

TIMSS 2015

**Wyniki międzynarodowego
badania osiągnięć czwartoklasistów
w matematyce i przyrodzie**

International Association for the Evaluation
of Educational Achievement

Instytut Badań Edukacyjnych

TIMSS 2015

**Wyniki międzynarodowego
badania osiągnięć czwartoklasistów
w matematyce i przyrodzie**

Redakcja

Krzysztof Konarzewski i Krzysztof Bulkowski

Warszawa 2016

Międzynarodowy zespół TIMSS 2015

International Association for the Evaluation of Educational Achievement, Amsterdam
Anne-Berit Kavli (przewodnicząca)

TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College
Ina V. S. Mullis, Michael O. Martin (dyrektorzy)

IEA Data Processing and Research Center, Hamburg
Heiko Sibberns (dyrektor)

Statistics Canada, Ottawa
Marc Joncas (starszy metodolog)

Educational Testing Service, Princeton
Matthias von Davier (główny specjalista ds. badań)

Krajowy zespół TIMSS 2015

Instytut Badań Edukacyjnych, Warszawa
Krzysztof Bulkowski (kierownik przetwarzania danych)
Aleksandra Chlebowicz
Alina Cichomska (kierowniczką Zespołu Współpracowników Terenowych IBE)
Monika Czajkowska
Wojciech Grajkowski
Jerzy Janowicz
Marcin Karpiński
Joanna Kaźmierczak (kierowniczką zespołu)
Krzysztof Konarzewski (koordynator krajowy)
Jacek Lech
Dominik Marszał
Margarzeta Orzechowska
Barbara Ostrowska
Agnieszka Sułowska
Małgorzata Zambrowska

Badanie TIMSS 2015 w Polsce sfinansowało Ministerstwo Edukacji Narodowej.

© 2016 Instytut Badań Edukacyjnych
Przedruk w całości lub w części wyłącznie za zgodą Instytutu Badań Edukacyjnych.
Cytowanie oraz wykorzystywanie danych jedynie z podaniem źródła.

Recenzent: dr hab. Jacek Haman
Projekt okładki: Beata Czapska

ISBN 978-83-65115-86-7

Wydawca: Instytut Badań Edukacyjnych
ul. Górczewska 8, 00-180 Warszawa
tel. (+48 22) 241 71 00
fax (+48 22) 241 71 11
www.ibe.edu.pl
ibe@ibe.edu.pl

Druk i oprawa: ViW Studio
ul. Szkolna 3, 05-530 Dobiesz

Podziękowania

Krajowy zespół TIMSS 2015 składa podziękowania dyrektorom szkół wylosowanych do badania, nauczycielom badanych oddziałów, rodzicom uczniów oraz uczniom biorącym udział w badaniu za owocną współpracę, która umożliwiła zebranie wiarygodnych informacji o procesie i wynikach edukacji w Polsce.

Niniejszy raport opiera się na wewnętrznych raportach TIMSS & PIRLS International Study Center:

TIMSS 2015 International Results in Mathematics

TIMSS 2015 International Results in Science

oraz na bazie danych z polskiej części badania TIMSS 2015.

Odpowiedzialność za opinie wyrażone w tym raporcie spoczywa wyłącznie na jego autorach.

Spis treści

NAJWAŻNIEJSZE FAKTY.....	8
Rozdział 1. PODSTAWOWE INFORMACJE O BADANIU	
<i>Krzysztof Bulkowski i Joanna Kaźmierczak</i>	11
Polski kontekst badania.....	11
Organizatorzy i wykonawcy badania	13
Osoby badane	13
Przebieg badania	18
Narzędzia badawcze	18
Rozdział 2. OSIĄGNIĘCIA MATEMATYCZNE	
<i>Krzysztof Konarzewski</i>	27
Średnie i rangi	27
Miary zróżnicowania.....	28
Osiągnięcia szczegółowe	30
Wyniki pomiaru a program kształcenia	34
Wyniki pomiaru a płeć i wiek uczniów	36
Zmiany osiągnięć matematycznych	38
Rozdział 3. OSIĄGNIĘCIA PRZYRODNICZE	
<i>Krzysztof Konarzewski</i>	41
Osiągnięcia szczegółowe.....	43
Wyniki pomiaru a program kształcenia	43
Wyniki pomiaru a płeć uczniów.....	48
Zmiany osiągnięć przyrodniczych.....	49
Rozdział 4. UCZEŃ I JEGO RODZINA	
<i>Krzysztof Konarzewski</i>	51
Uczeń i jego rodzina	51
Status socjoekonomiczny rodziny	51
Wiedza początkowa	52
Postawy dziecka wobec szkoły i uczenia się.....	55
Rozdział 5. SZKOŁA I PROCES KSZTAŁCENIA	
<i>Krzysztof Konarzewski</i>	59
Wielkość oddziałów klasowych	60
Zróżnicowanie szkół i oddziałów klasowych.....	61
Zasoby materialne szkoły	62
Zasoby kadrowe szkoły	66
Kultura szkoły.....	66
Czas nauczania.....	68
Metodyka nauczania.....	68
Rozdział 6. TEST TIMSS z BLISKA	
<i>Marcin Karpiński i Dominik Marszał</i>	77
Klucz punktowania odpowiedzi	73
Wyniki zbiorów zadań Zeszytu 2.....	77
Komentarz dydaktyczny	77
Zeszyt 2	85
LITERATURA	119

Najważniejsze fakty

TIMSS to międzynarodowy program mierzenia osiągnięć uczniów w matematyce i przyrodoznawstwie. Program wystartował w 1995 r. i powtarza pomiary co cztery lata. Pomiary obejmują dwie populacje uczniów: dziesięcioletnich w czwartym roku nauki i czternastoletnich w ósmym.

Polska po raz pierwszy wzięła udział w badaniu TIMSS w 2011 r. Pomiarowi poddaliśmy wtedy osiągnięcia uczniów w wieku średnio 9,9 roku pod koniec klasy trzeciej. W 2015 r. byli to uczniowie w wieku średnio 10,7 roku kończący klasę czwartą. Zmianę klasy podyktowała reforma wieku obowiązkowego: po jej wprowadzeniu trzecioklasiści byłiby za młodzi. Ponieważ uczniowie badani w 2015 r. byli starsi i kształceni o rok dłużej, porównanie średnich wskaźników osiągnięć w latach 2011 i 2015 nie dostarcza informacji o zmianach efektywności kształcenia w Polsce.

W 2015 r. badaniem w Polsce objęto 4747 uczniów z 254 oddziałów w 150 szkołach podstawowych. Stanowili oni reprezentatywną próbę populacji czwartoklasistów. Przeprowadzono też ankiety wśród rodziców badanych uczniów, dyrektorów szkół i nauczycieli matematyki i przyrody.

Osiągnięcia polskich uczniów

Średnia osiągnięć matematycznych polskich dzieci wynosi 535 punktów i jest istotnie wyższa od międzynarodowej średniej wzorcowej. Wśród 49 krajów Polska zajmuje 17. miejsce razem z Finlandią i Litwą.

Średnia osiągnięć przyrodniczych polskich dzieci wynosi 547 punktów i jest istotnie wyższa od międzynarodowej średniej wzorcowej. Wśród 47 krajów Polska zajmuje dziewiąte miejsce.

Co polscy uczniowie potrafią zrobić ze zdobytą wiedzą? W matematyce najsłabiej wypadają w zadaniach wymagających odtwarzania wiadomości, najlepiej zaś – w rozwiązywaniu problemów. W rankingu zastosowań problemowych Polska znalazła się na 11. miejscu – wyprzedziło nas tylko dziewięć krajów ze ścisłej czołówki oraz (nieistot-

nie) Dania. Podobne zjawisko zaobserwowano u trzecioklasistów w 2011 r.

W wiedzy przyrodniczej polskich uczniów obserwuje się niewielki niedobór względny wiadomości i zastosowań problemowych przy niewielkiej nadwyżce zastosowań typowych.

Postawy wobec szkoły i nauki

W rankingu postawy wobec szkoły polscy uczniowie znaleźli się na 46. miejscu – tylko o trzy miejsca wyższym od japońskich, ostatnich na liście. W Europie są na samym końcu wraz z Francuzami i Czechami. Najbardziej lubią szkołę dzieci w Indonezji.

Porównanie wyników badania TIMSS w latach 2011 i 2015 ujawnia, że masowej poprawie osiągnięć szkolnych towarzyszy równie masowe pogorszenie się postaw wobec uczenia się matematyki i przyrody. Na przykład w 2011 r. nasi uczniowie znaleźli się na 34. miejscu (za wszystkimi krajami europejskimi) pod względem obiektywnie zmierzonej wiedzy matematycznej, a jednocześnie zajęli pierwsze miejsce w samoocenie swojej wiedzy z tego przedmiotu. W 2015 r. polscy uczniowie przesunęli się w górę o 17 miejsc w rankingu wiedzy, a zarazem spadli o 32 miejsca w rankingu samooceny.

Dziesięcioletni uczniowie czują się najbezpieczniej w szkołach Kazachstanu, a najmniej bezpiecznie w Republice Południowej Afryki. W rankingu krajów według średnich ocen bezpieczeństwa wystawionych przez uczniów polska szkoła sytuje się na wysokiej, siódmej pozycji. Taką samą pozycję zajmowała w 2011 r., gdy badaliśmy trzecioklasistów.

Warunki kształcenia w Polsce

W przeciętnym oddziale klasy czwartej jest 18–19 uczniów. Najmniejsze oddziały prowadzą szkoły na wsi (średnio 17-osobowe), a największe – w miastach średniej wielkości (21-osobowe). Na jednego nauczyciela przypada u nas średnio 10 uczniów – mniej tylko w Omanie, Gruzji, Kuwejcie i Norwe-

gii. Z większą liczbą uczniów niż u nas pracuje nauczyciel na Litwie (13) i Słowacji (15), w Czechach (19) i w Rosji (20). Najwięcej uczniów na jednego nauczyciela (32) przypada w szkołach Republiki Południowej Afryki.

W porównaniu z 2011 r. zmalały różnice średnich osiągnięć uczniów między szkołami i oddziałami klasowymi. Być może jest to skutek wzrostu standaryzacji kształcenia za sprawą nowej podstawy programowej z 2008 r.

W 16 proc. szkół prowadzących przynajmniej dwa oddziały klasy czwartej stwierdzono duże różnice między oddziałami pod względem pochodzenia społecznego uczniów (SES). Może to świadczyć o polityce celowego dobierania uczniów do oddziałów klasowych. Średnie osiągnięć w tych szkołach nie odbiegają od średnich w pozostałych szkołach.

Nauczyciele

Kobiety stanowią 92 proc. nauczycieli w klasie czwartej.

Na świecie nauczyciele z pełnym wykształceniem uniwersyteckim II stopnia (odpowiednikiem naszego magisterium) nauczają matematyki 26 proc. uczniów, a przyrody – 28 proc. W Polsce, odpowiednio, 97 i 100 proc. Więcej magistrów jest tylko w szkołach na Słowacji.

Pod względem odczuwanego obciążenia zawodowego polscy nauczyciele znaleźli się na przedostatnim miejscu. Łżej pracuje się tylko ich kolegom w Gruzji.

Średni poziom satysfakcji zawodowej jest, niezależnie od nauczanego przedmiotu, najwyższy w Iranie i Katarze, najniższy zaś w Japonii i Francji. Polscy nauczyciele zajmują w tym rankingu zaskakująco niskie miejsca: matematycy (wraz z matematykami z Danii) 46., a przyrodniczy 44. W analogicznym rankingu przed czterema laty polscy nauczyciele edukacji wczesnoszkolnej zajęli ósme miejsce.

Skutki przejścia do drugiego etapu kształcenia

Klasę trzecią oddziela od czwartej wyraźny próg. Nie brak rodziców, nauczycieli i badaczy, którzy twierdzą, że spowalnia on postępy dzieci w nauce. Porównanie wyników badania z lat 2011 i 2015 pozwala zweryfikować ten sąd.

Wiedza czwartoklasistów z matematyki i przyrody jest znacznie większa niż dawnych trzecioklasistów w tym samym wieku. To efekt dodatkowego roku nauki w bardziej zaawansowanym trybie.

W klasie trzeciej chłopcy istotnie wyprzedzali dziewczynki w matematyce i wiedzy o przyrodzie. W klasie czwartej obie różnice zniknęły.

Czwartoklasiści czują się równie bezpiecznie w swoich oddziałach jak dawni trzecioklasiści.

Skutki reformy wieku obowiązku szkolnego

Badani przez nas czwartoklasiści zaczęli edukację w roku szkolnym 2011/2012. Według ustawy z 2009 r. obniżającej wiek obowiązku szkolnego z siedmiu do sześciu lat był to ostatni rok okresu przejściowego, w którym o posłaniu sześciolatka do klasy pierwszej mogli decydować rodzice. Pozytywną decyzję podjęło wtedy 19 proc. rodziców sześciolatków.

Dzieci, które rozpoczęły naukę w wieku sześciu lat, mają po czterech latach edukacji takie same osiągnięcia w matematyce jak dzieci, które poszły do szkoły jako siedmiolatki, ale istotnie niższe w przyrodoznawstwie.

Rodzice sześciolatków niżej niż rodzice siedmiolatków oceniali umiejętności, z jakimi ich dzieci rozpoczynały naukę w klasie pierwszej. Można stąd wnosić, że postawa rodziców wobec obniżenia wieku obowiązku szkolnego była nie mniej ważną przesłanką decyzji o posłaniu sześciolatka do szkoły niż ich ocena gotowości szkolnej dziecka. Pozytywną decyzję częściej podejmowali rodzice z wyższych warstw społecznych.

Podstawowe informacje o TIMSS 2015

Krzysztof Bulkowski i Joanna Kaźmierczak

TIMSS to akronim powstały z nazwy Trends in International Mathematics and Science Study, którą można rozwinąć jako Tendencje w międzynarodowym badaniu osiągnięć w matematyce i przyrodzie. Nazwa oznacza międzynarodowy program cyklicznego (co cztery lata) pomiaru osiągnięć szkolnych uczniów dziesięcio- i czternastoletnich. Polska uczestniczy jedynie w pomiarze osiągnięć dziesięciolatków.

Programowi przyświecają trzy cele:

- dostarczenie rzetelnych i porównywalnych danych o osiągnięciach szkolnych w populacji uczniów z każdego kraju
- dostarczenie rzetelnej informacji o dynamice, czyli zmianach poziomu osiągnięć w każdym kraju, który uczestniczył w badaniu więcej niż raz
- dostarczenie danych „kontekstowych”, czyli informacji o warunkach, w których przebiega nauczanie i uczenie się w każdym kraju.

Program wystartował w 1995 r. Polska po raz pierwszy wzięła w nim udział w 2011 r. (Konarzewski, 2012). Edycja 2015, której wyniki są przedmiotem niniejszego raportu, była w Polsce wyjątkowa, ponieważ przypadła w środku reformy wieku obowiązku szkolnego. Należy pokrótce ją przypomnieć.

Polski kontekst badania

Obowiązek szkolny dla dzieci sześciolatków wprowadziła nowelizacja ustawy o systemie oświaty z dnia

19 marca 2009 r. Obowiązek miał wejść w życie od 1 września 2012 r. Wcześniej, w okresie przejściowym (lata 2009–2011), o rozpoczęciu nauki przez sześciolatkę mogli decydować jego rodzice.

W styczniu 2012 r. parlament odroczył termin powszechnego obowiązku o dwa lata, do 1 września 2014 r. W sierpniu 2013 r. uchwalono kolejną nowelizację, zgodnie z którą 1 września 2014 r. obowiązkowo zaczynały edukację szkolną sześciolatki urodzone w pierwszej połowie 2008 r. (wraz z siedmiolatkami urodzonymi w 2007 roku), a 1 września 2015 r. – pozostałe dzieci z tego rocznika.

Powód kolejnych nowelizacji był jasny: reforma nie biegła w myśl planu rządowego. Porażkę unaocznia Tabela 1.1. Jak widać, w pierwszym roku reformy do szkoły dotarło nieco ponad 4 proc. sześciolatków

Tabela 1.1 Procent sześciolatków posłanych do klasy pierwszej

Okres	Rok szkolny	Procent	Procent dziewczynek
Przed reformą	2008/2009	1,0	58,2
	2009/2010	4,3	58,6
Pierwszy okres przejściowy	2010/2011	9,4	58,1
	2011/2012	19,0	57,4
Drugi okres przejściowy	2012/2013	17,1	57,5
	2013/2014	15,5	57,2
Okres obowiązkowy	2014/2015	44,9	53,2
	2015/2016	73,4	51,2

Źródło: Główny Urząd Statystyczny (2009–2016)

– zaledwie o kilka punktów procentowych więcej niż w okresie sprzed reformy. W pierwszym okresie przejściowym odsetki rosły powoli – do 19 w roku szkolnym 2011/2012, gdy badani przez nas czwartoklasiści zaczęli swoją karierę szkolną. Było jasne, że wśród rodziców reforma nie wzbudza entuzjazmu. Bolesnie potwierdził to następny okres przejściowy. Rząd mógł ustąpić lub okazać siłę. Wybrano drugą możliwość kosztem półtorakrotnego zwiększenia naboru dzieci w dwóch kolejnych latach szkolnych. Pod naciskiem społecznym nowy rząd wyłoniony po wyborach parlamentarnych w 2015 r. skutecznie przeprowadził zniesienie obowiązku szkolnego dla sześciolatków i powrót do stanu sprzed 2009 r.

W siedmioletnim okresie wprowadzania w życie reformy z 2009 r. do szkół trafiały sześciolatki razem z siedmiolatkami. W niektórych szkołach tworzone oddziały homogeniczne wiekowo, w innych sześciolatki uczyły się razem z siedmiolatkami w tej samej izbie lekcyjnej. Dla badaczy oświaty jest to samorzutny eksperyment naturalny, który może rzucić nowe światło na stary spór o wyższość oddziałów wieloletnich nad jednorocznikowymi.

Pierwszy i drugi etap kształcenia

W 2011 r., gdy Polska po raz pierwszy przystąpiła do TIMSS, pomiarowi osiągnięć poddaliśmy uczniów klasy trzeciej, ponieważ byli oni – wliczając obowiązkową klasę przygotowawczą („zerówkę”) – w czwartym roku edukacji. W innych krajach, z których większość edukowała dzieci począwszy od lat sześciu, testowano czwartoklasistów. Różnica jednego roku formalnej edukacji szkolnej stawiała polskie dzieci na przegranej pozycji. Gdy w 2013 r. zaczynaliśmy przygotowania do TIMSS 2015, Polska znajdowała się w środku reformy obowiązku szkolnego i testowanie trzecioklasistów straciło uzasadnienie. Dlatego zdecydowaliśmy się, wzorem innych krajów, objąć pomiarem uczniów klasy czwartej. Jednak zmiana klasy trzeciej na czwartą implikuje w Polsce coś więcej niż dodatkowy rok nauki – zmianę etapu kształcenia.

Warto przypomnieć, że reforma systemu oświaty zainicjowana przez rząd Jerzego Buzka powołała do życia sześcioklasową szkołę podstawową podzieloną na dwa etapy: kształcenia zintegrowanego (klasy od pierwszej do trzeciej) i kształcenia przedmiotowego (klasy od czwartej do szóstej). W pierwszym etapie, dziś nazywanym wczesnoszkolnym, nauczanie odbywa się w formie zintegrowanej, co w praktyce znaczy, że treść kształcenia nie jest podzielona na przedmioty nauczania. Główną postacią w oddziale klasowym jest jeden nauczyciel-wychowawca. Buduje on program

kształcenia, tak by znalazły się w nim: edukacja polonistyczna, społeczna, przyrodnicza, matematyczna, muzyczna i plastyczna, zajęcia techniczne i komputerowe, język obcy nowożytny oraz wychowanie fizyczne¹. Prym w programie edukacji wczesnoszkolnej wiodą język polski i matematyka. Nauczyciel dba o adaptację dzieci do warunków szkolnych i poczucie bezpieczeństwa. Służą temu izby lekcyjne składające się z dwóch części: edukacyjnej i rekreacyjnej. W wielu szkołach edukacja wczesnoszkolna nie podlega tradycyjnemu reżimowi dzwonek szkolnych. Poszczególne zajęcia rozpoczyna i kończy nauczyciel według własnego uznania, gdy widzi, że uczniowie potrzebują przerwy lub zmiany aktywności. Ocenianie postępów w nauce ma charakter opisowy: zamiast wystawiać cenzurki ze stopniami, nauczyciel opisuje mocne i słabe strony każdego ucznia. Ten etap kształcenia nie kończy się żadnym sprawdzianem zewnętrznym.

W drugim etapie pojawiają się przedmioty nauczania: język polski, historia i społeczeństwo, język obcy nowożytny, matematyka, przyroda, muzyka, plastyka, zajęcia techniczne, zajęcia komputerowe i wychowanie fizyczne. Każdego przedmiotu naucza inny nauczyciel. Kontakt ucznia z wychowawcą jest, w porównaniu z pierwszym etapem, znacznie rzadszy. W niektórych szkołach uczniowie nie spędzają całego dnia w tej samej izbie lekcyjnej, lecz przechodzą z jednej pracowni przedmiotowej do drugiej. Lekcje rozpoczyna i kończy dzwonek.

Od początku klasy czwartej na lekcjach częściej pojawiają się różnorodne sprawdziany wiedzy, więcej też jest prac domowych, i to nie zawsze do wykonania na następny dzień, lecz za tydzień. Dziecko musi nauczyć się pracy podług rozkładu przedmiotów w tygodniowym planie lekcji i dostosować się do swoich dla każdego nauczyciela wymagań i sposobów nauczania, czyli na nowo odnaleźć się w swojej szkole.

Ocenianie postępów w nauce odbywa się za pomocą stopni od niedostatecznego po celujący. Stawia je z osobna każdy nauczyciel. Różnice kryteriów oceny próbuje ograniczyć szkolny system oceniania. Niezależnie od ocen szkolnych – bieżących, semestralnych i końcowych – nad uczniami wisi² egzamin zewnętrzny pod koniec klasy szóstej. Choć w założeniu przyjazny i niemający konsekwencji dla dalszej edukacji dziecka, bywa on odczuwany przez uczniów i rodziców, a za ich

¹ W większości szkół lekcje języka obcego są prowadzone przez innego, wyspecjalizowanego nauczyciela. Specjalistom powierza się też często zajęcia muzyczne, plastyczne i wf.

² Dokładniej: wisiał nad badanymi przez nas uczniami, ale zlikwidowała go nowelizacja ustawy o systemie oświaty z dnia 23 czerwca 2016.

pośrednictwem także przez nauczycieli jako zagrożenie. Przypuszcza się, że ma to negatywny wpływ na wyznaczanie celów i sposobów kształcenia, choć nie ma badań, które by to potwierdzały.

Z porównania dwóch etapów kształcenia wynika, że klasę trzecią oddziela od czwartej wyraźny próg. Nie brak rodziców, badaczy, a nawet nauczycieli (głównie edukacji wczesnoszkolnej), którzy twierdzą, że spowalnia on postępy dzieci w nauce, ponieważ zmusza je do angażowania zasobów psychicznych w readaptację do szkoły zamiast w opanowywanie nowych umiejętności. Porównanie wyników referowanego badania z wynikami z 2011 r. pozwoli zweryfikować ten sąd.

Organizatorzy i wykonawcy badania

Inicjatorem i organizatorem badania jest założone w 1958 r. Międzynarodowe Towarzystwo Oceniania Osiągnięć Szkolnych (International Association for the Evaluation of Educational Achievement, IEA). Jest to niezależne konsorcjum krajowych organizacji badawczych i agencji rządowych z prawie 70 krajów członkowskich ze stałym sekretariatem w Amsterdamie. IEA prowadzi wiele międzynarodowych badań edukacyjnych. Oprócz badania TIMSS do najbardziej znanych należą: Międzynarodowe badanie kompetencji komputerowych i informacyjnych (International Computer and Information Literacy Study, ICILS), Międzynarodowe badanie edukacji obywatelskiej (International Civic and Citizenship Education Study, ICCS), Międzynarodowe badanie postępów w czytaniu (Progress in International Reading Literacy Study, PIRLS) oraz Badanie kształcenia i doskonalenia zawodowego nauczycieli matematyki (Teacher Education and Development Study in Mathematics, TEDS-M).

Odpowiedzialność za plan, narzędzia i przebieg badania TIMSS spoczywa na Międzynarodowym Ośrodku Badawczym przy Wydziale Edukacji w Koleżu Bostońskim (Lynch School of Education in Boston College). Ośrodek ten ściśle współpracuje z sekretariatem IEA w Amsterdamie, który sprawuje ogólny nadzór nad badaniem, w tym weryfikuje przekłady narzędzi badawczych, Ośrodkiem Przetwarzania Danych w Hamburgu (Data Processing Center), który sprawdza i analizuje dane przesyłane przez uczestniczące kraje, kanadyjskim urzędem statystycznym (Statistics Canada) w Ottawie odpowiedzialnym za dobór prób szkół i oddziałów klasowych oraz Agencją Testowania Edukacyjnego w Princeton (Educational Testing Service), która doradza w sprawie metod psychometrycznych i dostarcza oprogramowania pozwalającego skalować wyniki testowania.

W poszczególnych krajach badanie prowadzą zespoły krajowe na własny koszt i odpowiedzialność. Każdy kraj powołuje krajowego koordynatora badania (NRC), który dba o dostosowanie założeń do realiów swojego kraju oraz dotrzymanie procedur i terminów. W każdym kraju działa również kontroler jakości powołany i opłacany przez IEA. Przygląda się on przede wszystkim przebiegowi badań w terenie. Wszelkie odstępstwa od procedury są odnotowywane w międzynarodowym raporcie z badania, a poważne uchybienia mogą spowodować, że kraj w ogóle nie zostanie w nim ujęty.

W Polsce badanie przeprowadził Instytut Badań Edukacyjnych w Warszawie na zlecenie Ministerstwa Edukacji Narodowej i z jego środków finansowych. Skład zespołu badawczego jest podany na stronie redakcyjnej. Zbieranie danych w szkołach wzorowo przeprowadzili współpracownicy terenowi IBE pod kierunkiem dr Aliny Cichomskiej.

Osoby badane

W badaniu wzięły udział dzieci z 49 krajów. Większość tych krajów to państwa, ale niektóre (np. Anglia lub Irlandia Północna) to części składowe państw. Wybrane cechy krajów uczestniczących w badaniu przedstawia Tabela 1.2. Na pewno przydadzą się w interpretowaniu wyników badania.

Pomiarem osiągnięć objęto dzieci, które chodziły do szkoły i spełniały dwa dodatkowe warunki:

- kończyły czwarty rok nauki na szczeblu ISCED 1
- ich średni wiek w chwili testowania nie był niższy niż 9,5 roku.

Drugi z tych warunków zapobiega testowaniu zbyt młodych dzieci. Byłoby tak w Anglii i w Nowej Zelandii, bo dzieci w tych krajach zaczynają naukę w wieku pięciu lat. Dlatego pomiar, zamiast w czwartej, przeprowadzono tam w klasie piątej. W Belgii, Czechach, Francji, Hiszpanii, Słowenii i we Włoszech dzieci zaczynają naukę w roku kalendarzowym, w którym kończą 6 lat, więc pod koniec czwartego roku nauki mają około 10 lat. W Finlandii, Szwecji, Bułgarii i na Litwie w badaniu wzięły udział uczniowie klasy czwartej, ale że w tych krajach obowiązek szkolny zaczyna się rok później, to średnia ich wieku w dniu testowania wyniosła prawie 11 lat. Ogółem średnia wieku uczniów, którzy wzięli udział w badaniu TIMSS 2015, waha się od 9,6 roku (w Omanie) do 11,5 roku (w Południowej Afryce) wokół mediany 10,2. W Polsce wyniosła 10,7 roku i przewyższyła średnie wieku w 38 krajach.

Badania TIMSS prowadzi się metodą reprezentacyjną, co znaczy, że w każdym kraju obejmuje się pomiarem uczniów z losowej próby, która reprezentuje

Tabela 1.2. Wybrane cechy krajów uczestniczących w badaniu TIMSS 2015

Kraj	Ludność	Po-wierzchnia	Gęstość zaludnienia	Ludność miejska	Długość życia	Śmier-telność niemow-ląt	GNI per capita	Wydatki oświato-we	Skolaryzacja netto w szkole		Liczba uczniów na nauczyciela w szkole	
									podsta-wowej	średniej	podsta-wowej	średniej
Anglia	54	133	411	81	80	4	37 900	6	100	98	18	16
Arabia Saudyjska	31	2 150	14	83	74	13	51 320	5	96	88	11	11
Australia	23	7 741	3	89	82	3	44 700	5	97	87	16	12
Bahrajn	1		1 769	89	77	6	37 680	3	–	–	12	10
Belgia (flamandzka)	11	31	371	98	81	3	44 090	6	99	95	11	9
Bułgaria	7	111	67	74	75	10	16 840	4	95	88	18	13
Chile	18	756	24	89	81	7	21 320	5	92	89	20	21
Chorwacja	4	57	76	59	77	4	20 910	4	89	95	14	8
Cypr	1	9	125	67	80	3	29 190	7	97	95	13	10
Czechy	11	79	136	73	78	3	28 740	4		0	19	12
Dania	6	43	133	88	81	3	46 850	9	98	90	–	–
Federacja Rosyjska	144	17 098	9	74	70	8	22 160	4	96	–	20	9
Finlandia	6	338	18	84	81	2	40 630	7	100	94	13	13
Francja	66	549	121	79	82	4	40 100	6	99	99	18	13
Gruzja	4	70	79	53	75	11	9 080	2	99	92	9	7
Hiszpania	47	506	93	79	83	4	33 490	4	100	96	13	11
Holandia	17	42	501	90	81	3	48 860	6	99	92	12	14
Hong Kong	7	1	6 897	100	84	2	56 570	4	95	88	14	13
Indonezja	255	1 911	140	53	69	23	10 190	3	90	75	16	15
Iran	78	1 745	48	73	75	13	16 590	3	99	81	26	18
Irlandia	5	70	67	63	81	3	42 830	6	95	100	16	14
Irlandia Północna	2	14	133	64	80	5	37 900	6	98	87	21	15
Japonia	127	378	349	93	84	2	38 120	4	100	99	17	12
Jordania	7	89	74	83	74	15	11 910	9	88	85	18	9
Kanada	36	9 985	4	82	82	5	44 350	5	99	–	14	14
Katar	2	12	187	99	79	7	134 420	4	92	95	11	10
Kazachstan	17	2 725	6	53	72	13	21 710	3	87	96	17	9
Korea Południowa	50	100	517	82	82	3	33 650	5	96	96	17	16
Kuwejt	4	18	211	98	75	7	79 850	–	92	83	9	8
Litwa	3	65	47	67	74	3	26 390	18	98	94	13	8
Maroko	34	447	76	60	74	24	7 290	5	98	56	26	–
Niemcy	81	357	232	75	81	3	47 460	5	98	–	12	13
Norwegia	5	385	14	80	82	2	67 100	7	100	95	9	9
Nowa Zelandia	5	268	17	86	81	5	36 200	7	98	96	15	14
Oman	4	310	14	80	77	10	33 690	4	91	89	7	–
Polska	38	313	124	61	77	5	24 430	5	97	92	10	10
Południowa Afryka	54	1 219	45	64	57	34	12 700	6	99	87	32	25
Portugalia	10	92	114	63	81	3	28 370	5	99	93	13	10
Serbia	7	88	82	55	76	6	13 040	4	96	92	16	9
Singapur	6	1	7 737	100	83	2	80 270	3	100	98	17	13
Słowacja		49	113	54	81	6	27 410	4	–	–	15	11
Słowenia	2	20	102	50	81	2	30 360	6	97	95	17	10
Stany Zjednoczone	319	9 832	35	81	79	6	55 900	5	93	88	14	15
Szwecja	10	447	24	86	82	2	46 870	8	99	94	10	10
Tajwan	23	36	646	–	80	4	–	4	98	96	13	15
Turcja	76	784	99	73	75	12	19 560	3	93	88	20	20
Węgry	10	93	109	71	76	5	23 960	5	91	93	10	10
Włochy	61	301	207	69	83	3	35 450	4	98	95	12	11
Zjednoczone Emiraty Arabskie	9	84	109	85	77	6	67 720	1	91	81	19	13

Objaśnienia: ludność w milionach, powierzchnia w tysiącach km², gęstość zaludnienia w osobach na km², ludność miejska jako procent populacji, długość życia oczekiwana przy urodzeniu w latach, śmiertelność niemowląt na tysiąc żywych urodzeń, GNI – dochód narodowy brutto z korektą siły nabywczej w dolarach USA, wydatki oświatowe jako procent PKB, skolaryzacja netto w procentach.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

krajową populację uczniów. Ogólne zasady definiowania krajowych populacji oraz dobierania próby zostały uzgodnione na poziomie międzynarodowym. Zostawiają one nieco swobody krajowym zespołom badawczym. Po pierwsze, dopuszczają wyłączenie z definicji populacji (a więc i z operatu losowania³) dzieci mniejszości narodowych niemówiące językiem urzędowym. Skorzystała z tego jedynie Gruzja, wskutek czego w ich operacie losowania znalazło się tylko 90 proc. gruzińskich czwartoklasistów. Niepełne, 79-procentowe pokrycie populacji mamy również w Kanadzie, która badaniem objęła jedynie 5 prowincji.

Po drugie, zasady międzynarodowe pozwalają na usunięcie z operatu losowania szkół, które nie kształcą uczniów ze zdefiniowanej populacji, lub szkół, do których dotarcie byłoby bardzo kosztowne. Zwykle dotyczy to szkół:

- dla uczniów ze specjalnymi potrzebami edukacyjnymi (to wykluczenie objęło w Polsce 3,4 tys. uczniów)
- bardzo małych (w Polsce były to 782 szkoły mające mniej niż pięciu uczniów w klasie czwartej, łącznie 1,6 tys. uczniów)
- znajdujących się w odległych i trudno dostępnych miejscach kraju (w Polsce ich nie było).

Wszystkie te wykluczenia objęły w Polsce 1,4 proc. populacji czwartoklasistów.

Oprócz ograniczeń na poziomie populacji krajowej zasady międzynarodowe dopuszczają, po trzecie, wykluczenie pojedynczych uczniów z wylosowanych oddziałów klasowych, jeśli:

- są niepełnosprawni sensorycznie, ruchowo lub umysłowo (w Polsce, nie chcąc ich dyskryminować, umożliwiliśmy im pracę nad testem w towarzystwie rówieśników, ale nie włączaliśmy ich wyników do krajowej bazy danych)
- nie znają języka, w którym prowadzi się testowanie.

W Polsce na tej podstawie wykluczono z bazy danych 2,6 proc. uczniów.

Ogółem wykluczenia szkół i uczniów w wylosowanych szkołach nie powinny przekroczyć pięciu proc. populacji. Nie udało się spełnić tego wymogu w 12 krajach, szczególnie w Serbii (11,3 proc.), Singapurze (10,1 proc.) i Danii (7,5 proc.). W Polsce udało się ograniczyć wykluczenia do czterech procent. Znaczący to, że populacja czwartoklasistów, o której wnioskujemy na podstawie wyników badania, liczy 359,7 tys. dzieci.

³ Operatem losowania nazywa się spis jednostek badanej populacji lub obiektów zawierających jednostki badanej populacji. W naszym wypadku są to szkoły podstawowe prowadzące klasę czwartą.

Dobór próby

We wszystkich krajach zastosowano podobną metodę doboru próby opartą na dwustopniowym schemacie losowania warstwowo-zespołowego. Decyzję o podziale operatu losowania na warstwy podejmuje krajowy zespół badawczy, kierując się troską o zapewnienie w próbie proporcjonalnej reprezentacji uczniów należących do wyznaczonych warstw lub chęcią zwiększenia dokładności oszacowań. W Polsce zastosowaliśmy dwie zmienne, które posłużyły do wyodrębnienia 15 warstw jawnych. Pierwszą z nich była lokalizacja szkoły: na wsi, w mieście do 20 tys. mieszkańców, w mieście od 20 do 100 tys. mieszkańców i w mieście powyżej 100 tys. mieszkańców. Drugą był średni wynik, jaki osiągnęli uczniowie każdej szkoły w sprawdzianie CKE przeprowadzanym pod koniec klasy szóstej. Obie zmienne pozwoliły podzielić operat na 15 warstw.

W każdej z tych warstw losowano szkoły z prawdopodobieństwem proporcjonalnym do wielkości szkoły (w Polsce miarą wielkości szkoły była liczba uczniów w klasie czwartej). Tak wylosowane szkoły nazywa się próbą główną. Każdej szkole z próby głównej przypisano losowo dwie szkoły zastępcze⁴, na wypadek gdyby dyrektor szkoły z próby głównej nie zgodził się na udział w badaniu. Szkoły te stanowiły pierwszą i drugą próbę rezerwową. W każdej wylosowanej szkole, której dyrektor zgodził się na udział w badaniu, losowano z równym prawdopodobieństwem dwa oddziały spośród wszystkich oddziałów klasy czwartej (jeśli szkoła prowadziła tylko jeden oddział, obywało się bez losowania). Badaniem obejmowano wszystkich uczniów w każdym z wylosowanych oddziałów pod warunkiem pisemnej zgody ze strony ich rodziców. Oddziały, w których większość uczniów nie przystąpiła do badania z powodu sprzeciwu rodziców, nie mogły być zastępowane innymi.

Poziom realizacji próby

Ze 150 szkół wylosowanych do próby głównej 13 odmówiło wzięcia udziału w badaniu. Spotkaliśmy się też z odmową ze strony jednej szkoły z pierwszej próby rezerwowej. W wylosowanych 254 oddziałach

⁴ W każdej warstwie szkoły uporządkowano według miary ich wielkości i wybierano je za pomocą losowania systematycznego z losowym punktem startowym i ustalonym interwałem losowania. Każdej szkole z próby głównej przypisywano szkołę następującą po niej w operacie jako pierwszą zastępczą, a poprzedzającą ją w operacie – jako drugą zastępczą. Obie szkoły zastępcze są zatem najbardziej podobne do szkoły z próby głównej pod względem cech użytych w schemacie doboru.

szkół, które zgodziły się na udział, uczyło się 5297 uczniów. Z powodu niepełnosprawności lub niedostatecznej znajomości języka polskiego wykluczaliśmy 118. Rodzice 352 uczniów nie wyrazili zgody na testowanie ich dzieci, a 80 dzieci nie wzięło udziału w badaniu z różnych powodów losowych.

Tabela 1.3 przedstawia poziom realizacji próby w poszczególnych krajach obliczony jako ważony iloraz liczby uczniów biorących udział w badaniu do liczby uczniów ze wszystkich szkół i oddziałów wylosowanych do badania (po uwzględnieniu szkół z prób rezerwowych)⁵. Realizację próby uznaje się za zadowalającą, gdy:

- w badaniu wzięło udział co najmniej 85 proc. wylosowanych szkół, co najmniej 95 proc. wylosowanych oddziałów w szkołach i co najmniej 85 proc. uczniów w oddziałach, które przystąpiły do badania
- poziom realizacji badania wśród uczniów (trzecia kolumna tabeli) wyniósł co najmniej 75 proc.

W Polsce oba te warunki zostały spełnione z nadkładem. W badaniu wzięło udział 91 proc. szkół z próby głównej i 100 proc. po uwzględnieniu szkół rezerwowych, 100 proc. oddziałów w szkołach, które wyraziły zgodę na badanie, i 92 proc. uczniów z wylosowanych oddziałów. W sześciu krajach warunki zostały spełnione (a w Irlandii Północnej niemal spełnione) dopiero po uwzględnieniu szkół z prób rezerwowych. Niespełnienie warunków jest odnotowane w przypisach do tabel przedstawiających główne wyniki badania.

Ostatecznie w bazie międzynarodowej znalazły się dane ponad 336 tys. uczniów z niemal 10 tys. szkół z 49 krajów. W Polsce zebraliśmy dane od 4747 uczniów z 254 oddziałów klasy czwartej 150 szkół podstawowych. Ponadto badaniem ankietowym objęto dyrektorów wylosowanych szkół, nauczycieli uczących matematyki i przyrody w wylosowanych oddziałach klasowych oraz rodziców uczniów poddanych pomiarowi osiągnięć. W Polsce ankietę wypełniło 150 dyrektorów lub wicedyrektorów, 390 nauczycieli oraz 4636 rodziców.

Ważenie danych

Próba losowa jest reprezentatywna, jeśli każdy obiekt w próbie reprezentuje tę samą liczbę obiektów w populacji. Odwrotność tej liczby to oczywiście prawdopodobieństwo wylosowania obiektu z populacji do

⁵ W razie odmowy zarówno ze strony szkoły z próby głównej, jak i jej szkół zastępczych, a także odmowy ze strony całego oddziału lub wykluczenia go z bazy danych z powodu opuszczenia sesji testowej przez ponad połowę uczniów, wskaźnik ten jest obniżany.

próby. Jeśli schemat lub proces losowania nie zapewnia „równego traktowania” wszystkich obiektów, to równość przywraca się za pomocą ważenia. Polega ono na przypisaniu każdemu obiektowi w próbie liczby, która mówi, ile obiektów w populacji reprezentuje ten obiekt. Obiekt, który miał mniejsze niż inne obiekty prawdopodobieństwo znalezienia się w próbie, reprezentuje więcej obiektów w populacji. Waga „dowartościowuje” jego wkład w oszacowanie parametrów populacyjnych. To samo (*mutatis mutandis*) dotyczy obiektów, które miały większe prawdopodobieństwo wylosowania.

Najlepiej wyjaśnić to na przykładzie z badania TIMSS. Zaczniemy od szkół. Wagi przypisuje się szkołom osobno w każdej warstwie jawnej. W polskim operacie losowania było 7646 szkół wiejskich, do których chodziło 139052 czwartoklasistów. W wylosowanej próbie znalazło się 56 takich szkół. Na jedną szkołę przypada więc średnio 139052/56 uczniów. Powiedzmy, że szkole w Kożuszkach w trzech oddziałach klasy czwartej uczy się 60 uczniów. Każdy z nich reprezentuje zatem 41,385 uczniów ze wsi. Ta liczba jest wagą tej szkoły.

Waga oddziału klasowego informuje o liczbie oddziałów w danej szkole reprezentowanych przez ten oddział. Jeśli w szkole w Kożuszkach wylosowano dwa oddziały: IIIa i IIIc, waga każdego wynosi $\frac{3}{2}$. Waga ucznia w oddziale, który w całości przystąpił do badania, wynosi 1 (każdy uczeń reprezentuje samego siebie). Jeśli jednak w 20-osobowym oddziale IIIc jeden uczeń nie przyszedł na sesję testową, waga pozostałych równa się $\frac{20}{19}$. Widać, że w badaniu TIMSS wagi mają podstawę probabilistyczną z poprawką na niepełną realizację badania⁶.

Ogólna waga przypisana każdemu uczniowi to iloczyn powyższych wag (szkoły, oddziału i ucznia). Waga ucznia klasy IIIc w Kożuszkach to zatem 65,344. Z tą wagą jego wynik testowania wchodzi do wszystkich analiz. Ignorowanie wag mści się błędnymi oszacowaniami. Dla przykładu: ważona średnia testu matematycznego TIMSS w Polsce wynosi 534,77. Gdyby zignorować wagi, wyniosłaby 537,25. Przesunęlibyśmy się o dwa miejsca w górę. Szkoda, że za sprawą niepoprawnej statystyki.

Przebieg badania

Badanie przeprowadzono w szkołach od połowy kwietnia do końca maja 2015 r. w indywidualnym, do-

⁶ Podobną poprawkę stosuje się przy wagach szkół, gdy jakaś szkoła i obie jej szkoły zastępcze odmówiły udziału w badaniu, lub przy wagach oddziałów, gdy w jakimś oddziale do badania przystąpiła mniej niż połowa uczniów.

