiwanie Internetu, w celu wykonania pracy związanej z nauką w szkole (np. przygotowania eseju lub prezentacji).	11.7	23.6	32.1	21.4	11.2	100
Pobieranie i wysyłanie lub przeglądanie materiałów ze strony internetowej swojej szkoły (np. planu lekcji czy materiałów do zajęć lekcyjnych).	31.5	28.1	19.8	11.9	8.6	100
Korzystanie z portali społecznościowych do porozumiewania się z nauczycielami (np. Facebook).	63.2	11	9.5	7.9	8.4	100
Przeszukiwanie Internetu w związku z tym co było na lekcjach, np. szukając wyjaśnień.	17.1	26.6	31.6	16.9	7.8	100
Pobieranie na urządzenie mobilne aplikacji edukacyjnych.	52.1	21.2	13.9	7.6	5.3	100
Korzystanie z e-maila do porozumiewania się z innymi uczniami w związku z nauką.	47.6	20.2	17.9	9	5.3	100
Pobieranie na urządzenie mobilne aplikacji edukacyjnych o tematyce nauk przyrodniczych.	59.7	16.4	12	7	4.9	100
Korzystanie z e-maila do porozumiewania się z nauczycielami i przesyłanie im prac domowych lub innych materiałów związanych z nauką.	51.5	24.8	12.3	7.1	4.3	100

Tabela 7.3. Częstość deklarowanych czynności w szkole (Jak często korzystasz z urządzeń cyfrowych do wykonywania następujących czynności w szkole?)

	Nigdy prawie nigdy	lub Raz dwa razy w miesiącu	lub raz dwa razy w tygodniu	lub prawie codziennie	codziennie	
Czatowanie w szkole (np. Gadu-Gadu).	63.1	8.8	9.5	8.9	9.6	100
Używanie poczty elektronicznej (e-mail) w szkole.	66.6	16.2	9.4	4.1	3.7	100
Poszukiwanie w Internecie informacji potrzebnych do nauki.	32.7	22.8	25	13	6.5	100
Przesyłanie, ściągnięcie lub przeglądanie materiałów ze szkolnej strony internetowej (np. intranet).	62.1	15	11.6	6.2	5.1	100
Umieszczanie własnych materiałów na stronie internetowej szkoły.	80.3	7.8	6	3	2.8	100
Przeprowadzanie symulacji komputerowych w szkole.	78.3	9.3	6.4	3.2	2.8	100
Uczenie się i utrwalanie wiedzy, np. z języków obcych lub matematyki.	53.1	17.7	16.2	8.1	4.9	100
Odrabianie prac domowych na szkolnym komputerze.	74.4	11	7.2	4	3.3	100
Do pracy w grupie i porozumiewania się z innymi uczniami używając szkolnych komputerów.	66.8	15.5	9.7	4.2	3.9	100