Tabela 1.3. Realizacja badania TIMSS 2015

Kraj	Liczba szkół biorących udział w badaniu ^a	Liczba uczniów biorących udział w badaniu	Procent realizacji próby ^b	Średnia wieku uczniów w dniu testowania
Anglia	147 (5)	4006	96	10,1
Arabia Saudyjska	189 (11)	4337	93	10,0
Australia	287 (2)	6057	94	10,0
Bahrajn	182 (0)	4146	99	9,9
Belgia (flamandzka)	153 (36)	5404	95*	10,1
Bułgaria	149 (1)	4228	93	10,8
Chile	179 (18)	4756	88	10,2
Chorwacja	163 (2)	3985	94	10,6
Cypr	148 (0)	4125	98	9,8
Czechy	159 (0)	5202	95	10,4
Dania	193 (80)	3710	86*	10,9
Federacja Rosyjska	208 (0)	4921	98	10,8
Finlandia	158 (1)	5015	97	10,8
Francja	164 (5)	4873	97	9,9
Gruzja	153 (2)	3919	98	9,7
Hiszpania	358 (1)	7764	95	9,9
Holandia	129 (55)	4515	83*	10,0
Hong Kong	132 (9)	3600	76*	10,1
Indonezja	230 (0)	4025	99	10,4
Iran	248 (0)	3823	99	10,2
Irlandia	149 (0)	4344	96	10,4
Irlandia Północna	118 (18)	3116	71**	10,4
Japonia	148 (5)	4383	97	10,5
Jordania	254 (0)	7861	96	9,8
Kanada	441 (38)	12 283	80*	9,9
Katar	211 (0)	5194	99	10,1
Kazachstan	171 (6)	4702	97	10,3
Korea Południowa	149 (0)	4669	97	10,5
Kuwejt	166 (0)	3593	90	9,7
Litwa	225 (2)	4529	94	10,7
Maroko	358 (0)	5068	99	10,3
Niemcy	204 (5)	3948	95	10,4
Norwegia	140 (0)	4329	89	10,7
Nowa Zelandia	174 (27)	6322	90	10,0
Oman	300 (4)	9105	97	9,6
POLSKA	150 (13)	4747	92	10,7
Południowa Afryka	297 (4)	10 932	98	11,5
Portugalia	217 (24)	4693	92	9,9
Serbia	160 (2)	4036	96	10,7
Singapur	179 (0)	6517	96	10,4
Słowacja	198 (5)	5773	97	10,4
Słowenia	148 (4)	4445	93	9,8
Stany Zjednoczone	250 (22)	10 029	81*	10,2
Szwecja	144 (0)	4142	95	10,8
Tajwan	150 (1)	4291	99	10,2
Turcja	242 (0)	6456	98	9,9
Węgry	144 (1)	5036	96	10,7
Włochy	164 (28)	4373	94	9,7
Zjednoczone Emiraty Arabskie	558 (0)	21 177	97	9,8

^a W nawiasie podano liczbę szkół rezerwowych w łącznej liczbie szkół.

^b Po uwzględnieniu szkół rezerwowych.

* Poziom realizacji próby spełnia wymagania jedynie po uwzględnieniu szkół rezerwowych.

** Poziom realizacji próby niemal spełnia wymagania jedynie po uwzględnieniu szkół rezerwowych.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

godnym dla każdej szkoły terminie⁷. Na 2–3 tygodnie przed ustalonym terminem badania rodzice lub opiekunowie prawni uczniów z wylosowanych oddziałów otrzymywali pisemną informację o badaniu TIMSS 2015, przeznaczoną dla nich ankietę oraz formularz zgody na udział dziecka w badaniu. Wypełnione formularze i ankiety w zaklejonych kopertach zbierał wyznaczony nauczyciel i przekazywał prowadzącemu badanie. Przed główną sesją testową prowadzący badanie rozdawał dyrektorowi szkoły i nauczycielom wylosowanych oddziałów przeznaczone dla nich ankiety i prosił o zwrot po wypełnieniu.

Sesja testowa zaczynała się od poinformowania dzieci o międzynarodowym charakterze badania oraz rozdania imiennie oznaczonych zeszytów testowych. Każdy zeszyt otwierały dwustronicowe *Wskazówki*. Prowadzący czytał je głośno wraz z uczniami i sprawdzał, czy właściwie wykonują przykładowe zadania. Następnie dzieci przystępowały do pracy nad pierwszą częścią zeszytu, a po 15-minutowej przerwie – nad drugą. Nad każdą częścią mogły pracować nie dłużej niż 36 min.

Z dwóch części zeszytu TIMSS jedna odnosiła się do matematyki, druga do przyrody. We wszystkich częściach znajdowały się pytania dwojakiego rodzaju. Jedne wymagały wyboru jednej z kilku opcji, inne – sformułowania krótkiej odpowiedzi. Dzieci zaznaczały właściwe opcje lub pisały swoje odpowiedzi w zeszycie testowym.

Po zakończeniu sesji testowej i dłuższej przerwie dzieci wypełniały ankietę ucznia. Ją również otwierały *Wskazówki*, które prowadzący czytał wraz z dziećmi, śledził ich odpowiedzi na przykładowe pytania i dostarczał dodatkowych wyjaśnień, gdy czegoś nie rozumiały. W razie potrzeby procedura zezwalała na głośne czytanie dzieciom pytań ankiety.

Jeśli w badaniu wzięło udział mniej niż 90 proc. uczniów, prowadzący organizował sesję uzupełniającą dla dzieci nieobecnych w dniu sesji głównej.

Narzędzia badawcze

Dwa pierwsze cele badania wymienione na początku tego rozdziału wymagały narzędzi pomiaru osiągnięć szkolnych zwanych testami, trzeci został osiągnięty za pomocą ankiet zaadresowanych do dyrektorów szkół, nauczycieli, uczniów i ich rodziców.

⁷ Warto dodać, że w maju 2014 r. poprzedziło je badanie pilotażowe, w którym wzięło udział 1 056 uczniów klasy czwartej z 34 losowo dobranych szkół. Badanie pilotażowe miało na celu przetestowanie narzędzi oraz wszystkich procedur realizacyjnych.

Zaczynamy od testów osiągnięć. W Polsce są one dobrze znane dzięki działalności Centralnej Komisji Egzaminacyjnej. Jest jednak wielka różnica między egzaminowaniem uczniów w jednym kraju a szacowaniem poziomów osiągnięć w wielu krajach. Testy CKE informują o wiedzy każdego ucznia na tle pozostałych uczniów, natomiast test TIMSS – o wiedzy wszystkich uczniów w każdym kraju na tle innych krajów. Egzamin musi dawać wyniki pozwalające porównywać wiedzę dwóch dowolnych uczniów, to zaś wymaga pełnej standaryzacji sytuacji egzaminacyjnej. Z tego właśnie powodu uczniowie w całym kraju dostają w sesji egzaminacyjnej identyczny test. Ponieważ czas egzaminu jest ograniczony, test nie może być zbyt długi. Aplikowany szóstoklasistom test osiągnięć CKE składał się w 2015 r. z 19 zadań i 27 pytań⁸. Trudno uznać, że tak ograniczony zbiór zadań trafnie reprezentuje ogół oczekiwanych wyników kształcenia. Ponieważ jednak chcemy z możliwie największą precyzją wyznaczyć miejsce każdego ucznia na skali osiągnięć, godzimy się poświęcić szerokość wiedzy o wykształceniu populacji uczniów.

Badanie TIMSS zmierza do zdobycia jak najpełniejszej wiedzy o osiągnięciach populacji uczniów i dlatego używa się w nim dłuższych testów. Test osiągnięć matematycznych składa się ze 169, a przyrodniczych – z 168 pytań. Jest oczywiste, że tak długiego testu nie byłoby w stanie wykonać żadne dziecko. Dlatego dzieli się go na podzbiory („bloki”) i składa z nich wiele różnych zeszytów testowych, które zostaną rozdane grupom uczniów. Z zasady każdy blok występuje w dwóch różnych zeszytach. W badaniu TIMSS było 14 zeszytów, a w każdym od 39 do 51 pytań. Zeszyty przydzielano uczniom według z góry ustalonego planu. Takie postępowanie nazywa się próbkowaniem macierzowym. Na marginesie zauważmy, że ma ono pożądaną skutek uboczny: eliminuje ściąganie na sesji testowej.

Drugi cel badania wymaga, by wyniki pomiarów w kolejnych edycjach dało się umieścić na tej samej skali, to zaś jest możliwe, tylko jeśli sąsiadujące ze sobą pomiary mają wspólne pytania. Dlatego zastosowany w 2015 r. test osiągnięć matematycznych zawiera 94 pytania z testu z 2011 r., a test osiągnięć przyrodniczych – 97 takich pytań. Wielokrotne wykorzystywanie tych samych zadań wymaga zadbania, by nie dostały się w niepowołane ręce.

⁸ Zadaniem nazywa się niezależną i samodzielną treściowo jednostką testu, pytaniem zaś – niezależnie punktowaną jednostką testu. Liczba pytań pokrywa się z liczbą elementarnych zmiennych w bazie danych. Każde zadanie zawiera jedno pytanie lub więcej.

Struktura testów osiągnięć

Potocznie testem nazywa się przypadkowy, byle nie-wykraczający poza program kształcenia zbiór zadań aplikowany uczniom. Testy IEA są jednak budowane według przemyślanego planu, który pozwala uzasadnić obecność każdego zadania, czyli odpowiedzieć na pytanie: dlaczego znalazło się w teście?

Zadania w testach IEA są sklasyfikowane pod dwoma względami: treści przedmiotowej i funkcji podmiotowej. Dwa względy podziału dostarczają odpowiedzi na dwa pytania o trafność pomiaru – merytoryczną: czy test obejmuje wszystkie działy programu kształcenia? i psychologiczną: czy na podstawie odpowiedzi ucznia można zrekonstruować stan jego umysłu?

Pierwsze pytanie odnosi się do treści kształcenia, a ostatecznie – do struktury wiedzy naukowej, która jest źródłem tej treści. Drugie – do teorii uczenia się. Głównym pojęciem tej teorii jest oczekiwany wynik kształcenia. Wszelki wynik oczekiwany definiujemy jako stan rzeczy będący przedmiotem czyjś dążenia. Równie dobrze można go nazwać celem. Proces kształcenia – zarówno w skali pojedynczej lekcji, jak i całego etapu – zaczyna się od postawienia celu. Poprawnie sformułowany cel nie mówi, co ma robić nauczyciel (np. rozwijać, kształtować, pobudzać), ani co mają robić uczniowie (np. słuchać, przepisywać, recytować), lecz czym wszystko to, co robią nauczyciel i uczniowie, powinno się skończyć. Na przykład umiejętnością dodawania ułamków o różnych mianownikach, odróżniania organizmów od przedmiotów nieożywionych lub wnioskowania o motywach postępowania bohatera opowiadania.

Wyniki kształcenia ujmuje się w dwóch perspektywach – zasobów (X) i zastosowań (Y). Ta konwencja ma głęboki sens – pozwala zdefiniować wynik kształcenia jako możliwość robienia y za pomocą x . Symbol x oznacza element zasobu, symbol y – użytek, jaki można zrobić z tego elementu. „Potrafię zrobić kompas z igły” – chwali się harcerz. Sama igła należy do zawartości harcerskiego plecaka, ale pojęcie igły – do umysłowego zasobu wiedzy o świecie. Na to pojęcie składa się wiedza o cechach charakterystycznych igły oraz wiedza – gotowa lub potencjalna – o jej zastosowaniach. Tę dwoistość dobrze pojmują uczniowie, gdy pytają nauczycielkę: Do czego może się przydać to, czego pani każe się nam uczyć?

Test osiągnięć matematycznych. Na treść przedmiotową tego testu składa się wiedza o liczbie (arytmetyka), obiektach geometrycznych (geometria) i sposobach przedstawiania danych (statystyka opisowa). Dobór zagadnień nie odzwierciedla „najmniejszego wspólnego mianownika” krajowych programów

kształcenia, lecz stanowi uzgodnioną przez ekspertów treść edukacji dla przyszłości.

Na wiedzę o liczbie składa się osiem zagadnień:

1. Liczby całkowite dodatnie, system pozycyjny, porządkowanie liczb.
2. Dodawanie, odejmowanie, mnożenie i dzielenie liczb całkowitych.
3. Wielokrotności i dzielniki; liczby parzyste i nieparzyste.
4. Ułamki zwykłe (jako części całości lub zbioru albo miejsca na osi liczbowej).
5. Dodawanie i odejmowanie ułamków zwykłych, porównywanie i porządkowanie ułamków.
6. Ułamki dziesiętne, w tym wartość pozycyjna, porządkowanie, dodawanie i odejmowanie ułamków.
7. Równania (znajdowanie brakującej liczby, przedstawianie prostych sytuacji za pomocą równań).
8. Sekwencje liczbowe (rozwijanie sekwencji, znajdowanie brakujących elementów w sekwencji).

Na wiedzę o figurach i miarach geometrycznych składa się siedem zagadnień:

1. Odcinki: mierzenie, szacowanie długości, proste równoległe i prostopadłe.
2. Porównywanie i rysowanie kątów.
3. Wykorzystywanie nieformalnych układów współrzędnych (np. szachownicy) do określania położenia punktów na płaszczyźnie.
4. Podstawowe właściwości typowych figur geometrycznych.
5. Obroty i odbicia względem osi.
6. Związki między figurami dwuwymiarowymi i przestrzennymi.
7. Wyznaczanie i szacowanie pola powierzchni, obwodu i objętości.

Sposoby przedstawiania danych obejmują:

1. Odczytywanie i przedstawianie danych w tabelach i na wykresach.
2. Wyciąganie wniosków z przedstawionych danych.

Funkcji podmiotowych jest trzy: wiedzieć (*knowing*), stosować (*applying*) i rozumować (*reasoning*). Celem pierwszej jest sprawdzenie, czy dane pojęcie (wiadomość lub umiejętność) znajduje się w zasobie umysłowym ucznia (X). Zadania z tej grupy mierzą znajomość konwencji matematycznych (np. że $3 \cdot 5 = 5 \cdot 3 = 5 + 5 + 5 = 3 + 3 + 3 + 3 + 3 = 15$), obiektów (np. liczby parzystej, ułamka, trójkąta, wykresu słupkowego) i procedur (np. dodawania dużych liczb, szacowania wartości wyrażenia, grupowania przedmiotów ze względu na wspólną cechę).

Dwie następne funkcje dotyczą zastosowań zasobów matematycznych (Y). Funkcja „stosować” ogranicza się do zastosowań, które zgodnie z programem kształcenia powinny istnieć w umyśle ucznia jako goto-

we schematy postępowania – wystarczy wybrać właściwy spośród wielu innych (co nie zawsze jest trywialne) i odtworzyć (co też nie zawsze udaje się bezbłędnie). Źródłem takich schematów są zadania-ćwiczenia często wykonywane na lekcjach matematyki. Najczęściej chodzi w nich o znalezienie modelu matematycznego dla sytuacji opisanej w języku naturalnym albo wybór obiektu reprezentującego taką sytuację.

Funkcja „rozumować” dotyczy zastosowań, które nie istnieją w umyśle ucznia jako gotowe schematy i muszą być wytworzone w chwili testowania. Zadania, które wymagają takich zastosowań, tradycyjnie nazywa się problemami. Nauczyciel od czasu do czasu stawia na lekcjach takie zadania-problemy, ale nie wraca do nich po wielokroć, więc zachowują one swą nietypowość. Jest jasne, że rozwiązywanie problemów wymaga bardziej złożonych operacji umysłowych niż rozpoznanie typu zadania („Aha, to na dzielenie”) i zastosowanie wyćwiczonego schematu postępowania. Trzeba wykryć regularność w dostarczonych danych i wykorzystać ją do wyjścia poza te dane. Trzeba wysnuć wnioski z podanych przesłanek lub poszukać mocnej przesłanki dla przypuszczenia, słowem – pomyśleć. Powodzenie w takich zadaniach dobrze świadczy nie tylko o uczniu, ale też o jego nauczycielu, najwyraźniej bowiem potrafił się on od czasu do czasu oderwać od algorytmów i zająć z uczniami rozwiązywaniem problemów.

Test osiągnięć przyrodniczych. Na treść przedmiotową tego testu składa się wiedza o życiu (biologia), materii nieożywionej (fizyka) i Ziemi (geografia).

Wiedza o życiu obejmuje siedem zagadnień:

1. Charakterystyka organizmów, główne grupy organizmów (np. ssaki, ptaki, owady, rośliny nasienne).
2. Budowa i funkcjonowanie głównych układów ciała człowieka i innych organizmów roślinnych i zwierzęcych.
3. Cykle życiowe pospolitych roślin i zwierząt (np. ludzi, motyli, żab, roślin nasiennych).
4. Cechy wrodzone i nabyte.
5. Cechy fizyczne, zachowanie i strategie przeżycia organizmów w różnych środowiskach.
6. Zależności w typowych biocenozach (proste łańcuchy pokarmowe, relacje drapieżnik–ofiara, wpływ człowieka na środowisko).
7. Zdrowie człowieka (np. przenoszenie chorób zakaźnych i zapobieganie im, oznaki dobrego zdrowia, objawy choroby, zdrowe odżywianie, aktywność fizyczna).

Wiedza o materii nieożywionej składa się z dziewięciu zagadnień:

1. Stany skupienia materii (ciała stałe, ciecze, gazy) i ich właściwości (objętość, kształt), zmiany stanu skupienia w wyniku ogrzewania i schładzania.

2. Klasyfikacja przedmiotów i materiałów na podstawie właściwości fizycznych (np. ciężaru, objętości, przewodnictwa cieplnego i elektrycznego, przyciągania przez magnes).
3. Mieszanki, sposoby rozdzielania mieszanin (np. przesiewanie, filtrowanie, odparowanie, użycie magnesu).
4. Przemiany chemiczne znane z codziennego życia (np. rozkład, spalanie, rdzewienie, gotowanie).
5. Naturalne źródła (np. Słońce, woda, wiatr) i formy energii (np. elektryczność, ciepło) oraz ich zastosowanie w praktyce (np. ogrzewanie i chłodzenie domów, oświetlenie).
6. Światło i dźwięk w życiu codziennym (np. cienie i odbicia, wibracja jako źródło dźwięku).
7. Elektryczność, proste obwody elektryczne (np. przewodniki prądu, przemiany elektryczności w dźwięk lub światło, obwody otwarte i zamknięte).
8. Właściwości magnesu (np. zasada przyciągania się biegunów różnoimiennych i odpychania jednoimiennych, przyciąganie niektórych ciał).
9. Siły, które wpływają na ruch przedmiotów (np. grawitacja, siły pchania i ciągnięcia).

Wiedza o Ziemi zawiera siedem zagadnień:

1. Podstawowe składniki krajobrazu (np. góry, równiny, pustynie, rzeki, oceany), sposoby ich wykorzystania przez człowieka (np. uprawa roli, irygacja, zabudowa).
2. Woda na Ziemi (występowanie, obieg wody: parowanie, chmury, deszcz, rosa).
3. Zmiany warunków pogodowych w kolejnych dniach i porach roku w różnych obszarach geograficznych.
4. Skamieliny zwierzęce i roślinne (powstawanie, znaczenie dla nauki).
5. Układ Słoneczny (Słońce, Ziemia, Księżyc, inne planety), ruch ciał (planety obiegają Słońce, Księżyc obiega Ziemię).
6. Wpływ ruchu obrotowego Ziemi na pory dnia i długość cienia.
7. Wpływ ruchu obiegowego Ziemi wokół Słońca na pory roku.

Funkcje podmiotowe zostały zdefiniowane podobnie jak w teście matematycznym. Zasoby wiedzy przyrodniczej to znajomość terminologii naukowej i faktów, umiejętności opisanie obserwacji, dobierania przykładów pojęć ogólnych, a także posługiwania się przyrządami naukowymi, w tym dokonywania pomiarów. Zastosowania podzielono na typowe, przedstawiane dziecku w procesie nauczania, i problemowe, niećwiczone na lekcjach. W obu wypadkach chodzi o wyjaśnianie naukowe, czyli odpowiedź na pytanie: dlaczego coś jest takie a takie, zachowuje się tak a tak,

nadaje się do tego a tego? W odróżnieniu od wyjaśnień potocznych, wyjaśnianie naukowe wymaga odwołania się do pojęć podstawowych, czyli – jak się mówi po angielsku – leżących u podstawy (*underlying*) zjawisk. Tak rozumiane wyjaśnianie i jego pochodne: przewidywanie, ocenianie i znajdowanie zastosowania praktycznego są istotą kształcenia przyrodniczego. Są też jego uzasadnieniem. Nie ma lepszej odpowiedzi na pytanie sceptycznego ucznia: Dlaczego miałbym się uczyć o przyrodzie? niż ta: Żebyś umiał wyjaśnić to, co dzieje się wokół ciebie.

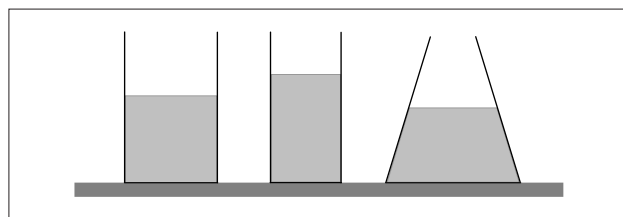
Różnica między zastosowaniami typowymi i problemowymi polega głównie na złożoności operacji umysłowych. W zastosowaniach problemowych nie wystarczy sięgnąć do gotowych schematów odpowiedzi, lecz należy zebrać i powiązać ze sobą różne informacje, rozważyć różne możliwości, sprawdzić różne przypuszczenia. Tu mieszczą się też pierwsze próby tzw. metody naukowej, czyli postępowania, które prowadzi do odkryć w prawdziwej nauce. Metoda naukowa wymaga jasnego sformułowania problemu (zwykle w postaci pytania dopełnienia: dlaczego, jak, ile itp.), zaplanowania i przeprowadzenia eksperymentu, opisanie wyników i wyciągnięcia wniosków. W polskiej edukacji przyrodniczej pojawiła się całkiem niedawno.

Na dwóch następnych stronach znajduje się tabelaryczne przedstawienie struktury obu testów z podaniem liczb pytań w każdej części testu oraz przykładowych zadań z puli udostępnionej przez organizatorów badania.

Skalowanie wyników

Plonem pracy każdego ucznia nad testem jest długi ciąg punktów przyznanych za każdą odpowiedź, czyli liczb: zer, jedynek i dwójek. Jak przekształcić ten ciąg, nazywany wektorem odpowiedzi, w pojedynczy wynik testowania? W polskim systemie egzaminów zewnętrznych robi się to przez zwykłe sumowanie. Taki wynik nazywa się surowym. Ma on swoje zalety – jest zrozumiały i otwarty na kontrolę społeczną. Ale wad ma więcej. Główna to splątanie właściwości testu i kompetencji osób badanych.

Jest intuicyjnie zrozumiałe, że wynik y pewnej osoby i w pewnym teście t zależy nie tylko od kompetencji θ tej osoby, lecz także od trudności b (i innych cech) samego testu. Symbolicznie: $y_{it} = f(\theta_i, b_t)$. Rysunek poniżej przedstawia trzy naczynia wypełnione cieczą. Niech nieobserwowalnej kompetencji θ odpowiada objętość cieczy, obserwowalnemu wynikowi y – wysokość cieczy, a trudności testu – średnica naczynia. Sens pomiaru zasadza się na możliwości oszacowania objętości na podstawie wysokości, ale jak to



zrobić, skoro nie potrafimy niezależnie oszacować średnicy? Tylko małe dzieci nie mają wątpliwości: ich zdaniem najwięcej soku jest w środkowym naczyniu, a najmniej w prawym.

Pół biedy, jeśli wszystkim osobom aplikujemy ten sam test, wtedy bowiem b jest stałą. Jeśli jednak tę samą kompetencję mierzymy różnymi testami – a tak jest co roku z testami egzaminacyjnymi – to taki sam wynik, powiedzmy w testach historii z lat 2015 i 2016, nie musi oznaczać takiej samej kompetencji, bo oba testy mogą się różnić (i najczęściej się różnią) trudnościami.

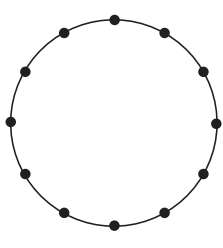
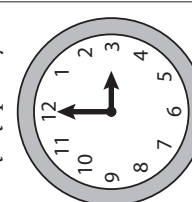
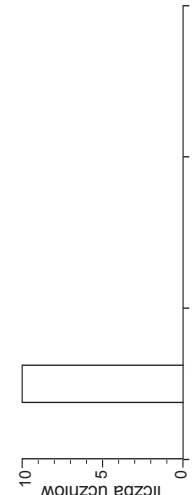
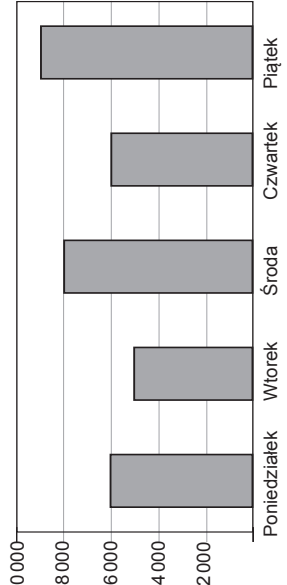
W badaniach IEA stosuje się metodę skalowania opartą na teorii odpowiedzi testowej (IRT). Wielką zaletą IRT jest to, że pozwala, by tak rzec, rozwiązać jedno równanie z dwiema niewiadomymi (θ i b), a nawet z większą ich liczbą. Na gruncie dwóch założeń: jednowymiarowości testu i lokalnej niezależności pytań metoda zapewnia niezmienniczość parametrów pytań i rozkładu kompetencji. To, że parametry pytań (np. trudność) nie zależą od rozkładu zdolności badanych osób, a rozkład zdolności nie zależy od parametrów pytań,





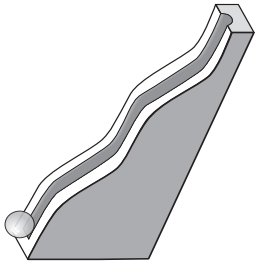
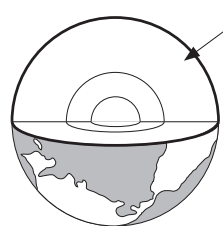
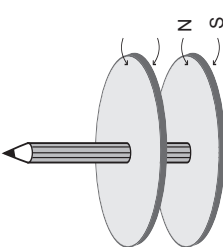
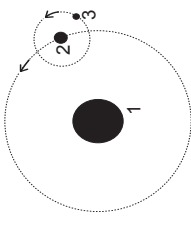
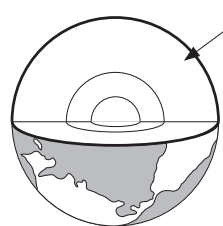
[...] jest istotą teorii odpowiedzi testowej i umożliwia takie ważne zastosowania jak zrównywanie [wyników], tworzenie banków zadań, badanie stronniczości pytania i testowanie adaptujące się do wiedzy testowanego (Hambleton, Swaminathan i Rogers, 1991, s. 25).

Kalibracja pytań, czyli określenie ich parametrów, to wstęp do dalszych analiz. Ze względu na drugi cel badania – dostarczenie informacji o zmianach średniej osiągnięć w każdym kraju, który uczestniczył w badaniu więcej niż raz – stosuje się kalibrację jednoczesną (*concurrent*) na pytaniach z obecnej i poprzedniej edycji badania – a trzeba pamiętać, że ponad połowa z nich jest wspólna – i na ważonych odpowiedziach dzieci ze wszystkich 41 krajów, które brały udział w obu edycjach. Podczas kalibracji pytania opuszczone traktuje się jako niezadane, ale przy obliczaniu wyniku testowania – jako odpowiedzi niepoprawne.

W zależności od formy pytania stosuje się jeden z trzech modeli IRT:

- do pytań wyboru – trzyparametrowy (obejmujący trudność, dyskryminację i podatność na zgadywanie)

Zastosowania problemowe, 33 (19%) pytań	Zastosowania typowe, 72 (43%) pytania	Zasoby matematyczne, 64 (38%) pytań																							
<p>Beata wybrała 4-cyfrową liczbę: Cyfrą setek jest 7. Cyfra tysięcy jest większa niż cyfra setek. Cyfra jedności jest mniejsza niż cyfra setek. Którą liczbę wybrała Beata?</p> <p>(A) 2708 (B) 4733 (C) 8726 (D) 9718</p>	<p>Janek miał 16 jabłek. Oddał 4 jabłka. Pozostałe włożył po równo do 2 koszyków. Ile jabłek jest w każdym koszyku?</p> <p>(A) 6 (B) 8 (C) 10 (D) 12</p>	<p>Trzy tysiące dwadzieścia trzy zapisujemy jako:</p> <p>(A) 323 (B) 3023 (C) 30023 (D) 300023</p>																							
<p>W tym kole narysuj trójkąt, który ma wszystkie boki równe.</p> 	<p>O godzinie 3:00 wskazówki zegara tworzą kąt prosty. Wskaż inną godzinę, o której wskazówki tworzą kąt prosty.</p> <p>(A) 3:15 (B) 3:45 (C) 9:00 (D) 9:45</p> 	<p>Uczniowie wycinają figury z papieru. Nauczycielka podnosi jedną i mówi: „Ta figura to trójkąt”. Które zdanie z poniższych musi być prawdziwe?</p> <p>(A) Figura ma trzy boki. (B) Figura ma ką prosty. (C) Figura ma równe boki. (D) Figura ma równe kąty.</p>																							
<p>Pan nauczyciel zapytał uczniów, co robią po lekcjach. Po każdej odpowiedzi stawiał kreskę w odpowiednim miejscu tabeli. Oto wyniki dotyczące trzech zajęć wybieranych przez uczniów:</p> <table border="1" data-bbox="973 224 1117 761"> <thead> <tr> <th>Zajęcia</th> <th>Kreski</th> </tr> </thead> <tbody> <tr> <td>Gra w piłkę</td> <td> </td> </tr> <tr> <td>Oglądanie telewizji</td> <td> </td> </tr> <tr> <td>Czytanie</td> <td> </td> </tr> </tbody> </table> <p>Nauczyciel chciał pokazać na rysunku, ilu uczniów wybrało każde zajęcie. Dokończ rysunek. Narysuj i podpisz dwa pozostałe słupki.</p> <p style="text-align: center;">Zajęcia po lekcjach</p> 	Zajęcia	Kreski	Gra w piłkę		Oglądanie telewizji		Czytanie		<p>W tabeli poniżej podano wagę i długość wielkich węży.</p> <table border="1" data-bbox="877 784 1069 1366"> <thead> <tr> <th>Gatunek węża</th> <th>Waga (w kilogramach)</th> <th>Długość (w metrach)</th> </tr> </thead> <tbody> <tr> <td>Boa dusiciel</td> <td>27</td> <td>4</td> </tr> <tr> <td>Python birmański</td> <td>90</td> <td>5 do 7</td> </tr> <tr> <td>Anakonda zielona</td> <td>227</td> <td>6 do 9</td> </tr> <tr> <td>Kobra królewska</td> <td>9</td> <td>4</td> </tr> </tbody> </table> <p>A. Kuba zobaczył węża, który miał 8 metrów długości. Jaki to mógł być gatunek węża? B. Naima zobaczyła węża, który miał 6 metrów długości i ważył około 80 kilogramów. Jaki to mógł być gatunek węża?</p>	Gatunek węża	Waga (w kilogramach)	Długość (w metrach)	Boa dusiciel	27	4	Python birmański	90	5 do 7	Anakonda zielona	227	6 do 9	Kobra królewska	9	4	<p>Rysunek pokazuje, ile osób odwiedziło stronę internetową „Znajdź odpowiedź”.</p> <p>Odwiedziny strony internetowej „Znajdź odpowiedź”</p>  <p>Ile osób odwiedziło tę stronę we środę?</p>
Zajęcia	Kreski																								
Gra w piłkę																									
Oglądanie telewizji																									
Czytanie																									
Gatunek węża	Waga (w kilogramach)	Długość (w metrach)																							
Boa dusiciel	27	4																							
Python birmański	90	5 do 7																							
Anakonda zielona	227	6 do 9																							
Kobra królewska	9	4																							
Arytmetyka, 89 (53%) pytań	Geometria, 56 (33%) pytań	Statystyka, 24 (14%) pytania																							

<p>Zastosowania problemowe, 37 (21%) pytań</p>	<p>Rysunek 1</p>  <p>Rysunek 2</p>  <p>Rysunki 1 i 2 przedstawiają te same oczy w różnych warunkach zewnętrznych. Czym różnią się warunki zewnętrzne na obu rysunkach?</p> <p>Ⓐ Na rysunku 1 jest silniejsze oświetlenie. Ⓑ Na rysunku 2 jest silniejsze oświetlenie. Ⓒ Na rysunku 1 jest wyższa temperatura. Ⓓ Na rysunku 2 jest wyższa temperatura.</p>																				
<p>Zastosowania typowe, 67 (38%) pytań</p>	<p>Organizmy przedstawione na rysunku żyją na pustyni. Adam zaczął rysować łańcuch pokarmowy organizmów przedstawionych na rysunku. Umieścił już w łańcuchu trawę i owada, ponieważ wie, że owady żywią się nasionami traw.</p> <p>Uzupełnij łańcuch pokarmowy, wpisując trzy brakujące organizmy.</p> <p>trawa (z nasionami) → owad → _____ → _____</p> 																				
<p>Zasoby przyrodnicze, 72 (41%) pytań</p>	<p>Na rysunku widać cztery zwierzęta. Wpisz w tabeli nazwy tych zwierząt obok środowiska, w którym najczęściej występują.</p> <table border="1" data-bbox="454 1433 598 2004"> <thead> <tr> <th>Środowisko</th> <th>Zwierzę</th> </tr> </thead> <tbody> <tr> <td>Deszczowy las tropikalny</td> <td></td> </tr> <tr> <td>Pustynia</td> <td></td> </tr> <tr> <td>Ocean</td> <td></td> </tr> <tr> <td>Trawiasta sawanna</td> <td></td> </tr> </tbody> </table> 	Środowisko	Zwierzę	Deszczowy las tropikalny		Pustynia		Ocean		Trawiasta sawanna											
Środowisko	Zwierzę																				
Deszczowy las tropikalny																					
Pustynia																					
Ocean																					
Trawiasta sawanna																					
<p>Biologia, 79 (45%) pytań</p>	<p>Marysia przeprowadziła doświadczenie z solą i wodą. Wyniki tego doświadczenia przedstawia tabela.</p> <table border="1" data-bbox="702 224 885 795"> <thead> <tr> <th>Ilość soli</th> <th>Ilość wody</th> <th>Temperatura wody</th> <th>Czy mieszano roztwór?</th> </tr> </thead> <tbody> <tr> <td>15 gramów</td> <td>50 ml</td> <td>25°C</td> <td>Tak</td> </tr> <tr> <td>30 gramów</td> <td>100 ml</td> <td>25°C</td> <td>Tak</td> </tr> <tr> <td>45 gramów</td> <td>150 ml</td> <td>25°C</td> <td>Tak</td> </tr> <tr> <td>60 gramów</td> <td>200 ml</td> <td>25°C</td> <td>Tak</td> </tr> </tbody> </table> <p>Czego Marysia chciała się dowiedzieć z tego doświadczenia?</p> <p>Ⓐ Ile soli rozpuści się w różnej ilości wody. Ⓑ Ile soli rozpuści się w różnej temperaturze. Ⓒ Czy mieszanie przyspiesza rozpuszczanie się soli. Ⓓ Czy mieszanie opóźnia rozpuszczanie się soli.</p>	Ilość soli	Ilość wody	Temperatura wody	Czy mieszano roztwór?	15 gramów	50 ml	25°C	Tak	30 gramów	100 ml	25°C	Tak	45 gramów	150 ml	25°C	Tak	60 gramów	200 ml	25°C	Tak
Ilość soli	Ilość wody	Temperatura wody	Czy mieszano roztwór?																		
15 gramów	50 ml	25°C	Tak																		
30 gramów	100 ml	25°C	Tak																		
45 gramów	150 ml	25°C	Tak																		
60 gramów	200 ml	25°C	Tak																		
<p>Fizyka, 64 (36%) pytań</p>	<p>Marek położył szklaną kulkę na górze stromej torze, jak widać na rysunku.</p>  <p>Kulka stacza się po torze. Podaj nazwę siły, która porusza kulkę.</p>																				
<p>Geografia, 33 (19%) pytań</p>	<p>Rysunek przedstawia budowę Ziemi. Zewnętrzna warstwa nazywa się skorupą. Wymień dwa składniki skorupy Ziemi.</p> 																				
<p>Zastosowania problemowe, 37 (21%) pytań</p>	<p>Dwa magnesy pierścieniowe nasadzono na ołówkę, jak widać na rysunku.</p>  <p>Dołny magnes odpycha górny magnes. Bieguny dolnego magnesu zostały oznaczone.</p> <p>Oznacz bieguny górnego magnesu.</p>																				
<p>Zastosowania typowe, 67 (38%) pytań</p>	<p>Tabela przedstawia pogodę w czterech różnych miejscach.</p> <table border="1" data-bbox="1173 828 1324 1400"> <thead> <tr> <th>Miejsce</th> <th>Temperatura</th> <th>Chmury</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>5 °C</td> <td>są</td> </tr> <tr> <td>B</td> <td>-5 °C</td> <td>nie ma</td> </tr> <tr> <td>C</td> <td>-5 °C</td> <td>są</td> </tr> <tr> <td>D</td> <td>5 °C</td> <td>nie ma</td> </tr> </tbody> </table> <p>W którym miejscu najprawdopodobniej pada śnieg?</p> <p>Ⓐ W miejscu A. Ⓑ W miejscu B. Ⓒ W miejscu C. Ⓓ W miejscu D.</p>	Miejsce	Temperatura	Chmury	A	5 °C	są	B	-5 °C	nie ma	C	-5 °C	są	D	5 °C	nie ma					
Miejsce	Temperatura	Chmury																			
A	5 °C	są																			
B	-5 °C	nie ma																			
C	-5 °C	są																			
D	5 °C	nie ma																			
<p>Zastosowania problemowe, 37 (21%) pytań</p>	<p>Rysunek przedstawia ciała niebieskie: Ziemię, Księżyc i Słońce. Każde ciało jest oznaczone numerem. Strzałki wskazują kierunek, w którym każde ciało się porusza.</p>  <p>Wpisz właściwy numer obok każdego ciała (1, 2 lub 3).</p> <p>Ziemia to ciało numer _____ Księżyc to ciało numer _____ Słońce to ciało numer _____</p>																				
<p>Zastosowania typowe, 67 (38%) pytań</p>	<p>Rysunek przedstawia budowę Ziemi. Zewnętrzna warstwa nazywa się skorupą. Wymień dwa składniki skorupy Ziemi.</p> 																				

- do pytań krótkiej odpowiedzi ocenianych w skali 0–1 – dwuparametrowy (obejmujący trudność i dyskryminację)
- do pytań krótkiej odpowiedzi ocenianych w skali 0–2 – model częściowego uznania (*partial credit*). Model trzyparametrowy ma postać:

$$P(\theta) = c + (1 - c) \frac{e^{Da(\theta - b)}}{1 + e^{Da(\theta - b)}}.$$

W tym wzorze a , b i c to parametry pytania (a – dyskryminacji, b – trudności, c – podatności na zgadywanie), D można uważać za stałą, a e to podstawa logarytmów naturalnych. Jeśli znamy parametry konkretnego pytania, to potrafimy obliczyć funkcję prawdopodobieństwa $P(\theta)$, czyli dowiedzieć się, jakie jest prawdopodobieństwo poprawnej odpowiedzi na to pytanie w zależności od kompetencji badanego.

Losowanie wartości prawdopodobnych (*plausible values*). Załóżmy, że odpowiedzi pewnego ucznia na skalibrowane pytania testu tworzą wektor $\mathbf{u} = u_1, u_2, \dots, u_n$. Na gruncie założenia o lokalnej niezależności pytań łączne prawdopodobieństwo warunkowe wektora odpowiedzi jest iloczynem prawdopodobieństw warunkowych poszczególnych pytań:

$$f(u_1, u_2, \dots, u_n | \theta) = \prod_{j=1}^n P_j^{u_j} Q_j^{1-u_j}$$

Czynnik $P_j = P(U_j | \theta)$ ma zastosowanie do pytań, na które uczeń odpowiedział poprawnie, a czynnik $Q_j = 1 - P(U_j | \theta)$ – do pytań, na które odpowiedział niepoprawnie. Wynikiem testowania tego ucznia jest wartość θ , przy której funkcja $f(\mathbf{u})$ osiąga maksimum.

Niestety, takie punktowe oszacowanie θ jest niezadowolające, ponieważ nie uwzględnia nieuniknionej niepewności, jaką jest obarczony każdy pomiar zmiennych ukrytych. Żeby oszacować tę niepewność i włączyć ją do oszacowania błędów standardowych, stosuje się skomplikowaną metodę wartości prawdopodobnych⁹. Punktem wyjścia w tej metodzie jest rozkład *a posteriori* kompetencji θ osoby, która „wytworzyła” wektor \mathbf{u} , czyli ciąg liczb będących wynikiem jej pracy nad testem (Wu, 2005):

$$P(\theta | \mathbf{u}) = \frac{f(\mathbf{u} | \theta) g(\theta)}{\int f(\mathbf{u} | \theta) g(\theta) d\theta}$$

Ten rozkład mówi, ile wynosi prawdopodobieństwo każdej z możliwych wartości kompetencji θ

w wypadku tej osoby. Należy zwrócić uwagę, że rozkład uwzględnia nie tylko znaną nam już funkcję $f(\mathbf{u} | \theta)$, lecz także funkcję $g(\theta)$, będącą rozkładem θ w populacji, do której należy ta osoba. Rozkład $P(\theta | \mathbf{u})$ jest zatem syntezą wiedzy o wykonaniu testu przez konkretną osobę i wiedzy o przynależności tej osoby do konkretnej populacji.

Powiedzmy, że test wykonuje dwóch uczniów: jeden z małego, drugi z dużego miasta, i obaj uzyskują taki sam wektor odpowiedzi. Gdybyśmy posłużyli się tylko funkcją $f(\mathbf{u} | \theta)$, to obu przypisałibyśmy taki sam wynik θ . Ale średnia θ w populacji uczniów z małych miast jest mniejsza niż w populacji wielkomiejskiej. Szacunkowy wynik ucznia z małego miasta leży powyżej średniej jego populacji, a szacunkowy wynik ucznia z dużego miasta – poniżej średniej jego populacji. Wiadomo, że niepewność wnioskowania o parametrach populacji jest tym większa, im odleglejszych od średniej obszarów dotyczy. Żeby zmniejszyć niepewność, „przesuwamy” indywidualne oszacowania w kierunku średniej; zachowując różnicę, redukujemy jej wielkość. W efekcie średnia rozkładu $P(\theta | \mathbf{u})$ pierwszego ucznia będzie nieco niższa niż maksimum jego funkcji $f(\mathbf{u} | \theta)$, a średnia rozkładu drugiego ucznia – nieco wyższa.

Rozróżnienie dwóch subpopulacji (takie jak małomiasteczkowej i wielkomiejskiej) można włączyć do modelu populacji $g(\theta)$. Przestaje być on wtedy rozkładem normalnym i staje się mieszaniną dwóch rozkładów:

$$g(\theta) \sim N(\mu + \beta y, \sigma^2),$$

gdzie μ jest średnią pierwszego rozkładu, β jest różnicą między średnimi, a σ^2 to (z założenia identyczna) wariancja każdego rozkładu. Zmienna Y , zwana regresorem lub zmienną warunkującą (*conditioning*), przyjmuje tu dwie wartości: 0 – jeśli uczeń żyje w małym mieście, i 1 – jeśli żyje w dużym. Im większa jest wartość β , tym większa będzie korekta rozkładów θ obu uczniów, ale gdy β jest bliska zera, czyli gdy θ jest stochastycznie niezależna od Y , korekty nie ma. Dlatego oskarżanie metody o produkowanie pozornych zależności nie jest uzasadnione.

Uczniowie z powyższego przykładu, zwłaszcza pierwszy z nich, uznałiby korektę za niesprawiedliwą; dlatego tej metody nie można stosować w skalowaniu wyników egzaminu zewnętrznego. Ale gdy celem jest rzetelne oszacowanie parametrów populacji, ze średnią krajową na czele, ta metoda, zwana Bayesowską, sprawdza się doskonale.

Listę zmiennych Y można dowolnie poszerzać. W badaniach IEA włącza się do rachunków płeć, wiek i oddział klasowy ucznia, a także informacje

⁹ Opis metody w języku naturalnym dają Mathias von Davieri in. (2009). O pomiarze cech ukrytych traktuje książka pod redakcją Artura Pokropka (2015).

uzyskane z ankiet ucznia i jego rodziców. Informacje te redukuje się metodą analizy głównych składowych osobno dla każdego kraju. Wykorzystuje się tyle pierwszych składowych, ile potrzeba do odtworzenia 90 proc. zróżnicowania wszystkich zmienionych w analizie. Ostatecznie dla każdego ucznia uzyskuje się rozkład θ uwzględniający zarówno wektor jego odpowiedzi (wraz z parametrami pytań), jak i wartości wszystkich zmiennych warunkujących (wraz z ich współczynnikami regresji w populacji). Z tego rozkładu pobiera się losowo pięć wartości – takich, które są prawdopodobnymi oszacowaniami jego kompetencji θ . Te pięć wartości łącznie reprezentują kompetencję tego ucznia we wszystkich dalszych analizach, a także mieszczą w sobie informację o niepewności towarzyszącej wszelkim oszacowaniom punktowym.

Szacowanie błędów. Jak wiadomo, wyniki każdego badania metodą reprezentacyjną z użyciem narzędzi mierzących zmienne nieobserwowalne są obciążone dwojakim błędem. Pierwszy to błąd próbkowania, czyli różnica między statystyką z próby a parametrem w populacji. Byłoby naiwnością zakładać, że średni wynik testowania niespełnia 5 tys. uczniów, którzy wzięli udział w polskim badaniu, będzie identyczny ze średnim wynikiem wszystkich 360 tys. uczniów klasy czwartej w Polsce. Nigdy, rzecz jasna, nie dowiemy się, ile wynosi średnia w tej populacji, ale możemy się dowiedzieć, z jak dużymi odchyleniami średniej z próby od średniej w populacji powinniśmy się liczyć.

Wzór na wielkość błędu oszacowania średniej μ w populacji na podstawie średniej m z próby można znaleźć w każdym podręczniku statystyki:

$$\hat{\sigma}_\mu = \frac{s}{\sqrt{n}}$$

W tym wzorze s to odchylenie standardowe zmiennej, której średnią chcemy oszacować, a n to liczba próbek w próbie. Niestety, ten prosty wzór ma zastosowanie jedynie do próbek pobranych w losowaniu nieograniczonym indywidualnym. W badaniach IEA stosuje się wielostopniowe losowanie warstwowo-zespołowe, toteż błędy próbkowania szacuje się inaczej – metodą wielokrotnego pobierania podpróbek z próby (*jackknife repeated replication*).

Wszystkie szkoły z krajowej próby dzieli się na 75 stref, po dwie podobne szkoły w każdej. Zaczynając od pierwszej strefy, dane z jednej szkoły się usuwa, a z drugiej podwaja. Osiąga się to przez modyfikację wag: wagi uczniów z pierwszej szkoły zostają utożsamione z zerem, a wagi uczniów z drugiej – podwojone. Z tak spreparowanej próby oblicza się średnią

ważoną m_1 i powtarza całą operację dla pozostałych stref. Wariancja próbkowania jest dana wzorem:

$$s_m^2 = \sum_{r=1}^R (m_r - m)^2$$

w którym R to liczba powtórzeń (75), a m to średnia w całej próbie. Błąd próbkowania średniej to pierwiastek kwadratowy z tej wartości. Pamiętajmy, że każdy uczeń zamiast jednego wyniku ma pięć wyników prawdopodobnych. Jeśli zastosujemy metodę *jackknife* do każdego wyniku z osobna, otrzymamy pięć oszacowań wariancji próbkowania: od s_1^2 do s_5^2 . Można udowodnić, że najlepszym estymatorem wariancji próbkowania jest średnia z tych pięciu wariancji. W podobny sposób oblicza się błędy innych estymatorów.

Drugi rodzaj błędu to błąd pomiaru, czyli różnica między oszacowaną a prawdziwą wartością kompetencji osoby badanej. By określić jego wielkość, oblicza się średnią z każdej wartości możliwej: od m_1 do m_5 . Można udowodnić, że najlepszym estymatorem wariancji pomiaru jest wariancja tych pięciu średnich.

Łączny błąd standardowy oszacowania średniej w populacji to pierwiastek kwadratowy z wyrażenia:

$$s_m^2 = \frac{\sum_{i=1}^n s_i^2}{n} + \left(1 + \frac{1}{n}\right) \cdot \frac{\sum_{i=1}^n (m_i - m)^2}{n-1}$$

Symbol n oznacza tu liczbę wartości prawdopodobnych (5). Tak obliczone wartości znajdują się w tabelach 2.1 i 3.1 w nawiasie przy średniej każdego kraju.

O pożytku z wartości prawdopodobnych łatwo się przekonać – wystarczy porównać oszacowania średniej i odchylenia standardowego wyników testu matematycznego. Obliczone z użyciem wartości prawdopodobnych parametry te wynoszą w Polsce, odpowiednio, 534,8 i 71,3. W zaokrągleniu do liczb całkowitych podaje je Tabela 2.1. Gdybyśmy utożsamili wynik każdego ucznia ze średnią jego pięciu wartości prawdopodobnych, to średnia krajowa wynosiłaby nadal 534,8, ale odchylenie standardowe – tylko 67,8. Łatwo zrozumieć, skąd bierze się ta różnica: punktowe oszacowanie kompetencji nie uwzględnia nieuniknionego „rozmycia” indywidualnych oszacowań wskutek błędu pomiaru.

Transformacja skali. Wartości prawdopodobne są wyrażone w skali θ o średniej 0 i odchyleniu standardowym 1, takiej samej w każdej edycji badania. Z punktu widzenia drugiego celu badania taka skala jest bezużyteczna. Żeby porównywać ze sobą kolejne roczniki uczniów z tego samego kraju, trzeba ich wyniki umieścić na tej samej skali. Łatwo spełnić ten

warunek: wystarczy stosować ciągle ten sam test. To jednak zablokowałoby modernizację testów w miarę postępu teorii pomiaru osiągnięć. Badania dzieci testem sprzed 10 lat nie dałoby się obronić. Dlatego w każdej nowej edycji badania usuwa się część starych zadań i dodaje nowe, zawsze jednak pozostawiając część wspólną. Dzięki części wspólnej, czyli zadaniom „łącznikowym”, w każdej edycji badania można przekształcić nową skalę θ w skalę wzorcową, zdefiniowaną w badaniu TIMSS 1995. Żeby pozbyć się liczb ujemnych i części dziesiętnych, średnią tej skali ustalono na 500, a odchylenie standardowe na 100.

Jak przekłada się rozkład wyników z jednej skali na inną? Czytelnik pamięta, że kalibrację testu prowadzi się łącznie na zbiorze pytań z obecnej i poprzedniej edycji badania. Na podstawie nowo skalibrowanych pytań oblicza się nowy rozkład θ na danych z poprzedniej edycji i porównuje się go ze starym (z założenia w skali wzorcowej). Sens tego porównania można wysłowić pytaniem: Jak wypadłyby wyniki tych samych dzieci, gdyby wykonały oba testy – z poprzedniej i obecnej edycji? Jest jasne, że każda różnica w średnich odzwierciedlałaby wyłącznie różnicę trudności (i innych parametrów) pytań obu testów. Tę różnicę łatwo usunąć: wystarczy znaleźć funkcję liniową, która przekształca nowy rozkład w stary. Obliczywszy funkcję, stosujemy ją do rozkładu θ w obecnej edycji, a to daje nam pewność, że pomiary w tej i poprzedniej edycji są wyrażone w tej samej skali. Różnica między średnimi obu pomiarów w danym kraju jest miarą zmiany, jaka zaszła między dwoma rocznikami uczniów w tym kraju.

Ankiety

Pozostaje omówić drugi rodzaj narzędzi pomiaru – ankiety. Dane, których dostarczają, dają wgląd w kontekst osiągnięć szkolnych.

1. Dyrektor szkoły wypełniał ankietę o zasobach swojej szkoły, zasadach pracy dydaktycznej i wychowawczej oraz o własnej roli zawodowej.
2. Nauczyciele badanych oddziałów klasowych wypełniali ankietę o kulturze swojej szkoły, organizacji swojego oddziału, programie i metodach nauczania matematyki i przyrody, sposobach oceniania osiągnięć, a także o swoim wykształceniu i karierze zawodowej.
3. Rodzice lub prawni opiekunowie uczniów objętych badaniem wypełniali ankietę dotyczącą rozwoju dziecka, wczesnej edukacji domowej, zasobów wspierających uczenie się dziecka (z obowiązkowym we wszystkich badaniach osiągnięć pytaniem o liczbę książek w domu) i opinii o szkole ich dziec-

ka. Ważną częścią ankiety były pytania o wykształcenie, zawód i pozycję na rynku pracy – pozwalają one zbudować wskaźnik statusu socjoekonomicznego (SES) rodziny.

4. Uczeń odpowiadał na pytania ankiety o zasoby domowe wspierające uczenie się, zwyczaje panujące w jego szkole, a także metody nauczania matematyki i przyrody. Ważną częścią ankiety stanowiły skale postaw wobec szkoły oraz nauczania i uczenia się matematyki i przyrody.

Ponadto koordynatorzy krajowi opracowywali charakterystyki swoich systemów oświatowych i opisywali programy i praktykę edukacji w szkole podstawowej. Te informacje złożyły się na odrębne wydawnictwo: *TIMSS 2015 Encyclopedia: Educational policy and curriculum in mathematics and science* (Mullis i in., 2016), dostępne na stronie internetowej <http://timssandpirls.bc.edu>.

Osiągnięcia matematyczne

Krzysztof Konarzewski

Wyniki pomiaru osiągnięć szkolnych często sprowadza się do jednej liczby: średniej arytmetycznej. Średnie pozwalają uporządkować zbiorowości, dla których zostały obliczone, czyli utworzyć ich ranking. Ale ranking oparty jedynie na średniej mówi mniej, niżby się zdawało.

Po pierwsze, nie informuje o wielkości różnic między porównywanymi szkołami, a przecież różnica między szkołami na pierwszej i drugiej pozycji rankingu może być tak mała, że da się ją przypisać zwykłemu przypadkowi – wtedy chełpić się pierwszym miejscem nie wypada. Po drugie, ranking nic nie mówi o zróżnicowaniu wyników wewnątrz zbiorowości, a przecież jest ważne, czy na wysoką średnią zapracowała większość uczniów, czy głównie niewielka grupka olimpijczyków. No i oczywiście ranking ignoruje stan początkowy, czyli kapitał wykształcenia, z jakim uczniowie wchodzi do szkoły.

Pełniejsza wiedza o badanych zbiorowościach kryje się w rozkładach wyników testowania. Dlatego omawianie osiągnięć matematycznych badanych populacji zaczniemy wprawdzie od średnich i rang poszczególnych krajów, ale zaraz przejdziemy do miar zróżnicowania ich rozkładów: odchylenia standardowego, punktów centylowych i progów wykonania. Na zakończenie zapoznamy się ze zmianami osiągnięć matematycznych, jakie zaszły w poszczególnych krajach od 1995 r. Polski w tym zestawieniu nie ma, bo wprawdzie nasz kraj po raz pierwszy wziął udział w badaniu TIMSS w 2011 r., ale w obecnej edycji

poddał pomiarom dzieci nieco starsze i z wyższej niż poprzednio klasy, więc porównanie wyników byłoby niemiarodajne.

Średnie i rangi

Podstawowe wyniki testowania osiągnięć matematycznych przedstawia Tabela 2.1. W trzeciej kolumnie znajdują się oszacowania średnich krajowych. Według nich uporządkowano kraje – od Singapuru (z najwyższą średnią) do Kuwejtu (z najniższą). Średnia osiągnięć matematycznych polskich dzieci wynosi 535 punktów.

Czy dokładnie 535 punktów? Tego nie możemy być pewni, ponieważ każde oszacowanie parametru populacyjnego jest obciążone błędem zdefiniowanym w Rozdziale 1. Tabela podaje wielkość tego błędu w nawiasie obok średniej. Znając błąd, możemy wyznaczyć przedział, który obejmuje (z prawdopodobieństwem 0,95) średnią populacji krajowej. Ilustruje go czarny prostokąt na rozkładzie centylowym. Prostokąt jest zbudowany według prostej reguły: lewy kraniec to punktowe oszacowanie średniej minus podwojony błąd standardowy, prawy – plus podwojony błąd. Widzimy, że w szerokości czarnego prostokąta przoduje Bułgaria. Im szerszy jest ten prostokąt, tym więcej porównań danego kraju ze światem (jeśli wolno tak nazwać nielosową próbę 49 krajów) i z poszczególnymi krajami nie znajduje rozstrzygnięcia.

Wyniki pierwszego z tych porównań – istotność różnic między średnimi krajowymi a średnią międzynarodowego rozkładu wzorcowego ustalonego w pierwszym badaniu TIMSS w 1995 r. – znajdują się w drugiej kolumnie Tabeli 2.1. Kraje, których średnia jest istotnie wyższa, są oznaczone strzałką skierowaną ku górze, a kraje, których średnia jest istotnie niższa – ku dołowi. Brak strzałki oznacza nieistotność różnicy. Średnia Polski jest istotnie wyższa od średniej wzorcowej.

Porównania drugiego rodzaju dotyczą par krajów. Wskutek nieprecyzyjnego pomiaru osiągnięcia bułgarskich dzieci są statystycznie nieodróżnialne od osiągnięć dzieci z 12 innych krajów. Bahrajn i Holandia, które mogą się pochlubić najmniejszym błędem standardowym, mają średnie nieodróżnialne odpowiednio tylko od jednego i pięciu krajów. Błąd naszego oszacowania jest nieco większy i nie pozwala ustalić naszej pozycji względem ośmiu krajów. Istotnie wyprzedziło nas 12 krajów – od Singapuru do Belgii oraz Portugalia. Za nami zostało 28 krajów – Czechy i wszystkie kraje w tabeli od Bułgarii do Kuwejtu.

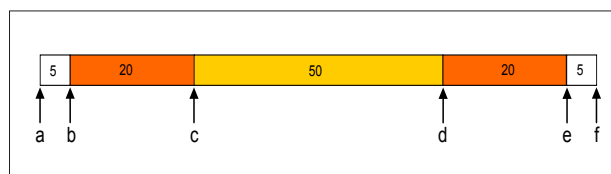
Czy stąd, że wyprzedziło nas 12 krajów, wynika, że zajmujemy 13. miejsce w rankingu? Nie, ponieważ często jest więcej krajów mogących aspirować do tego miejsca. Lepszą miarą położenia jest ranga, czyli miejsce w uporządkowaniu średnich osiągnięć od największej do najmniejszej. Jeśli kilka krajów ma taką samą średnią, to wszystkim przypisuje się tę samą rangę równą medianie ich liczb porządkowych. Wśród 48 innych krajów Polska ma 17. rangę i dzieli ją z Finlandią i Litwą.

Trzeba dodać, że ranga ma charakter względny, zależy bowiem od tego, które kraje uczestniczą w badaniu. Jeśli w którejś edycji nie bierze udziału kraj z czołówki tabeli, to pozycje wszystkich krajów o niższych osiągnięciach wzrastają, choćby nic się zmieniło w ich praktyce kształcenia. Ważna jest też liczba krajów – określona ranga, np. dziesiąta, ma inne znaczenie w zbiorze 20 niż 40 krajów. Celem badania IEA nie jest ranking krajów. Gdyby badanie miało pełnić tę funkcję, musiałoby dostosowywać wyniki testowania do rozmaitych zmiennych ubocznych, np. do średnich wieku populacji krajowych – bo jak zobaczymy, wiek dzieci jest związany z ich osiągnięciami. Kraje biedniejsze mogłyby zażądać handicapu dla krajów bogatszych, kraje podzwrotnikowe – poprawki na upał... i jasna interpretacja wyniku pomiaru zostałaby bezpowrotnie utracona. Szowinistom, dla których każda dziedzina życia publicznego jest polem rywalizacji międzynarodowej, wypada przypomnieć, że społeczeństwa nie po to utrzymują swoje systemy oświaty, by nie dać się wyprzedzić innym społeczeństwom.

Miary zróżnicowania

Pierwszą miarą zróżnicowania jest odchylenie standardowe, czyli pierwiastek kwadratowy ze średniej kwadratów odległości każdego wyniku od średniej krajowej. Odchylenie standardowe powinno się podawać, ilekroć podaje się średnią arytmetyczną. W Tabeli 2.1 znajduje się w czwartej kolumnie. Widać, że najmniej zróżnicowane populacje mają – podobnie jak w 2011 r. – Holandia, Belgia i Niemcy, najbardziej zaś – Jordania, Zjednoczone Emiraty Arabskie i Kuwejt.

Druga statystyka – punkty centylowe – pozwala głębiej wniknąć w naturę zróżnicowania populacji. Każdy rozkład krajowy można przedstawić jak na Rysunku 2.1.



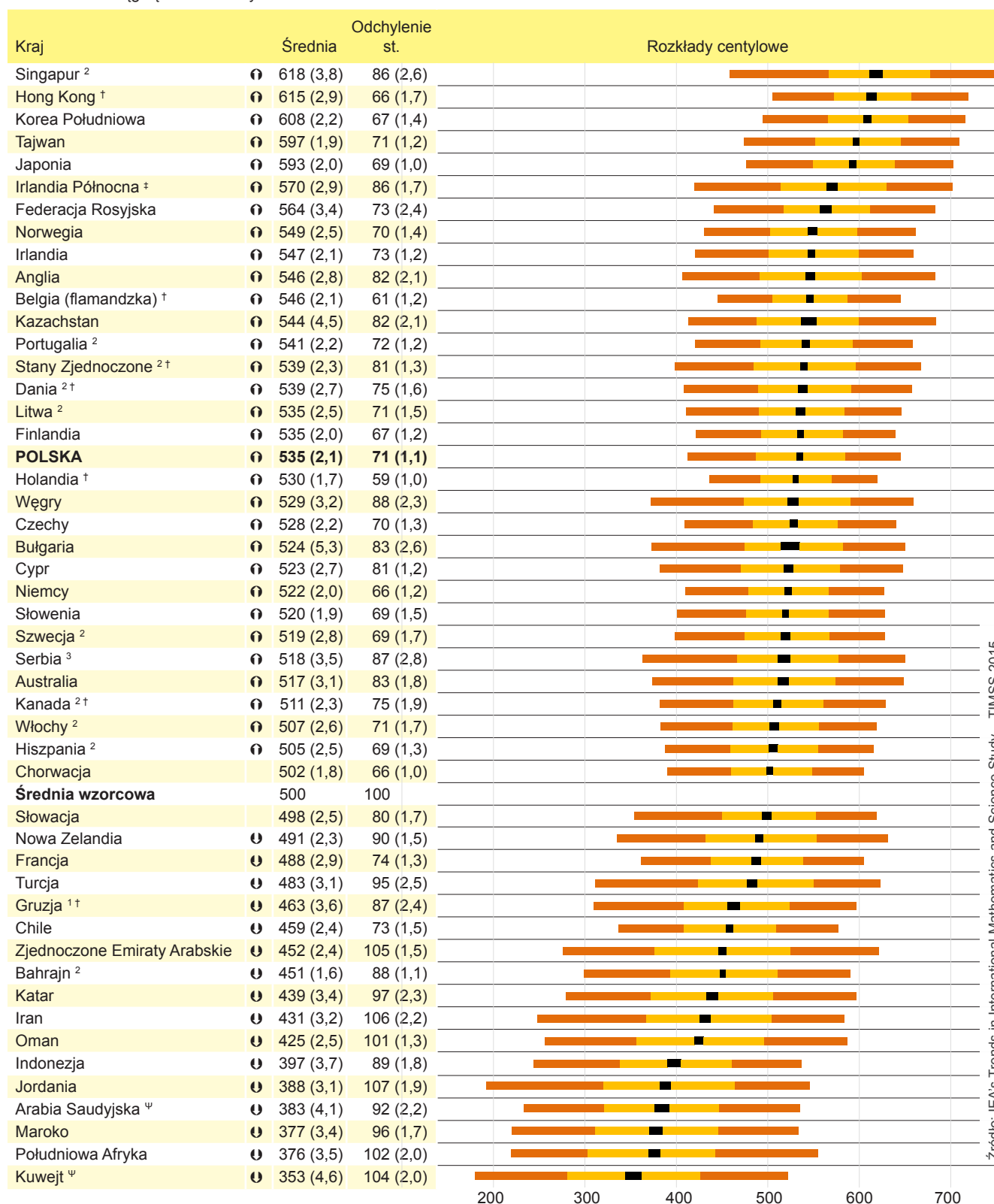
Rysunek 2.1. Punkty centylowe

Punkt a to najniższy, a punkt f – najwyższy wynik w próbie krajowej. Tabela ich nie podaje, bo dla interpretacji nie mają żadnego znaczenia. Znaczenie ma natomiast punkt b – wynik na skali, który odcina pięć procent najniższych wyników. Wiedząc, gdzie leży ten punkt, wiemy, jak słabi są najsłabsi uczniowie w danym kraju. Widzimy np., że na Węgrzech dzieci wykonały test tylko trochę (statystycznie nieistotnie) gorzej niż w Polsce, ale najsłabsi uczniowie węgierscy nie przekraczają granicy 372 punktów, są więc znacznie słabsi niż najsłabsi polscy (412 punktów). Na drugim krańcu mamy punkt e : mówi on, gdzie zaczynają się wyniki pięciu procent najmocniejszych uczniów. Nasi najmocniejsi mają powyżej 645 punktów i ustępują najmocniejszym uczniom węgierskim (powyżej 660), ale i im daleko daleko do dzieci z Singapuru (powyżej 746).

Wartości b i e zależą oczywiście od położenia całego rozkładu na osi wyników, ale łatwo się pozbyć tej zależności: wystarczy wziąć odległości między nimi a punktem środkowym rozkładu, czyli medianą (M). Te odległości można uznać za miary dyskryminacji oświatowej: negatywnej ($M - b$) i pozytywnej ($e - M$).

Najmniejsza dyskryminacja negatywna panuje w Holandii (95 punktów, ranga 1), a największa w Jordanii (208 punktów, ranga 49). Polska i Włochy mają rangę 18,5, co daje im miejsce wśród krajów o przeciętnej dyskryminacji tego rodzaju. Z najmniejszą dyskryminacją pozytywną spotykamy się znów w Holandii (89 punktów), a z największą – w RPA

Tabela 2.1. Osiągnięcia matematyczne



Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

W nawiasach błędy standardowe. z powodu zaokrąglenia liczb niektóre wyniki mogą wydawać się niespójne.

ⓘ Średnia krajowa istotnie wyższa niż średnia wzorcowa

Ⓜ Średnia krajowa istotnie niższa niż średnia wzorcowa

^ψ Średnia osiągnięć mało wiarygodna, ponieważ odsetek uczniów uzyskujących wyniki nieprzewyższające poziomu zgadywania leży w przedziale 15–25.

¹ Definicja populacji krajowej nie uwzględnia pełnej definicji międzynarodowej.

² Operat losowania pokrywa 90–95% populacji krajowej.

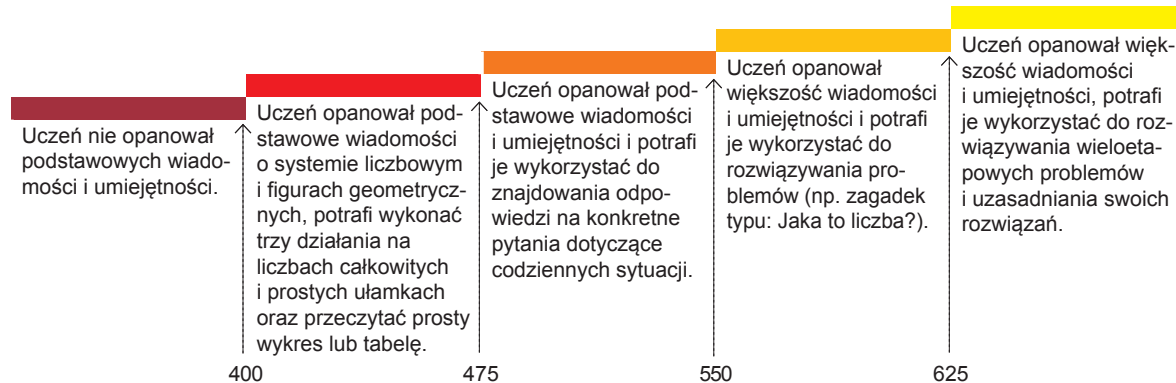
³ Operat losowania pokrywa mniej niż 90% (ale przynajmniej 77%) populacji krajowej.

[†] Poziom realizacji próby spełnia wymaganie jedynie po wykorzystaniu zapasowych szkół.

[‡] Poziom realizacji próby niemal spełnia wymaganie jedynie po wykorzystaniu zapasowych szkół.

Tabela 2.2. Poziomy osiągnięć matematycznych: definicje i procenty uczniów, którzy się na nich znaleźli

Świat 2011 ^a	10	21	41	24	4
Świat 2015 ^a	7	18	39	30	6
Polska 2011 ^b	10	21	41	24	4
Polska 2015 ^c	4	16	36	34	10



^a Przez „świat” należy rozumieć kraje uczestniczące w każdej edycji badania. Ponieważ kraje nie zostały dobrane losowo, nie można użyć ich wyników do oszacowania parametrów populacji krajów na Ziemi. Stopień pokrywania się list krajów z lat 2011 i 2015 można ocenić na podstawie Tabeli 2.8.

^b Uczniowie klasy trzeciej.

^c Uczniowie klasy czwartej.

(185 punktów). Polska i trzy inne kraje (Hiszpania, Litwa i Irlandia) mają rangę 10,5, co lokuje je w grupie niskiej dyskryminacji.

Na koniec zauważmy, że różnica szerokości lewego i prawego czerwonego prostokąta jest przybliżoną miarą skośności rozkładu. W Europie rozkłady węgierski, bułgarski i serbski mają wydłużony ogon w lewej (niższej) części rozkładu. Może to świadczyć o istnieniu w tych krajach jakichś zaniedbanych oświatowo mniejszości etnicznych.

Trzecia statystyka to progi osiągnięć. Na ustalonej od 1995 r. skali wyników TIMSS określono cztery wartości progowe: 400, 475, 550 i 625, które pozwalają zdefiniować pięć przedziałów. Autorzy testów IEA dokładają starań, by każdy przedział miał własną charakterystykę jakościową, czyli by mówił o uczniu więcej, niż tylko ile zadań uczeń wykonał poprawnie. Te charakterystyki są kumulatywne: uczeń, który pokonał dany próg, umie wszystko, co umieją uczniowie na niższych progach, oraz coś, czego oni już nie umięją (Tabela 2.2). Jeśli ponumerujemy przedziały od 1 do 5, nawiążemy do zwyczajowej skali stopni szkolnych w Polsce od jedynki do piątki. Jedynekę w TIMSS dostaje uczeń, który nie przekroczył progu 400 punktów, dwójkę – który nie przekroczył 475 punktów itd.

Tabela 2.2 informuje o procentach uczniów, którzy znaleźli się na każdym poziomie. Widać postęp na świecie: w ciągu czterech lat procent „jedykowiczów” zmniejszył się z 10 do 7, a procent „piątkowiczów” wzrósł z 4 do 6. W Polsce zmiana na lepsze jest jeszcze większa, ale chęć się nią nie należy, ponieważ w 2015 r. testowaliśmy dzieci zarazem starsze i kształcone o rok dłużej niż w 2011 r.

Liczby w Tabeli 2.3 pokazują to samo w inny, mniej intuicyjny sposób: mówią mianowicie, ile dzieci (w procentach) przekroczyło kolejne progi, choć to, ile przekroczyło, wydaje się mniej ważne niż to, ile z nich utknęło za progiem i nie poszło dalej. Tabela porządkuje kraje według odsetka dzieci, które przekroczyły najwyższy próg – od Singapuru do czterech krajów arabskojęzycznych i Indonezji, w których takich dzieci nie było. W tym rankingu pozycja Polski jest odrobinę niższa niż w rankingu według średnich osiągnięć.

Osiągnięcia szczegółowe

Test osiągnięć matematycznych ma złożoną strukturę. Pora zobaczyć, jak wypadły wyniki jego części (subtestów). Przedstawiają je tabele 2.4 i 2.5. Dla każdego subtestu tabele podają wartość średniej oraz różnicę między tą wartością a średnią krajową z całego testu.

Zacznijmy od klasyfikacji zadań według treści przedmiotowej. Dane w Tabeli 2.4 wykazują, że profil polskiej populacji jest wybitnie płaski – w żadnym subteście wyniki nie odbiegają od średniej z całego testu. Podobnie płaski profil mają tylko trzy inne kraje z czołówki rankingu. Jeśli z osiągnięć uczniów wnioskować o programach nauczania, to widać, że i u nich, i u nas są one treściowo zrównoważone. W pozostałych krajach mamy do czynienia z przewagą któregoś działu. W Singapurze i Gruzji dominuje wiedza o liczbie, w Belgii i Danii – wiedza o figurach i miarach itd.

W klasyfikacji według funkcji podmiotowej (Tabela 2.5) profil polskich uczniów przypomina profil

Tabela 2.3. Progi i poziomy osiągnięć matematycznych

Kraj	Odsetki uczniów, którzy pokonali próg				Odsetki uczniów na każdym poziomie
	625	550	475	400	
Singapur ²	50 (2,1)	80 (1,7)	93 (0,9)	99 (0,3)	
Hong Kong [†]	45 (2,0)	84 (1,3)	98 (0,4)	100 (0,1)	
Korea Południowa	41 (1,3)	81 (1,0)	97 (0,4)	100 (0,1)	
Tajwan	35 (1,4)	76 (1,0)	95 (0,4)	100 (0,2)	
Japonia	32 (1,1)	74 (1,0)	95 (0,4)	99 (0,1)	
Irlandia Północna [‡]	27 (1,3)	61 (1,5)	86 (1,1)	97 (0,6)	
Federacja Rosyjska	20 (1,8)	59 (1,8)	89 (1,1)	98 (0,4)	
Anglia	17 (1,2)	49 (1,5)	80 (1,2)	96 (0,7)	
Kazachstan	16 (1,8)	47 (2,6)	80 (1,5)	96 (0,5)	
Stany Zjednoczone ^{2†}	14 (0,8)	47 (1,1)	79 (1,0)	95 (0,5)	
Irlandia	14 (1,1)	51 (1,6)	84 (1,0)	97 (0,4)	
Norwegia	14 (1,1)	50 (1,6)	86 (1,0)	98 (0,4)	
Węgry	13 (0,9)	44 (1,5)	75 (1,5)	92 (0,9)	
Portugalia ²	12 (0,9)	46 (1,3)	82 (1,1)	97 (0,4)	
Dania ^{2†}	12 (0,9)	46 (1,6)	80 (1,3)	96 (0,6)	
Serbia ³	10 (0,8)	37 (1,4)	72 (1,6)	91 (1,2)	
Bułgaria	10 (1,3)	40 (2,6)	75 (2,1)	92 (1,3)	
Litwa ²	10 (0,9)	44 (1,5)	81 (1,1)	96 (0,5)	
POLSKA	10 (0,7)	44 (1,4)	80 (1,0)	96 (0,4)	
Belgia (flamandzka) [†]	10 (0,8)	47 (1,5)	88 (0,9)	99 (0,3)	
Cypr	10 (0,7)	39 (1,5)	74 (1,3)	93 (0,6)	
Australia	9 (0,8)	36 (1,6)	70 (1,3)	91 (0,9)	
Finlandia	8 (0,7)	43 (1,3)	82 (1,0)	97 (0,4)	
Czechy	8 (0,7)	38 (1,4)	78 (1,1)	96 (0,5)	
Nowa Zelandia	6 (0,5)	26 (0,9)	59 (1,2)	84 (0,9)	
Słowenia	6 (0,5)	34 (1,4)	75 (1,2)	95 (0,5)	
Kanada ^{12†}	6 (0,5)	31 (1,1)	69 (1,2)	92 (0,8)	
Niemcy	5 (0,5)	34 (1,3)	77 (1,1)	96 (0,6)	
Szwecja ²	5 (0,5)	34 (1,6)	75 (1,6)	95 (0,8)	
Emiraty Arabskie	5 (0,4)	18 (0,8)	42 (1,0)	68 (0,9)	
Turcja	5 (0,5)	25 (1,2)	57 (1,3)	81 (1,1)	
Włochy ²	4 (0,5)	28 (1,3)	69 (1,4)	93 (0,8)	
Słowacja	4 (0,4)	26 (1,1)	65 (1,4)	88 (0,9)	
Holandia [†]	4 (0,6)	37 (1,3)	83 (1,0)	99 (0,3)	
Hiszpania ²	3 (0,4)	27 (1,1)	68 (1,4)	93 (0,8)	
Chorwacja	3 (0,4)	24 (1,1)	67 (1,2)	93 (0,6)	
Katar	3 (0,5)	13 (1,1)	36 (1,4)	65 (1,4)	
Francja	2 (0,3)	21 (1,3)	58 (1,8)	88 (1,0)	
Gruzja ^{1†}	2 (0,6)	15 (1,4)	47 (1,7)	78 (1,6)	
Oman	2 (0,3)	11 (0,6)	32 (1,1)	60 (1,1)	
Bahrajn ²	2 (0,2)	13 (0,5)	41 (0,8)	73 (0,8)	
Iran	1 (0,3)	11 (0,7)	36 (1,1)	65 (1,4)	
Południowa Afryka	1 (0,3)	5 (0,7)	17 (1,0)	39 (1,4)	
Chile	1 (0,2)	10 (0,7)	42 (1,4)	78 (1,5)	
Arabia Saudyjska ^ψ	0 (0,2)	3 (0,7)	16 (1,2)	43 (1,7)	
Maroko	0 (0,1)	3 (0,5)	17 (1,1)	41 (1,6)	
Jordania	0 (0,1)	5 (0,6)	21 (1,1)	50 (1,2)	
Indonezja	0 (0,1)	3 (0,4)	20 (1,2)	50 (1,8)	
Kuwejt ^ψ	0 (0,1)	3 (0,5)	12 (1,2)	33 (1,7)	
Międzynarodowa mediana	6	36	75	93	

Kolejność i kolory poziomów jak w Tabeli 2.2.
Znaczenia symboli jak w Tabeli 2.1.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 2.4. Wyniki pomiaru osiągnięć matematycznych według treści przedmiotowej

Kraj	Liczby		Obiekty geometryczne		Przedstawianie danych	
	Średnia	Różnica*	Średnia	Różnica*	Średnia	Różnica*
Singapur ²	630 (4,1)	12 (1,1) \circ	608 (4,2)	-10 (1,5) \ominus	600 (4,1)	-18 (1,7) \ominus
Hong Kong [†]	616 (3,0)	2 (1,4)	617 (3,4)	2 (1,9)	610 (3,9)	-4 (2,9)
Korea Południowa	610 (2,6)	2 (1,4)	610 (2,3)	3 (1,8)	606 (2,6)	-2 (1,3)
Tajwan	599 (1,8)	3 (1,1) \circ	597 (2,9)	0 (2,1)	591 (2,1)	-6 (1,3) \ominus
Japonia	592 (1,9)	-1 (1,0)	601 (2,4)	9 (1,3) \circ	593 (2,6)	0 (1,3)
Irlandia Północna [‡]	574 (3,1)	4 (1,0) \circ	566 (3,3)	-4 (2,0) \ominus	566 (3,7)	-4 (2,4)
Federacja Rosyjska	567 (3,3)	3 (1,2) \circ	557 (4,3)	-7 (1,5) \ominus	572 (3,6)	8 (1,1) \circ
Norwegia	542 (2,4)	-7 (1,1) \ominus	559 (3,5)	10 (1,8) \circ	565 (3,0)	16 (1,2) \circ
Irlandia	551 (2,3)	3 (1,6)	542 (2,6)	-6 (1,7) \ominus	548 (2,9)	0 (2,3)
Anglia	547 (3,2)	1 (1,6)	542 (3,3)	-4 (1,6) \ominus	552 (3,2)	6 (2,0) \circ
Belgia (flamandzka) [†]	543 (2,1)	-2 (0,8) \ominus	563 (2,3)	18 (1,3) \circ	523 (3,0)	-23 (2,4) \ominus
Kazachstan	552 (4,0)	7 (1,3) \circ	540 (5,8)	-5 (2,0) \ominus	524 (5,3)	-20 (2,1) \ominus
Portugalia ²	541 (2,1)	-1 (0,9)	539 (2,6)	-2 (1,0) \ominus	546 (2,8)	5 (1,9) \circ
Stany Zjednoczone ^{2†}	546 (2,2)	7 (0,8) \circ	525 (2,6)	-14 (0,8) \ominus	540 (2,8)	1 (2,1)
Dania ^{2†}	535 (2,7)	-4 (1,4) \ominus	555 (3,2)	16 (1,5) \circ	526 (3,4)	-13 (2,3) \ominus
Finlandia	532 (2,1)	-3 (1,0) \ominus	539 (2,5)	4 (1,7) \circ	541 (3,4)	6 (2,6) \circ
Litwa ²	538 (2,6)	3 (1,1) \circ	525 (3,1)	-10 (2,2) \ominus	540 (3,6)	4 (2,3)
POLSKA	534 (2,3)	0 (1,1)	534 (2,5)	-1 (1,7)	538 (2,8)	3 (2,0)
Holandia [†]	532 (2,1)	2 (1,4)	522 (1,9)	-8 (1,2) \ominus	538 (3,4)	9 (2,6) \circ
Węgry	531 (3,0)	2 (0,9) \circ	536 (3,6)	7 (1,6) \circ	512 (3,6)	-17 (1,2) \ominus
Czechy	528 (2,4)	0 (1,1)	531 (2,5)	3 (0,9) \circ	525 (3,0)	-4 (1,7) \ominus
Bułgaria	530 (4,6)	5 (1,4) \circ	525 (5,9)	0 (2,0)	504 (7,6)	-20 (3,1) \ominus
Cypr	529 (2,5)	6 (0,9) \circ	523 (2,9)	0 (1,3)	507 (3,8)	-16 (2,6) \ominus
Niemcy	515 (2,1)	-6 (0,9) \ominus	531 (2,5)	9 (1,5) \circ	534 (2,6)	13 (1,4) \circ
Słowenia	512 (1,8)	-8 (0,9) \ominus	530 (2,1)	10 (1,6) \circ	540 (3,2)	20 (2,3) \circ
Szwecja ²	514 (2,6)	-4 (1,4) \ominus	523 (3,3)	4 (1,7) \circ	529 (3,9)	10 (2,8) \circ
Serbia ³	524 (3,4)	6 (1,0) \circ	502 (3,8)	-16 (1,8) \ominus	516 (3,8)	-2 (2,3)
Australia	509 (3,1)	-8 (0,7) \ominus	527 (3,3)	10 (1,6) \circ	532 (3,6)	15 (2,1) \circ
Kanada ^{12†}	503 (2,2)	-7 (0,9) \ominus	517 (2,4)	6 (0,7) \circ	528 (2,5)	18 (1,0) \circ
Włochy ²	510 (2,4)	3 (0,9) \circ	503 (2,8)	-4 (1,0) \ominus	498 (2,9)	-9 (1,6) \ominus
Hiszpania ²	505 (2,5)	-1 (1,0)	503 (2,8)	-3 (1,5)	508 (3,1)	3 (1,5) \circ
Chorwacja	498 (1,8)	-4 (1,1) \ominus	512 (2,3)	10 (1,5) \circ	498 (3,0)	-5 (2,1) \ominus
Słowacja	502 (2,4)	4 (1,6) \circ	491 (2,5)	-8 (1,2) \ominus	496 (3,9)	-3 (2,6)
Nowa Zelandia	486 (2,7)	-5 (1,0) \ominus	489 (2,8)	-2 (1,9)	506 (2,9)	15 (2,0) \circ
Francja	484 (3,0)	-4 (1,7) \ominus	503 (3,0)	15 (2,0) \circ	475 (3,1)	-13 (1,7) \ominus
Turcja	489 (3,2)	6 (1,2) \circ	475 (3,0)	-9 (0,9) \ominus	476 (3,4)	-8 (1,3) \ominus
Gruzja ^{1†}	483 (3,6)	20 (1,1) \circ	428 (4,6)	-35 (2,2) \ominus	434 (4,4)	-29 (1,9) \ominus
Chile	456 (2,7)	-3 (1,2) \ominus	459 (3,0)	1 (1,8)	463 (3,2)	4 (2,2)
Zjednoczone Emiraty Arabskie	456 (2,3)	4 (0,8) \circ	441 (2,7)	-10 (0,8) \ominus	453 (2,4)	1 (0,8)
Bahrajn ²	453 (1,6)	2 (0,9) \circ	447 (2,0)	-4 (1,1) \ominus	454 (2,3)	3 (1,8)
Katar	447 (3,3)	8 (1,6) \circ	423 (4,5)	-16 (2,1) \ominus	435 (3,9)	-4 (1,7) \ominus
Iran	436 (3,2)	4 (1,2) \circ	428 (3,4)	-4 (1,6) \ominus	416 (3,2)	-16 (1,8) \ominus
Oman	424 (2,6)	-2 (1,0)	430 (2,9)	4 (1,9) \circ	413 (2,6)	-12 (1,5) \ominus
Indonezja	400 (3,6)	2 (0,9) \circ	394 (4,3)	-3 (1,8)	385 (4,2)	-13 (1,9) \ominus
Jordania	388 (3,1)	-1 (1,1)	395 (3,1)	6 (1,8) \circ	382 (3,3)	-7 (1,5) \ominus
Arabia Saudyjska ^ψ	385 (4,0)	1 (1,8) \circ	381 (4,9)	-3 (3,1)	365 (4,0)	-19 (2,5) \ominus
Maroko	382 (2,8)	4 (0,9) \circ	385 (3,4)	7 (1,7) \circ	351 (3,7)	-27 (1,4) \ominus
Południowa Afryka	379 (3,4)	3 (0,9) \circ	360 (3,6)	-16 (1,1) \ominus	381 (3,9)	5 (1,8) \circ
Kuwejt ^ψ	357 (4,5)	4 (1,2) \circ	338 (4,9)	-16 (1,4) \ominus	345 (5,3)	-8 (2,4) \ominus

W nawiasach błędy standardowe. Z powodu zaokrąglenia liczb niektóre wyniki mogą się wydawać niespójne.

Kolejność krajów według Tabeli 2.1.

* Różnica między średnim wynikiem subtestu i całego testu w danym kraju.

\circ Średni wynik subtestu istotnie wyższy niż średni wynik całego testu w kraju.

\ominus Średni wynik subtestu istotnie niższy niż średni wynik całego testu w kraju.

Pozostałe symbole jak w Tabeli 2.1.

Tabela 2.5. Wyniki pomiaru osiągnięć matematycznych według funkcji podmiotowej

Kraj	Zasoby		Zastosowania typowe		Zastosowania problemowe	
	Średnia	Różnica [*]	Średnia	Różnica [*]	Średnia	Różnica [*]
Singapur ²	631 (4,0)	13 (1,4) Ⓢ	619 (4,0)	2 (1,0)	603 (4,5)	-15 (1,4) Ⓢ
Hong Kong [†]	618 (3,1)	4 (1,3) Ⓢ	621 (3,1)	6 (1,3) Ⓢ	600 (3,2)	-15 (1,5) Ⓢ
Korea Południowa	627 (2,9)	19 (1,4) Ⓢ	595 (2,1)	-13 (1,2) Ⓢ	619 (2,5)	11 (2,0) Ⓢ
Tajwan	620 (2,3)	24 (1,9) Ⓢ	593 (2,1)	-3 (1,5) Ⓢ	576 (3,1)	-21 (2,0) Ⓢ
Japonia	601 (2,4)	9 (1,3) Ⓢ	589 (2,1)	-4 (1,2) Ⓢ	595 (2,7)	2 (1,9)
Irlandia Północna [‡]	582 (3,9)	11 (1,6) Ⓢ	575 (3,2)	5 (1,2) Ⓢ	550 (3,3)	-21 (1,9) Ⓢ
Federacja Rosyjska	556 (3,4)	-7 (1,0) Ⓢ	566 (3,7)	3 (1,7)	570 (4,0)	6 (1,8) Ⓢ
Norwegia	544 (3,1)	-5 (1,9) Ⓢ	550 (2,6)	1 (1,1)	556 (2,9)	7 (2,2) Ⓢ
Irlandia	554 (2,9)	7 (2,2) Ⓢ	549 (2,2)	1 (1,2)	535 (2,7)	-12 (1,7) Ⓢ
Anglia	554 (3,3)	8 (1,5) Ⓢ	544 (3,2)	-2 (1,7)	540 (3,2)	-6 (2,0) Ⓢ
Belgia (flamandzka) [†]	554 (2,3)	8 (0,8) Ⓢ	544 (2,2)	-2 (1,1)	536 (2,7)	-10 (1,4) Ⓢ
Kazachstan	546 (4,4)	1 (1,3)	541 (4,9)	-4 (1,3) Ⓢ	553 (4,6)	9 (1,6) Ⓢ
Portugalia ²	548 (2,6)	6 (1,9) Ⓢ	540 (2,4)	-2 (1,2)	532 (2,3)	-10 (1,3) Ⓢ
Stany Zjednoczone ^{2†}	547 (2,3)	8 (1,2) Ⓢ	537 (2,4)	-2 (1,0)	531 (2,5)	-9 (1,3) Ⓢ
Dania ^{2†}	536 (3,3)	-3 (1,6)	538 (2,8)	-1 (1,7)	548 (3,2)	9 (2,0) Ⓢ
Finlandia	532 (2,5)	-3 (1,1) Ⓢ	537 (2,7)	1 (1,3)	534 (2,8)	-1 (1,4) Ⓢ
Litwa ²	530 (2,2)	-5 (1,4) Ⓢ	536 (2,1)	1 (1,0)	540 (3,1)	5 (2,2)
POLSKA	517 (2,4)	-18 (1,0) Ⓢ	541 (2,1)	6 (0,7) Ⓢ	546 (2,3)	11 (1,3) Ⓢ
Holandia [†]	521 (1,8)	-9 (0,8) Ⓢ	531 (1,7)	1 (1,4)	543 (2,6)	13 (2,4) Ⓢ
Węgry	532 (3,1)	3 (1,2) Ⓢ	526 (3,3)	-3 (1,0) Ⓢ	529 (3,6)	0 (1,5)
Czechy	519 (2,5)	-9 (1,2) Ⓢ	528 (2,4)	0 (0,9)	544 (3,0)	16 (1,9) Ⓢ
Bułgaria	527 (5,1)	3 (1,7)	523 (5,6)	-2 (1,8)	521 (5,8)	-4 (1,8)
Cypr	519 (2,8)	-4 (1,7) Ⓢ	529 (2,8)	6 (1,6) Ⓢ	519 (3,1)	-4 (1,6) Ⓢ
Niemcy	524 (2,3)	2 (0,9) Ⓢ	515 (2,2)	-6 (1,2) Ⓢ	535 (2,4)	13 (1,6) Ⓢ
Słowenia	517 (1,9)	-3 (1,2) Ⓢ	521 (2,1)	1 (0,8)	524 (2,2)	4 (1,2) Ⓢ
Szwecja ²	501 (3,4)	-18 (1,8) Ⓢ	521 (2,7)	3 (0,9) Ⓢ	542 (3,3)	23 (1,5) Ⓢ
Serbia ³	513 (3,5)	-5 (1,7) Ⓢ	521 (3,4)	3 (1,4) Ⓢ	517 (3,8)	-1 (1,8)
Australia	509 (3,5)	-8 (1,6) Ⓢ	521 (3,0)	4 (1,2) Ⓢ	523 (3,0)	6 (1,7) Ⓢ
Kanada ^{12†}	505 (2,4)	-5 (0,7) Ⓢ	510 (2,3)	0 (0,6)	521 (2,4)	10 (0,7) Ⓢ
Włochy ²	511 (2,9)	4 (1,0) Ⓢ	504 (2,5)	-3 (1,7)	503 (3,3)	-4 (3,0)
Hiszpania ²	505 (2,4)	0 (1,2)	505 (2,4)	0 (0,9)	502 (2,5)	-3 (0,9) Ⓢ
Chorwacja	502 (1,9)	0 (1,2)	499 (1,9)	-4 (1,2) Ⓢ	507 (2,1)	5 (1,1) Ⓢ
Słowacja	491 (2,4)	-8 (1,3) Ⓢ	497 (2,5)	-2 (1,0)	515 (2,9)	17 (1,8) Ⓢ
Nowa Zelandia	475 (2,6)	-15 (1,3) Ⓢ	497 (2,5)	6 (1,5) Ⓢ	504 (2,7)	13 (1,3) Ⓢ
Francja	484 (2,8)	-4 (0,9) Ⓢ	488 (3,1)	0 (1,0)	491 (3,4)	3 (2,2)
Turcja	491 (3,4)	8 (1,5) Ⓢ	482 (3,5)	-1 (2,0)	466 (3,5)	-17 (2,0) Ⓢ
Gruzja ^{1†}	466 (4,0)	3 (1,8)	461 (4,1)	-2 (2,4) Ⓢ	452 (4,4)	-11 (2,1) Ⓢ
Chile	449 (2,8)	-10 (1,8) Ⓢ	462 (2,4)	4 (1,2)	466 (2,3)	7 (1,2) Ⓢ
Zjednoczone Emiraty Arabskie	453 (2,7)	1 (1,0)	452 (2,5)	1 (0,7)	445 (2,4)	-6 (1,0) Ⓢ
Bahrajn ²	453 (1,8)	2 (1,2) Ⓢ	450 (1,6)	-1 (0,9)	447 (2,0)	-4 (1,4) Ⓢ
Katar	444 (3,4)	5 (1,2) Ⓢ	434 (3,5)	-5 (1,0) Ⓢ	431 (4,4)	-8 (3,0) Ⓢ
Iran	429 (3,2)	-3 (1,7)	435 (2,9)	3 (1,4) Ⓢ	426 (3,3)	-5 (1,8) Ⓢ
Oman	422 (2,7)	-3 (1,0) Ⓢ	428 (2,4)	2 (1,2)	420 (2,4)	-6 (1,1) Ⓢ
Indonezja	395 (4,2)	-3 (1,8)	397 (3,6)	0 (1,0)	397 (3,5)	-1 (1,2)
Jordania	389 (3,1)	1 (0,9)	388 (3,1)	0 (1,1)	385 (3,3)	-3 (1,9)
Arabia Saudyjska ^ψ	374 (4,6)	-10 (2,0) Ⓢ	382 (4,5)	-2 (2,1)	383 (4,3)	-1 (2,4)
Maroko	377 (3,7)	-1 (1,2)	375 (3,6)	-3 (0,9) Ⓢ	379 (3,6)	2 (1,7)
Południowa Afryka	378 (3,6)	2 (0,6) Ⓢ	377 (3,4)	1 (0,7)	369 (3,5)	-7 (0,9) Ⓢ
Kuwejt ^ψ	354 (4,5)	1 (1,6)	348 (4,8)	-5 (1,7) Ⓢ	332 (5,0)	-21 (1,5) Ⓢ

W nawiasach błędy standardowe. Z powodu zaokrąglenia liczb niektóre wyniki mogą się wydawać niespójne.

Kolejność krajów według Tabeli 2.1.

* Różnica między średnim wynikiem subtestu i całego testu w danym kraju.

Ⓢ Średni wynik subtestu istotnie wyższy niż średni wynik całego testu w kraju.

Ⓢ Średni wynik subtestu istotnie niższy niż średni wynik całego testu w kraju.

Pozostałe symbole jak w Tabeli 2.1.

Szwedów i Czechów: zastosowania problemowe przeważają nad typowymi, czemu towarzyszy rekordowy deficyt względny zasobów. Ciekawe, że podobnie było w 2011 r., mimo że wtedy testowanych trzecioklasistów kształcili nauczyciele edukacji wczesnoszkolnej. W obecnej edycji większą niż nasza nadwyżkę zastosowań problemowych mają, oprócz Szwecji i Czech, cztery inne kraje: Słowacja, Niemcy, Holandia i Nowa Zelandia. W rankingu zastosowań problemowych Polska znalazła się na 11. pozycji – wyprzedziło nas tylko dziewięć krajów ze ścisłej czołówki oraz (nieistotnie) Dania.

Wyniki pomiaru a program kształcenia

Program często uważa się za główny czynnik osiągnięć uczniów, ale kierunek wpływu programu na osiągnięcia pozostaje sporny. Szukając sposobów poprawy osiągnięć, jedni zalecają rozbudowę, inni – redukcję treści programowych. Pierwsi krytykują istniejący program za to, że jest za mało ambitny, drudzy – że jest „przeładowany”. I jedni, i drudzy zgadzają się jednak, że test osiągnięć powinien być dostosowany do programu kształcenia. Widząc niskie wyniki testowania, są skłonni oskarżyć test TIMSS o to, że wykraczał poza program realizowany w kraju, czyli odwoływał się do zagadnień „nieprzerabianych” w szkole.

Hipotezę, że na wynik testowania wpływa niedostosowanie testu do programu, łatwo sprawdzić – wystarczy, by każdy kraj wskazał zadania testu mieszczące się w programie kształcenia dla klasy, w której znajdowały się badane dzieci, i pytania wykraczające poza ten program. W Polsce było to trudne z powodu nieistnienia centralnego programu kształcenia klasy czwartej. Zadania podjęli się eksperci będący doświadczonymi nauczycielami matematyki i przyrody, zaznajomionymi z najpopularniejszymi programami i praktyką znanych im nauczycieli klasy czwartej.

Przypomnijmy, że cały test matematyczny składał się z 169 pytań dających maksymalnie 178 punktów¹. Wśród tych pytań nasi eksperci znaleźli jedynie 22 pytania wyraźnie wykraczające poza programy nauczania matematyki w klasach od pierwszej do czwartej.

- Dziewięć dotyczyło wiedzy o liczbie i wymagało rozwiązania równań z parametrem, działania na ułamkach o różnych mianownikach, znajdowania wielokrotności liczby i zaokrąglania liczb całkowitych.
- Osiem dotyczyło geometrii i wymagało znajdowania osi symetrii, porównywania brył, identyfikowania ich siatek i obliczania pola trójkąta.

¹ Nie znaczy to oczywiście, że każdy uczeń odpowiadał na 169 pytań (por. Rozdział 1).

- Pięć dotyczyło statystyki opisowej i wymagało odczytywania nietypowych wykresów (np. ekstrapolacji danych na wykresie rozproszenia).

Liczba punktów możliwych do zdobycia za zadania mieszczące się w naszych programach kształcenia wyniosła 153 (86 proc. wszystkich).

Siła technologii IRT – szacowanie kompetencji ucznia niezależnie od składu testu – tu akurat okazuje się słabością, ponieważ interesuje nas różnica wyniku testowania w zależności od włączenia do testu lub wyłączenia niektórych pytań. Dlatego dla każdego kraju obliczono średni procent poprawnych odpowiedzi w całym teście i średni procent poprawnych odpowiedzi w lokalnym zbiorze zadań mających uzasadnienie w lokalnym programie nauczania.

W próbie krajów średni procent poprawnych odpowiedzi w całym teście wahał się od 21 (w Kuwejcie) do 75 (w Hong Kongu) wokół średniej 50,1 z odchyleniem standardowym 13,0. Widać, że dzieci zdobyły średnio połowę punktów możliwych do uzyskania. Porównajmy z tymi wynikami średni procent poprawnych odpowiedzi w lokalnych zbiorach zadań. Wahał się on od 21 (w Kuwejcie) do 76 (w Singapurze i Korei) wokół średniej 51,9 z odchyleniem standardowym 13,5. Jak widać, średnie różnią się zaledwie o 1,7 punktu procentowego. Różnica waha się od 0 (w 12 krajach) do 6 (w Bułgarii i na Słowacji) z odchyleniem standardowym 1,7. Oba agregaty są skorelowane ze sobą prawie doskonale ($r = 0,99$). Widać, że ograniczenie zbioru zadań nie wnosi niemal nic nowego do naszej wiedzy o osiągnięciach matematycznych.

Przejdźmy do Polski. Średni procent poprawnych odpowiedzi w całym teście równa się 55, co daje nam 18. rangę z Finlandią i Węgrami. Średni procent poprawnych odpowiedzi w lokalnym zbiorze zadań równa się 56, co daje rangę 20,5 wraz pięcioma innymi krajami (Finlandią, Węgrami, Czechami, Holandią i Niemcami). Wyłączenie z testu 22 zadań, jakoby nieznanymi polskim uczniom, podwyższyło nasz wynik zaledwie o 1 punkt procentowy, ale obniżyło naszą rangę o 2,5.

Jakie wnioski można wyciągnąć z opisanej analizy? Pierwszy jest oczywisty: niepełne dostosowanie testu IEA do lokalnych programów kształcenia praktycznie nie wpływa na wyniki testowania. Drugi – że związek między programem kształcenia a tym, co uczniowie potrafią zademonstrować na egzaminie czy w badaniu osiągnięć, jest mniej ścisły, niż by się wydawało. Jednym z czynników moderujących ten związek jest coś, co można nazwać kulturą nauczania matematyki. Kultura ta wpływa na podejście uczniów do nieznanymi im zadań. Jeśli nauczyciel przedstawia matematykę jako narzędzie rozwiązywania problemów, to uczeń próbuje nadać sens nowym zadaniom i często udaje się mu je

Tabela 2.6 Osiągnięcia matematyczne dziewczynek i chłopców

Kraj	Dziewczynki	Chłopcy	Różnica	
Arabia Saudyjska	405 (4,4)	363 (6,5)	43 (7,7)	♀
Oman ^ψ	436 (3,0)	415 (2,8)	22 (3,0)	♀
Jordania	399 (3,3)	379 (4,9)	20 (5,8)	♀
Południowa Afryka	384 (3,8)	368 (4,4)	15 (4,2)	♀
Bahrain ²	459 (1,7)	443 (2,3)	15 (2,5)	♀
Kuwejt ^ψ	359 (5,4)	347 (5,6)	12 (6,2)	♀
Iran	437 (4,5)	426 (4,5)	10 (6,3)	
Indonezja	403 (4,0)	393 (3,9)	10 (2,7)	♀
Finlandia	540 (2,3)	531 (2,6)	9 (2,9)	♀
Bułgaria	527 (5,7)	522 (5,1)	5 (2,9)	
Norwegia	551 (2,6)	547 (3,1)	4 (2,9)	
Singapur ²	620 (3,9)	616 (4,3)	4 (3,0)	
Zjednoczone Emiraty Arabskie	453 (3,9)	450 (3,4)	3 (5,4)	
Gruzja	465 (3,9)	461 (4,4)	3 (4,0)	
Serbia ³	520 (3,7)	517 (4,7)	3 (4,7)	
Katar	440 (4,1)	438 (4,9)	3 (5,9)	
Litwa ²	537 (2,8)	534 (3,1)	2 (3,3)	
Kazachstan	546 (4,6)	543 (4,8)	2 (2,8)	
Maroko	378 (3,5)	377 (3,9)	1 (2,8)	
Szwecja ²	519 (3,2)	518 (3,2)	1 (3,0)	
Federacja Rosyjska	564 (3,7)	564 (3,7)	1 (2,8)	
Japonia	593 (2,0)	593 (2,5)	0 (2,3)	
Chile	458 (2,8)	459 (3,0)	-1 (3,2)	
POLSKA	534 (2,3)	536 (2,7)	-1 (2,5)	
Turcja	482 (3,2)	484 (3,5)	-2 (2,7)	
Irlandia Północna [‡]	569 (3,8)	571 (3,1)	-2 (3,8)	
Nowa Zelandia	489 (2,8)	492 (2,6)	-2 (2,8)	
Niemcy	520 (2,4)	523 (2,3)	-3 (2,3)	
Irlandia	545 (2,6)	549 (2,9)	-4 (3,4)	
Słowenia	518 (2,1)	522 (2,4)	-4 (2,6)	
Tajwan	594 (2,2)	599 (2,3)	-6 (2,5)	♂
Belgia (flamandzka)	543 (2,4)	549 (2,4)	-6 (2,4)	♂
Węgry	526 (3,4)	532 (3,8)	-6 (3,4)	
Francja	485 (3,2)	491 (3,2)	-6 (2,8)	♂
Dania ^{2†}	536 (3,1)	542 (3,0)	-6 (2,8)	♂
Anglia	543 (3,0)	549 (3,3)	-6 (2,9)	♂
Cypr	520 (2,9)	526 (3,1)	-6 (2,7)	♂
Stany Zjednoczone ^{2†}	536 (2,3)	543 (2,6)	-7 (1,9)	♂
Czechy	525 (3,0)	532 (2,5)	-7 (3,2)	♂
Korea Południowa	604 (2,3)	612 (2,5)	-7 (1,9)	♂
Holandia [†]	526 (1,8)	534 (2,2)	-8 (2,2)	♂
Australia	513 (3,1)	522 (3,9)	-9 (3,5)	♂
Kanada ^{2†}	506 (2,5)	515 (2,6)	-9 (2,1)	♂
Hong Kong [†]	609 (3,8)	619 (2,8)	-10 (3,3)	♂
Portugalia ²	536 (2,4)	547 (2,5)	-11 (2,2)	♂
Słowacja	493 (3,0)	504 (2,6)	-11 (2,6)	♂
Hiszpania ²	499 (2,7)	511 (2,7)	-12 (2,4)	♂
Chorwacja	496 (2,1)	508 (2,3)	-12 (2,7)	♂
Włochy ²	497 (2,7)	517 (3,0)	-20 (2,7)	♂
Średnia międzynarodowa	505 (0,5)	505 (0,5)		

♀ Średnia dziewczynek istotnie wyższa niż średnia chłopców.

♂ Średnia dziewczynek istotnie niższa niż średnia chłopców.

Pozostałe symbole jak w Tabeli 2.1.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

rozwiązać. Jeśli jednak matematykę utożsamia się ze zbiorem definicji, procedur i twierdzeń, to wystarczy nowy termin (np. os symetrii), by uczeń opuścił ręce: „Nie wiem, co to jest, tego jeszcze nie przerabialiśmy”. Byłoby interesujące zobaczyć, jaka kultura nauczania matematyki dominuje w Bułgarii i na Słowacji.

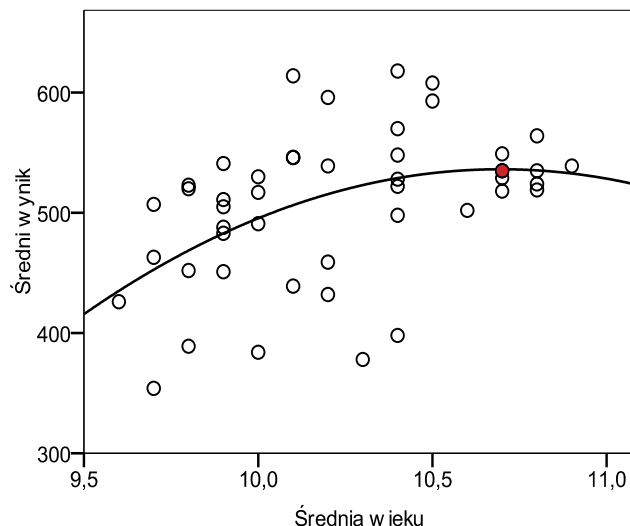
Wyniki pomiaru a płeć i wiek uczniów

W populacji międzynarodowej różnica między średnimi osiągnięć w matematyce dziewczynek i chłopców wynosi zero (Tabela 2.6). W poprzedniej edycji TIMSS wyniosła jeden punkt na korzyść chłopców. Nie warto byłoby o tym wspominać, gdyby nie silniejszy trend ku redukcji różnicy płciowej w osiągnięciach przyrodniczych.

Brak istotnej różnicy w osiągnięciach matematycznych odnotowano w 23 krajach, wśród nich w Polsce². Chłopcy wypadli istotnie lepiej w 18 krajach – szczególnie we Włoszech, Chorwacji i Hiszpanii. W siedmiu krajach dziewczynki przewyższyły chłopców. Są to – jak w poprzednich edycjach TIMSS – kraje arabskojęzyczne (Arabia Saudyjska, Oman, Jordania i Bahrain), a także Południowa Afryka, Indonezja i Finlandia. Widać, że na różnice osiągnięć matematycznych dziewczynek i chłopców silniej wpływa kultura niż natura.

Przejdźmy do wieku uczniów. Według założeń IEA międzynarodowa populacja uczniów, z której pobiera się próby krajowe, obejmuje dzieci średnio w wieku 10 lat znajdujące się w czwartym roku nauki na pierwszym szczeblu klasyfikacji ISCED. Życie jest jednak bardziej złożone, więc rozpiętość krajowych średnich wieku wynosi aż 1,9 roku: od 9,6 w Omanie do 11,5 w Południowej Afryce. Czy różnice średniej wieku mają związek z osiągnięciami matematycznymi? Odpowiedź jest twierdząca. Jeśli usunąć z analizy nietypową RPA, która testowała najstarsze dzieci w próbie krajów i osiągnęła drugi wynik od końca, średnie wieku pozwalają odtworzyć 21 proc. zróżnicowania średnich osiągnięć. Miejsca poszczególnych krajów w układzie współrzędnych przedstawia Rysunek 2.2. Polskę reprezentuje czerwone kółeczko.

Rysunku nie można, rzecz jasna, interpretować psychologicznie, w istocie bowiem mówi on nie o dzieciach, lecz o wieku startu szkolnego w różnych systemach oświatowych. Krzywoliniowość związku sugeruje istnienie optimum wiekowego. Z rysunku wyznaczyć go nie można, bo nie we wszystkich krajach rok szkolny zaczyna się, jak u nas, w ostatnich miesiącach roku



Rysunek 2.2. Średnie krajowe wyników testu matematycznego w zależności od średnich wieku uczniów

Czerwone kółeczko oznacza Polskę.

Na podstawie: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

kalendrzowego, ale można z danych pochodzących z krajów europejskich³. Analiza indywidualnych danych z 25 krajów Europy uczestniczących w badaniu TIMSS 2011 wykazała, że optymalny dla osiągnięć matematycznych w klasie czwartej wiek startu szkolnego wynosi średnio ok. 6,6 roku, co odpowiada normie posyłania do klasy pierwszej dzieci, które kończą 6 lat do 30 czerwca włącznie (Konarzewski, 2013).

Wiek uczniów w Polsce

Zagadnienie wieku w próbie polskich uczniów wymaga szerszego omówienia, ponieważ w obecnej edycji badania ma on niezwykle rozpiętość – waha się (w latach życia) od 8,25 do 13,25 wokół średniej 10,68 z odchyleniem standardowym 0,44. Jest to oczywiście skutek polskiej reformy wieku startu szkolnego zainicjowanej ustawą z 2009 r. Tabela 2.7 pokazuje rozkład rocznika w polskiej próbie TIMSS 2015. Widać, że w próbie znalazło się siedmioro dzieci, które poszły do szkoły wcześniej, niż to przewidywała reforma, i 106 dzieci opóźnionych w nauce według standardu sprzed reformy (dzieci te albo skorzystały z odroczenia obowiązku szkolnego, albo powtarzały którąś klasę). Trzon próby to dzieci, które poszły do szkoły jako siedmiolatki. Dzieci, które miały na starcie sześć lat, jest 18 proc.

² Warto dodać, że w Polsce nie odnotowano też różnicy w żadnym subteście osiągnięć szczegółowych.

³ Pod warunkiem, że wiek startu w różnych krajach jest losowy, dzieci nie uczą się umiejętności szkolnych przed rozpoczęciem nauki, a umiejętności uczniów przyrastają w trakcie nauki w stałym tempie.

Tabela 2.7. Rozkład rocznika polskich uczniów w TIMSS 2015

Rok urodzenia	Liczba uczniów	Procent
2002	12	0,3
2003	94	2,0
2004	3791	79,9
2005	843	17,8
2006	6	0,1
2007	1	0,0
Ogółem	4747	100,0

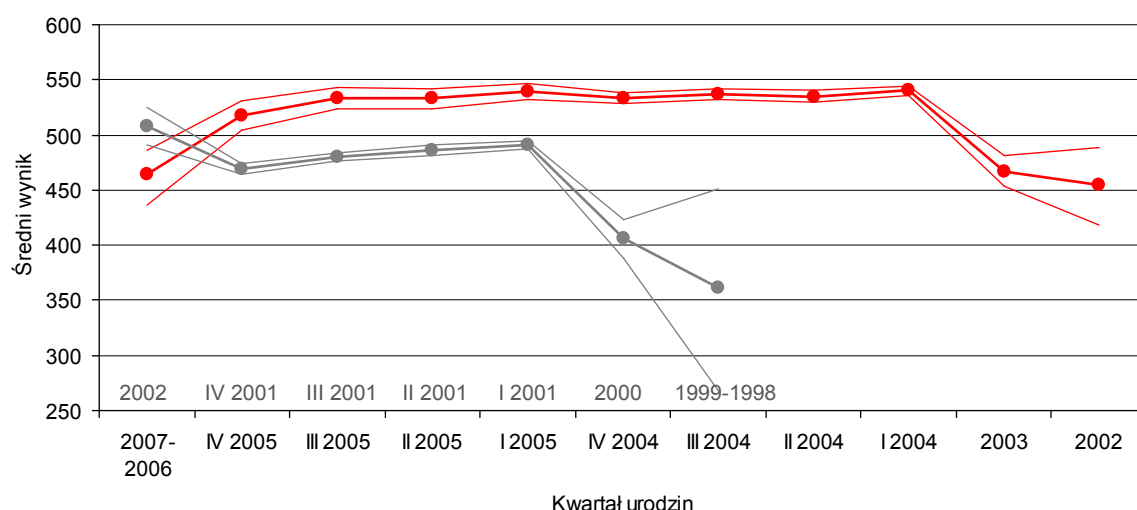
Czy wiek polskich uczniów wiąże się z ich osiągnięciami matematycznymi w klasie czwartej? – to pytanie, do którego teraz przechodzimy. Odpowiedź musimy jednak poprzedzić ważnym ostrzeżeniem. Próba 4747 uczniów reprezentuje populację czwartoklasistów w Polsce, ale podpróba 843 uczniów z rocznika 2005 z pewnością nie reprezentuje swojej kohorty wiekowej, a podpróba 3791 uczniów z rocznika 2004 reprezentuje swoją kohortę jedynie w przybliżeniu. Jest tak dlatego, że w 2011 r., kiedy przeważająca większość dzieci testowanych w 2015 r. rozpoczynała swoją karierę szkolną, do klasy pierwszej poszło tylko ok. 19 proc. sześciolatków – reszta bowiem pozostała w przedszkolu lub w „zerówce” – oraz ok. 90 proc. siedmiolatków – bo pozostałe 10 proc. rozpoczęło naukę rok wcześniej. Najważniejsze jest to, że w latach 2010–2011 sześciolatki szły do szkoły na mocy decyzji rodziców, a ta w przeważającej większości wypadków musiała się opierać na intuicyjnej (rodziciel-

skiej) lub fachowej (psychologicznej) ocenie gotowości szkolnej dziecka. Jeśli oceniający nie byli pewni, czy dziecko poradzi sobie w szkole, to najczęściej nie zaczynało ono nauki. Dlatego można przypuszczać, że przeciętny sześciolatek w klasie pierwszej znajdował się na wyższym poziomie rozwoju umysłowego, społecznego i fizycznego niż przeciętny sześciolatek w kraju.

Teraz możemy przejść do odpowiedzi na postawione pytanie. Przedstawia ją Rysunek 2.3. Na osi odciętych odłożono kwartał urodzenia. Kolorem czerwonym oznaczono średnie warunkowe wyników testowania w 2015 r., a szarym – w 2011 r., gdy testowani wtedy trzecioklasiści byli w takim samym wieku co młodsi rocznik czwartoklasistów z 2015 r. Cienkie linie wokół grubszych oznaczają granice 95-procentowych przedziałów ufności. Dwie średnie różnią się istotnie od siebie, jeśli ich przedziały nie mają wspólnych wartości (nie nachodzą na siebie).

Komentarz trzeba zacząć od ogólnego stwierdzenia, że zmienna wieku uczniów w niewielkim stopniu odtwarza wyniki pomiaru osiągnięć w matematyce: redukuje wariację wewnątrz oddziałów o mniej niż 1 proc., a między oddziałami o 4 proc.⁴ Rolę wieku w szkole podstawowej należy doceniać, ale nie przeceniać. Przechodzimy do Rysunku 2.3. Wynika z niego, po pierwsze, że obecni czwartoklasiści umieją

⁴ HLM z kontrolą płci, lokalizacji szkoły i SES na obu poziomach.



Rysunek 2.3. Średnie wyniki testu matematycznego TIMSS 2011 i 2015 w Polsce w zależności od kwartału urodzenia uczniów

Średnie warunkowe oszacowano za pomocą ważonej analizy wariancji przy kontroli statusu socjoekonomicznego (opisanego w Rozdziale 4), a błędy standardowe – metodą *bootstrap*. Cienkie linie przedstawiają granice 95-procentowych przedziałów ufności. Skrajne kwartały połączono z powodu zbyt małej liczby uczniów. Etykiety nad osią odciętych odnoszą się do szarej krzywej.

znacznie więcej niż dawni trzecioklasiści w tym samym wieku⁵. Wzrost osiągnięć to efekt dodatkowego roku nauki w bardziej zaawansowanym trybie.

Po wtóre, wśród czwartoklasistów z rocznika 2005 nie ma efektu kwartału urodzenia (osiągnięcia dzieci urodzonych w ostatnim i w pierwszym kwartale tego roku nie różnią się istotnie), a wśród dawnych trzecioklasistów jest (Konarzewski, 2013). Skoro klasa czwarta (kształcona już według nowej podstawy programowej z 2008 r.) eliminuje efekt wcześniej obserwowany jeszcze w ostatniej klasie gimnazjum (Dolata i Pokropek, 2012), to trzeba uznać, że ma wyższą efektywność dydaktyczną niż dawna klasa trzecia.

Po trzecie, efekt kwartału urodzenia niemal nie występuje wśród czwartoklasistów z obu roczników⁶. Inaczej mówiąc: dzieci, które rozpoczęły naukę w wieku sześciu lat, mają po czterech latach edukacji takie same osiągnięcia w matematyce jak dzieci, które poszły do szkoły jako siedmiolatki. Choć ten wynik zapewne ucieszy zwolenników reformy wieku obowiązku szkolnego z 2009 r., to powinni oni pamiętać, że nie ma pewności, czy odtworzyłyby się w 2020 r., kiedy w klasie czwartej byłyby już tylko dzieci posłane do szkoły w wieku sześciu lat.

Zmiany osiągnięć matematycznych

Zmiana osiągnięć, czyli różnica między średnimi późniejszego i wcześniejszego pomiaru, wydaje się najwartościowszą, bo nieuwarunkowaną przez okoliczności zewnętrzne informacją o efektywności krajowego szkolnictwa. Wzrost, spadek czy stagnacja? – to mówi więcej niż średnia i pozycja w rankingu.

Tabela 2.8 pokazuje średnie pomiaru w poszczególnych edycjach oraz informuje o istotności statystycznej różnic między nimi za pomocą liter w indeksie górnym. Różnica między dwiema średnimi jest istotna, jeśli nie mają one wspólnych liter. Spójrzmy np. na Anglię. Widzimy, że w pierwszych trzech edycjach średnie wyników istotnie rosły, a w dwóch następnych ustabilizowały się na poziomie 2007 r.

Analiza danych w tabeli pozwala zauważyć kilka interesujących faktów. Po pierwsze, znaczną zmienność składu krajów w poszczególnych edycjach. Tylko 12 krajów, wśród nich Słowenia i Węgry, uczestniczyło we wszystkich pięciu edycjach. Dwadzieścia pięć

krajów, wśród nich Polska, przystąpiły do badania TIMSS po 1995 r. (większość z nich w 2011 r., w którym zbiegły się edycje TIMSS i PIRLS) i wzięły udział we wszystkich następnych edycjach. Osiem krajów zachowało się niekonsekwentnie, opuszczając jedną lub więcej edycji, tak jakby ich opinię o użyteczności danych TIMSS kształtowały bieżące okoliczności polityczne.

Drugi fakt to znaczna zmienność dynamiki średnich. Wyłączając z analizy kraje mające tylko dwa pomiary, zauważamy trzy typy zmian.

- Monotoniczny wzrost (z dopuszczeniem jednego niewielkiego spadku) ujawnił się w 17 krajach. Najlepszym przykładem jest tu Słowenia, która w ciągu 20 lat małymi krokami podwyższyła średnią o 84 proc. swojego odchylenia standardowego.
- Monotoniczny spadek (z dopuszczeniem jednego niewielkiego wzrostu) ujawnił się w dwóch krajach: Holandii, która w ciągu 20 lat obniżyła średnią o 36 proc. swojego odchylenia standardowego, i Austrii.
- Wahania w górę i w dół ujawniły się w trzech krajach: w Czechach i na Węgrzech (nagła zapaść w 2007 r.) oraz Norwegii (spadek w 2003 r.).
- Stabilizacja jest obecna w czterech krajach: Belgii, Słowacji, we Włoszech i na Litwie (z wyłączeniem uczniów rosyjsko- i polskojęzycznych).

Widać, że najliczniejsza kategoria to monotoniczny wzrost. Warto dodać, że wzrostowi średniej może towarzyszyć:

- spadek zróżnicowania wyników – w ośmiu krajach: na Cyprze, w Australii, Anglii, Hong Kongu, Irlandii, Portugalii, Singapurze i Słowenii
- wzrost zróżnicowania wyników – w pięciu krajach: na Tajwanie, w Danii, Iranie, Korei Południowej i Szwecji
- stabilizacja lub wahania zróżnicowania wyników – w Gruzji, Japonii, Federacji Rosyjskiej i Stanach Zjednoczonych.

⁵ Z wyjątkiem garstki najmłodszych, którzy rozpoczęli naukę przed terminem, ale byli oni średnio aż o cztery miesiące młodsi niż najmłodszy w 2011 r.

⁶ Jedynie dwie różnice osiągają istotność statystyczną: między grupą z czwartego kwartału 2005 a grupami z trzeciego i pierwszego kwartału 2004.

Tabela 2.8. Osiągnięcia matematyczne w kolejnych edycjach badania TIMSS

Kraj	1995	2003	2007	2011	2015
Anglia	484 (3,3) ^c	531 (3,7) ^b	541 (2,9) ^a	542 (3,5) ^a	546 (2,9) ^a
Arabia Saudyjska				410 (5,2) ^b	384 (4,0) ^a
Armenia		456 (3,5) ^a		452 (3,5) ^a	
Australia	495 (3,4) ^b	499 (3,9) ^b	516 (3,5) ^a	516 (2,9) ^a	517 (3,1) ^a
Austria	531 (2,9) ^b		505 (2,0) ^a	508 (2,6) ^a	
Bahrajn				436 (3,2) ^b	451 (1,6) ^a
Belgia (flamandzka)		551 (1,8) ^a		549 (1,9) ^a	546 (2,1) ^a
Chile				462 (2,3) ^a	459 (2,4) ^a
Chorwacja				490 (1,9) ^b	502 (1,7) ^a
Cypr	475 (3,2) ^c	510 (2,4) ^b			523 (2,7) ^a
Czechy	541 (3,1) ^d		486 (2,8) ^c	511 (2,4) ^b	528 (2,2) ^a
Dania			523 (2,4) ^b	537 (2,6) ^a	539 (2,7) ^a
Federacja Rosyjska		532 (4,7) ^b	544 (4,9) ^b	542 (3,7) ^b	564 (3,4) ^a
Finlandia				545 (2,4) ^b	535 (2,0) ^a
Gruzja			438 (4,3) ^c	450 (3,7) ^b	463 (3,6) ^a
Hiszpania				482 (2,8) ^b	505 (2,5) ^a
Holandia	549 (3,0) ^c	540 (2,1) ^b	535 (2,1) ^{ba}	540 (1,7) ^b	530 (1,7) ^a
Hong Kong	557 (4,0) ^e	575 (3,2) ^d	607 (3,6) ^{cab}	602 (3,4) ^b	614 (2,9) ^a
Iran	387 (5,0) ^c	389 (4,2) ^c	402 (4,1) ^b	431 (3,5) ^a	432 (3,2) ^a
Irlandia	523 (3,5) ^b			527 (2,6) ^b	548 (2,2) ^a
Irlandia Północna				562 (2,8) ^a	570 (2,9) ^a
Japonia	567 (1,9) ^c	565 (1,6) ^c	568 (2,1) ^c	585 (1,7) ^b	593 (1,9) ^a
Katar				413 (3,4) ^b	439 (3,4) ^a
Kazachstan				501 (4,5) ^b	544 (4,5) ^a
Korea Południowa	581 (1,8) ^b			605 (1,9) ^a	608 (2,2) ^a
Kuwejt				342 (3,6) ^a	328 (3,6) ^a
Maroko				335 (4,0) ^b	378 (3,0) ^a
Litwa		534 (2,8) ^a	530 (2,4) ^a	534 (2,4) ^a	536 (2,7) ^a
Niemcy			525 (2,1) ^b	528 (2,2) ^b	522 (2,0) ^a
Norwegia	476 (3,0) ^c	451 (2,3) ^c	473 (2,5) ^b	495 (2,8) ^a	493 (2,3) ^a
Nowa Zelandia	469 (4,4) ^c	493 (2,2) ^b	492 (2,3) ^{ba}	486 (2,6) ^a	491 (2,3) ^{ba}
Oman				385 (2,9) ^b	426 (2,5) ^a
Portugalia	442 (3,9) ^c			532 (3,4) ^b	541 (2,2) ^a
Serbia				516 (3,0) ^a	518 (3,5) ^a
Singapur	590 (4,5) ^c	594 (5,6) ^{cb}	599 (3,7) ^{cb}	606 (3,2) ^b	618 (3,8) ^a
Słowacja			496 (4,5) ^a	507 (3,8) ^a	498 (2,5) ^a
Słowenia	462 (3,1) ^e	479 (2,6) ^d	502 (1,8) ^c	513 (2,2) ^b	520 (1,9) ^a
Stany Zjednoczone	518 (2,9) ^d	518 (2,4) ^{dc}	529 (2,4) ^b	541 (1,8) ^a	539 (2,3) ^a
Szwecja			503 (2,5) ^b	504 (2,0) ^b	519 (2,8) ^a
Tajwan		564 (1,8) ^c	576 (1,7) ^b	591 (2,0) ^a	596 (1,8) ^a
Tunezja		339 (4,7) ^b	327 (4,5) ^b	359 (3,9) ^a	
Turcja				469 (4,7) ^b	483 (3,1) ^a
Węgry	521 (3,6) ^a	529 (3,1) ^a	510 (3,5) ^{cb}	515 (3,4) ^b	529 (3,2) ^a
Włochy		503 (3,7) ^a	507 (3,1) ^a	508 (2,6) ^a	507 (2,6) ^a
Zjednoczone Emiraty Arabskie				434 (2,0) ^b	452 (2,4) ^a

W nawiasach błędy standardowe. z powodu zaokrąglenia liczb niektóre wyniki mogą się wydawać niespójne.

Dwie średnie niemające takich samych liter w superskrypcie istotnie różnią się od siebie.

Tabela uwzględnia kraje, które brały udział przynajmniej w dwóch edycjach badania TIMSS. Brak Polski jest następstwem zmiany klasy, w której byli uczniowie poddani pomiarowi – z trzeciej w 2011 r. na czwartą w 2015 r.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Osiągnięcia przyrodnicze

Krzysztof Konarzewski

Podstawowe wyniki pomiaru osiągnięć przyrodniczych zawiera Tabela 3.1, a objaśnienia – Rozdział 2 poświęcony osiągnięciom matematycznym. W drugiej kolumnie Tabeli 3.1 znajdują się oszacowania średnich krajowych. Według nich uporządkowano kraje – od Singapuru (z najwyższą średnią) do Kuwejtu (z najniższą). Średnia osiągnięć przyrodniczych polskich dzieci wynosi 547 punktów.

Obok każdej średniej podano w nawiasie błąd standardowy oszacowania, a określony na tej podstawie przedział ufności zilustrowano czarnym prostokątem na rozkładzie centylowym. Największym błędem jest obciążone oszacowanie średniej Kuwejtu (6,2 punktu), a wśród krajów europejskich – Bułgarii (5,9). Polska z błędem standardowym 2,4 ustępuje pod tym względem tylko 10 krajom. Świadczy to o względnej homogeniczności populacji polskich uczniów, a także o poprawności samego pomiaru.

Średnia Polski jest istotnie wyższa od wzorcowej średniej międzynarodowej. W porównaniach par krajów polskie wyniki są nieodróżnialne od wyników w sześciu innych krajach. Siedem krajów – od Singapuru do Finlandii – wyprzedza nas istotnie, a my wyprzedzamy 33 kraje: Norwegię, Anglię i wszystkie od Czech do końca listy. Wśród 46 innych krajów Polska ma dziewiątą rangę – o osiem miejsc wyższą niż w osiągnięciach matematycznych.

Zróżnicowanie wyników pomiaru zostanie ukazane za pomocą trzech statystyk. Pierwsza to odchylenie standardowe. Z tabeli 3.1 wynika, że najmniej zróż-

nicowaną populację mają – podobnie jak w 2011 r. – Holandia, flamandzkojęzyczna część Belgii i Chorwacja. Na drugim biegunie mieszczą się kraje arabskojęzyczne, a także Iran i Indonezja.

Druga statystyka to punkty centylowe. Lewy kraniec prostokąta na Rysunku 2.1 pokazuje, gdzie kończą się osiągnięcia najsłabszych pięciu procent uczniów, a prawy – gdzie zaczynają osiągnięcia pięciu procent najsilniejszych. Środkowy żółty prostokąt wyznacza przedział, w którym mieści się połowa uczniów o przeciętnych wynikach. Pięcioprocentowa grupa względnie najsłabszych uczniów w swoim kraju jest bezwzględnie najsłabsza w Kuwejcie, ale trudno się temu dziwić, skoro ten kraj ma też najniższą średnią. Bardziej miarodajną informacją o zróżnicowaniu wyników są odległości między punktami b i e na Rysunku 2.1 a medianą. W przednim rozdziale uznaliśmy je za miary negatywnej i pozytywnej dyskryminacji oświatowej.

Najmniejszą różnicę, czyli najmniejszą dyskryminację negatywną, notujemy w Portugalii (102 punkty i pierwsza ranga); Holandia tym razem jest druga (106 punktów). Największą różnicę (217 punktów) mamy w Zjednoczonych Emiratach Arabskich. Polska wraz z Tajwanem, Danią i Czechami ma pod tym względem rangę 16,5. W unikaniu dyskryminacji pozytywnej przoduje Holandia (90 punktów); na drugim biegunie jest Kuwejt (205). Polska wraz z Japonią i Irlandią ma 100 punktów różnicy i rangę 10. Mimo wielu różnic między nauczaniem i uczeniem się ma-

Tabela 3.1 Osiągnięcia przyrodnicze

Kraj	Średnia	Odchylenie st.	Rozkłady centylowe
Singapur ²	♣ 590 (3,7)	85 (2,6)	
Korea Południowa	♣ 589 (2,0)	62 (0,9)	
Japonia	♣ 569 (1,8)	65 (1,0)	
Federacja Rosyjska	♣ 567 (3,2)	69 (1,9)	
Hong Kong [†]	♣ 556 (2,9)	70 (1,4)	
Tajwan	♣ 555 (1,8)	68 (1,1)	
Finlandia	♣ 554 (2,3)	64 (1,7)	
Kazachstan	♣ 550 (4,4)	85 (2,5)	
POLSKA	♣ 547 (2,4)	69 (1,4)	
Stany Zjednoczone ^{2†}	♣ 546 (2,2)	80 (1,3)	
Słowenia	♣ 542 (2,4)	69 (1,5)	
Węgry	♣ 542 (3,3)	83 (2,6)	
Szwecja ²	♣ 540 (3,6)	73 (2,5)	
Norwegia	♣ 537 (2,7)	63 (1,5)	
Anglia	♣ 536 (2,4)	70 (1,7)	
Bułgaria	♣ 535 (5,9)	95 (3,6)	
Czechy	♣ 534 (2,4)	70 (1,4)	
Chorwacja	♣ 533 (2,1)	62 (1,1)	
Irlandia	♣ 529 (2,8)	69 (1,8)	
Niemcy	♣ 528 (2,4)	70 (1,3)	
Litwa ²	♣ 527 (2,5)	69 (1,2)	
Dania ^{2†}	♣ 527 (2,1)	69 (1,3)	
Kanada ^{12†}	♣ 525 (2,5)	73 (1,6)	
Serbia ³	♣ 524 (3,7)	81 (3,4)	
Australia	♣ 524 (2,9)	76 (1,9)	
Słowacja	♣ 520 (2,6)	85 (1,9)	
Irlandia Północna [†]	♣ 520 (2,2)	69 (1,5)	
Hiszpania ²	♣ 518 (2,6)	69 (1,5)	
Holandia [†]	♣ 517 (2,7)	60 (1,3)	
Włochy ²	♣ 516 (2,6)	66 (1,3)	
Belgia (flamandzka) [†]	♣ 511 (2,3)	62 (1,3)	
Portugalia ²	♣ 508 (2,2)	60 (1,0)	
Nowa Zelandia	♣ 505 (2,7)	85 (1,6)	
Średnia wzorcowa	500	100	
Francja	♣ 487 (2,7)	73 (1,4)	
Turcja	♣ 483 (3,3)	92 (2,5)	
Cypr	♣ 481 (2,6)	76 (1,4)	
Chile	♣ 478 (2,7)	73 (1,4)	
Bahrajn ²	♣ 459 (2,6)	105 (1,7)	
Zjednoczone Emiraty Arabskie	♣ 451 (2,8)	121 (1,5)	
Gruzja ^{1†}	♣ 451 (3,7)	87 (2,5)	
Katar	♣ 436 (4,1)	111 (2,2)	
Oman	♣ 431 (3,1)	119 (1,7)	
Iran	♣ 421 (4,0)	102 (2,9)	
Indonezja	♣ 397 (4,8)	102 (2,0)	
Arabia Saudyjska	♣ 390 (5,0)	115 (2,9)	
Maroko ^ψ	♣ 352 (4,4)	120 (2,3)	
Kuwejt ^ψ	♣ 337 (6,2)	126 (2,0)	

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

W nawiasach błędy standardowe. Z powodu zaokrąglenia liczb niektóre wyniki mogą wydawać się niespójne.

♣ Średnia krajowa istotnie wyższa niż średnia wzorcowa

♣ Średnia krajowa istotnie niższa niż średnia wzorcowa

ψ Średnia osiągnięć mało wiarygodna, ponieważ odsetek uczniów uzyskujących wyniki nieprzewyższające poziomu zgadywania leży w przedziale 15–25.

¹ Definicja populacji krajowej nie uwzględnia pełnej definicji międzynarodowej.

² Operat losowania pokrywa 90–95% populacji krajowej.

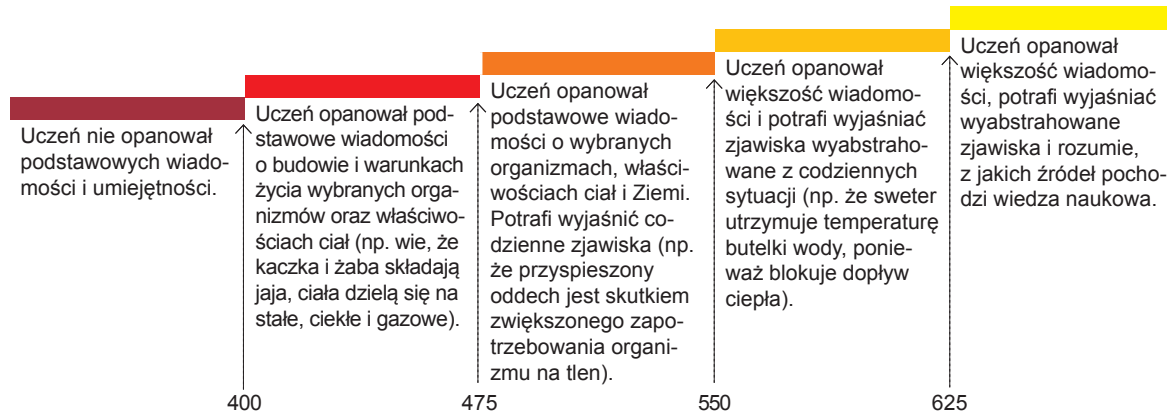
³ Operat losowania pokrywa mniej niż 90% (ale przynajmniej 77%) populacji krajowej.

[†] Poziom realizacji próby spełnia wymaganie jedynie po wykorzystaniu zapasowych szkół.

[‡] Poziom realizacji próby niemal spełnia wymaganie jedynie po wykorzystaniu zapasowych szkół.

Tabela 3.2. Poziomy osiągnięć przyrodniczych: definicje i i procenty uczniów, którzy się na nich znaleźli

Świat 2011 ^a	8	20	40	27	5
Świat 2015 ^a	5	18	38	32	7
Polska 2011 ^b	9	24	38	24	5
Polska 2015 ^c	3	12	34	39	12



^a Przez „świat” należy rozumieć kraje uczestniczące w każdej edycji badania. Ponieważ kraje nie zostały dobrane losowo, nie można użyć ich wyników do oszacowania parametrów populacji krajów na Ziemi. Stopień pokrywania się list krajów z lat 2011 i 2015 można ocenić na podstawie Tabeli 2.7.

^b Uczniowie klasy trzeciej.

^c Uczniowie klasy czwartej.

tematyki i przyrody w polskich szkołach pozycja naszego kraju w dyskryminacji negatywnej (matematyka: 18,5, przyroda: 16,5) i pozytywnej (matematyka: 10,5, przyroda: 10) jest niemal taka sama. Czy coś powstrzymuje polskie szkoły przed skróceniem dystansu między najsłabszymi uczniami a peletonem? I czy względnie mały dystans między peletonem a czołówką to skutek wyboru, czy rutyny?

Przechodzimy do trzeciej statystyki: progów osiągnięć. Cztery progi: 400, 475, 550 i 625 punktów wyznaczają pięć przedziałów. Tabela 3.2 zawiera definicje przedziałów i odsetki uczniów na świecie i w Polsce, którzy je zajmują. Tabela 3.3 podaje szczegółowe dane o każdym kraju. Kolejność krajów w tabeli ustalono według odsetka uczniów pokonujących najwyższy próg. Listę otwiera Singapur (37 proc.), a zamyka Kuwejt (1 proc.). Polska z 12-procentową grupą prymusów spada z 9. pozycji w uporządkowaniu pod względem średniej w całym teście na 12.

Osiągnięcia szczegółowe

Tabele 3.4 i 3.5. przedstawiają osiągnięcia przyrodnicze w subtrestach opisanych w Rozdziale 1. Dla każdego subtestu tabele podają wartość średniej oraz różnicę między tą wartością a średnią krajową w całym teście. Różnice statystycznie istotne oznaczają strzałki – skierowane ku górze, gdy wynik w subtście jest większy niż średnia w całym teście, lub ku dołowi, gdy jest mniejszy.

Zaczynamy od klasyfikacji zadań według treści przedmiotowej. Dane w Tabeli 3.4 sugerują, że

w nauczaniu i uczeniu się przyrody w klasie czwartej polskich szkół podstawowych dominują zagadnienia biologiczne kosztem fizycznych i geograficznych. Że treść kształcenia przyrodniczego można rozłożyć bardziej równomiernie, przekonuje przykład Rosji, Chorwacji i Emiratów Arabskich. Nie rozkład decyduje jednak o sumie wiedzy opanowanej przez uczniów, ponieważ uczniowie w Rosji lokują się w ścisłej czołówce, Chorwaci w środku tabeli, a mieszkańcy Emiratów na końcu. Z drugiej strony liderzy rankingu – Singapurczycy – mają wielki deficyt względny wiedzy geograficznej.

W klasyfikacji według funkcji podmiotowej (Tabela 3.5) polscy uczniowie wykazują niewielkie niedobory zasobów i zastosowań problemowych przy niewielkiej nadwyżce zastosowań typowych. Triumf pomysłowości nad wiadomościami, który kontemplowaliśmy w osiągnięciach matematycznych Szwedów i Polaków, tu już się nie powtórzył.

Wyniki pomiaru a program kształcenia

Pozostaje omówić związek między wynikami testowania a udziałem pytań wykraczających poza krajowy program nauczania. Jak wiemy z poprzedniego rozdziału, w celu oszacowania tego związku obliczono dla każdego kraju średni procent poprawnych odpowiedzi w całym teście i średni procent poprawnych odpowiedzi w lokalnym zbiorze zadań mających uzasadnienie w lokalnym programie nauczania.

Test przyrodniczy składał się z 168 pytań, które pozwalały zdobyć maksymalnie 180 punktów.

Tabela 3.3. Progi i poziomy osiągnięć przyrodniczych

Kraj	Odsetki uczniów, którzy pokonali próg				Poziomy osiągnięć
	625	550	475	400	
Singapur ²	37 (2,0)	71 (1,8)	90 (1,1)	97 (0,5)	
Korea Południowa	29 (1,6)	75 (1,1)	96 (0,5)	100 (0,1)	
Federacja Rosyjska	20 (1,6)	62 (2,1)	91 (1,0)	99 (0,3)	
Japonia	19 (0,9)	63 (1,3)	93 (0,5)	99 (0,2)	
Kazachstan	19 (1,7)	49 (2,5)	81 (1,4)	96 (0,6)	
Hong Kong [†]	16 (1,2)	55 (1,8)	88 (1,1)	98 (0,4)	
Bułgaria	16 (1,5)	49 (2,5)	77 (2,2)	90 (1,5)	
Stany Zjednoczone ^{1†}	16 (0,8)	51 (1,2)	81 (0,9)	95 (0,5)	
Tajwan	14 (0,7)	55 (1,2)	88 (0,8)	98 (0,3)	
Węgry	14 (1,1)	50 (1,5)	81 (1,6)	94 (0,9)	
Finlandia	13 (0,9)	54 (1,4)	89 (0,9)	98 (0,4)	
POLSKA	12 (0,9)	51 (1,4)	85 (1,3)	97 (0,4)	
Szwecja ²	11 (1,1)	47 (2,1)	82 (1,5)	96 (0,8)	
Słowenia	11 (0,9)	48 (1,3)	84 (1,0)	97 (0,5)	
Anglia	10 (0,8)	43 (1,5)	81 (1,2)	97 (0,5)	
Słowacja	9 (0,6)	39 (1,4)	74 (1,2)	91 (0,8)	
Czechy	9 (0,7)	43 (1,5)	81 (1,1)	96 (0,6)	
Serbia ³	8 (0,7)	40 (1,5)	77 (1,7)	93 (1,1)	
Australia	8 (0,7)	39 (1,6)	75 (1,4)	94 (0,8)	
Niemcy	7 (0,5)	39 (1,6)	78 (1,2)	96 (0,6)	
Kanada ^{12†}	7 (0,5)	38 (1,2)	77 (1,3)	95 (0,7)	
Norwegia	7 (1,0)	44 (1,8)	85 (1,1)	98 (0,6)	
Irlandia	7 (0,9)	41 (1,8)	80 (1,7)	96 (0,7)	
Litwa ²	7 (0,7)	39 (1,6)	78 (1,3)	96 (0,5)	
Dania ^{2†}	7 (0,6)	39 (1,5)	78 (1,3)	96 (0,5)	
Nowa Zelandia	6 (0,6)	32 (1,1)	67 (1,4)	88 (0,8)	
Emiraty Arabskie	6 (0,4)	22 (0,8)	46 (1,0)	67 (0,9)	
Chorwacja	6 (0,7)	41 (1,4)	83 (1,1)	98 (0,4)	
Irlandia Północna [‡]	5 (0,5)	34 (1,3)	76 (1,4)	95 (0,6)	
Hiszpania	5 (0,5)	34 (1,3)	74 (1,6)	95 (0,7)	
Oman	4 (0,4)	16 (0,8)	38 (1,2)	61 (1,0)	
Bahrajn ²	4 (0,4)	19 (0,9)	47 (1,2)	72 (1,1)	
Turcja	4 (0,5)	24 (1,1)	58 (1,4)	82 (1,2)	
Włochy ²	4 (0,5)	32 (1,5)	75 (1,7)	95 (0,7)	
Katar	3 (0,5)	15 (1,2)	39 (1,7)	64 (1,6)	
Holandia [†]	3 (0,4)	30 (1,5)	76 (1,4)	97 (0,6)	
Belgia (flamandzka)	3 (0,4)	27 (1,5)	73 (1,4)	96 (0,6)	
Francja	2 (0,3)	20 (1,2)	58 (1,6)	88 (1,1)	
Portugalia	2 (0,3)	25 (1,2)	72 (1,6)	96 (0,6)	
Cypr	2 (0,3)	18 (1,1)	56 (1,4)	86 (1,0)	
Chile	2 (0,2)	16 (1,2)	53 (1,4)	85 (1,2)	
Gruzja ^{1†}	1 (0,6)	12 (1,3)	41 (1,7)	74 (1,7)	
Arabia Saudyjska	1 (0,3)	8 (0,9)	25 (1,4)	48 (1,8)	
Iran	1 (0,3)	9 (0,8)	33 (1,5)	61 (1,7)	
Maroko ^ψ	1 (0,2)	5 (0,6)	17 (1,2)	35 (1,5)	
Indonezja	1 (0,2)	6 (0,7)	24 (1,8)	51 (2,1)	
Kuwejt ^ψ	1 (0,2)	4 (0,6)	15 (1,4)	33 (1,9)	
Mediana międzynarodowa	7	39	77	95	

Kolejność i kolory poziomów jak w Tabeli 3.2.
Znaczenia symboli jak w Tabeli 3.1.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 3.4. Wyniki pomiaru osiągnięć przyrodniczych według treści przedmiotowej

Kraj	Wiedza o życiu		Wiedza o materii nieożywionej		Wiedza o Ziemi	
	Średnia	Różnica*	Średnia	Różnica*	Średnia	Różnica*
Singapur ²	606 (4,4)	16 (1,3) ⊖	603 (3,7)	13 (1,1) ⊖	546 (3,7)	-44 (1,4) ⊖
Korea Południowa	581 (1,9)	-8 (1,1) ⊖	598 (2,1)	8 (1,5) ⊖	590 (4,1)	1 (3,9)
Japonia	556 (2,2)	-13 (1,5) ⊖	587 (2,7)	18 (2,6) ⊖	562 (2,5)	-6 (1,4) ⊖
Federacja Rosyjska	568 (3,1)	1 (1,4)	567 (3,6)	0 (2,2)	562 (4,7)	-5 (2,8)
Hong Kong [†]	550 (3,7)	-6 (1,6) ⊖	555 (3,5)	-2 (2,1)	574 (3,1)	18 (1,7) ⊖
Tajwan	545 (2,0)	-10 (1,4) ⊖	568 (1,9)	13 (1,5) ⊖	555 (2,6)	0 (1,8)
Finlandia	556 (2,6)	2 (2,0)	547 (2,2)	-6 (1,6) ⊖	560 (2,6)	6 (2,1) ⊖
Kazachstan	545 (4,1)	-5 (1,3) ⊖	559 (5,9)	9 (1,9) ⊖	542 (5,4)	-8 (2,5) ⊖
POLSKA	556 (2,6)	9 (1,8) ⊖	540 (2,1)	-7 (1,5) ⊖	540 (2,6)	-7 (1,3) ⊖
Stany Zjednoczone ^{2†}	555 (2,3)	9 (1,0) ⊖	537 (2,6)	-8 (1,0) ⊖	539 (2,4)	-7 (1,3) ⊖
Słowenia	544 (2,4)	2 (2,0)	546 (2,4)	4 (2,2)	530 (4,1)	-12 (2,2) ⊖
Węgry	550 (3,4)	8 (1,0) ⊖	534 (3,5)	-8 (0,9) ⊖	535 (4,0)	-7 (2,6) ⊖
Szwecja ²	539 (3,3)	-1 (1,3)	534 (3,6)	-6 (1,5) ⊖	551 (4,1)	11 (2,3) ⊖
Norwegia	546 (2,6)	8 (1,2) ⊖	522 (2,8)	-15 (1,8) ⊖	549 (3,8)	11 (2,2) ⊖
Anglia	536 (2,5)	0 (1,4)	540 (2,7)	4 (1,8) ⊖	527 (3,3)	-9 (2,0) ⊖
Bułgaria	541 (6,3)	6 (1,9) ⊖	529 (6,5)	-6 (2,0) ⊖	532 (6,9)	-4 (3,6)
Czechy	538 (2,0)	4 (1,6) ⊖	531 (2,4)	-4 (1,4) ⊖	531 (3,0)	-3 (1,5)
Chorwacja	531 (2,6)	-3 (1,4)	535 (2,9)	2 (2,0)	535 (3,5)	2 (2,3)
Irlandia	531 (2,7)	1 (2,4)	525 (3,3)	-5 (1,5) ⊖	536 (2,9)	7 (1,4) ⊖
Niemcy	528 (2,0)	-1 (1,2)	532 (2,5)	4 (2,0)	519 (4,0)	-10 (2,9) ⊖
Litwa ²	527 (3,0)	-1 (2,5)	535 (2,5)	8 (2,0) ⊖	515 (3,7)	-12 (2,5) ⊖
Dania ^{2†}	534 (2,4)	7 (1,5) ⊖	516 (2,7)	-11 (1,6) ⊖	530 (3,1)	3 (2,2)
Kanada ^{1,2†}	535 (2,6)	11 (1,1) ⊖	518 (2,5)	-7 (0,9) ⊖	513 (2,9)	-12 (1,9) ⊖
Serbia ³	531 (3,8)	6 (2,2) ⊖	529 (3,8)	4 (2,1) ⊖	496 (4,8)	-29 (2,8) ⊖
Australia	531 (3,1)	7 (1,4) ⊖	516 (2,7)	-8 (1,1) ⊖	520 (3,3)	-4 (1,9) ⊖
Słowacja	517 (3,0)	-3 (1,6)	526 (3,4)	5 (2,1) ⊖	513 (2,9)	-7 (1,4) ⊖
Irlandia Północna [†]	521 (2,7)	1 (1,7)	514 (2,6)	-6 (1,6) ⊖	522 (3,0)	2 (2,1)
Hiszpania ²	523 (2,6)	5 (1,8) ⊖	507 (2,9)	-11 (1,5) ⊖	519 (3,1)	1 (2,1)
Holandia [†]	525 (2,7)	8 (1,8) ⊖	504 (2,6)	-13 (1,3) ⊖	520 (3,0)	3 (2,2)
Włochy ²	519 (2,7)	2 (1,9)	513 (2,8)	-3 (1,5)	510 (3,6)	-6 (2,4) ⊖
Belgia (flamandzka) [†]	513 (2,4)	1 (1,1)	506 (3,2)	-5 (2,0) ⊖	512 (2,9)	1 (1,2)
Portugalia ²	508 (2,1)	0 (0,9)	502 (3,0)	-6 (2,0) ⊖	512 (2,5)	4 (1,8) ⊖
Nowa Zelandia	511 (2,8)	6 (1,4)	497 (2,5)	-8 (1,2) ⊖	506 (3,4)	0 (2,0)
Francja	490 (3,1)	2 (1,2)	482 (2,7)	-6 (0,9) ⊖	484 (4,7)	-3 (2,8)
Turcja	472 (3,3)	-11 (1,1) ⊖	496 (3,3)	13 (1,5) ⊖	480 (3,3)	-4 (1,6) ⊖
Cypr	481 (2,8)	0 (0,9)	486 (2,6)	5 (1,2) ⊖	462 (3,5)	-19 (1,7) ⊖
Chile	487 (2,6)	10 (1,2) ⊖	466 (2,9)	-12 (2,2) ⊖	464 (3,4)	-13 (2,4) ⊖
Bahrajn ²	455 (2,9)	-4 (1,6) ⊖	465 (3,3)	6 (1,6) ⊖	448 (3,3)	-11 (3,2) ⊖
Zjednoczone Emiraty Arabskie	449 (3,3)	-2 (1,2)	453 (3,0)	2 (0,7) ⊖	448 (3,4)	-3 (1,7)
Gruzja ^{††}	459 (4,1)	7 (1,6) ⊖	438 (4,7)	-13 (1,7) ⊖	441 (4,3)	-10 (1,6)
Katar	436 (4,3)	0 (1,7)	436 (4,7)	-1 (2,4)	427 (5,0)	-9 (3,5)
Oman	426 (3,2)	-5 (2,1) ⊖	435 (3,5)	4 (1,8) ⊖	423 (3,6)	-8 (2,4) ⊖
Iran	417 (4,5)	-4 (2,7)	423 (5,0)	2 (2,2)	408 (4,8)	-13 (3,7) ⊖
Indonezja	387 (5,1)	-10 (1,9) ⊖	405 (5,5)	8 (2,0) ⊖	384 (5,6)	-13 (2,8) ⊖
Arabia Saudyjska	382 (4,8)	-9 (2,4) ⊖	390 (5,6)	-1 (2,0)	395 (4,8)	5 (1,9) ⊖
Maroko ^ψ	351 (4,2)	-2 (1,8)	357 (5,7)	5 (3,0)	290 (6,3)	-62 (3,2) ⊖
Kuwejt ^ψ	332 (6,6)	-6 (2,8) ⊖	325 (6,6)	-12 (3,6) ⊖	334 (6,4)	-4 (5,4)

W nawiasach błędy standardowe. Z powodu zaokrąglenia liczb niektóre wyniki mogą się wydawać niespójne.

Kolejność krajów według Tabeli 3.1.

* Różnica między średnim wynikiem subtestu i całego testu w danym kraju.

⊖ Średni wynik subtestu istotnie wyższy niż średni wynik całego testu w kraju.

⊕ Średni wynik subtestu istotnie niższy niż średni wynik całego testu w kraju.

Pozostałe symbole jak w Tabeli 3.1.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 3.5. Wyniki pomiaru osiągnięć przyrodniczych według funkcji podmiotowej

Kraj	Zasoby		Zastosowania typowe		Zastosowania problemowe	
	Średnia	Różnica*	Średnia	Różnica*	Średnia	Różnica*
Singapur ²	574 (4,1)	-16 (1,3) U	599 (4,0)	9 (1,3) U	605 (3,6)	15 (1,8) U
Korea Południowa	582 (2,2)	-7 (1,2) U	593 (1,9)	4 (1,8) U	594 (2,2)	5 (1,6) U
Japonia	543 (2,3)	-25 (1,3) U	576 (1,8)	7 (0,8) U	594 (1,9)	25 (1,6) U
Federacja Rosyjska	568 (3,9)	1 (1,9)	568 (3,3)	1 (1,3)	560 (3,7)	-7 (2,5) U
Hong Kong [†]	562 (3,0)	5 (1,9) U	554 (3,3)	-3 (1,6)	552 (4,1)	-4 (2,5)
Tajwan	557 (2,5)	2 (1,6)	553 (2,6)	-2 (1,8)	558 (3,2)	2 (2,3)
Finlandia	556 (3,1)	2 (1,6)	553 (2,4)	-1 (1,9)	552 (2,3)	-2 (1,5)
Kazachstan	551 (5,0)	1 (2,4)	547 (4,6)	-3 (2,1)	552 (4,5)	2 (1,5)
POLSKA	543 (2,5)	-4 (1,1) U	554 (2,8)	7 (2,2) U	542 (3,2)	-5 (2,0) U
Stany Zjednoczone ^{2†}	548 (2,5)	2 (0,7)	546 (2,2)	0 (1,3)	541 (2,7)	-4 (1,4) U
Słowenia	541 (2,6)	-2 (1,4)	546 (2,8)	3 (1,9)	538 (2,6)	-4 (2,2)
Węgry	550 (3,8)	8 (1,2) U	539 (3,4)	-3 (1,1) U	533 (3,9)	-9 (1,6) U
Szwecja ²	538 (3,8)	-2 (1,1)	540 (3,4)	0 (1,5)	542 (3,8)	2 (3,0)
Norwegia	532 (3,0)	-5 (1,6) U	541 (2,9)	4 (1,0) U	537 (3,9)	-1 (2,8)
Anglia	533 (2,6)	-2 (1,3)	538 (2,7)	2 (1,3)	539 (2,8)	3 (1,7)
Bułgaria	551 (6,4)	15 (2,2) U	536 (6,2)	0 (1,2)	507 (6,4)	-29 (1,5) U
Czechy	544 (3,0)	10 (1,9) U	528 (2,1)	-6 (1,6) U	529 (2,4)	-6 (1,4) U
Chorwacja	534 (2,9)	1 (1,7)	530 (2,2)	-3 (1,7)	535 (2,4)	2 (2,5)
Irlandia	529 (2,4)	-1 (2,1)	531 (2,3)	2 (1,6)	526 (3,2)	-3 (1,4) U
Niemcy	527 (2,9)	-1 (1,5)	529 (2,4)	0 (1,0)	531 (2,3)	3 (1,8)
Litwa ²	523 (3,0)	-4 (2,4)	526 (2,4)	-1 (1,2)	537 (3,0)	10 (2,4) U
Dania ^{2†}	524 (2,6)	-3 (1,7)	529 (2,4)	2 (1,3)	525 (2,9)	-1 (2,7)
Kanada ^{12†}	523 (3,0)	-2 (1,8)	527 (2,4)	3 (0,9) U	524 (2,5)	0 (1,3)
Serbia ³	526 (3,9)	2 (1,4)	522 (4,5)	-3 (1,8)	520 (3,9)	-4 (2,9)
Australia	523 (3,3)	-1 (1,7)	522 (2,7)	-1 (1,3)	527 (3,0)	4 (1,6)
Słowacja	529 (3,3)	9 (1,5) U	517 (2,8)	-4 (1,1) U	507 (3,4)	-13 (2,3) U
Irlandia Północna [†]	518 (2,9)	-1 (1,7)	519 (2,9)	-1 (1,9)	520 (2,6)	0 (1,7)
Hiszpania ²	522 (3,3)	4 (2,0) U	514 (3,2)	-4 (2,0) U	516 (2,6)	-2 (1,2)
Holandia [†]	508 (2,4)	-9 (1,3) U	519 (2,5)	2 (1,4)	526 (2,9)	9 (2,3) U
Włochy ²	520 (3,1)	4 (1,4)	513 (3,1)	-3 (1,3) U	511 (3,6)	-5 (2,2) U
Belgia (flamandzka) [†]	498 (2,7)	-14 (1,3) U	513 (2,5)	2 (0,8) U	526 (2,9)	14 (2,0) U
Portugalia ²	507 (2,8)	-1 (2,5)	508 (1,9)	0 (1,7)	506 (1,9)	-2 (2,4)
Nowa Zelandia	504 (2,8)	-2 (2,2)	502 (3,1)	-3 (1,5) U	514 (2,4)	8 (1,0) U
Francja	482 (3,8)	-6 (2,0) U	494 (3,1)	6 (1,6) U	481 (2,8)	-6 (1,4) U
Turcja	478 (3,0)	-6 (1,2) U	486 (3,1)	3 (1,3) U	483 (3,3)	0 (1,8)
Cypr	467 (3,2)	-14 (2,1) U	489 (3,4)	8 (1,9) U	490 (3,6)	8 (2,6) U
Chile	477 (3,2)	0 (1,9)	476 (3,0)	-2 (1,4)	477 (2,6)	-1 (2,5)
Bahrajn ²	455 (2,6)	-3 (1,8)	462 (3,0)	3 (2,0)	455 (3,0)	-4 (2,1)
Zjednoczone Emiraty Arabskie	453 (3,3)	2 (1,0) U	452 (3,2)	1 (1,0)	444 (3,0)	-7 (1,2) U
Gruzja ^{1†}	460 (4,2)	8 (1,6) U	449 (4,8)	-2 (2,2)	426 (4,0)	-25 (1,6) U
Katar	437 (4,5)	1 (2,5)	431 (4,6)	-6 (1,8) U	433 (4,4)	-3 (2,2)
Oman	421 (3,3)	-10 (2,1) U	435 (2,9)	4 (1,7) U	431 (3,0)	0 (1,3)
Iran	416 (4,1)	-5 (2,4) U	418 (4,5)	-3 (3,3)	422 (4,9)	1 (2,5)
Indonezja	397 (4,9)	1 (2,3)	392 (5,3)	-5 (2,9)	389 (5,5)	-7 (1,9) U
Arabia Saudyjska	394 (5,3)	4 (2,5)	388 (4,7)	-2 (2,3)	365 (5,4)	-25 (4,2) U
Maroko ^ψ	331 (5,2)	-21 (2,4) U	357 (4,4)	5 (2,0) U	353 (4,5)	1 (2,4)
Kuwejt ^ψ	343 (6,4)	6 (2,3) U	325 (7,2)	-13 (3,2) U	297 (8,1)	-41 (4,4) U

W nawiasach błędy standardowe. Z powodu zaokrąglenia liczb niektóre wyniki mogą się wydawać niespójne. Kolejność krajów według Tabeli 3.1.

* Różnica między średnim wynikiem subtestu i całego testu w danym kraju.
U Średni wynik subtestu istotnie wyższy niż średni wynik całego testu w kraju.
U Średni wynik subtestu istotnie niższy niż średni wynik całego testu w kraju.
 Pozostałe symbole jak w Tabeli 3.1,

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 3.6. Osiągnięcia przyrodnicze dziewczynek i chłopców

Kraj	Dziewczynki	Chłopcy	Różnica 2015	Różnica 2011
Arabia Saudyjska	431 (5,3)	352 (7,7)	79 (8,9) ♀	48 (11,0) ♀
Bahrajn ²	478 (3,0)	439 (3,6)	39 (4,0) ♀	23 (7,0) ♀
Oman	447 (3,4)	415 (3,5)	32 (3,1) ♀	34 (3,8) ♀
Kuwejt ^ψ	352 (7,5)	323 (7,6)	30 (9,1) ♀	53 (8,6) ♀
Katar	448 (4,7)	424 (6,0)	24 (7,2) ♀	26 (6,5) ♀
Zjednoczone Emiraty Arabskie	459 (4,4)	445 (4,0)	14 (6,4) ♀	18 (5,3) ♀
Finlandia	560 (2,3)	548 (2,9)	12 (2,5) ♀	0 (3,0)
Iran	427 (5,2)	415 (5,6)	11 (7,4)	-2 (8,8)
Maroko ^ψ	358 (5,1)	347 (5,1)	11 (5,1) ♀	9 (4,4) ♀
Bułgaria	540 (6,3)	531 (5,9)	8 (2,9) ♀	-
Szwecja ²	544 (4,1)	536 (3,5)	8 (2,7) ♀	-4 (3,0)
Indonezja	401 (5,2)	393 (5,3)	8 (4,2)	-
Kazachstan	552 (4,5)	547 (4,7)	5 (2,7) ♀	-8 (3,0) ♂
Gruzja	453 (3,9)	449 (4,6)	4 (4,1)	9 (3,9) ♀
Nowa Zelandia	507 (3,2)	504 (3,0)	3 (3,1)	-1 (3,2)
Litwa ²	529 (2,9)	526 (3,1)	3 (3,4)	-1 (2,6)
Serbia ³	526 (3,7)	523 (4,9)	3 (4,6)	-3 (3,9)
Kanada ^{1 2†}	526 (2,3)	524 (3,0)	2 (2,1)	-
Belgia (flamandzka)	512 (2,6)	511 (2,6)	2 (2,4)	-11 (2,9) ♂
POLSKA	548 (2,6)	546 (3,0)	1 (2,8)	-6 (2,8) ♂
Turcja	484 (3,3)	483 (4,0)	1 (3,1)	4 (3,8)
Australia	524 (3,3)	523 (3,4)	1 (3,4)	0 (3,9)
Holandia [†]	517 (2,8)	517 (3,0)	1 (2,4)	-10 (2,1) ♂
Anglia	536 (3,0)	535 (2,6)	1 (2,8)	1 (3,1)
Norwegia	538 (3,1)	537 (3,1)	1 (3,2)	-4 (3,1)
Singapur ²	590 (3,7)	590 (4,2)	0 (2,8)	-4 (2,7)
Federacja Rosyjska	567 (3,1)	567 (3,8)	0 (2,7)	1 (2,4)
Francja	487 (3,1)	487 (2,9)	0 (2,4)	-
Irlandia Północna [‡]	520 (3,0)	520 (2,8)	0 (3,7)	1 (3,8)
Cypr	481 (2,8)	481 (2,9)	0 (2,6)	-
Chile	477 (2,9)	478 (3,4)	-1 (3,3)	-12 (2,9) ♂
Niemcy	572 (2,7)	529 (2,7)	-2 (2,4)	-12 (2,5) ♂
Chorwacja	532 (2,8)	534 (2,2)	-2 (2,8)	-5 (2,7)
Stany Zjednoczone ^{2†}	544 (2,4)	548 (2,5)	-4 (2,0) ♂	-10 (1,5) ♂
Dania ^{2†}	525 (2,4)	529 (2,6)	-4 (2,8)	-2 (3,0)
Japonia	567 (2,0)	571 (2,3)	-4 (2,4)	-5 (2,8)
Irlandia	527 (3,1)	532 (3,5)	-5 (3,6)	1 (5,5)
Hiszpania ²	515 (2,9)	521 (2,9)	-6 (2,7) ♂	-10 (2,8) ♂
Słowenia	539 (2,4)	546 (3,1)	-7 (2,7) ♂	-6 (3,2)
Portugalia ²	504 (2,5)	511 (2,4)	-7 (2,2) ♂	-5 (3,2)
Węgry	538 (3,5)	545 (3,9)	-7 (3,1) ♂	-5 (2,9)
Czechy	530 (2,8)	538 (2,7)	-8 (2,6) ♂	-15 (2,6) ♂
Słowacja	516 (3,2)	524 (2,7)	-8 (2,7) ♂	-8 (2,7) ♂
Tajwan	551 (2,2)	559 (2,4)	-9 (2,9) ♂	-7 (2,3) ♂
Włochy ²	512 (3,0)	521 (2,8)	-9 (2,5) ♂	-7 (2,9) ♂
Hong Kong [†]	551 (3,9)	561 (3,3)	-10 (3,9) ♂	-6 (2,5) ♂
Korea Południowa	584 (2,3)	595 (2,3)	-11 (2,4) ♂	-8 (2,3) ♂
Średnia międzynarodowa	508 (0,5)	504 (0,6)		

W nawiasach błędy standardowe.

Z powodu zaokrąglenia liczb niektóre wyniki mogą się wydawać niespójne.

♀ Średni wynik dziewczynek istotnie wyższy niż średni wynik chłopców.

♂ Średni wynik dziewczynek istotnie niższy niż średni wynik chłopców.

Pozostałe symbole jak w Tabeli 2.1.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Wśród tych pytań nasi eksperci znaleźli 74 wyraźnie wykraczające poza typowe programy nauczania o przyrodzie w klasach od pierwszej do czwartej. Według tego kryterium do lokalnego zbioru nie weszło 39 z 64 pytań z zakresu fizyki, 19 z 33 pytań z zakresu geografii i 16 z 63 pytań z biologii. Jak łatwo obliczyć, w lokalnym zbiorze ostało się 75 proc. pytań biologicznych, 42 proc. pytań geograficznych i 39 proc. pytań fizycznych. Liczba punktów możliwych do zdobycia za zadania mieszczące się w naszych programach kształcenia wyniosła 111 (59 proc. wszystkich), znacznie mniej niż analogicznym zbiorze zadań matematycznych (86 proc.).

Porównajmy procent poprawnych odpowiedzi polskich uczniów w całym teście i w zbiorze zadań dostosowanym do programu nauczania przyrody. Pierwszy wynik: 57 proc. i dziesiąta ranga razem z Bułgarami i Amerykanami. Drugi: 60 proc. i ósma ranga razem z Bułgarami i mieszkańcami Hong Kongu. Test dostosowany do programu daje nam przyrost trzech punktów procentowych i awans o dwa miejsca – więcej niż w pomiarze osiągnięć matematycznych. Dlaczego? Czy wskutek usunięcia większej liczby pytań, czy może dlatego, że matematyka jest wszędzie, a z przyrodoznawstwem uczeń obcuje głównie w szkole?

W próbie krajów średni procent poprawnych odpowiedzi w całym teście wahał się od 25 w Kuwejcie do 67 w Singapurze wokół średniej 49,7 z odchyleniem standardowym 9,8. Porównajmy z tymi wynikami średni procent poprawnych odpowiedzi w lokalnych zbiorach zadań. Wahał się on od 26 w Kuwejcie do 81 w Singapurze wokół średniej 51,8 z odchyleniem standardowym 11,2. Widać, że obie średnie różnią się jedynie o 2,1 punktu procentowego. W dwóch krajach: Arabii Saudyjskiej i Zjednoczonych Emiratach Arabskich różnica była ujemna, ale jej maksimum (14 punktów w Singapurze) znacznie przewyższyło maksimum różnicy w analogicznej analizie osiągnięć matematycznych. Można dodać, że tuż za Singapurem ulokowała się Korea z różnicą dziewięciopunktową i Japonia z ośmiopunktową. Te różnice tłumaczą odrobinę mniejszy współczynnik korelacji między dwoma agregatami ($r = 0,98$).

Wyniki powyższej analizy sugerują, że przyrodoznawstwo ma ściślejszy związek ze szkołą niż matematyka: prawdopodobieństwo, że uczeń nie będzie wiedział tego, czego mu szkoła nie powie, jest w przyrodoznawstwie większe niż w matematyce. Mimo to prawomocność testu przyrodniczego nie jest zagrożona, ponieważ „oczyszczenie” zbioru pytań niemal nic nie zmienia w naszej wiedzy o osiągnięciach przyrodniczych.

Wyniki pomiaru a płeć uczniów

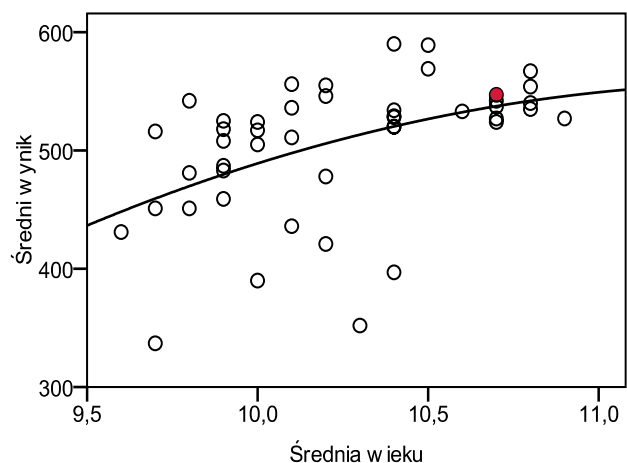
Porównanie osiągnięć przyrodniczych dziewczynek i chłopców w populacji międzynarodowej ujawnia istotną, czteropunktową przewagę tych pierwszych – wyraźnie większą niż dwupunktowa w 2011 r. Podobną zmianę widzimy w Polsce: przed czterema laty chłopcy istotnie wyprzedzili dziewczynki aż sześcioma punktami, a w 2015 r. dziewczynki odrobinę (nieistotnie) wyprzedziły chłopców. Byłaby to zmiana emancypacyjna?

Dane z kolejnych edycji TIMSS świadczą o silnym i dość stabilnym zróżnicowaniu krajów pod względem pozycji dziewczynek w szkolnym przyrodoznawstwie. W arabskojęzycznych krajach Bliskiego Wschodu pozycja dziewczynek zawsze była wysoka zarówno w szkole podstawowej, jak i średniej. W wysoko rozwiniętych krajach środkowej i południowej Europy, a także w krajach Dalekiego Wschodu pozycja dziewczynek była niska. W krajach północnej Europy i byłego imperium brytyjskiego pozycja dziewczynek była zbliżona do pozycji chłopców. Tabela 3.6 przedstawia wyniki dziewczynek i chłopców w obecnej edycji TIMSS, a także przypomina różnice między wynikami dziewczynek i chłopców zarejestrowane w poprzedniej edycji. Analiza tych danych prowadzi do wniosku, że w 2015 r. podział sprzed czterech lat z grubsza się odtworzył (współczynnik korelacji Spearmana między różnicami osiągnięć dziewczynek i chłopców w 42 krajach wynosi 0,67), ale ze znamienym przesunięciem. W 23 krajach pozycja dziewczynek wzrosła średnio o 8,5 punktu, a tylko w 17 krajach spadła, i tylko średnio o 3,9 punktu. To przesunięcie na korzyść dziewczynek – zbyt powszechne, żeby było przypadkowe – może mieć różne przyczyny. Nie można wykluczyć, że wśród nich jest rosnąca niezgoda na zamykanie kobiet w tradycyjnej roli płciowej.

Wyniki pomiaru a wiek uczniów

Analiza regresji średnich osiągnięć na średni wiek testowanych dzieci (z wyłączeniem RPA i Jordanii, których dane nie zostały zamieszczone w międzynarodowym raporcie) przynosi współczynnik R^2 równy 0,23. Krzywoliniowość, którą obserwowaliśmy w poprzednim rozdziale, tu mniej rzuca się w oczy (Rysunek 3.1).

Związek wieku z osiągnięciami przyrodniczymi w Polsce badamy w sposób opisany w poprzednim rozdziale. Na osi odciętych Rysunku 3.2 odłożono kwartały urodzenia. Kolorem czerwonym oznaczono średnie warunkowe wyników testowania w 2015 r., a szarym – w 2011 r., gdy testowani wtedy trzeciokla-



Rysunek 3.1. Średnie krajowe wyników testu przyrodniczego w zależności od średnich wieku uczniów

Czerwone kółeczko oznacza Polskę.

Na podstawie: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

siści byli w takim samym wieku co młodszy rocznik czwartoklasistów z 2015 r.

Ogółem zmienna wieku uczniów w niewielkim stopniu odtwarza wyniki pomiaru osiągnięć przyrodniczych: redukuje wariancję wewnątrz oddziałów o mniej niż 1 proc., a między oddziałami o 10 proc.¹ – więcej niż w wypadku matematyki.

Rysunek 3.2 pokazuje, po pierwsze, że obecni czwartoklasiści umieją znacznie więcej niż dawni

¹ HLM z kontrolą płci, lokalizacji szkoły i SES na obu poziomach.

trzecioklasiści w tym samym wieku². Wzrost osiągnięć to efekt dodatkowego roku nauki w bardziej zaawansowanym trybie.

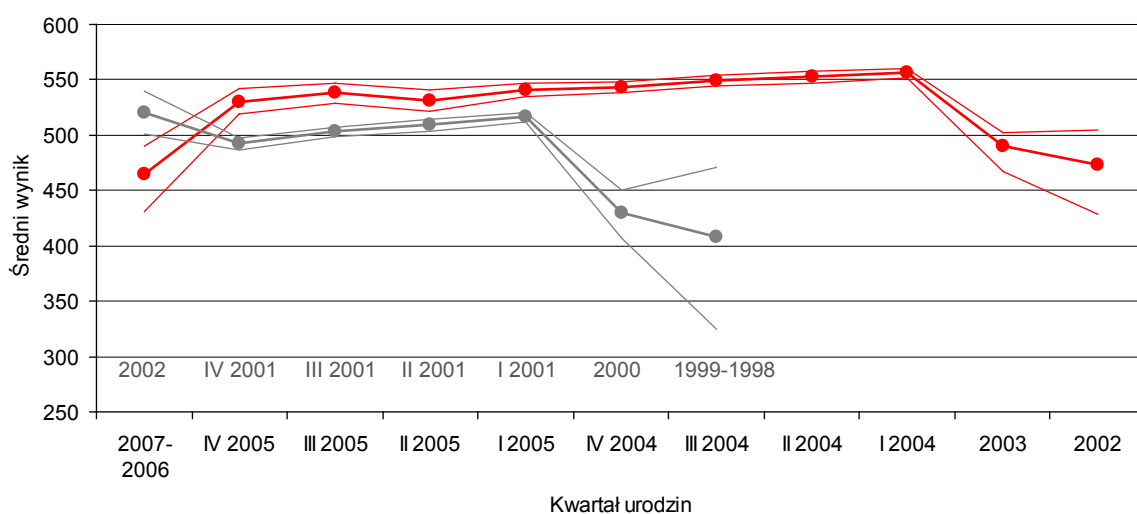
Po wtóre, wśród czwartoklasistów z rocznika 2005 nie ma efektu kwartału urodzenia (osiągnięcia dzieci urodzonych w ostatnim i w pierwszym kwartale tego roku nie różnią się istotnie), a wśród trzecioklasistów w 2011 był. Entuzjazm studzi jednak obecność tego efektu u czwartoklasistów z rocznika 2004.

Po trzecie, inaczej niż w wypadku matematyki, średni wynik czwartoklasistów z rocznika 2004 jest wyższy niż z rocznika 2005. Różnica – 6,5 punktu, czyli 0,09 odchylenia standardowego – jest statystycznie istotna, ale tak mała, że nie ma praktycznego znaczenia. Ponieważ w zdrowej szkole efekt względnego wieku zmniejsza się z roku na rok (np. Oshima i Domaleski, 2006), zapewne i ta różnica wkrótce zniknie zupełnie.

Zmiany osiągnięć przyrodniczych

Tabela 3.7 przedstawia różnice między kolejnymi pomiarami osiągnięć przyrodniczych w 45 krajach, które brały udział w przynajmniej dwóch edycjach badania w latach 1995–2015. Danymi ze wszystkich pięciu pomiarów dysponuje jedynie 12 krajów. Biorąc pod uwagę wyłącznie kraje, które mają przynajmniej trzy pomiary, zauważamy cztery rodzaje zmian.

² Z wyjątkiem najmłodszych, którzy rozpoczęli naukę przed terminem, ale ta grupa była w 2015 r. średnio o cztery miesiące młodsza niż analogiczna w 2011 r.



Rysunek 3.2. Średnie wyniki testu przyrodniczego TIMSS 2011 i 2015 w Polsce w zależności od kwartału urodzenia uczniów

Średnie warunkowe oszacowano za pomocą ważonej analizy wariancji przy kontroli statusu socjoekonomicznego (opisanego w Rozdziale 4), a błędy standardowe – metodą *bootstrap*. Cienkie linie przedstawiają granice 95-procentowych przedziałów ufności.

Skrajne kwartały połączono z powodu zbyt małej liczby uczniów.

Etykiety nad osią odciętych odnoszą się do szarej krzywej.

Tabela 3.7. Osiągnięcia przyrodnicze w kolejnych edycjach badania TIMSS

Kraj	1995	2003	2007	2011	2015
Anglia	528 (3,2) ^b	540 (3,5) ^a	542 (2,8) ^a	529 (3,0) ^a	536 (2,4) ^a
Arabia Saudyjska				429 (5,5) ^b	390 (5,0) ^a
Armenia		437 (4,3) ^b		416 (3,8) ^a	
Australia	521 (3,7) ^{ba}	521 (4,3) ^{ba}	527 (3,3) ^b	516 (2,9) ^a	524 (2,9) ^{ba}
Austria	538 (3,6) ^b		526 (2,5) ^a	532 (2,8) ^a	
Bahrajn				449 (3,5) ^b	459 (2,6) ^a
Belgia (flamandzka)		518 (1,9) ^b		509 (2,0) ^a	511 (2,3) ^a
Chile				480 (2,5) ^a	478 (2,7) ^a
Chorwacja				516 (2,2) ^b	533 (2,1) ^a
Cypr	450 (3,4) ^b	480 (2,4) ^a			481 (2,6) ^a
Czechy	532 (3,1) ^a		515 (3,0) ^b	511 (2,4) ^a	528 (2,2) ^a
Dania			517 (2,9) ^b	528 (2,8) ^a	527 (2,1) ^a
Federacja Rosyjska		526 (5,3) ^c	546 (5,0) ^b	552 (3,4) ^b	567 (3,2) ^a
Finlandia				570 (2,6) ^b	554 (2,3) ^a
Gruzja			418 (4,6) ^b	455 (3,9) ^a	451 (3,7) ^a
Hiszpania				505 (3,1) ^b	518 (2,6) ^a
Holandia	530 (3,2) ^c	525 (2,0) ^c	523 (2,6) ^{ac}	531 (2,2) ^b	517 (2,7) ^a
Hong Kong	508 (3,4) ^e	542 (3,0) ^{cd}	554 (3,5) ^a	535 (3,7) ^{bc}	556 (2,9) ^a
Iran	380 (4,6) ^e	414 (4,2) ^d	436 (4,4) ^c	453 (3,8) ^b	421 (4,0) ^{ad}
Irlandia	515 (3,5) ^b			516 (3,3) ^b	529 (2,8) ^a
Irlandia Północna				517 (2,5) ^a	520 (2,2) ^a
Japonia	553 (1,7) ^e	543 (1,5) ^d	548 (2,1) ^{ced}	559 (1,9) ^b	569 (1,8) ^a
Katar				394 (4,3) ^b	436 (4,1) ^a
Kazachstan				495 (5,1) ^b	550 (4,4) ^a
Korea Południowa	576 (2,1) ^b			587 (2,1) ^a	589 (2,0) ^a
Kuwejt				347 (4,8) ^b	315 (5,0) ^a
Maroko				264 (4,4) ^b	352 (4,4) ^a
Litwa		512 (2,6) ^b	514 (2,4) ^b	515 (2,4) ^b	530 (2,7) ^a
Niemcy			528 (2,4) ^a	528 (2,9) ^a	528 (2,4) ^a
Norwegia	504 (3,7) ^d	466 (2,6) ^c	477 (3,5) ^b	494 (2,5) ^a	493 (2,2) ^a
Nowa Zelandia	505 (5,4) ^{ab}	520 (2,4) ^{bc}	504 (2,7) ^a	497 (2,4) ^b	505 (2,7) ^a
Oman				377 (4,3) ^b	431 (3,1) ^a
Portugalia	452 (4,1) ^c			522 (3,8) ^b	508 (2,2) ^a
Serbia				516 (3,1) ^a	524 (3,7) ^a
Singapur	523 (4,8) ^c	565 (5,5) ^b	587 (4,1) ^a	583 (3,4) ^a	590 (3,7) ^a
Słowacja			526 (4,8) ^{ab}	532 (3,7) ^b	520 (2,6) ^a
Słowenia	464 (3,1) ^e	490 (2,6) ^d	518 (1,9) ^{bc}	520 (2,6) ^b	542 (2,4) ^a
Stany Zjednoczone	542 (3,4) ^{ab}	536 (2,5) ^b	539 (2,7) ^b	544 (2,1) ^a	546 (2,2) ^a
Szwecja			525 (2,9) ^b	533 (2,8) ^a	540 (3,6) ^a
Tajwan		551 (1,8) ^b	557 (2,0) ^a	552 (2,2) ^{ab}	555 (1,8) ^{ab}
Tunezja		314 (5,7) ^b	318 (5,9) ^b	346 (5,3) ^a	
Turcja				463 (4,7) ^b	483 (3,3) ^a
Węgry	508 (3,4) ^c	530 (2,8) ^b	536 (3,4) ^{ab}	534 (3,7) ^{ab}	542 (3,3) ^a
Włochy		516 (3,8) ^a	535 (3,2) ^c	524 (2,7) ^b	516 (2,6) ^a
Zjednoczone Emiraty Arabskie				428 (2,5) ^b	451 (2,8) ^a

W nawiasach błędy standardowe. Z powodu zaokrąglenia liczb niektóre wyniki mogą się wydawać niespójne. Dwie średnie niemające takich samych liter w superskrypcie istotnie różnią się od siebie. Brak Polski jest następstwem zmiany klasy, w której byli uczniowie – z trzeciej w 2011 r. na czwartą w 2015 r.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

- Monotoniczny wzrost (z dopuszczeniem jednego niewielkiego spadku) nastąpił w 15 krajach. Najlepszym przykładem jest tu Słowenia, która w ciągu 20 lat małymi krokami podwyższyła średnią o ponad jedno odchylenie standardowe. Podobną dynamiką mogą się poszczycić Rosja i Singapur.
- Monotoniczny spadek (z dopuszczeniem jednego

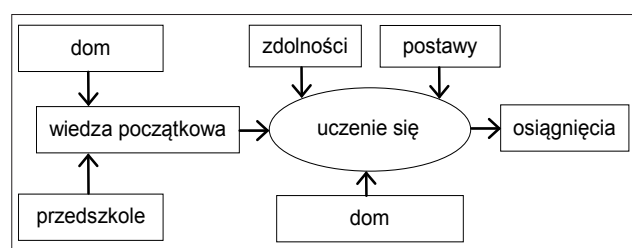
niewielkiego wzrostu) obserwujemy w dwóch krajach: Belgii i Holandii.

- Wahania w górę i w dół zachodziły w trzech krajach: Czechach, Norwegii i we Włoszech.
- Stabilizacja trwała w pięciu krajach: Australii, USA, Niemczech, na Tajwanie i Słowacji. Widać, że najliczniejsza kategoria zmiany to monotoniczny wzrost.

Uczeń i jego rodzina

Krzysztof Konarzewski

W badaniach IEA, podobnie jak w innych badaniach osiągnięć szkolnych, uczeń jest traktowany jednowymiarowo jako istota ucząca się. Nie ma w tym nic dziwnego – to uczenie się bowiem, a nie nauczanie, jest pierwotnym źródłem wszelkich wartości edukacyjnych, wśród nich wartości mierzalnych testami osiągnięć.



Jak pokazuje rysunek, proces uczenia się w szkole czerpie z zasobów wiedzy zakumulowanej przez ucznia w okresie przedszkolnym. Jednym z możliwych źródeł tej wiedzy jest dom rodzinny, innym – przedszkole. Uczenie się w szkole zasilają też zainteresowanie i pomoc okazywane dziecku przez domowników oraz jego zasoby osobowościowe, do których należą m.in. zdolności i postawy wobec uczenia się. Wybrany uwarunkowaniem procesu uczenia jest poświęcony niniejszy rozdział.

Status socjoekonomiczny rodziny

Uczeń w szkole jest najpierw dzieckiem w rodzinie i podlega jej przemożnym wpływom, dlatego opis

ucznia należy zacząć od charakterystyki jego domu rodzinnego. Wśród pojedynczych charakterystyk żadna nie mówi o rodzinie więcej niż status socjoekonomiczny (SES). Jest to zmienna ilościowa, ale wyposażona w „nadznaczenie” – jej przedziały są bowiem skorelowane z jakościowo różnymi kombinacjami wartości rozmaitych zmiennych psycho- i socjologicznych, które za Pierrem Bourdieu (2005) nazywa się habitusem. Podobnie jak w poprzednich edycjach badań IEA w Polsce SES został zdefiniowany jako wspólna wariancja¹ ośmiu zmiennych porządkowych należących do czterech kategorii:

- wykształcenie rodziców (podstawowe, zasadnicze, średnie, policealne, wyższe I stopnia, wyższe II stopnia, doktorat)
- status zawodowy rodziców (osoby bez zawodu, robotnicy niewykwalifikowani, robotnicy wykwalifikowani, urzędnicy, właściciele przedsiębiorstw, specjaliści, kadra kierownicza)
- zatrudnienie rodziców (brak zatrudnienia, zatrudnienie w niepełnym wymiarze czasu pracy lub dorywczo, zatrudnienie w pełnym wymiarze)
- wyposażenie mieszkania (liczba miejsc i urządzeń, z których korzysta dziecko).

SES jest zmienną standaryzowaną o średniej 0 i odchyleniu standardowym 1. Współczynniki ko-

¹ Spójność zmiennych jest zadowalająca (α Cronbacha równa się 0,79, a wskaźnik czynnikowy z analizy głównych składowych zachowuje 41 proc. wariancji zmiennych).

relacji ze wskaźnikami osiągnięć matematycznych ($r = 0,44$) i przyrodniczych (0,43) sugerują, że pominięcie SES w równaniach opisujących osiągnięcia szkolne grozi poważnym obciążeniem oszacowań parametrów.

Przekonuje o tym niemal każda próba odpowiedzi na pytanie o uwarunkowania osiągnięć szkolnych. Tytułem przykładu rozważmy pytanie: Czy osiągnięciom szkolnym sprzyja edukacja przedszkolna? Informacje od rodziców pozwalają określić rozkład lat spędzonych przez badane dzieci w przedszkolu. Przedstawia go Tabela 4.1. Warto zauważyć, że tylko 15 proc. dzieci nie chodziło do przedszkola – znacznie mniej niż w 2011 r. (24 proc.). Tabela pokazuje też średnie SES rodzin uczniów. Widać wyraźną tendencję: status rodziny rośnie wraz liczbą lat pobytu dziecka w przedszkolu ($R^2 = 0,22$, podobnie było w 2011 r.: 0,24). W Polsce posyłali dzieci do przedszkola na kilka lat głównie rodzice z wyższych warstw społecznych.

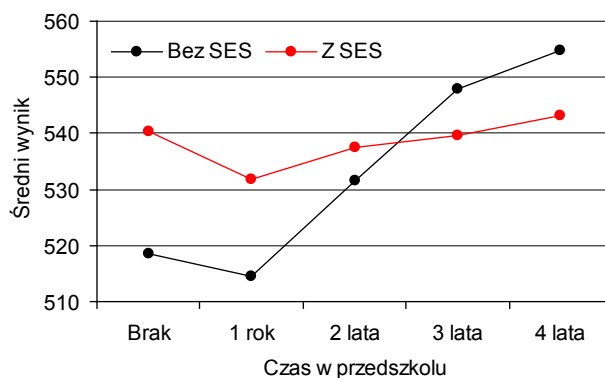
Tabela 4.1. Rozkład procentowy pobytu w przedszkolu i średnie SES rodziny ($n = 4613$)

Okres spędzony w przedszkolu	Procent dzieci	Średnia SES rodziny
Cztery lata	18	0,42 (0,03)
Trzy lata	36	0,30 (0,03)
Dwa lata	18	-0,21 (0,04)
Rok lub mniej	13	-0,63 (0,04)
Nie chodziło do przedszkola	15	-0,79 (0,04)

Ważona ANOVA, błędy standardowe oszacowane metodą *bootstrap*

Zobaczymy, czy lata spędzone w przedszkolu różnicują wskaźniki osiągnięć matematycznych w klasie czwartej. Czarna krzywa na Rysunku 4.1 pokazuje, że gdy w analizie uwzględnić jedynie zmienną lat, odpowiedź jest twierdząca. Wszystkie różnice między osiągnięciami dzieci niechodzących do przedszkola lub chodzących tylko przez rok a pozostałymi są statystycznie istotne. Jeśli jednak do równania regresji włączymy średnią SES, wszystkie średnie stają się nieodróżnialne od siebie, więc odpowiedź jest przecząca: długość kariery przedszkolnej przeciętnego dziecka nie ma żadnego związku z jego osiągnięciami w klasie czwartej.

Którą odpowiedź przyjąć? Pierwsza jest z pewnością błędna, ale druga wcale nie jest pewna, ponieważ można podejrzewać, że uwzględnienie innych zmiennych mogłoby zmodyfikować i ten układ zależności. Jedną z takich zmiennych jest jakość edukacji przedszkolnej. Być może o korzyściach decyduje nie tyle liczba lat spędzonych w przedszkolu, ile to, czym są wypełnione. Liczba lat mogłaby wpływać na karierę szkolną w interakcji z jakością edukacji: długi pobyt



Rysunek 4.1. Średnie osiągnięcia matematyczne w zależności od liczby lat spędzonych przez ucznia w przedszkolu

Ważona analiza regresji na wartościach prawdopodobnych z oszacowaniem błędów standardowych metodą *jackknife* za pomocą aplikacji IDB Analyser 3.2.21.

w dobrym przedszkolu mogłyby podwyższać osiągnięcia szkolne, w marnym – obniżać. Czy jednak tak jest, nie wiemy, ponieważ nie dysponujemy niezależną miarą jakości edukacji w przedszkolach, do których chodziły badane dzieci.

Ogólny wniosek z powyższych analiz brzmi: każda próba odpowiedzi na ważne pytania oświatowe musi opierać się na całościowej teorii zjawiska i przemyślanym modelu statystycznym. Zestawianie ze sobą kilku zmiennych, które badacz akurat ma pod ręką, często prowadzi do odkryć niepewnych. Jeśli miałyby wpłynąć na praktykę lub choćby na opinię publiczną, to lepiej powstrzymać się przed ich ogłaszaniem.

Wiedza początkowa

O roli wiadomości i umiejętności, z którymi dziecko zaczyna naukę w klasie pierwszej, pisano wiele (np. Duncan i in., 2007). Dość wspomnieć, że są one głównym składnikiem tzw. gotowości szkolnej, o której było głośno w Polsce w okresie dyskusji nad zmianą wieku obowiązku szkolnego. W badaniu TIMSS nie mamy oczywiście wskaźnika samej wiedzy początkowej, lecz tylko jej ocenę dokonaną przez rodziców w odpowiedzi na 11 pytań typu: „Gdy Pani/Pana dziecko zaczynało naukę w klasie pierwszej, to jak dobrze umiało...” – np. czytać pojedyncze słowa, zdania, opowiadania, samodzielnie liczyć do 100. Wobec wysokiej spójności odpowiedzi (w Polsce $\alpha = 0,90$) utworzono z nich metodą IRT skalę o średniej 10 i odchyleniu standardowym 2. Średnie wskaźniki zawiera w drugiej kolumnie Tabela 4.2. Najwyżej ocenili wiedzę początkową swoich dzieci rodzice w Korei Południowej, najniżej – na Słowacji. Polska ze średnią 10,5 ma rangę 11,5.

Korzystając z polskich danych, spróbujemy ujawnić rodzinne czynniki wiedzy początkowej, czyli

Tabela 4.2. Średnie wiedzy początkowej, edukacji domowej i postaw rodziców wobec matematyki i przyrodznawstwa

Kraj	Wiedza początkowa	Edukacja domowa	Postawy rodziców
Anglia			
Arabia Saudyjska	10,2 (0,06)	9,7 (0,07)	10,2 (0,06)
Australia			
Bahrajn	11,2 (0,02)	10 (0,03)	10,6 (0,02)
Belgia (flamandzka)	8,7 (0,03)	9,5 (0,03)	8,9 (0,03)
Bułgaria	9,4 (0,09)	9,7 (0,12)	10,1 (0,08)
Chile	9,9 (0,04)	10,2 (0,05)	10,4 (0,03)
Chorwacja	10,5 (0,03)	10,8 (0,03)	8,2 (0,03)
Cypr	9,8 (0,03)	10,5 (0,04)	10,3 (0,02)
Czechy	9,4 (0,03)	10,7 (0,03)	9,2 (0,03)
Dania	9 (0,03)	9,9 (0,04)	10,2 (0,03)
Federacja Rosyjska	9,9 (0,05)	11,3 (0,04)	10,1 (0,03)
Finlandia	9,9 (0,03)	9,7 (0,03)	9,7 (0,04)
Francja	9,5 (0,03)	10,1 (0,03)	9,4 (0,05)
Gruzja	9,3 (0,04)	10,6 (0,04)	10,3 (0,06)
Hiszpania	10,7 (0,04)	10,3 (0,03)	10,5 (0,03)
Holandia			
Hong Kong	11 (0,04)	9,2 (0,04)	9,7 (0,05)
Indonezja	10,1 (0,08)	9,2 (0,08)	11,4 (0,06)
Iran			
Irlandia	11,6 (0,04)	11,1 (0,05)	10,5 (0,04)
Irlandia Północna		11,5 (0,06)	10,4 (0,05)
Japonia	10,7 (0,03)	9,2 (0,03)	7,5 (0,03)
Jordania			
Kanada	10,3 (0,04)	10,7 (0,05)	10,2 (0,04)
Katar	10,6 (0,04)	9,8 (0,04)	10,6 (0,04)
Kazachstan	10,4 (0,05)	11,1 (0,06)	11,4 (0,04)
Korea Południowa	12 (0,04)	10,4 (0,04)	8,7 (0,03)
Kuwejt	10,1 (0,05)	9,6 (0,05)	10,6 (0,05)
Litwa	10,2 (0,03)	10,3 (0,04)	10,3 (0,03)
Maroko	9,1 (0,10)	7,7 (0,12)	10,4 (0,08)
Niemcy	8,9 (0,03)	10,3 (0,04)	9,1 (0,04)
Norwegia			
Nowa Zelandia	9,2 (0,04)	11,2 (0,05)	10,2 (0,05)
Oman	10,4 (0,03)	9,4 (0,03)	10,6 (0,03)
POLSKA	10,5 (0,03)	10,9 (0,04)	10,2 (0,04)
Południowa Afryka			
Portugalia	9,3 (0,04)	10,1 (0,03)	10,9 (0,02)
Serbia	10,4 (0,05)	10,9 (0,06)	10,3 (0,05)
Singapur	11,4 (0,05)	9,8 (0,04)	10,7 (0,03)
Słowacja	8,6 (0,04)	10,8 (0,04)	9,3 (0,04)
Słowenia	8,8 (0,03)	10,7 (0,04)	8,8 (0,04)
Stany Zjednoczone			
Szwecja	10,1 (0,04)	9,8 (0,03)	9,9 (0,04)
Tajwan	11 (0,03)	9 (0,05)	9,2 (0,03)
Turcja	9,1 (0,09)	9 (0,07)	10,7 (0,05)
Węgry	8,8 (0,04)	10,6 (0,03)	9,7 (0,04)
Włochy	9 (0,04)	10,5 (0,04)	9,3 (0,03)
Zjednoczone Emiraty Arabskie	10,7 (0,02)	9,9 (0,03)	10,4 (0,02)

W nawiasie błędy standardowe. Puste komórki oznaczają, że kraj nie dostarczył porównywalnych danych.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

wskazać zmienne, które ją istotnie prognozują. W grę wchodzi cztery zmienne niezależne.

1. Edukacja rodzicielska przed rozpoczęciem szkoły. Jest to skala o średniej 10 i odchyleniu standardowym 2 utworzona metodą IRT z odpowiedzi rodziców na 16 pytań typu: „Zanim Pani/Pana dziecko zaczęło naukę w klasie pierwszej, jak często ktoś z domowników...” – np. bawił się z nim z klockami z literami alfabetu, głośno odczytywał z dzieckiem szyldy i napisy, pisał z dzieckiem litery lub słowa, pisał z dzieckiem liczby. Spójność odpowiedzi była zadowalająca (w Polsce $\alpha = 0,85$). Średnie wskaźnika znajdują się w trzeciej kolumnie Tabeli 4.2. Najintensywniej edukują swoje dzieci rodzice w Irlandii Północnej, najmniej intensywnie – w Maroku. Polska ze średnią 10,9 ma wysoką rangę 6,5.
2. Liczba lat spędzonych w przedszkolu. W raporcie międzynarodowym brak porównywalnych danych. Polskie dane zawiera Tabela 4.1.
3. Aspiracje rodziców dotyczące końcowego wykształcenia dziecka określone na podstawie odpowiedzi na pytanie: „Jakiego wykształcenia życzyłaby sobie Pani/życzyłby sobie Pan dla swojego dziecka?”. W raporcie międzynarodowym brak porównywalnych danych. W polskiej próbie udzielono następujących odpowiedzi (w procentach): zasadniczego zawodowego – 2,6, średniego – 12,1, policealnego – 4,9, licencjackiego – 25,8 magisterskiego – 54,7. Ten rozkład jest bardzo podobny do analogicznego rozkładu w 2011 r.
4. Postawa rodziców wobec matematyki i przyrodniczości. Jest to skala o średniej 10 i odchyleniu standardowym 2 utworzona metodą IRT z ustosunkowania się rodziców do ośmiu stwierdzeń typu: „Nauki przyrodnicze i techniczne pomagają rozwiązywać problemy świata” lub „Matematyka ma zastosowania w codziennym życiu”. Spójność odpowiedzi była zadowalająca (w Polsce $\alpha = 0,83$). Średnie wskaźnika znajdują się w czwartej kolumnie Tabeli 4.2. Najwyżej cenią matematykę i nauki przyrodnicze rodzice w Kazachstanie i Indonezji, najniżej – w Japonii, a z bliższych nam krajów – w Chorwacji. Polska ze średnią 10,2 lokuje się w środku listy: ma rangę 22 wraz z Nową Zelandią, Kanadą, Danią i Arabią Saudyjską².

² Na liście zmiennych niezależnych powinno się też znaleźć wyposażenie domu w dobra sprzyjające uczeniu się dzieci, ale odpowiednia skala w raporcie międzynarodowym ma tak synkretyczny charakter (łączy liczbę książek, dostęp do internetu i pokój dziecienny z wykształceniem i zawodem rodziców), że nie sposób dociec, o czym w istocie świadczy.

Tabela 4.3. Uwarunkowania wiedzy początkowej postrzeganej przez rodziców ($n = 4528$)

Zmienna	Współczynnik	p
Stała	6,21 (0,67)	
Edukacja rodzicielska ^a	0,55 (0,03)	0,001
Lata spędzone w przedszkolu ^b		
Jeden	-0,18 (0,10)	0,074
Dwa	-0,25 (0,10)	0,007
Trzy	-0,21 (0,09)	0,022
Cztery	-0,17 (0,10)	ni.
Pożądane wykształcenie dla dziecka ^c		
Średnie	0,79 (0,17)	0,001
Policealne	0,58 (0,21)	0,011
Wyższe I st.	1,08 (0,17)	0,001
Wyższe II st.	1,28 (0,17)	0,001
Postawa rodziców ^b	0,06 (0,03)	0,041
Płeć dziecka ^d	-0,11 (0,06)	0,054
Wiek (w latach)	0,32 (0,06)	0,001
SES	-0,11 (0,04)	0,005

Ważona analiza regresji, w nawiasie błędy standardowe oszacowane metodą *bootstrap*. Uwaga: wśród 14 indeksów warunkowych dwa są większe od 15, w tym jeden większy od 30.

^a Zmienne standaryzowane.

^b Kategorią odniesienia jest: Nie chodziło do przedszkola.

^c Kategorią odniesienia jest: Wykształcenie gimnazjalne lub zasadnicze zawodowe.

^d Kategorią odniesienia jest: Dziewczynka.

5. Zmienne kontrolowane: płeć dziecka i rocznik (2004 lub 2005) oraz SES rodziny.

Wyniki analizy przedstawia Tabela 4.3. Widać, co następuje.

1. SES różnicuje postrzeganą wiedzę początkową dziecka ujemnie. To oczywiście nie znaczy, że przeciętne dziecko z rodziny o niskim SES umiało na starcie szkolnym więcej niż z rodziny o wysokim SES, lecz tylko że po wytrąceniu rzeczywistych czynników sprawczych SES stracił swoje „nadzaczucie” i mierzy już tylko „gołą” zamożność, władzę i prestiż rodziców, które same w sobie nie muszą sprzyjać rozwojowi potomstwa.
2. Zasobom wiedzy początkowej dziecka towarzyszą pozytywne postawy rodziców wobec nauki i wykształcenia oraz ich praktyczny wyraz: wczesna edukacja rodzicielska.
3. Dzieci chodzące do przedszkola mają mniejszą wiedzę początkową niż dzieci, które do przedszkola nie chodziły. Jeszcze jedna zagadka dla teoretyków edukacji przedszkolnej.
4. Postrzegana przez rodziców wiedza początkowa sześciolatków na starcie szkolnym była mniejsza niż siedmiolatków.

Ostatni z tych efektów, pozornie banalny, jest najbardziej interesujący, ponieważ rzuca nowe świa-

tło na powody posyłania do szkoły sześciolatków w okresie ustawowego *laissez faire*. Zwykle przypuszcza się, że rodzice podejmowali taką decyzję, gdy ich dziecko umiało dość dużo, by ich zdaniem mogło poradzić sobie w szkole. Gdyby jednak tak było, efekt wieku nie byłby się ujawnił. Najwyraźniej wielu rodziców zdecydowało się posłać swojego sześciolatka do szkoły, choć wydawało się im, że umie on mniej niż przeciętny siedmiolatek. Dlaczego to zrobili? Najwyraźniej dlatego, że uwierzyli w potencjał rozwojowy swojego dziecka i uznali, że szkoła jest bardziej sprzyjającym środowiskiem rozwojowym niż przedszkole.

Ponieważ próba sześciolatków w szkole nie była losowa, nie możemy wykluczyć, że ich wiedza początkowa była większa niż wiedza sześciolatków, które w 2011 r. zostały w przedszkolu. Dlatego ostateczna konkluzja musi być ostrożna: Postawa rodziców wobec obniżenia wieku obowiązku szkolnego wydaje się nie mniej ważną przesłanką decyzji o posłaniu sześciolatka do szkoły niż ich ocena gotowości szkolnej swojego dziecka. Wiemy, że tę decyzję częściej podejmowali rodzice z wyższych warstw społecznych, więc wolno przypuszczać, że w tych warstwach postawa proedukacyjna przeważała nad tradycjonalistyczną postawą ochrony dzieciństwa. Wiemy też z danych przytoczonych w dwóch poprzednich rozdziałach, że zwolennicy postawy proedukacyjnej nie zawiedli się w swoich rachubach.

Postawy dziecka wobec szkoły i uczenia się

Narzędziami pomiaru były cztery krótkie skale postaw: wobec szkoły oraz wobec matematyki i przyrody, lekcji matematyki i przyrody oraz samego siebie jako uczącego się matematyki i przyrody. Skalowanie przeprowadzono metodą IRT, przyjmując średnią 10 i odchylenie standardowe 2.

Na skalę postawy wobec szkoły składa się siedem pozycji: Lubię chodzić do szkoły; Czuję się związana/związany ze swoją szkołą; Jestem dumna/dumny, że chodzę do tej szkoły; W szkole czuję się bezpiecznie; Nauczyciele traktują mnie sprawiedliwie; Dużo uczę się w szkole; Lubię spotykać się w szkole z koleżankami i kolegami z klasy. Skala jest krótka, ale spójna ($\alpha = 0,80$). Treść skal postaw wobec matematyki i przyrody oraz współczynniki spójności wewnętrznej podaje Tabela 4.4., a średnie postaw w poszczególnych krajach – Tabela 4.5. Widzimy, że najbardziej lubią szkołę dzieci w Indonezji, a najmniej – w Japonii. Polscy uczniowie zajęli 46. miejsce, tylko o trzy miejsca wyżej od japońskich. W Europie są na samym końcu wraz z Francuzami i Czechami.

Pod względem postawy wobec matematyki pierwsze miejsce zajmują dzieci w Turcji, a ostatnie – w Korei. Pod względem postawy wobec nauk przyrodniczych pierwsi są uczniowie z Portugalii, a ostatni – z Jordanii. Nie kontynuujemy jednak porównań międzynarodowych, z których niewiele wynika, lecz

Tabela 4.4. Treść skal postaw

Wobec przedmiotu	Wobec lekcji przedmiotu	Wobec siebie
Lubię uczyć się matematyki/przyrody. Marzę o tym, żeby nie trzeba było uczyć się matematyki/przyrody. Matematyka/wiedza o przyrodzie jest nudna. Na lekcjach matematyki/przyrody uczę się wielu ciekawych rzeczy. Lubię matematykę/ wiedzę o przyrodzie. Lubię wszystkie zajęcia, w których pojawiają się liczby./Lubię robić doświadczenia przyrodnicze. Lubię rozwiązywać zadania z matematyki./Lekcje przyrody uczą mnie, jak jest urządzony świat. Czekam niecierpliwie na lekcje matematyki/przyrody. Matematyka/przyroda to jeden z moich ulubionych przedmiotów.	Wiem, czego pani/pan od matematyki/przyrody wymaga ode mnie. Rozumiem wszystko, co mówi pani/pan od matematyki/przyrody. Interesuje mnie to, co mówi pani/pan od matematyki/przyrody. Pani/pan od matematyki/przyrody daje mi ciekawe zadania. Pani/pan od matematyki/przyrody jasno odpowiada na moje pytania. Pani/pan od matematyki/przyrody dobrze tłumaczy matematykę/wiedzę o przyrodzie. Pani/pan od matematyki/przyrody pozwala mi pokazać, co umiem. Pani/pan od matematyki/przyrody stosuje różne sposoby, żeby nam pomóc uczyć się. Gdy zrobię błąd, pani/pan od matematyki/przyrody mówi mi, jak go uniknąć w przyszłości. Pani/pan od matematyki/przyrody słucha tego, co mam do powiedzenia.	Zazwyczaj matematyka/przyroda dobrze mi idzie. Matematyka/przyroda jest dla mnie trudniejsza niż dla większości uczniów w mojej klasie. Po prostu nie jestem dobra/dobry z matematyki/przyrody. Szybko uczę się matematyki/przyrody. Denerwuję się przy matematyce. Jestem dobra/dobry w rozwiązywaniu trudnych zadań z matematyki. Pani/pan od matematyki/przyrody mówi, że jestem dobra/dobry z matematyki/przyrody. Matematyka/przyroda to dla mnie najtrudniejszy przedmiot. Gubię się w zadaniach z matematyki/przyrody.
$\alpha = 0,93/0,90$	$\alpha = 0,91/0,93$	$\alpha = 0,89/0,83$

Tabela 4.5. Średnie postaw wobec szkoły oraz matematyki i przyrody

Kraj	Matematyka				Przyroda		
	Szkoła	Wobec przedmiotu	Wobec lekcji	Wobec siebie	Wobec przedmiotu	Wobec lekcji	Wobec siebie
Anglia	10,2 (0,06)	10,1 (0,05)	10,1 (0,06)	10,1 (0,05)	9,8 (0,06)	9,9 (0,06)	9,5 (0,05)
Arabia Saudyjska	10,4 (0,07)	10,3 (0,05)	10,3 (0,06)	10,1 (0,06)	10,2 (0,06)	10,2 (0,06)	10,0 (0,06)
Australia	9,8 (0,05)	9,5 (0,04)	9,7 (0,05)	9,7 (0,03)	10,0 (0,05)	9,7 (0,05)	9,7 (0,04)
Bahrajn	10,0 (0,04)	10,4 (0,04)	10,4 (0,05)	10,2 (0,05)	10,7 (0,03)	10,5 (0,04)	10,3 (0,04)
Belgia(flamandzka)	9,8 (0,06)	9,2 (0,05)	9,7 (0,06)	9,7 (0,04)	9,5 (0,05)	9,7 (0,05)	9,8 (0,05)
Bułgaria	10,9 (0,06)	10,4 (0,06)	11,2 (0,06)	10,5 (0,06)	10,7 (0,06)	11,1 (0,06)	10,8 (0,07)
Chile	10,2 (0,06)	10,0 (0,06)	10,3 (0,07)	9,6 (0,04)	9,6 (0,06)	10,2 (0,05)	9,3 (0,05)
Chorwacja	9,5 (0,06)	9,3 (0,05)	9,9 (0,07)	10,1 (0,04)	9,8 (0,06)	10,2 (0,06)	10,4 (0,05)
Cypr	9,8 (0,06)	10,3 (0,05)	10,4 (0,06)	10,5 (0,05)	9,4 (0,08)	9,7 (0,08)	9,7 (0,07)
Czechy	9,2 (0,05)	9,5 (0,04)	9,4 (0,05)	9,6 (0,03)	9,5 (0,05)	9,5 (0,06)	9,6 (0,05)
Dania	9,8 (0,06)	9,6 (0,05)	9,1 (0,06)	10,1 (0,04)	9,5 (0,09)	9,0 (0,07)	9,8 (0,05)
Federacja Rosyjska	10,0 (0,06)	10,2 (0,04)	10,3 (0,06)	9,7 (0,04)	10,1 (0,05)	10,4 (0,05)	9,9 (0,05)
Finlandia	10,0 (0,05)	9,2 (0,04)	9,4 (0,04)	9,8 (0,03)	9,2 (0,05)	9,4 (0,04)	9,7 (0,03)
Francja	9,3 (0,05)	10,1 (0,04)	9,8 (0,05)	10,0 (0,03)	9,6 (0,06)	9,6 (0,05)	9,4 (0,05)
Gruzja	9,7 (0,06)	10,5 (0,05)	9,9 (0,06)	10,3 (0,06)	10,0 (0,05)	10,0 (0,06)	10,1 (0,05)
Hiszpania	10,5 (0,06)	9,9 (0,05)	10,7 (0,07)	10,0 (0,04)	10,1 (0,06)	10,6 (0,07)	10,0 (0,05)
Holandia	10,0 (0,06)	9,3 (0,04)	9,6 (0,04)	10,3 (0,04)	9,5 (0,06)	9,6 (0,04)	9,8 (0,04)
Hong Kong	9,0 (0,07)	9,5 (0,05)	9,2 (0,06)	9,3 (0,05)	10,1 (0,05)	9,4 (0,06)	9,3 (0,04)
Indonezja	11,5 (0,04)	10,7 (0,04)	10,8 (0,05)	9,7 (0,04)	10,5 (0,06)	10,6 (0,06)	9,8 (0,06)
Iran	10,2 (0,07)						
Irlandia	10,2 (0,05)	9,6 (0,05)	10,2 (0,06)	10,2 (0,04)	10,2 (0,06)	10,0 (0,06)	9,8 (0,05)
Irlandia Północna	10,2 (0,06)	9,5 (0,05)	10,2 (0,07)	9,9 (0,04)	10,2 (0,05)	10,0 (0,06)	9,7 (0,04)
Japonia	8,9 (0,05)	9,2 (0,04)	8,2 (0,05)	9,1 (0,03)	10,0 (0,05)	8,1 (0,05)	9,3 (0,03)
Jordania	10,9 (0,08)	11,0 (0,06)	10,9 (0,07)	10,5 (0,06)			
Kanada		9,6 (0,04)	10,1 (0,04)	9,9 (0,04)	9,9 (0,05)	10,1 (0,04)	9,8 (0,04)
Katar	9,7 (0,06)	10,2 (0,06)	10,1 (0,06)	10,1 (0,05)	10,4 (0,05)	10,3 (0,05)	10,2 (0,05)
Kazachstan	10,9 (0,07)	11,0 (0,05)	10,4 (0,08)	10,6 (0,06)	10,5 (0,06)	10,2 (0,07)	10,5 (0,06)
Korea Południowa	9,5 (0,06)	8,9 (0,03)	8,2 (0,05)	9,1 (0,03)	9,5 (0,05)	8,4 (0,06)	9,1 (0,03)
Kuwejt	10,3 (0,06)	10,3 (0,06)	10,3 (0,07)	10,3 (0,05)	10,6 (0,06)	10,6 (0,06)	10,4 (0,05)
Litwa	10,1 (0,05)	10,2 (0,04)	10,1 (0,05)	9,9 (0,04)	10,2 (0,05)	10,2 (0,05)	10,0 (0,04)
Maroko	11,3 (0,05)	10,9 (0,04)	10,8 (0,06)	10,0 (0,05)	10,7 (0,07)	10,6 (0,06)	10,0 (0,05)
Niemcy	9,5 (0,06)	9,5 (0,05)	9,8 (0,05)	10,1 (0,05)	10,0 (0,06)	9,9 (0,05)	10,1 (0,04)
Norwegia	10,5 (0,05)	9,8 (0,05)	9,9 (0,05)	10,5 (0,05)	10,0 (0,07)	10,0 (0,05)	10,3 (0,05)
Nowa Zelandia	10,0 (0,05)	9,8 (0,04)	9,7 (0,04)	9,5 (0,03)	10,1 (0,05)	9,6 (0,05)	9,3 (0,03)
Oman	10,7 (0,05)	11,0 (0,04)	10,7 (0,05)	10,1 (0,04)	10,9 (0,04)	10,7 (0,04)	10,4 (0,05)
POLSKA	9,1 (0,05)	9,4 (0,05)	9,5 (0,06)	9,8 (0,04)	9,6 (0,06)	9,8 (0,06)	9,9 (0,05)
Południowa Afryka	10,1 (0,07)	10,1 (0,04)	10,1 (0,06)	9,3 (0,03)			
Portugalia	11,3 (0,05)	10,6 (0,05)	10,9 (0,04)	9,6 (0,05)	11,3 (0,05)	10,8 (0,04)	10,2 (0,05)
Serbia	10,3 (0,06)	10,1 (0,06)	11,0 (0,06)	10,5 (0,05)	10,0 (0,07)	10,9 (0,05)	10,5 (0,06)
Singapur	9,5 (0,03)	9,6 (0,03)	9,3 (0,04)	9,2 (0,05)	10,1 (0,04)	9,4 (0,04)	9,2 (0,03)
Słowacja	9,7 (0,05)	9,8 (0,05)	10,0 (0,06)	9,9 (0,04)	9,5 (0,05)	10,0 (0,06)	9,8 (0,04)
Słowenia	9,5 (0,06)	9,4 (0,06)	9,6 (0,06)	9,9 (0,04)	9,4 (0,05)	9,7 (0,05)	9,7 (0,04)
Stany Zjednoczone	9,9 (0,04)	9,7 (0,04)	10,2 (0,04)	10,0 (0,03)	10,3 (0,04)	10,3 (0,04)	10,0 (0,03)
Szwecja	9,9 (0,06)	9,5 (0,05)	9,5 (0,04)	10,2 (0,05)	9,7 (0,05)	9,4 (0,04)	10,0 (0,04)
Tajwan	9,1 (0,04)	8,9 (0,05)	9,2 (0,05)	8,9 (0,03)	10,2 (0,06)	9,5 (0,06)	9,8 (0,04)
Turcja	10,8 (0,05)	11,3 (0,04)	10,7 (0,05)	10,4 (0,05)	11,1 (0,04)	10,7 (0,04)	10,8 (0,05)
Węgry	9,9 (0,05)	9,7 (0,05)	10,4 (0,05)	10,1 (0,04)	9,7 (0,05)	10,5 (0,05)	10,1 (0,05)
Włochy	9,7 (0,05)	10,1 (0,05)	9,8 (0,05)	10,1 (0,04)	10,1 (0,04)	9,9 (0,05)	9,9 (0,04)
Zjednoczone Emiraty Arabskie	9,9 (0,04)	10,4 (0,03)	10,1 (0,04)	10,0 (0,03)	10,5 (0,04)	10,3 (0,04)	10,1 (0,03)

Wszystkie skale mają średnią 10 i odchylenie standardowe 2.

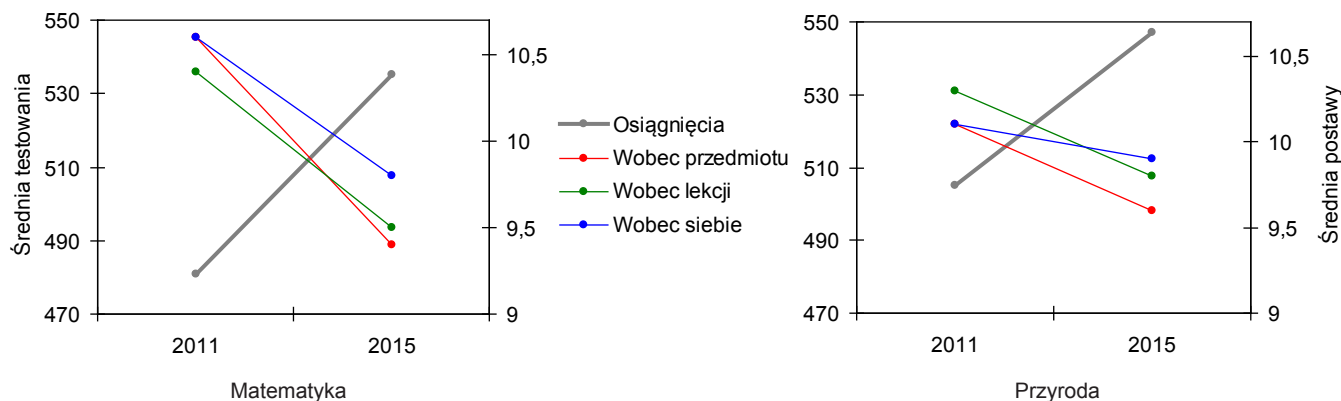
W nawiasie błędy standardowe.

Puste komórki oznaczają, że kraj nie dostarczył porównywalnych danych.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 4.6. Wyniki testowania i pomiaru postaw w dwóch kohortach polskich uczniów

Przedmiot	Rok pomiaru	Osiągnięcia		Wobec przedmiotu		Wobec lekcji		Wobec siebie	
		Średnia	Ranga	Średnia	Ranga	Średnia	Ranga	Średnia	Ranga
Matematyka	2011	481 (2,2)	34	10,6 (0,03)	7	10,4 (0,03)	8	10,6 (0,03)	1
	2015	535 (2,1)	17	9,4 (0,05)	40,5	9,5 (0,06)	40	9,8 (0,04)	33,5
Przyroda	2011	505 (2,6)	29	10,1 (0,04)	15,5	10,3 (0,04)	10	10,1 (0,04)	22,5
	2015	547 (2,4)	9	9,6 (0,06)	36	9,8 (0,06)	30	9,9 (0,05)	23



Rysunek 4.2. Wyniki testowania i pomiaru postaw w dwóch kohortach polskich uczniów

raczej porównajmy dwie kohorty polskich dzieci z lat 2011 i 2015. Wystarczy rzut oka na dane w Tabeli 4.6, by odkryć ciekawą regularność: masywnej poprawie osiągnięć towarzyszy równie masywne pogorszenie się postaw wobec uczenia się. Na przykład w 2011 r. nasi uczniowie znaleźli się na 34. miejscu (za wszystkimi krajami europejskimi) pod względem obiektywnie zmierzonej wiedzy matematycznej, a jednocześnie zajęli pierwsze miejsce w samoocenie swojej wiedzy z tego przedmiotu. W 2015 r. polscy uczniowie przesunęli się w górę o 17 miejsc w rankingu wiedzy, a zarazem spadli o 32 miejsca w rankingu samooceny. Ta i inne regularności są tak intrygujące, że dodatkowo zilustrowano je na Rysunku 4.2. Jak je wyjaśnić? Rozważmy trzy hipotezy.

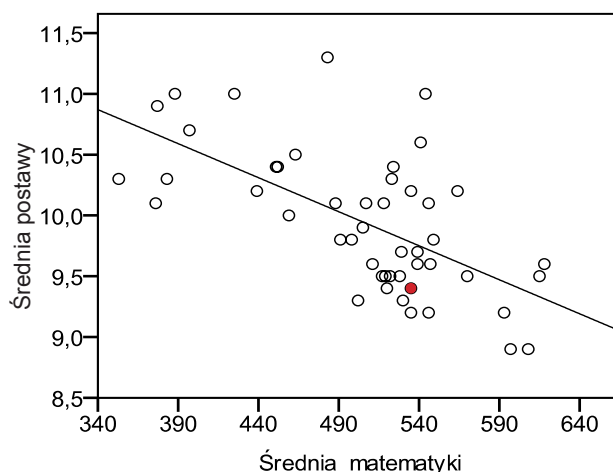
Pierwsza odwołuje się do ogólnej prawidłowości rozwojowej, zgodnie z którą wraz z wiekiem rośnie krytycyzm i samokrytycyzm dzieci. Ostatnio potwierdził ją Paweł Grygiel (2016) na przykładzie dziecięcych ocen stosunków z rówieśnikami w oddziale klasowym. Ponieważ w 2015 r. testowano dzieci średnio ponad dziewięć miesięcy starsze niż w 2011 r., omawiana hipoteza wydaje się niebezpieczna. Przeczy jej jednak fakt, że zaobserwowany wzrost krytycyzmu zależy od przedmiotu nauczania: jest znacznie silniejszy wobec matematyki niż przyrody. W samoocenie wiedzy przyrodniczej ranga dzieci z obu kohort jest identyczna. Trudno przypuścić, by ogólna prawidłowość rozwojowa była tak selektywna.

Następna hipoteza obciąża odpowiedzialnością za pogorszenie się postaw drugi etap kształcenia, w który weszły dzieci testowane w 2015 r. Powiada się, że przejście od radosnego uczenia się zintegrowanej wiedzy pod macierzyńskim okiem jednej nauczycielki, bez terroru stopni i perspektywy egzaminu zewnętrznego, do uczenia się wielu przedmiotów pod kierunkiem różnych nauczycieli, z których każdy śrubuje wymagania, tak jakby jego przedmiot był jedyny lub najważniejszy – jest traumatyczne. Przeciążonemu pracą i stresem dziecku uczenie się przestaje sprawiać przyjemność, a to znajduje wyraz w jego opiniach, że matematyka i przyroda nie dają się lubić, z nauczycielami trudno się porozumieć, a jego wiedza pozostawia wiele do życzenia. Przeciw tej hipotezie przemawia jednak fakt, że wiedza czwartoklasistów okazała się znacznie większa niż trzecioklasistów. Jeśli reżim dydaktyczny drugiego etapu kształcenia jest wrogi dzieciństwu, to jak wyjaśnić rozwój poznawczy podlegających mu dzieci?

Trzecia hipoteza jest przeciwieństwem drugiej: odpowiedzialnością za pogorszenie się postaw obciąża etap kształcenia zintegrowanego. Matematyka, której atrakcyjność i stopień opanowania oceniali trzecioklasiści, nie była matematyką *tout court*, lecz jej szczególną odmianą – uproszczoną, ludyczną. Gdy w klasie czwartej pojawiła się zwyczajna, dzieci zaczęły doświadczać trudności, które skorygowały ich nadmiernie optymistyczne przekonania.

Trzecia hipoteza wydaje się najbardziej uzasadniona. Jeśli jest prawdziwa, to nasuwa się ogólny wniosek, że dobre samopoczucie ucznia w szkole nie powinno być głównym celem edukacji. Uczenie się to praca, a nie zabawa. *Litterarum radices amarae, fructus dulce* – przypomina Cynceron. Kto zabiega o bieżącą atrakcyjność zajęć kosztem odroczonej korzyści, ten działa na szkodę ucznia i społeczeństwa.

Ale nie rozgrzeszajmy pochopnie nauczycieli matematyki i innych przedmiotów w drugim etapie kształcenia. Rysunek 4.3 przedstawia kraje biorące udział w TIMSS 2015 w podwójnym uporządkowaniu: według średniej wyników z matematyki i średniej postawy wobec matematyki. Widać ujemną korelację ($R = -0,62$) między obiema średnimi. Tam, gdzie osiągnięcia matematyczne są najniższe, np. w Maroku, uczniowie bardzo lubią matematykę. Tam, gdzie są najwyższe, np. w Korei, entuzjazm jest mniejszy. Ujemna korelacja tłumaczyłaby pozycję Polski (czerwone kółeczko), gdyby nie fakt, że jest kilkanaście krajów (wśród nich Kazachstan i Portugalia), w których uczniowie mają zbliżone do polskich osiągnięcia w matematyce, a jednocześnie bardziej pozytywną niż polska postawę wobec matematyki. To



Rysunek 4.3. Średnie krajowe postawy wobec matematyki w zależności od średnich wyników w matematyce

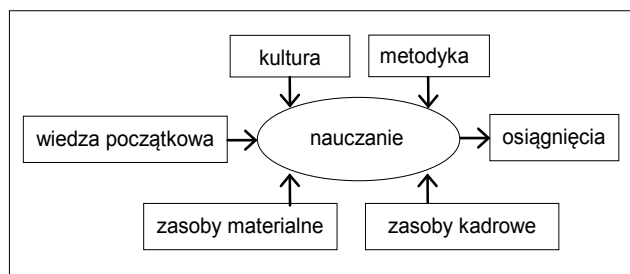
Na podstawie: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

nie powinno dziwić: za korelacją nie stoi przecież bezwyjątkowe prawo natury, lecz określona forma praktyki społecznej. Polska szkoła podstawowa ma jeszcze wiele do zrobienia, by zidentyfikować źródła nieproduktywnego ucisku i skutecznie je zablokować.

Szkoła i proces kształcenia

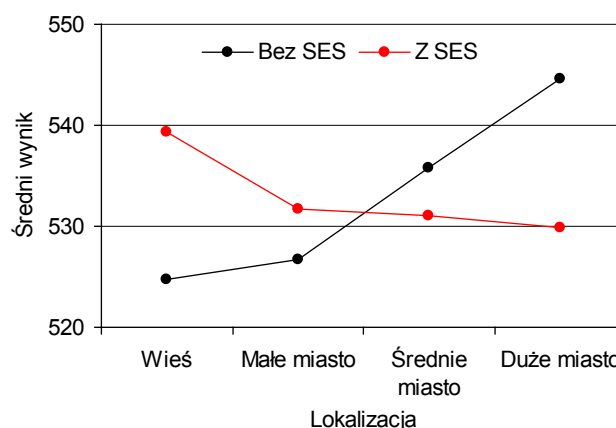
Krzysztof Konarzewski

Uczeniu się po stronie ucznia odpowiada nauczanie po stronie szkoły. Część ankietowa badania TIMSS daje wgląd w możliwe uwarunkowania tego procesu. Jeden z układów hipotetycznych zależności pokazuje rysunek.



Niestety, taki układ łatwiej narysować, niż oszacować związki między jego składnikami. W międzynarodowych raportach IEA zestawia się średnie krajowe poszczególnych skal ze średnimi osiągnięciami w tych krajach, co stanowi substytut analizy korelacyjnej. Nie da się jednak ze zbioru współczynników korelacji między izolowanymi zmiennymi niezależnymi (czynnikami) i zmienną zależną złożyć teorii osiągnięć szkolnych. Jest tak dlatego, że czynniki często bywają skorelowane ze sobą – więc nie można wykluczyć, że współczynnik statystycznie istotny skrywa rzeczywistą niezależność – albo wchodzi w interakcje, co powoduje, że statystycznie nieistotny współczynnik może skrywać rzeczywistą zależność.

Zakres niepewności wnioskowania dobrze ukazuje próba odpowiedzi na pytanie o związek lokalizacji



Rysunek 5.1. Średnie wyniki testu matematycznego w zależności od lokalizacji szkoły

Dwupoziomowa analiza regresji na ważonych wartościach prawdopodobnych przy kontroli płci i wieku na pierwszym poziomie.

szkoły z osiągnięciami jej uczniów, a właściwie z jej poziomem kształcenia, o którym pytający wnioskuje na podstawie osiągnięć. Pytanie wydaje się statystycznie proste. W próbie mamy 56 szkół wiejskich, 18 szkół w małych miastach (liczących do 20 tysięcy mieszkańców), 36 szkół w miastach średniej wielkości (od 20 do 100 tysięcy mieszkańców) i 40 szkół w dużych miastach (powyżej 100 tys. mieszkańców). Wydawałoby się, że wystarczy oszacować średnie osiągnięć w tych podzbiorach szkół. Wynik takiej analizy osiągnięć matematycznych ilustruje czarna krzywa na Rysunku 5.1. Uczniowie z dużych miast istotnie, a ze średnich miast prawie istotnie ($p < 0,10$) górują nad uczniami szkół

wiejskich. Wniosek sam się nasuwa: im większy ośrodek, tym lepsze szkoły. Wystarczy jednak włączyć do równania na drugim poziomie średnią szkolną SES, by wniosek okazał się fałszywy. Średnia SES ma zasadniczo inny sens niż indywidualny SES rodziny ucznia – mówi o otoczeniu społecznym, w którym działa szkoła¹. Uwzględnienie średniej SES sprawia, że wszystkie różnice osiągnięć stają się nieistotne, a uczniowie ze szkół wiejskich – jak pokazuje czerwona linia – nawet przewyższają (nieistotnie) uczniów z dużego miasta². Skoro zniwelowanie różnic otoczenia niweluje różnice osiągnięć, to nasuwa się wniosek, że przewaga uczniów ze szkół wielkomiejskich nie jest skutkiem wyższego poziomu kształcenia, lecz wyłącznie tego, że większa ich liczba ma zamożnych, wykształconych i ustabilizowanych zawodowo rodziców. Rzekome upośledzenie szkół wiejskich byłoby w istocie upośledzeniem polskich wsi i miasteczek – relatywnie biednych i źle wykształconych, ale same te szkoły byłyby nie gorsze niż szkoły w wielkich miastach.

Ale i ten wniosek nie jest pewny, ponieważ średnia SES może być skorelowana z poziomem nauczania wskutek prostego dostosowania się szkoły do oczekiwań jej otoczenia. Rodziny z wyższych warstw społecznych mogą wywierać nacisk na organ prowadzący szkołę, by więcej w nią inwestował, i na samą szkołę, by bardziej dbała o osiągnięcia uczniów. Im więcej byłoby takich rodzin w otoczeniu szkoły, tym skuteczniejszy byłby ten nacisk. Wtedy jednak statystyczne wyrównanie średnich SES niwelowałoby zarazem jakąś część różnic poziomu nauczania, a w następstwie sztucznie zmniejszało oszacowanie różnic osiągnięć.

Jedyny wniosek, którego możemy być pewni, brzmi: na gruncie danych, jakimi dysponujemy³, problem jest nierozwiązywalny. Podobną naturę ma wiele innych problemów często stawianych w dyskusji publicznej, np.: Czy liczba lat spędzonych przez dziecko w przedszkolu zwiększa szanse na sukces w szkole? lub Czy wielkość oddziału klasowego wpływa na osiągnięcia szkolne jego uczniów?⁴ Lepiej zatem po-

¹ Potwierdza to istotny związek ($R^2 = 0,27$) średnich szkolnych SES z odpowiedziami dyrektorów szkół na dwa pytania: o odsetek uczniów z ubogich i z zamożnych domów. Można dodać, że wkład pierwszego pytania w odtworzenie średniej SES był znacznie większy niż drugiego.

² Warto dodać, że średnia SES niweluje także różnice osiągnięć przyrodniczych w obecnym badaniu oraz wszystkie różnice osiągnięć w badaniach TIMSS i PIRLS z 2011 r.

³ I jakimi dysponuje Centralna Komisja Egzaminacyjna, która rutynowo informuje o wynikach testowania w czterech kategoriach miejscowości.

⁴ Z tym problemem zmierzli się Maciej Jakubowski i Paweł Sakowski (2006) za pomocą skomplikowanego

Tabela 5.1. Średnie liczby uczniów w oddziałach klasy czwartej w zależności od lokalizacji szkoły

Lokalizacja	Liczba szkół	Liczba oddziałów	Średnia liczba uczniów w oddziale
Wieś	56	74	16,8 (0,69)
Miasto do 20 tys. mieszkańców	18	34	20,2 (0,65)
Miasto od 20 do 100 tys. mieszkańców	36	69	20,8 (0,49)
Miasto powyżej 100 tys. mieszkańców	40	77	19,3 (0,52)
Razem	150	254	18,7 (0,31)

Ważona ANOVA, w nawiasie błędy standardowe oszacowane metodą bootstrap

wstrzymać się od rozpowszechniania informacji, które mogą być statystycznymi artefaktami.

Wielkość oddziałów klasowych

Główny Urząd Statystyczny (2015) podaje, że polska szkoła podstawowa ma średnio 179 uczniów (o 14 więcej niż w 2011 r.) i 18–19 uczniów w oddziale klasowym (tyle samo, ile w 2011 r.). W naszej próbie liczba testowanych uczniów w oddziale klasowym wahała się od 5 (szkoły z najmniejszymi oddziałami wolno było wyłączyć z operatu losowania) do 30 wokół średniej 18,7 z odchyleniem standardowym 5,0. Oddziałów, w których do testu przystąpiło więcej niż 26 dzieci, było 10 (4 proc.). Tabela 5.1 informuje o wielkości oddziałów w zależności od lokalizacji szkoły. Najmniejsze oddziały mamy w szkołach wiejskich, ale największe – w miastach średniej wielkości.

W badaniu PIRLS 2006, w którym obowiązywały takie same zasady doboru próby, testowaliśmy średnio 24 uczniów w oddziale. W badaniu TIMSS 2011 – 20 uczniów. Widać, że oddziały klasowe są coraz mniejsze – co bynajmniej nie przeszkadza niektórym dziennikarzom uskarżać się na przepełnienie izb klasowych. Zmniejszanie się liczby uczniów w oddziale to bez wątpienia odpowiedź polskiego systemu oświaty na dramatyczny spadek populacji uczniów.

Wyjątkowość naszego szkolnictwa podstawowego potwierdzają dane międzynarodowe. Wprawdzie w Tabeli 1.1 nie ma danych o liczbie uczniów w oddziale klasowym, ale są pokrewne – o liczbie uczniów przypadających na jednego nauczyciela. W Polsce wynosi ona 10 – mniej tylko w Omanie, Gruzji, Kuwejcie i Norwegii. Na drugim biegunie jest RPA, w której na jednego

schematu quasi-eksperymentalnego. Warto zapoznać się z ich doniesieniem, zanim zaatakuje się problem metodami korelacyjnymi.

Tabela 5.2. Podział wariancji osiągnięć w polskiej szkole podstawowej

Przedmiot i rok	szkół	Zróźnicowanie	
		oddziałów w jednej szkole	uczniów wewnątrz oddziału
Matematyka 2015 ^a	10,0	3,6	86,4
Matematyka 2011 ^b	11,8	3,1	85,1
Przyroda 2015 ^a	8,8	3,6	87,6
Przyroda 2011 ^b	10,9	3,0	86,1

Wariancję osiągnięć podzielono za pomocą trypoziomowej analizy regresji na ważonych wartościach prawdopodobnych za pomocą aplikacji HLM 7 S. Raudenbusha, A. Bryka i R. Congdona.

^a Klasa czwarta.

^b Klasa trzecia.

nauczyciela przypada 32 uczniów. W bliższych nam krajach nauczyciel pracuje z większą liczbą uczniów: 13 na Litwie, 15 na Słowacji, 19 w Czechach i 20 w Rosji.

Zróźnicowanie szkół i oddziałów klasowych

Jak bardzo zróźnicowane są polskie szkoły podstawowe pod względem osiągnięć uczniów w klasie czwartej? Dysponujemy próbą losową 150 szkół podstawowych, z których 104 prowadziły przynajmniej dwa oddziały klasy czwartej. Tabela 5.2 informuje, jaki procent całkowitego zróźnicowania osiągnięć w krajowej próbie uczniów pozwala odtworzyć podział na szkoły i oddziały klasowe wewnątrz szkół. Podaje też dla porównania wyniki identycznej analizy sprzed czterech lat. Widać, że zróźnicowanie międzyszkolne i międzyoddziałowe w 2015 r. (w klasie czwartej) jest mniejsze niż w 2011 r. (w klasie trzeciej), zwłaszcza pod względem osiągnięć przyrodniczych.

Czemu przypisać tę zmianę? Jedną z możliwych odpowiedzi brzmi: zwiększeniu standaryzacji kształcenia. Trzeba przypomnieć, że porównujemy osiągnięcia czwartoklasistów z 2015 r. i trzecioklasistów z 2011 r. Te populacje wielce się różniły: druga właśnie kończyła okres kształcenia zintegrowanego, pierwsza rozpoczęła okres kształcenia przedmiotowego. Co więcej, druga edukowała się pod panowaniem starej podstawy programowej (z 1999 r.), a pierwsza – nowej (z 2008 r.). Obie podstawy różnią się szczegółowością. Stara wytyczyła treść wczesnoszkolnego kształcenia matematycznego i przyrodniczego nadzwyczaj oszczędnie – za pomocą 14 ogólnych zagadnień. Nowa normuje nauczanie matematyki za pomocą 90, a przyrody – 119 wymagań szczegółowych. Nie ma wątpliwości: zakres swobody nauczycieli i autorów podręczników zmniejszył się. To, że zarazem zwiększyło się podobieństwo doświadczeń edukacyjnych

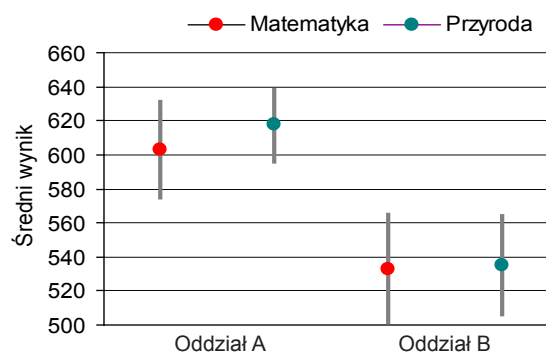
i wyników uczenia się w szkołach i oddziałach klasowych, świadczy o trafności określenia funkcji założonej podstawy programowej: stać na straży optymalnej, nie za małej i nie za dużej standaryzacji praktyki oświatowej w kraju (Konarzewski, 2004, s. 16).

Przyjrzyjmy się bliżej zróźnicowaniu międzyoddziałowemu. Duże różnice między oddziałami powstają albo wskutek różnic w metodyce nauczania, albo dlatego, że personel szkoły stara się je homogenizować, czyli tworzyć w taki sposób, by w każdym znalazły się dzieci o podobnych możliwościach umysłowych i podobnej motywacji do nauki. Gdy szkoła dzieli na oddziały dzieci, które dopiero zaczynają naukę, nie ma informacji o ich zdolnościach i motywacji (chyba że dostanie je z przedszkola), ale może się ich domyślać na podstawie statusu socjoekonomicznego ich rodzin. Ile szkół wykorzystuje tę informację do tworzenia homogenicznych oddziałów klasowych?

Rozumowanie prowadzące do odpowiedzi na to pytanie jest następujące. Jeśli dzieci są dzielone na oddziały losowo, to różnice SES między oddziałami są wyłącznie dziełem przypadku. Umiemy obliczyć prawdopodobieństwo przypadkowego powstania różnicy o określonej wielkości. Przyjmując zwyczajowy w naukach społecznych próg 0,05, uznajemy różnicę za nieprzypadkową, ergo wytworzoną intencjonalnie, jeśli jej prawdopodobieństwo jest mniejsze od 0,05. Innymi słowy: nieprzypadkowa jest różnica, która mogłaby się pojawić rzadziej niż raz na 20 losowań.

Przy tym kryterium intencjonalne różnice między oddziałami, czyli stosowanie homogenizacji, stwierdzamy w 16 proc. szkół prowadzących przynajmniej dwa oddziały klasy czwartej. Najwięcej takich szkół jest w miastach średniej wielkości (dzieci na 27), najmniej na wsi (jedna na 12).

Jedną z takich szkół przedstawia Rysunek 5.2. Oddziały A i B klasy czwartej dzieli przepaść. Różnica między średnimi osiągnięć matematycznych wynosi



Rysunek 5.2. Średnie wyniki testowania w dwóch oddziałach tej samej szkoły

Tabela 5.3. Osiągnięcia uczniów w szkołach stosujących homogenizację oddziałów klasowych według SES rodzin uczniów i niestosujących jej

Szkoły prowadzące przynajmniej dwa oddziały klasy czwartej	Liczba szkół	Liczba uczniów	Średnia osiągnięć	
			matematycznych	przyrodniczych
Stosujące homogenizację	17	661	533 (4,9)	544 (4,4)
Niestosujące homogenizacji	87	3386	537 (4,7)	550 (4,1)
Razem	104	4047	536 (4,5)	549 (4,2)

W nawiasach błędy standardowe. Średnie oszacowano za pomocą dwupoziomowej analizy regresji na ważonych wartościach prawdopodobnych przy kontroli płci i wieku oraz na drugim poziomie średniej SES i lokalizacji szkoły.

(w jednostkach lokalnego odchylenia standardowego) 1,0, a przyrodniczych 1,3. Osiągnięcia uczniów z oddziału A są nie tylko wyższe, ale i mniej zróżnicowane niż w oddziale B. Pokazują to pionowe odcinki reprezentujące oddziałowe odchylenia standardowe.

Homogenizacja oddziałów ma rzekomo podnosić osiągnięcia wszystkich uczniów, bo umożliwia dostosowanie tempa i metodyki nauczania do ich możliwości. Czy rzeczywiście podnosi? Jeśli homogenizacja służyłaby wszystkim uczniom, to średnia osiągnięć powinna być wyższa w szkołach, które ją stosują, niż w szkołach, które tego nie robią. Tabela 5.3 przekonuje, że tak nie jest. Różnice średnich wyników są statystycznie nieistotne. Skoro w szkole podstawowej dzieci podzielone na oddziały według głębszej myśli i dzieci podzielone „jak leci” uczą się – średnio rzecz biorąc – tak samo dobrze, to homogenizację można nazwać grą o sumie zerowej, w której wygrywają uczniowie uznani za obiecujących, a przegrywają uznani za „odpornych na wiedzę”.

Dane zebrane w badaniu TIMSS nie dają wglądu w okoliczności i motywy dobierania uczniów do oddziałów według SES ich rodzin, dlatego nie możemy być pewni, że tam, gdzie tak zrobiono, dobór losowy dałby lepsze rezultaty. Weryfikacja hipotezy o związku przyczynowym między organizacją oddziałów klasowych a osiągnięciami umysłowymi i społecznymi uczniów wymagałaby danych eksperymentalnych lub przynajmniej analiz kontrfaktycznych. Wolno jednak zamknąć tę analizę konkluzją, że dane, którymi dysponujemy, poddają w wątpliwość pedagogiczne i etyczne podstawy homogenizacji.

Zasoby materialne szkoły

Opinie o zasobach materialnych szkoły zebrano niezależnie od dyrektorów oraz nauczycieli matematyki i przyrody. Dyrektorzy odpowiadali na pytanie: W jakim stopniu na realizację programu nauczania w Pani/Pana szkole wpływa brak lub nieodpowiednia jakość ogólnych zasobów szkolnych (np. podręczników, ma-

teriałów biurowych, pomieszczeń i przestrzeni wokół budynku) oraz zasobów niezbędnych do nauczania matematyki (np. nauczycieli wyspecjalizowanych w nauczaniu matematyki, programów komputerowych do nauczania matematyki) i przyrody (np. nauczycieli wyspecjalizowanych w nauczaniu przyrody, wyposażenia pracowni przyrodniczej w urządzenia i materiały do przeprowadzania doświadczeń)? Odpowiedzi wyskalowano osobno dla zasobów matematycznych i przyrodniczych (Tabela 5.4), ale łącznie z częścią ogólną, więc korelacja średnich obu skal jest bardzo wysoka ($r = 0,98$). Według szacunków dyrektorów najlepiej wyposażone są szkoły w Korei, najgorzej w Turcji, a Polska może się poszczycić wysoką rangą 4 razem z Anglią i Czechami.

Zobaczymy, jak na podobne pytanie (Jak poważne są poniższe problemy w Pani/Pana szkole? – np. budynek szkoły wymaga gruntownego remontu, sale lekcyjne nie są sprzątane wystarczająco często, brakuje technicznych środków nauczania i wsparcia ze strony wykwalifikowanych pracowników technicznych) odpowiedzieli nauczyciele. Korelacja między średnimi szacunkami nauczycieli matematyki i przyrody okazała się jeszcze wyższa ($r = 0,99$). W opinii nauczycieli najlepiej wyposażone są szkoły Czechach, najgorzej w Indonezji, a Polska ma rangę 25 razem z Tajwanem i Kuwejtem.

Cieszyć się rangą 4 czy martwić 25? Współczynnik korelacji między średnimi krajowymi⁵ szacunków dyrektorów i nauczycieli to tylko 0,50. Najwyższą zgodność ocen zanotowano na Litwie i w Portugalii. Najbardziej przewyższyli dyrektorów optymizmem nauczyciele w Bahrajnie, Emiratach Arabskich i na Słowacji. W największy pesymizm popadli nauczyciele w Maroku i Indonezji. Polscy nauczyciele należą do pesymistów, ale umiarkowanych. W obliczu tych rozbieżności najrozsądniej zawiesić konkluzję, uzna-

⁵ W Polsce współczynnik korelacji między szacunkami dyrektora a średnią szacunków podległych mu nauczycieli wyniósł 0,35.

Tabela 5.4. Wyposażenie szkoły według dyrektorów i nauczycieli (średnie skali oraz odsetki uczniów)

Kraj	Matematyka		Przyroda		Szkoła ma pracownię*	Nauczyciel ma pomoc**
	Dyrektorzy	Nauczyciele	Dyrektorzy	Nauczyciele		
Anglia	11,2 (0,15)	10,8 (0,16)	11,0 (0,15)	10,8 (0,17)	8 (1,4)	62 (4,5)
Arabia Saudyjska	8,9 (0,15)	9,3 (0,14)	9,0 (0,16)	9,6 (0,15)	74 (3,3)	65 (3,5)
Australia	11,1 (0,15)	10,6 (0,12)	10,7 (0,14)	10,6 (0,13)	13 (2,1)	13 (2,1)
Bahrajn	9,0 (0,01)	10,6 (0,11)	9,0 (0,01)	10,8 (0,03)	85 (0,2)	72 (0,2)
Belgia (flamandzka)	10,3 (0,11)	10,0 (0,14)	10,2 (0,11)	10,0 (0,14)	2 (1,5)	62 (4,2)
Bułgaria	10,8 (0,16)	10,6 (0,18)	10,8 (0,14)	10,6 (0,18)	4 (1,6)	1 (0,5)
Chile	10,1 (0,18)	10,5 (0,17)	10,2 (0,18)	10,4 (0,16)	61 (4,0)	22 (3,6)
Chorwacja	10,2 (0,12)	9,5 (0,17)	10,3 (0,12)	9,5 (0,17)	22 (3,1)	19 (3,6)
Cypr	10,6 (0,21)	10,1 (0,16)	10,6 (0,22)	9,9 (0,17)	65 (4,1)	23 (3,9)
Czechy	11,1 (0,10)	11,1 (0,13)	11,1 (0,11)	11,2 (0,12)	40 (4,4)	4 (2,1)
Dania	10,6 (0,11)	9,3 (0,13)	10,4 (0,12)	9,2 (0,13)	51 (3,9)	15 (3,1)
Federacja Rosyjska	10,1 (0,14)	10,2 (0,11)	10,0 (0,15)	10,2 (0,11)	34 (4,2)	29 (3,1)
Finlandia	10,5 (0,11)	9,5 (0,11)	10,5 (0,10)	9,5 (0,12)	25 (4,0)	29 (3,7)
Francja	9,7 (0,14)	9,2 (0,14)	9,6 (0,14)	9,2 (0,13)	1 (0,9)	2 (0,9)
Gruzja	10,9 (0,13)	9,6 (0,17)	11,0 (0,13)	9,7 (0,18)	50 (4,3)	12 (3,0)
Hiszpania	10,6 (0,09)	10,3 (0,14)	10,6 (0,08)	10,2 (0,14)	32 (2,7)	21 (3,1)
Holandia	10,6 (0,15)	10,1 (0,14)	10,1 (0,12)	10,1 (0,14)	1 (1,5)	20 (4,4)
Hong Kong	9,9 (0,13)	10,4 (0,13)	9,9 (0,14)	10,4 (0,15)	38 (4,5)	42 (4,6)
Indonezja	9,1 (0,05)	8,0 (0,15)	9,2 (0,06)	8,0 (0,16)	16 (2,5)	9 (1,8)
Iran	8,5 (0,10)	9,1 (0,10)	8,6 (0,10)	9,1 (0,10)	37 (3,3)	11 (2,0)
Irlandia	10,1 (0,11)	10,4 (0,17)	10,0 (0,10)	10,4 (0,17)	1 (0,9)	8 (2,4)
Irlandia Północna	10,7 (0,16)	10,8 (0,19)	10,3 (0,14)	10,8 (0,19)	0 (0,0)	17 (4,0)
Japonia	10,2 (0,15)	9,6 (0,12)	10,3 (0,15)	9,7 (0,13)	97 (1,4)	28 (4,1)
Jordania	8,5 (0,15)	9,2 (0,12)				
Kanada	10,9 (0,12)	10,2 (0,09)	10,6 (0,10)	10,2 (0,09)	11 (2,0)	14 (1,9)
Katar	10,3 (0,20)	11,1 (0,14)	10,3 (0,20)	11,2 (0,13)	80 (2,4)	79 (2,8)
Kazachstan	10,3 (0,19)	10,5 (0,16)	10,3 (0,18)	10,5 (0,16)	35 (3,8)	60 (4,0)
Korea Południowa	12,5 (0,18)	10,8 (0,15)	12,7 (0,19)	10,9 (0,14)	99 (0,8)	89 (2,6)
Kuwejt	8,6 (0,23)	10,0 (0,16)	8,7 (0,24)	10,1 (0,15)	93 (2,5)	83 (3,7)
Litwa	10,2 (0,13)	10,0 (0,16)	10,1 (0,14)	10,0 (0,16)	2 (1,0)	4 (1,5)
Maroko	9,7 (0,09)	8,2 (0,15)	9,8 (0,08)	8,3 (0,12)	5 (1,4)	30 (2,7)
Niemcy	10,2 (0,10)	9,2 (0,13)	10,3 (0,10)	9,3 (0,13)	15 (2,4)	4 (1,4)
Norwegia	10,7 (0,11)	9,8 (0,13)	10,6 (0,10)	9,9 (0,15)	31 (3,9)	20 (3,8)
Nowa Zelandia	10,8 (0,12)	10,5 (0,11)	10,5 (0,12)	10,6 (0,11)	6 (1,6)	10 (2,4)
Oman	8,9 (0,13)	10,3 (0,11)	9,0 (0,14)	10,2 (0,10)	31 (2,6)	32 (3,0)
POLSKA	11,1 (0,16)	10,0 (0,14)	11,1 (0,15)	10,1 (0,15)	68 (3,9)	56 (4,1)
Południowa Afryka	9,0 (0,11)	8,6 (0,12)				
Portugalia	9,9 (0,10)	9,9 (0,14)	10,1 (0,10)	9,9 (0,14)	41 (4,0)	38 (3,8)
Serbia	9,6 (0,12)	9,3 (0,15)	9,7 (0,12)	9,3 (0,15)	18 (3,2)	32 (4,3)
Singapur	10,9 (0,00)	10,8 (0,09)	10,9 (0,00)	10,9 (0,10)	98 (0,0)	69 (0,0)
Stowacja	9,3 (0,12)	10,7 (0,13)	9,5 (0,11)	10,7 (0,13)	27 (3,3)	10 (2,2)
Słowenia	11,6 (0,15)	10,4 (0,14)	11,7 (0,15)	10,4 (0,14)	23 (3,7)	31 (4,3)
Stany Zjednoczone	10,8 (0,13)	10,3 (0,10)	10,7 (0,12)	10,4 (0,10)	29 (3,0)	22 (2,6)
Szwecja	10,7 (0,13)	9,3 (0,13)	10,6 (0,13)	9,3 (0,15)	31 (3,5)	28 (4,4)
Tajwan	10,5 (0,13)	10,1 (0,13)	10,6 (0,12)	10,0 (0,13)	94 (1,8)	90 (2,5)
Turcja	7,9 (0,12)	9,0 (0,12)	7,8 (0,14)	9,0 (0,12)	44 (2,9)	8 (1,8)
Węgry	9,5 (0,14)	9,1 (0,15)	9,6 (0,14)	9,3 (0,15)	12 (2,8)	17 (3,1)
Włochy	9,3 (0,07)	9,3 (0,12)	9,3 (0,07)	9,3 (0,12)	31 (3,8)	9 (2,5)
Zjednoczone Emiraty Arabskie	9,8 (0,11)	10,9 (0,09)	9,8 (0,12)	11,0 (0,08)	84 (1,7)	86 (1,6)
Średnia międzynarodowa					38 (0,4)	32 (0,5)

* Procent dzieci uczących się w szkołach, których dyrektorzy odpowiedzieli twierdząco na pytanie: Czy w Pani/Pana szkole jest pracownia przyrodnicza, z której mogą korzystać uczniowie klasy czwartej?

** Procent dzieci uczących się w szkołach, których dyrektorzy odpowiedzieli twierdząco na pytanie: Czy nauczyciel może liczyć na pomoc, gdy jego uczniowie wykonują doświadczenia?

W nawiasie błędy standardowe.

jąc, że wnioskowanie z opinii o faktach jest wysoce zawodne, ponieważ opinie kształtuje skomplikowany układ stosunków między zwierzchnikami dyrektora, dyrektorem i jego podwładnymi. Więcej o tym mogłaby powiedzieć teoria organizacji.

Dodatkowe pytania ankiety rzucają nieco światła na dwa szczegółowe zagadnienia: dostępności komputerów i pracowni przyrodniczych. Dostępność komputerów dla swoich uczniów oceniali nauczyciele matematyki i przyrody. Średnie krajowe ich ocen okazały się niemal identyczne ($r = 0,93$). W dostępności komputerów przoduje Nowa Zelandia, Dania i Gruzja – w tych krajach korzysta z nich na lekcjach matematyki i przyrody powyżej 75 proc. czwartoklasistów. Na drugim biegunie znajdują się Indonezja, Chorwacja, Maroko i RPA, gdzie z komputerów korzysta 10 proc. uczniów lub mniej. W Polsce ten odsetek wynosi 37, co daje nam, wraz z Hiszpanią, rangę 26,5.

Drugie zagadnienie oświetlają odpowiedzi dyrektorów na dwa pytania: Czy w Pani/Pana szkole jest pracownia przyrodnicza, z której mogą korzystać uczniowie klasy czwartej? oraz Czy nauczyciel może liczyć na pomoc, gdy jego uczniowie wykonują doświadczenia? Procenty dzieci uczących się w szkołach, których dyrektorzy odpowiedzi twierdząco na te pytania, znajdują się w dwóch ostatnich kolumnach Tabeli 5.4. Widać, że pracownie przyrodnicze są dostępne dla 80 i więcej proc. czwartoklasistów w ośmiu krajach azjatyckich. Na drugim biegunie sytuuje się m.in. osiem krajów europejskich (w tym Anglia i Irlandia Północna, Irlandia, Belgia i Francja), w których pracownia przyrodnicza jest dostępna dla mniej niż

Tabela 5.5. Rozkłady procentowe skrajnych odpowiedzi polskich nauczycieli na pytania skali obciążenia zawodowego

Pozycja skali	Matematycy (n = 199)		Przyrodnicy (n = 185)	
	Tak	Nie	Tak	Nie
Klasy są przepelnione	14	29	11	30
Program nauczania jest przeladowany.	19	7	16	8
Mam za dużo godzin lekcyjnych.	2	45	2	55
Brakuje mi czasu na przygotowanie się do lekcji.	4	36	3	35
Brakuje mi czasu na pracę z poszczególnymi uczniami.	14	9	11	13
Rodzice uczniów wywierają na mnie zbyt silny nacisk.	1	40	<1	50
Nie mogę nadążyć za zmieniającymi się programami nauczania.	3	43	3	47
Mam za dużo „papierkowej roboty”.	47	2	43	6

Tak – „Zdecydowanie tak”, Nie – „Zdecydowanie nie”

Tabela 5.6. Średnie skali satysfakcji zawodowej nauczycieli

Kraj	Matematycy	Przyrodnicy
Anglia	9,5 (0,20)	9,4 (0,19)
Arabia Saudyjska	10,3 (0,14)	10,2 (0,13)
Australia	10,2 (0,14)	10,3 (0,12)
Bahrajn	10,2 (0,10)	10,3 (0,08)
Belgia (flamandzka)	9,9 (0,14)	9,9 (0,14)
Bułgaria	9,8 (0,15)	9,7 (0,16)
Chile	10,7 (0,16)	10,7 (0,15)
Chorwacja	10,6 (0,13)	10,6 (0,13)
Cypr	10,2 (0,13)	9,6 (0,21)
Czechy	9,2 (0,15)	9,3 (0,16)
Dania	9,1 (0,17)	9,1 (0,16)
Federacja Rosyjska	9,9 (0,12)	9,9 (0,12)
Finlandia	9,8 (0,14)	9,8 (0,13)
Francja	8,9 (0,14)	8,9 (0,15)
Gruzja	10,7 (0,13)	10,7 (0,12)
Hiszpania	10,5 (0,15)	10,3 (0,14)
Holandia	9,9 (0,17)	9,9 (0,17)
Hong Kong	9,0 (0,19)	8,9 (0,20)
Indonezja	10,5 (0,11)	10,4 (0,12)
Iran	11,3 (0,10)	11,3 (0,10)
Irlandia	10,3 (0,18)	10,3 (0,18)
Irlandia Północna	10,3 (0,21)	10,3 (0,21)
Japonia	8,6 (0,15)	8,9 (0,16)
Jordania	9,6 (0,19)	
Kanada	10,2 (0,09)	10,1 (0,10)
Katar	11,0 (0,13)	11,1 (0,11)
Kazachstan	10,8 (0,13)	10,8 (0,13)
Korea	10,1 (0,17)	10,2 (0,18)
Kuwejt	10,2 (0,18)	10,3 (0,13)
Litwa	9,9 (0,16)	10,0 (0,16)
Maroko	10,0 (0,13)	10,1 (0,10)
Niemcy	9,8 (0,14)	9,9 (0,15)
Norwegia	9,8 (0,15)	9,9 (0,15)
Nowa Zelandia	9,8 (0,11)	9,8 (0,11)
Oman	10,8 (0,11)	10,7 (0,10)
POLSKA	9,1 (0,15)	9,0 (0,16)
Południowa Afryka	10,3 (0,15)	
Portugalia	10,1 (0,11)	10,1 (0,11)
Serbia	10,6 (0,14)	10,6 (0,14)
Singapur	9,3 (0,13)	9,2 (0,12)
Słowacja	9,7 (0,14)	9,7 (0,14)
Słowenia	10,1 (0,13)	10,1 (0,13)
Stany Zjednoczone	9,9 (0,12)	9,8 (0,13)
Szwecja	9,5 (0,17)	9,5 (0,16)
Tajwan	9,7 (0,16)	9,8 (0,16)
Turcja	10,3 (0,10)	10,3 (0,10)
Węgry	9,6 (0,14)	9,7 (0,13)
Włochy	9,5 (0,14)	9,6 (0,16)
Zjednoczone Emiraty Arabskie	10,7 (0,07)	10,6 (0,07)

W nawiasie błędy standardowe.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 5.7 Odsetki uczniów nauczanych matematyki i przyrody przez nauczycieli o podanym wykształceniu

Kraj	Matematyka				Przyroda			
	Wyższe II°	Wyższe I°	Policealne	Średnie	Wyższe II°	Wyższe I°	Policealne	Średnie
Anglia	13 (2,9)	86 (3,0)	1 (0,7)	0 (0,0)	12 (2,7)	87 (2,7)	1 (0,9)	0 (0,0)
Arabia Saudyjska	8 (2,0)	67 (4,2)	9 (2,6)	17 (3,5)	1 (0,1)	41 (3,8)	46 (4,2)	12 (2,8)
Australia	12 (2,6)	81 (3,2)	7 (1,9)	0 (0,0)	12 (2,7)	81 (3,3)	7 (1,9)	0 (0,0)
Bahrajn	12 (3,0)	87 (3,1)	1 (0,5)	0 (0,4)	24 (1,0)	73 (1,1)	3 (0,5)	0 (0,0)
Belgia (flamandzka)	1 (0,7)	98 (0,8)	1 (0,4)	0 (0,0)	1 (0,7)	98 (0,8)	1 (0,4)	0 (0,0)
Bułgaria	74 (3,4)	17 (2,5)	10 (2,2)	0 (0,0)	75 (3,6)	17 (2,7)	8 (2,1)	0 (0,0)
Chile	10 (2,7)	82 (3,6)	8 (2,3)	0 (0,0)	12 (3,0)	82 (3,6)	6 (2,1)	0 (0,2)
Chorwacja	0 (0,4)	41 (3,5)	58 (3,5)	0 (0,0)	0 (0,4)	41 (3,5)	58 (3,5)	0 (0,0)
Cypr	61 (3,6)	37 (3,6)	1 (0,7)	0 (0,0)	59 (4,3)	36 (4,1)	4 (1,9)	0 (0,0)
Czechy	94 (1,6)	2 (1,0)	0 (0,2)	4 (1,2)	92 (2,3)	2 (1,0)	0 (0,2)	6 (2,1)
Dania	4 (1,5)	87 (2,8)	3 (1,4)	7 (2,0)	6 (1,8)	88 (2,6)	2 (1,0)	5 (1,8)
Federacja Rosyjska	30 (4,5)	53 (4,3)	17 (2,8)	0 (0,0)	31 (4,6)	52 (4,4)	17 (2,8)	0 (0,0)
Finlandia	90 (1,8)	9 (1,7)	0 (0,1)	1 (0,6)	90 (1,8)	8 (1,7)	0 (0,1)	1 (0,5)
Francja	40 (3,7)	48 (3,8)	9 (2,3)	3 (1,2)	40 (3,8)	46 (4,0)	9 (2,3)	4 (1,6)
Gruzja	85 (3,3)	12 (3,0)	3 (1,4)	0 (0,0)	85 (3,0)	11 (2,7)	3 (1,5)	0 (0,3)
Hiszpania	4 (1,5)	33 (4,0)	61 (4,2)	2 (1,0)	5 (1,5)	34 (3,7)	58 (3,7)	2 (1,1)
Holandia	4 (2,0)	70 (4,2)	25 (4,0)	2 (0,3)	4 (2,0)	70 (4,2)	25 (4,0)	2 (0,3)
Hong Kong	30 (3,8)	66 (4,4)	4 (1,8)	0 (0,0)	39 (5,4)	54 (5,2)	7 (2,3)	0 (0,0)
Indonezja	2 (0,7)	85 (2,2)	5 (1,3)	8 (2,0)	2 (0,6)	85 (2,0)	5 (1,3)	8 (1,8)
Iran	7 (1,5)	55 (3,8)	28 (3,7)	10 (2,3)	7 (1,5)	55 (3,8)	28 (3,7)	10 (2,3)
Irlandia	13 (2,3)	84 (2,7)	3 (1,2)	1 (0,5)	13 (2,3)	84 (2,7)	3 (1,2)	1 (0,5)
Irlandia Północna	16 (3,3)	83 (3,4)	0 (0,0)	2 (0,9)	16 (3,3)	83 (3,4)	0 (0,0)	2 (0,9)
Japonia	4 (1,1)	90 (2,2)	7 (1,8)	0 (0,0)	6 (1,8)	89 (2,5)	5 (1,8)	0 (0,0)
Jordania	7 (2,0)	72 (3,9)	16 (3,3)	5 (1,8)				
Kanada	14 (2,0)	86 (2,0)	0 (0,0)	0 (0,0)	12 (1,9)	88 (1,9)	0 (0,0)	0 (0,0)
Katar	15 (2,3)	79 (2,5)	4 (1,4)	2 (1,4)	22 (4,1)	71 (4,2)	5 (1,6)	3 (1,5)
Kazachstan	1 (0,6)	78 (3,5)	14 (2,8)	8 (2,5)	1 (0,6)	78 (3,5)	14 (2,8)	8 (2,5)
Korea Południowa	21 (3,2)	72 (3,4)	6 (2,0)	0 (0,0)	24 (3,0)	70 (3,2)	7 (2,0)	0 (0,0)
Kuwejt	12 (2,7)	68 (3,8)	17 (2,8)	3 (1,2)	12 (4,2)	85 (4,3)	0 (0,4)	3 (1,6)
Litwa	21 (3,6)	74 (3,5)	5 (1,4)	0 (0,0)	21 (3,3)	74 (3,3)	5 (1,4)	0 (0,0)
Maroko	1 (0,6)	28 (3,7)	3 (1,2)	67 (3,7)	1 (0,7)	28 (4,0)	6 (1,6)	65 (3,9)
Niemcy	85 (1,8)	0 (0,0)	15 (1,8)	0 (0,0)	85 (1,7)	0 (0,3)	15 (1,8)	0 (0,0)
Norwegia	8 (2,7)	88 (3,0)	4 (1,4)	0 (0,0)	9 (2,5)	84 (3,2)	6 (2,0)	1 (0,7)
Nowa Zelandia	27 (2,3)	58 (2,6)	15 (2,0)	0 (0,0)	26 (2,3)	58 (2,7)	16 (2,0)	0 (0,0)
Oman	29 (2,9)	66 (3,2)	3 (1,2)	2 (1,0)	30 (2,8)	67 (2,8)	3 (1,0)	0 (0,2)
POLSKA	97 (1,3)	3 (1,3)	0 (0,0)	0 (0,0)	100 (0,0)	0 (0,0)	0 (0,0)	0 (0,0)
Południowa Afryka	1 (0,5)	46 (3,4)	49 (3,4)	4 (1,0)				
Portugalia	7 (1,8)	89 (2,3)	4 (1,3)	0 (0,0)	7 (1,8)	89 (2,3)	4 (1,3)	0 (0,0)
Serbia	12 (2,6)	39 (4,1)	48 (4,3)	1 (0,5)	12 (2,6)	39 (4,1)	48 (4,3)	1 (0,5)
Singapur	10 (1,6)	69 (2,6)	20 (2,1)	1 (0,6)	13 (1,7)	69 (2,3)	16 (2,0)	2 (0,7)
Słowacja	100 (0,0)	0 (0,0)	0 (0,0)	0 (0,0)	99 (0,5)	0 (0,0)	0 (0,0)	1 (0,5)
Słowenia	59 (3,2)	0 (0,4)	41 (3,2)	0 (0,0)	59 (3,2)	0 (0,4)	41 (3,2)	0 (0,0)
Stany Zjednoczone	53 (2,4)	47 (2,4)	0 (0,0)	0 (0,0)	52 (2,5)	48 (2,5)	0 (0,0)	0 (0,0)
Szwecja	11 (2,7)	76 (3,7)	9 (2,7)	4 (1,8)	10 (2,5)	80 (3,5)	6 (2,1)	4 (1,9)
Tajwan	39 (4,0)	61 (4,0)	0 (0,3)	0 (0,0)	36 (4,1)	63 (4,1)	1 (0,7)	0 (0,0)
Turcja	3 (1,1)	81 (2,6)	16 (2,5)	0 (0,0)	3 (1,1)	81 (2,6)	16 (2,5)	0 (0,0)
Węgry	1 (0,3)	99 (0,5)	0 (0,4)	0 (0,0)	2 (0,6)	98 (0,8)	1 (0,6)	0 (0,0)
Włochy	3 (1,4)	16 (3,4)	13 (2,6)	68 (3,6)	3 (1,5)	20 (3,6)	9 (2,1)	68 (3,9)
Zjednoczone Emiraty Arabskie	32 (2,7)	61 (2,6)	6 (1,3)	1 (0,3)	34 (2,5)	61 (2,6)	4 (1,0)	1 (0,3)
Średnia międzynarodowa	26 (0,3)	58 (0,4)	12 (0,3)	5 (0,2)	28 (0,4)	57 (0,4)	11 (0,3)	4 (0,2)

Wyższe II° – dyplom magistra, lekarza lub doktora, Wyższe I° – dyplom licencjata lub inżyniera, Policealne – studium nauczycielskie, Średnie – matura. W nawiasie błędy standardowe.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

10 proc. czwartoklasistów. Polskie szkoły podstawowe według ich dyrektorów udostępniają pracowni przyrodnicze 68 procentom uczniów, co daje nam wysoką 10. pozycję. Warto zauważyć, że według dyrektorów wiele polskich szkół zapewnia nauczycielom przeprowadzającym doświadczenia na lekcjach pomoc ze strony personelu szkoły. Ponieważ analogicznego pytania nie zadano nauczycielom, tej krzepiącej informacji nie sposób zweryfikować.

Zasoby kadrowe

W badaniu TIMSS nauczyciele są traktowani jako element środowiska uczniów, co znaczy, że rozkład każdej cechy nauczyciela wyraża się za pośrednictwem procentu uczniów nauczanych przez nauczycieli odpowiedniej kategorii. Wartości bezpośrednie można obliczyć tylko w polskiej bazie danych indywidualnych. W Polsce w badaniu wzięło udział 200 nauczycieli matematyki (88 proc. kobiet), 185 nauczycieli przyrody (95 proc. kobiet) i pięć nauczycielek obu tych przedmiotów – razem 390 osób, w tym 92 proc. kobiet. Średni wiek polskiego nauczyciela matematyki to 44,4 roku, przyrody – 45,6 roku, obu przedmiotów – 40,5 roku. Raport międzynarodowy dostarcza informacji porównawczych jedynie o stażu pracy. Średni staż nauczyciela matematyki i przyrody to 17 lat. W Polsce istotnie więcej – 19 i 21 lat.

Dane o wykształceniu nauczycieli zawiera Tabela 5.7. Na świecie nauczyciele z pełnym wykształceniem uniwersyteckim II stopnia (odpowiednikiem naszego magisterium) nauczają matematyki 26 proc., a przyrody 28 proc. uczniów. W Polsce, odpowiednio, 97 i 100 proc. Więcej magistrów jest tylko w szkołach na Słowacji.

Pod względem odczuwanego obciążenia zawodowego polscy nauczyciele znaleźli się na przedostatnim miejscu! Lżej pracuje się tylko ich kolegom w Gruzji. Tabela 5.5 pozwala wnikać w elementy polskiego optymizmu. Tylko w trzech sprawach: przeładowania programu nauczania, braku czasu na pracę z poszczególnymi uczniami i prowadzenia dokumentacji liczba zdecydowanych niezadowolonych góruje nad liczbą zdecydowanych zadowolonych. W pozostałych, w tym tak newralgicznych jak liczba uczniów w oddziale klasowym, liczba godzin lekcyjnych czy presja rodziców, przeważają zadowoleni. Wśród przyrodników jest ich nieco więcej niż wśród matematyków.

Ostatni rys do portretu nauczyciela wnosi skala satysfakcji zawodowej – krótka, siedmiopozycyjna (np. Jak często Pani/Pan czuje, że Pani/Pana praca ma sens i cel? Jak często Pani/Pan czuje chęć, by uczyć tak długo, jak to tylko będzie możliwe), ale wysoce rze-

telna (w Polsce $\alpha = 0,91$). Niezależnie od nauczanego przedmiotu średni poziom satysfakcji jest najwyższy w Iranie i Katarze, najniższy zaś w Japonii i Francji. Polscy nauczyciele zajmują w tym uporządkowaniu zaskakująco niskie pozycje: matematycy (wraz z matematykami z Danii) 45,5, a przyrodniczy 44. Warto dodać, że w analogicznym rankingu przed czterema laty polscy nauczyciele edukacji wczesnoszkolnej zajęli ósme miejsce.

Kultura szkoły

Kulturą szkoły nazywamy układ wynalezionych lub przejętych schematów współdziałania mających na celu rozwiązywanie powtarzalnych problemów o charakterze instrumentalnym i ekspresyjnym. Główny problem instrumentalny szkoły dotyczy ukierunkowania oddziaływań dydaktycznych. W badaniach TIMSS sprowadzono go do pytania: W jakim stopniu wszyscy uczestnicy życia w szkole: nauczyciele, uczniowie i ich rodzice podzielają przekonanie, że głównym celem edukacji są jak najwyższe osiągnięcia szkolne uczniów? Na to pytanie odpowiadali dyrektorzy i nauczyciele i częściowo rodzice.

Dyrektorzy i nauczyciele wypełniali 13- i 17-pozycyjną skalę nastawienia szkoły na osiągnięcia uczniów. Złożyły się na nią pytania o to, czy wszyscy nauczyciele podobnie pojmują cele nauczania i czy współpracują ze sobą w celu podniesienia wyników nauczania, czy rodzice angażują się w życie szkoły i czy wymagają od szkoły utrzymania wysokiego poziomu nauczania, czy uczniowie przykładają się do nauki i darzą szacunkiem rówieśników osiągających wysokie wyniki w nauce itp.⁶

Pamiętamy, że opinie dyrektorów i nauczycieli w sprawie zasobów materialnych szkoły znacznie się różniły. Tym razem jest inaczej. Średnie opinie nauczycieli matematyki i przyrody są niemal identyczne ($r = 0,99$) i wysoce zgodne ze średnimi opiniami dyrektorów ($r = 0,88$ i $0,90$). Ranking krajów otwiera Korea Południowa, tuż przed Irlandią Północną i Kazachstanem, zamyka zaś Maroko za Chile i Czechami. Polska znajduje się dokładnie pośrodku, wraz z Cyprem, Finlandią i Hong Kongiem.

Kulturę szkoły oceniali też rodzice za pomocą 8-pozycyjnej skali⁷. Pierwsze miejsce zajęły szkoły w Kazachstanie, ostatnie – w Japonii. Polskie i tajwańskie z rangą 31,5 znalazły się nisko. O czym to świad-

⁶ Spójność skal w Polsce była wysoka ($\alpha = 0,87$ i $0,89$).

⁷ Pięć krajów, m.in. Anglia, Australia i Stany Zjednoczone, nie dostarczyło danych potrzebnych do obliczenia średnich.

Tabela 5.8. Rozkłady procentowe i średnie skali poczucia bezpieczeństwa uczniów

Kraj	Procent uczniów atakowanych			Średnia
	prawie nigdy	raz na miesiąc	raz na tydzień	
Kazachstan	75 (1,1)	18 (0,8)	7 (0,6)	11,1 (0,07)
Korea	76 (1,0)	20 (0,8)	4 (0,4)	11,0 (0,05)
Serbia	73 (1,0)	19 (0,9)	8 (0,5)	10,9 (0,05)
Irlandia	73 (1,2)	20 (1,0)	6 (0,4)	10,8 (0,06)
Chorwacja	73 (1,2)	19 (0,9)	8 (0,6)	10,8 (0,06)
Gruzja	73 (1,1)	18 (0,7)	9 (0,7)	10,8 (0,05)
POLSKA	73 (1,0)	19 (0,8)	8 (0,5)	10,7 (0,05)
Japonia	68 (1,3)	23 (1,0)	8 (0,6)	10,6 (0,05)
Finlandia	71 (1,2)	22 (0,9)	7 (0,5)	10,5 (0,05)
Norwegia	70 (1,3)	23 (1,0)	7 (0,6)	10,5 (0,05)
Francja	65 (1,2)	26 (1,0)	8 (0,6)	10,4 (0,05)
Szwecja	65 (1,3)	28 (1,1)	7 (0,5)	10,3 (0,05)
Irlandia Północna	64 (1,5)	27 (1,1)	10 (0,7)	10,3 (0,06)
Czechy	60 (1,1)	28 (0,9)	12 (0,7)	10,2 (0,05)
Chile	60 (1,3)	24 (0,9)	16 (0,8)	10,1 (0,06)
Tajwan	58 (1,1)	29 (1,0)	13 (0,7)	10,1 (0,04)
Słowacja	57 (1,1)	30 (0,8)	13 (0,7)	10,1 (0,06)
Turcja	57 (1,1)	28 (0,8)	14 (0,7)	10,1 (0,05)
Holandia	59 (1,4)	31 (0,9)	10 (0,9)	10,0 (0,05)
Dania	58 (1,2)	32 (0,9)	10 (0,7)	10,0 (0,05)
Węgry	58 (1,3)	31 (1,1)	11 (0,7)	10,0 (0,05)
Słowenia	58 (1,0)	29 (0,9)	14 (0,8)	10,0 (0,05)
Portugalia	57 (1,0)	29 (0,9)	15 (0,9)	10,0 (0,04)
Niemcy	57 (1,3)	30 (0,9)	13 (0,7)	10,0 (0,05)
Litwa	56 (1,3)	31 (1,0)	13 (0,7)	9,9 (0,05)
Stany Zjednoczone	56 (0,8)	29 (0,5)	15 (0,5)	9,9 (0,04)
Cypr	55 (1,2)	29 (1,0)	16 (0,8)	9,9 (0,06)
Hong Kong	54 (1,4)	32 (1,1)	14 (0,9)	9,9 (0,05)
Bułgaria	54 (1,9)	30 (1,1)	16 (1,1)	9,9 (0,08)
Anglia	54 (1,3)	31 (1,1)	15 (0,8)	9,8 (0,05)
Jordania	52 (1,8)	26 (1,1)	21 (1,4)	9,8 (0,09)
Federacja Rosyjska	51 (1,3)	33 (0,9)	16 (0,6)	9,8 (0,05)
Kanada	53 (0,9)	30 (0,6)	17 (0,8)	9,7 (0,04)
Iran	50 (1,6)	32 (0,9)	18 (1,1)	9,7 (0,07)
Kuwejt	48 (1,2)	31 (0,8)	21 (0,9)	9,7 (0,06)
Włochy	50 (1,0)	35 (0,9)	15 (0,7)	9,6 (0,04)
Hiszpania	48 (1,0)	33 (0,6)	19 (0,8)	9,6 (0,05)
Belgia (flamandzka)	47 (1,3)	36 (0,9)	17 (0,8)	9,6 (0,05)
Singapur	47 (0,9)	34 (0,6)	19 (0,7)	9,5 (0,03)
Arabia Saudyjska	47 (1,7)	27 (1,1)	26 (1,3)	9,5 (0,08)
Maroko	44 (1,5)	35 (1,1)	21 (1,0)	9,5 (0,06)
Australia	45 (1,3)	36 (1,1)	20 (1,1)	9,4 (0,05)
Indonezja	44 (1,4)	31 (1,0)	25 (1,0)	9,4 (0,07)
Zjednoczone Emiraty Arabskie	43 (1,0)	31 (0,5)	26 (0,8)	9,4 (0,05)
Oman	42 (1,6)	33 (1,0)	25 (1,0)	9,4 (0,06)
Katar	43 (1,2)	28 (0,8)	28 (1,0)	9,3 (0,06)
Nowa Zelandia	40 (1,0)	36 (0,7)	24 (0,7)	9,2 (0,04)
Bahrajn	34 (0,7)	33 (0,6)	33 (0,7)	9,0 (0,03)
Południowa Afryka	23 (1,0)	34 (0,9)	44 (1,5)	8,5 (0,05)

W nawiasie błędy standardowe.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

czy, nie wiadomo, ponieważ zgodność uporządkowań szkół według ich personelu i według rodziców jest niska ($R = 0,30$). Rekordowa rozbieżność zachodzi w wypadku Korei – opinie rodziców przesunęły tamtejsze szkoły z pierwszego miejsca na przedostatnie.

Przejdźmy do ekspresyjnych aspektów kultury szkoły, czyli reguł współżycia społecznego. Sondują je dwie skale wykroczeń uczniowskich – jedna wypełniana przez dyrektorów, druga przez samych uczniów. Dyrektor oceniał dotkliwość 10 typowych wykroczeń uczniowskich, np. wagarowania, spóźniania się na lekcje, przeszkadzania w lekcjach, wandalizmu i różnych form agresji. Uczeń oceniał częstość (w cztero-punktowej skali: przynajmniej raz w tygodniu – raz lub dwa razy w miesiącu – kilka razy w roku – nigdy) ośmiu form zamachu na swoje bezpieczeństwo ze strony rówieśników: jak często ukradli mu coś, zmuszali do rzeczy, których nie chciał robić, zastraszały go, rozpuszczali o nim nie mile informacje, bili go (np. przewracali, uderzali, kopali), rozpuszczali kłamstwa na jego temat, nie pozwalali mu grać lub bawić się ze wszystkimi, wyśmiewali się z niego lub przezywali.

Według opinii dyrektorów najbardziej bezpieczne szkoły są w Korei, a najmniej – w Indonezji. Polska razem z Serbią, Słowenią i Węgrami ma w tym uporządkowaniu niską rangę: 32,5. Obraz kultury współżycia według uczniów przedstawia Tabela 5.8. Jak widać, uczniowie czują się najbezpieczniej w szkołach Kazachstanu, a najmniej bezpiecznie w RPA. W uporządkowaniu według średnich ocen wystawionych szkole przez uczniów polska szkoła sytuuje się na siódmej pozycji.

Które opinie są bliższe prawdy: dyrektorów czy uczniów? Trzeba pamiętać, że obie skale składały się częściowo z innych pozycji i że dyrektorzy występowali jako strażnicy ładu organizacyjnego, a uczniowie mówili o tym, co ich osobiście dotyka. Różnice perspektyw usprawiedliwiają rozbieżność ocen, dlatego niewysoka zgodność uporządkowań krajów według średnich obu skal ($R = 0,34$) nie powinna zaskakiwać⁸. Polska wraz z Kazachstanem i Serbią tworzy czołówkę krajów, w których uczniowie o wiele lepiej mówią o swojej szkole niż dyrektorzy. W RPA, Nowej Zelandii i Bahrajnie uczniowie mówią o szkole o wiele gorzej niż dyrektorzy. Jest możliwe, że w jednych szkołach uczniowie następczą trudności nauczycielom, ale wobec siebie są koleżeńscy, a w innych przykładnie odnoszą się do dorosłych, lecz nieprzyjaźnie do siebie. Czy polskie szkoły należą do pierwszej grupy? – o tym mogą powiedzieć tylko dalsze badania.

⁸ W Polsce współczynnik β ważonej regresji średnich ocen uczniowskich na ocenę dyrektora jest istotny, ale wynosi tylko 0,24.

Czas nauczania

Informacja o rocznym czasie nauczania w klasie czwartej (w godzinach zegarowych) pochodzi od dyrektorów szkół, a informacja o rocznym czasie nauczania matematyki i przyrody – od nauczycieli tych przedmiotów. Średnie czasu nauczania zawiera Tabela 5.9. Tym, co przede wszystkim rzuca się w oczy, jest ogromna rozpiętość całego czasu nauczania. W RPA jest to prawie 1200 godzin, na Litwie – niewiele ponad połowa tego czasu. Podstawą obliczeń była odpowiedź na pytanie „Ile wynosi cały czas nauczania, z wyłączeniem przerw, w typowym dniu szkolnym?”⁹. Ponieważ zabrakło definicji całego czasu nauczania, dyrektorzy mogli rozumieć go różnie – np. ograniczyć się do podania liczby godzin lekcyjnych lub wliczać też czas nauczania poza lekcjami. W Polsce wybraliśmy pierwszą, łatwiejszą do weryfikacji możliwość, pytając: „Jaki jest tygodniowy przydział godzin lekcyjnych dla klasy czwartej w Pani/Pana szkole?”. Na Litwie zapewne zrobiono podobnie.

Ramowy plan nauczania zobowiązuje polskie szkoły podstawowe do prowadzenia w klasach od czwartej do szóstej obowiązkowych zajęć edukacyjnych i zajęć z wychowawcą w wymiarze co najmniej 1849 godzin, czyli średnio 616 godzin rocznie. Średni czas nauczania według dyrektorów przewyższa to minimum o 22 proc. Według nauczycieli średni czas nauczania matematyki jest większy od minimum o 17 proc., a przyrody o 15 proc. Jeśli nasi dyrektorzy przeszacowali pełny czas nauczania, to niewiele.

Warto zauważyć, że w próbie międzynarodowej matematyka ma 18-procentowy, a przyroda 9-procentowy udział w całym czasie nauczania. W Polsce analogiczne odsetki to 15 i 11.

Niska pozycja Polski w rankingu czasu nauczania może budzić niepokój, jednak wydaje się, że nie ma po temu podstaw. „Goły” czas nauczania nie decyduje o osiągnięciach uczniów. Przekonują o tym współczynniki zgodności uporządkowań krajów według obu średnich: $-0,07$ dla matematyki i $-0,17$ dla przyrody, oba nieistotne.

Metodyka nauczania

Nie od dziś wiadomo, że ankieta nie jest najlepszym źródłem informacji na temat metod nauczania, ponieważ wymaga od nauczyciela wysoce przetworzonych i uogólnionych sądów, podczas gdy prawda o lekcji kryje się albo w szczegółach interakcji nauczyciela z uczniami, albo w nie do końca zwerba-

⁹ What is the total instructional time, excluding breaks, in a typical day?

Tabela 5.9. Roczny czas nauczania (w godzinach zegarowych)

Kraj	Cały	Matematyki	Przyrody
Anglia	994 (9,9)	189 (4,5)	61 (2,2)
Arabia Saudyjska	1080 (19,6)	148 (4,5)	77 (3,7)
Australia	1014 (8,4)	202 (3,5)	57 (1,5)
Bahrajn	976 (0,6)	159 (2,9)	103 (0,6)
Belgia (flamandzka)	955 (11,6)	218 (3,2)	
Bułgaria	707 (27,3)	105 (2,9)	42 (2,3)
Chile	1094 (16,9)	206 (6,4)	93 (2,5)
Chorwacja	778 (21,6)	124 (1,8)	82 (1,5)
Cypr	827 (12,4)	161 (5,5)	48 (0,9)
Czechy	771 (10,4)	125 (4,1)	38 (2,0)
Dania	1051 (11,2)	150 (3,1)	80 (2,3)
Federacja Rosyjska	661 (6,9)	106 (1,4)	49 (0,9)
Finlandia	737 (8,9)	115 (2,2)	82 (1,8)
Francja	858 (8,2)	193 (3,9)	56 (1,8)
Gruzja	743 (19,5)	138 (2,1)	80 (1,4)
Hiszpania	864 (10,2)	161 (2,3)	124 (2,6)
Holandia	1073 (16,2)		
Hong Kong	999 (13,1)	159 (4,7)	
Indonezja	1095 (20,9)	149 (5,0)	116 (4,0)
Iran	645 (6,4)	112 (2,3)	87 (3,0)
Irlandia	854 (0,0)	165 (2,4)	32 (0,7)
Irlandia Północna	962 (10,2)	215 (6,5)	38 (2,1)
Japonia	903 (3,7)	151 (1,1)	91 (0,5)
Jordania	931 (14,2)	133 (3,3)	
Kanada	951 (4,1)	196 (3,2)	81 (2,0)
Katar	1056 (16,1)	185 (4,6)	125 (4,4)
Kazachstan	813 (16,2)	132 (3,8)	58 (2,9)
Korea Południowa	712 (8,9)	100 (1,4)	76 (1,0)
Kuwejt	912 (27,9)	128 (4,4)	77 (3,5)
Litwa	629 (5,5)	111 (1,6)	53 (1,0)
Maroko	1054 (18,8)	172 (2,8)	54 (0,9)
Niemcy	820 (9,1)	147 (2,0)	61 (3,8)
Norwegia	817 (8,7)	117 (2,4)	59 (1,7)
Nowa Zelandia	923 (5,5)	163 (2,3)	43 (2,0)
Oman	962 (11,7)	148 (4,5)	123 (3,1)
POLSKA	752 (6,9)	112 (1,1)	84 (1,1)
Południowa Afryka	1199 (13,7)	227 (4,4)	
Portugalia	864 (8,5)	275 (4,0)	111 (3,8)
Serbia	737 (16,2)	154 (1,6)	75 (3,5)
Singapur	986 (0,0)	201 (1,6)	85 (1,4)
Słowacja	759 (8,1)	129 (2,1)	52 (2,0)
Słowenia	716 (7,2)	144 (1,2)	86 (1,3)
Stany Zjednoczone	1088 (9,2)	216 (4,1)	100 (3,7)
Szwecja	839 (10,6)	110 (2,3)	79 (1,8)
Tajwan	969 (14,4)	128 (4,3)	91 (1,9)
Turcja	847 (18,0)	120 (3,3)	83 (1,7)
Węgry	784 (11,8)	129 (2,5)	63 (1,7)
Włochy	1061 (20,5)	231 (4,5)	76 (1,6)
Zjednoczone Emiraty Arabskie	1009 (4,6)	162 (2,4)	111 (2,1)
Średnia międzynarodowa	894 (1,9)	157 (0,5)	76 (0,3)

W nawiasie błędy standardowe.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 5.10. Odsetki uczniów wykorzystujących komputery w wyszczególnionych celach oraz średnia skali eksperymentowania

Kraj	Matematyka			Przyroda				
	Badanie pojęć i twierdzeń	Ćwiczenie umiejętności	Poszukiwanie informacji	Ćwiczenie umiejętności	Poszukiwanie informacji	Obserwowanie lub eksperymentowanie	Wyjaśnianie zjawisk na podstawie symulacji	Średnia eksperymentowania
Anglia	49 (3,7)	52 (3,6)	45 (3,9)	43 (4,5)	69 (4,3)	42 (4,4)	54 (4,2)	10,1 (0,14)
Arabia Saudyjska	18 (2,9)	19 (3,0)	18 (2,9)	32 (3,5)	34 (3,5)	31 (3,4)	33 (3,5)	11,1 (0,12)
Australia	53 (3,8)	57 (3,9)	49 (3,8)	39 (3,2)	60 (3,4)	37 (3,3)	46 (3,5)	9,9 (0,12)
Bahrajn	19 (1,1)	17 (1,0)	20 (1,1)	28 (1,3)	34 (1,3)	29 (1,3)	29 (1,3)	10,7 (0,04)
Belgia (flamandzka)	14 (2,2)	32 (3,4)	20 (2,9)	62 (3,7)	75 (3,5)	18 (2,6)	29 (3,2)	8,3 (0,10)
Bułgaria	11 (2,5)	21 (3,6)	23 (3,6)	25 (3,7)	31 (3,8)	13 (2,8)	16 (3,1)	9,3 (0,13)
Chile	31 (3,8)	36 (4,2)	31 (4,1)	38 (4,3)	43 (4,0)	30 (4,0)	34 (4,2)	9,9 (0,17)
Chorwacja	1 (0,7)	2 (1,0)	1 (0,8)	9 (1,9)	10 (1,9)	7 (1,6)	6 (1,4)	10,1 (0,15)
Cypr	42 (3,0)	49 (2,9)	38 (3,2)	35 (4,4)	46 (4,7)	32 (4,2)	46 (4,8)	11,9 (0,13)
Czechy	14 (2,4)	27 (3,2)	19 (2,8)	26 (3,2)	30 (3,2)	18 (2,7)	20 (3,0)	9,0 (0,12)
Dania	52 (4,1)	79 (3,1)	45 (3,9)	50 (4,3)	82 (3,3)	38 (3,9)	49 (4,0)	9,2 (0,10)
Federacja Rosyjska	49 (4,1)	60 (3,7)	58 (4,0)	59 (4,2)	62 (3,6)	45 (3,5)	39 (3,4)	9,2 (0,13)
Finlandia	32 (3,3)	50 (3,5)	30 (3,3)	48 (4,0)	61 (3,3)	25 (3,1)	22 (2,8)	8,7 (0,09)
Francja	3 (1,2)	9 (2,3)	6 (1,8)	7 (1,8)	21 (3,3)	7 (1,6)	15 (2,9)	9,7 (0,11)
Gruzja	58 (4,6)	65 (4,1)	62 (4,4)	79 (3,6)	80 (3,4)	62 (4,8)	74 (3,7)	9,8 (0,15)
Hiszpania	20 (3,0)	28 (3,4)	26 (3,2)	27 (3,1)	36 (4,3)	19 (2,9)	21 (2,9)	9,1 (0,11)
Holandia	48 (4,1)	73 (3,7)	49 (3,9)	37 (4,9)	55 (4,8)	22 (4,1)	20 (3,6)	8,3 (0,11)
Hong Kong	33 (4,9)	35 (4,4)	29 (4,4)	29 (4,5)	37 (4,5)	33 (4,3)	27 (4,1)	9,0 (0,15)
Indonezja	5 (1,3)	5 (1,2)	5 (1,1)	4 (0,8)	4 (0,8)	4 (0,7)	4 (1,0)	10,6 (0,15)
Iran	17 (2,6)	19 (2,9)	17 (2,7)	18 (2,7)	22 (2,8)	21 (2,7)	20 (2,9)	12,1 (0,13)
Irlandia	31 (3,7)	34 (3,8)	27 (4,0)	22 (3,4)	36 (4,1)	14 (3,1)	22 (3,6)	10,0 (0,12)
Irlandia Północna	58 (3,9)	68 (3,8)	58 (5,0)	37 (4,5)	69 (4,2)	23 (4,5)	39 (4,8)	8,5 (0,13)
Japonia	10 (2,3)	14 (2,6)	12 (2,6)	13 (2,4)	29 (3,7)	12 (2,4)	35 (4,0)	11,3 (0,12)
Jordania	25 (3,1)	28 (3,3)	27 (3,2)					
Kanada	36 (2,6)	41 (2,6)	32 (2,5)	42 (2,9)	59 (2,9)	36 (2,7)	40 (3,0)	9,7 (0,09)
Katar	32 (3,5)	34 (3,3)	31 (3,3)	37 (3,9)	41 (3,8)	33 (3,8)	33 (3,6)	11,1 (0,15)
Kazachstan	38 (3,9)	39 (3,8)	39 (3,8)	48 (3,8)	48 (3,8)	48 (3,9)	46 (3,6)	10,9 (0,17)
Korea Południowa	7 (2,2)	8 (2,4)	8 (2,3)	14 (3,2)	19 (3,5)	18 (3,4)	16 (3,2)	11,4 (0,14)
Kuwait	12 (1,9)	12 (2,0)	13 (1,9)	31 (4,3)	33 (4,3)	31 (4,2)	32 (4,1)	11,2 (0,15)
Litwa	26 (3,7)	27 (3,7)	29 (3,9)	37 (4,0)	41 (3,9)	26 (3,5)	18 (3,2)	8,8 (0,13)
Maroko	2 (0,9)	2 (0,9)	2 (0,9)	5 (1,5)	8 (1,8)	7 (1,8)	8 (1,9)	11,1 (0,12)
Niemcy	23 (2,8)	43 (4,0)	29 (3,1)	22 (2,8)	47 (3,8)	10 (1,7)	14 (2,3)	8,9 (0,14)
Norwegia	44 (3,8)	54 (3,7)	40 (3,5)	47 (3,8)	59 (4,2)	42 (3,9)	36 (4,1)	8,4 (0,12)
Nowa Zelandia	78 (2,3)	86 (2,0)	76 (2,3)	46 (2,3)	87 (1,8)	52 (2,7)	55 (2,8)	9,5 (0,08)
Oman	10 (1,8)	11 (1,9)	10 (1,9)	14 (2,4)	14 (2,4)	13 (2,1)	12 (2,2)	12,1 (0,10)
POLSKA	18 (3,2)	28 (3,6)	23 (3,7)	36 (3,9)	35 (3,7)	29 (3,2)	31 (3,5)	9,6 (0,16)
Południowa Afryka	5 (1,6)	5 (1,3)	5 (1,4)					
Portugalia	13 (2,3)	15 (2,4)	17 (2,7)	18 (2,5)	22 (2,8)	13 (2,3)	13 (2,5)	9,2 (0,13)
Serbia	7 (1,6)	11 (2,3)	10 (2,3)	18 (3,1)	18 (3,0)	9 (2,1)	8 (1,8)	9,8 (0,15)
Singapur	30 (2,2)	34 (2,2)	28 (2,3)	36 (2,7)	43 (2,6)	35 (2,7)	31 (2,5)	10,7 (0,07)
Słowacja	19 (2,7)	21 (2,9)	21 (2,7)	33 (2,8)	35 (2,8)	22 (2,6)	26 (2,9)	10,2 (0,12)
Słowenia	9 (2,1)	12 (2,5)	12 (2,3)	15 (2,6)	20 (2,8)	11 (2,3)	17 (2,6)	9,4 (0,12)
Stany Zjednoczone	38 (3,0)	43 (3,2)	32 (2,5)	30 (2,4)	40 (2,7)	28 (2,6)	27 (2,4)	9,9 (0,11)
Szwecja	33 (3,4)	63 (4,1)	33 (3,7)	24 (3,6)	61 (3,9)	22 (3,7)	29 (3,9)	9,1 (0,14)
Tajwan	24 (3,3)	25 (3,4)	18 (3,2)	31 (4,0)	36 (4,0)	31 (3,9)	31 (3,8)	10,6 (0,11)
Turcja	25 (2,9)	26 (3,0)	27 (3,0)	31 (3,0)	32 (3,0)	31 (3,0)	25 (2,7)	11,0 (0,17)
Węgry	10 (2,3)	20 (2,9)	14 (2,8)	26 (3,5)	27 (3,6)	19 (3,3)	20 (3,2)	9,1 (0,08)
Włochy	25 (2,9)	29 (3,3)	25 (3,0)	34 (3,5)	41 (3,8)	33 (3,6)	34 (3,6)	10,0 (0,15)
Emiraty Arabskie	23 (1,4)	24 (1,5)	23 (1,4)	29 (2,1)	33 (2,0)	27 (2,0)	28 (2,1)	11,3 (0,08)

W nawiasie błędy standardowe.

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

Tabela 5.11. Odsetki uczniów w Polsce, którzy spotykają się z niższymi wymaganiami ze strony nauczyciela przyrody

Na lekcjach przyrody jak często wymaga Pani/Pan od uczniów w tym oddziale, żeby...	Na każdej lub prawie każdej lekcji	Na około połowie lekcji	Na niektórych lekcjach	Nigdy
obserwowali i opisywali zjawiska naturalne, np. pogodę lub wzrost rośliny?	25 (3,5)	41 (3,3)	34 (4,1)	–
przyglądali się doświadczeniom, które Pani/Pan przeprowadza?	19 (2,8)	30 (4,0)	50 (4,2)	1 (0,1)
obmyślali i planowali własne doświadczenia?	7 (2,3)	24 (3,5)	67 (3,9)	2 (1,2)
przeprowadzali własne doświadczenia?	5 (2,1)	22 (3,5)	72 (3,9)	1 (0,1)
przedstawiali dane pochodzące z doświadczeń?	9 (2,6)	22 (3,4)	69 (3,8)	–
interpretowali dane pochodzące z doświadczeń?	11 (2,3)	25 (3,4)	63 (3,5)	1 (0,1)
wykorzystywali dane z doświadczeń do uzasadnienia wniosków?	11 (2,2)	26 (3,4)	63 (3,6)	–
przewodzili badania w terenie?	2 (2,1)	16 (3,0)	80 (3,0)	2 (0,9)

IDBAnalyzer, odsetki ważone, w nawiasie błędy standardowe obliczone metodą *jackknife*

lizowanych założeniach, z którymi nauczyciel przystępuje do pracy. W ankiecie nauczyciela aż siedem pytań poświęcono metodyce, ale odpowiedzi najwyraźniej nie udało się wyskalować, ponieważ w raporcie międzynarodowym zamieszczono jedynie dane o sposobach wykorzystywania komputerów na lekcjach matematyki i przyrody oraz o częstości demonstracji i eksperymentów przyrodniczych. Dane te zawiera Tabela 5.10.

Porównanie krajowych odsetków uczniów, którzy na lekcjach wykorzystują komputery w celach wymienionych w główce tabeli, nie ujawnia żadnego wzoru, z czego można wnioskować, że wszystkie cele – zarówno ogólne: poszukiwanie informacji i ćwiczenie umiejętności, jak i specyficzne: badanie pojęć i twierdzeń, rejestrowanie danych z obserwacji lub eksperymentów oraz wyjaśnianie zjawisk na podstawie symulacji,

są równoprawnymi składnikami metodyki nauczania. Interesujący jest natomiast ujemny współczynnik zgodności uporządkowania krajów pod względem częstości używania komputerów na lekcjach przyrody i średniej eksperymentowania w kształceniu przyrodniczym ($R = -0,41$). Składniki tej średniej i rozkłady procentowe uczniów w Polsce przedstawia Tabela 5.11. Ze średnią 9,6 Polska zajmuje 29. miejsce w międzynarodowym rankingu. Otwierają go Iran i Oman, a zamykają Belgia i Holandia. Ujemny współczynnik sugeruje, że w kształceniu przyrodniczym kraje dokonują wyboru: albo komputery, albo eksperymenty. Na komputery stawiają przede wszystkim szkoły w Danii, Nowej Zelandii i Gruzji. Na eksperymentowanie – szkoły w Iranie, Omanie i Korei. Polskie szkoły unikają obu tych skrajności, choć do pełnej równowagi potrzebny by był niewielki wzrost eksperymentowania.

Test TIMSS z bliska

Marcin Karpiński i Dominik Marszał

Test osiągnięć TIMSS 2015 został szczegółowo opisany w Rozdziale 1, ale jest jasne, że żaden opis nie zastąpi pokazu. Na dalszych stronach pokazujemy odtajniony Zeszyt 2 – jeden z 14 zeszytów testowych użytych w badaniu – w takiej formie, w jakiej dostały go dzieci. Pokazujemy go, po to żeby czytelnik mógł sobie wyrobić opinię o jakości pomiaru, a nauczyciele matematyki i przyrody w klasie czwartej mogli zmierzyć osiągnięcia swoich uczniów. Z tą myślą poprzedziliśmy zeszyt testowy kilkoma informacjami.

Po pierwsze, podajemy klucz punktowania odpowiedzi. W odniesieniu do pytań wymagających wyboru klucz mówi, która opcja daje jeden punkt. W odniesieniu do pytań wymagających krótkiej odpowiedzi klucz mówi, za co można przyznać punkty, a za co nie. Tabele 6.1 i 6.2 informują o procentach uczniów udzielających poprawnej odpowiedzi na poszczególne pytania w Polsce i łącznie we wszystkich krajach.

Po zsumowaniu punktów uzyskanych przez dziecko nauczyciel otrzyma tzw. wynik surowy. Wynik ten może być użyty do porównania osiągnięć dzieci w oddziale klasowym, np. orzekania, że Jan lepiej opanował matematykę niż Piotr. To jednak niewiele, chcielibyśmy bowiem wiedzieć, czy Jan – prymus w tym oddziale – pozostałby prymusem także w innym oddziale, a jeszcze lepiej – w międzynarodowej populacji uczniów. To bardzo ważne, by móc odnieść wyniki lokalne do globalnych, ponieważ dopiero wtedy można adekwatnie ocenić poziom kształcenia

w oddziale, szkole, gminie itd. Dlatego po drugie, określamy na lokalnej skali wyników surowych przedziały odpowiadające progom osiągnięć ustalonym na międzynarodowej skali wyników. W Tabeli 6.3 podajemy przedziały wyników surowych, odpowiadające im międzynarodowe oceny w skali od jedynki do piątki oraz procent naszych uczniów znajdujących się w każdym przedziale. Dzięki ostatniej z tych informacji nauczyciel może bezpośrednio porównać rozkład ocen w swoim oddziale z rozkładem ocen w populacji i stwierdzić np., że u niego jest więcej piątek i czwórek, ale i więcej jedynek niż w całym kraju, co bezpośrednio dostarcza wskazówki, na co powinien zwrócić uwagę w nauczaniu. Niezależnie od tego w Tabeli 6.4 podajemy podstawowe statystyki krajowego rozkładu wyników surowych: wynik minimalny i maksymalny, średnią wyników oraz odchylenie standardowe. Można z nimi porównać średnie i odchylenia standardowe wyników szkolnych.

Po trzecie, komentujemy te zadania z Zeszytu 2, w których wyniki polskich uczniów odbiegają *in plus* lub *in minus* od średniej międzynarodowej, po to by pobudzić krytyczną refleksję nad rozwiązaniami przyjętymi w polskiej dydaktyce przyrodoznawstwa i matematyki.

Klucz punktowania odpowiedzi

Klucz otwierają zasady ogólne. W badaniu TIMSS stanowią one, co następuje.

1. Przedmiotem pomiaru są wiadomości i umiejętności matematyczne i przyrodnicze, a nie poprawność pisemnego komunikowania myśli.
2. Ciężar dowodu, że uczeń opanował mierzone wiadomości i umiejętności, spoczywa na uczniu. Sprawdzającemu nie wolno domyślać się, co uczeń chciał powiedzieć w swojej odpowiedzi. W szczególności zakazane jest posługiwanie się ogólną opinią o wiedzy ucznia w ocenianiu konkretnych odpowiedzi. Punkty przyznaje się na podstawie samej odpowiedzi – tak jak jest wyrażona w zeszycie testowym.
3. Odpowiedź przekreśloną, startą, z obcymi znakami, nieczytelną lub niezwiązaną z zadaniem uznaje się za niepoprawną.
4. Jeśli odpowiedź ucznia zawiera zarazem elementy poprawne i niepoprawne lub nieistotne dla zadania, należy zignorować elementy niepoprawne lub nieistotne i oceniać tylko część poprawną. Od tej ogólnej zasady należy odstąpić, gdy:
 - elementy niepoprawne lub nieistotne są sprzeczne z elementami poprawnymi, w szczególności gdy uczeń podaje odpowiedź i jednocześnie odpowiedź przeciwną¹
 - elementy nieistotne wyraźnie świadczą o tym, że uczeń nie rozumie zagadnienia postawionego w zadaniu
 - elementy niepoprawne występują w końcowej części odpowiedzi i wyraźnie świadczą o tym, że uczeń zmienił zdanie z poprawnego na niepoprawne
 - poprawna odpowiedź została przekreślona.

Punktacja

Zadanie 1 Odpowiedź poprawna: A (1 punkt)

Zadanie 2 Odpowiedzi poprawne (1 punkt)

Dwie różne pozycje z poniższej listy.

- Powietrze / Tlen
- Pokarm
- Energia
- Woda
- Miejsce do życia / Habitat / Siedlisko / Biotop

Przykłady:

Powietrze. Pożywienie.

Daj im pożywienie. Daj im wodę.

Uwaga 1: Jeśli na miejscach 1 lub 2 jest więcej niż jedna pozycja, wszystkie muszą być poprawne.

Uwaga 2: Jeśli uczeń wymienia na miejscu 1 pozycje niezbędne roślinom, a na miejscu 2 – zwierzętom, wtedy wymagamy, żeby na miejscu 1 były przynaj-

mniej dwie pozycje niezbędne roślinom, na miejscu 2 przynajmniej dwie pozycje niezbędne zwierzętom i żeby na obu miejscach znalazły się przynajmniej dwie takie same pozycje. Na miejscu 1 akceptujemy światło słoneczne i dwutlenek węgla.

Przykład:

Rośliny potrzebują światła słonecznego, pożywienia i wody. Zwierzęta potrzebują pożywienia, powietrza i wody.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności:

- tylko jedna pozycja niezbędna roślinom i zwierzętom; drugiej nie ma, jest redundantna lub niepoprawna
- dwie lub więcej poprawnych pozycji niezbędnych roślinom i zwierzętom oraz jedna lub więcej niepoprawnych.

Przykład:

Schronienia. Kogoś, kto się o nie zatroszczy.

Zadanie 3 Odpowiedź poprawna: A (1 punkt)

Zadanie 4 Odpowiedź poprawna: C (1 punkt)

Zadanie 5 Odpowiedzi poprawne (1 punkt)

Zarazają zarazki, które osoba chora rozsiewa w powietrzu LUB które osiadają na przedmiotach.

Przykłady:

Wiatr roznosi choroby.

Zarazki rozchodzą się w powietrzu, którym oddychamy.

Grypa unosi się w otaczającym powietrzu.

Kiedy kaszlesz, zarazki fruują w powietrzu i dostają się do nosa i ust innych ludzi.

Zarazki przywierają do ciebie.

Bakterie się roznoszą.

Gdy ktoś kaszle, rozpyla zarazki w powietrzu.

Gdy ktoś chory kaszle na coś, a ty tego dotykasz, też możesz zachorować.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności nieodnoszące się do przeniesienia chorób drogą powietrzną.

Przykłady:

Można złapać chorobę, jeśli się jest blisko chorych ludzi.

Ponieważ choroba tej osoby jest zaraźliwa.

Ludzie nie zakrywają ust.

Zadanie 6 Odpowiedzi poprawne (1 punkt)

Drugi kwadracik ORAZ wyjaśnienie, że bieguny jednoimienne / takie same odpychają się (uznajemy odpowiedź, że bieguny różnoimienne przyciągają się).

Uwaga: Jeśli żaden kwadracik nie został zaznaczony, należy ocenić odpowiedź na podstawie uzasadnienia.

¹ W Bostonie mówią: *If you make me choose, you lose.*

Przykłady:

Drugi kwadracik. Magnesy z takimi samymi literami koło siebie odpychają się.

Drugi kwadracik. Południowy pcha południowy od siebie.

Drugi kwadracik. Zacieniowane końce odpychają się.

Drugi kwadracik. Jeśli zbliżyć różne końce magnesów, to będą się przyciągać.

Drugi kwadracik. Magnesy przyciągają się, gdy N i S są koło siebie.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności:

- Pierwszy kwadracik bez uzasadnienia lub z uzasadnieniem.
- Drugi kwadracik bez uzasadnienia lub z niepoprawnym uzasadnieniem.

Przykład:

Magnesy się odpychają, jeśli N i S są blisko.

Zadanie 7 Odpowiedzi poprawne (1 punkt)

Bateria jest źródłem energii / prądu niezbędnych do świecenia żarówki.

Przykłady:

Światło potrzebuje skądś energii, żeby świecić.

Światło potrzebuje elektryczności.

Dzięki baterii płynie prąd.

W baterii jest napięcie, które robi prąd.

Odpowiedzi niepoprawne (0 punktów)

Inne.

Przykłady:

Bateria jest potrzebna, żeby żarówka pracowała.

Bo trzeba zamknąć przełącznik.

Zadanie 8 Odpowiedź poprawna: D (1 punkt)

Zadanie 9 Odpowiedzi poprawne (1 punkt)

Metal jest dobrym przewodnikiem elektryczności.

Przykłady:

Prąd będzie w nich łatwo płynął.

Żeby mogły przewodzić elektryczność.

Metalowe druty przewodzą prąd.

Przewód z metalu może przenieść więcej prądu niż inne przewody.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności mówiąca, że metal jest ciągliwy, że metalowe druty są elastyczne lub że przenoszą moc, ale bez informacji, że metale przewodzą prąd.

Przykłady:

Metalowe druty przenoszą większą moc.

Drut można zginać i się nie złamie.

Przewody metalowe są trwalsze.

Druty z metalu nie zrywają się.

Zadanie 10A Odpowiedź poprawna: D (1 punkt)

Zadanie 10B Odpowiedź poprawna: C (1 punkt)

Zadanie 11 Odpowiedź poprawna: B (1 punkt)

Zadanie 12 Odpowiedzi poprawne (1 punkt)

Światło słoneczne jest zbyt jasne w porównaniu ze światłem gwiazd (wymagamy porównania).

Przykłady:

W dzień patrzymy na słońce, a światło słońca jest tak jasne, że nie widać gwiazd.

Słońce jest zbyt jasne.

Nie jest dość ciemno, żeby widzieć gwiazdy.

Odpowiedzi niepoprawne (0 punktów)

Inne.

Przykłady:

Z powodu światła słonecznego. [brak porównania]

Jest dzień. [brak porównania]

Niebo jest jasne w dzień. [brak porównania]

Gwiazdy gasną w dzień.

Słońce zakrywa gwiazdy w ciągu dnia.

Gwiazdy krążą wokół Ziemi.

Gwiazdy odbijają światło słoneczne.

Zadanie 13 Odpowiedź poprawna: C (1 punkt)

Zadanie 14 Odpowiedzi poprawne (1 punkt)

Dwie pozycje z poniższej listy:

- Pocenie się / utrata wody
- Zaczerwienienie twarzy
- Szybkie tętno / szybkie bicie serca / serce mocniej pompuje krew
- Szybszy oddech / utrata tchu
- Zawrót głowy
- Pragnienie / poczucie odwodnienia
- Głód
- Zmęczenie / zwalnianie
- Ból mięśni / bolące nogi
- Kolka / kurcz

Odpowiedzi niepoprawne (0 punktów)

Tylko jedna pozycja LUB inne.

Zadanie 15 Odpowiedzi poprawne (1 punkt)

Doniczki 1 i 3 ORAZ uzasadnienie, które mówi o użyciu roślin tego samego gatunku z nawożeniem i bez.

Przykłady:

1 i 3 ORAZ W eksperymencie można zmieniać tylko jedną zmienną. Doniczki 1 i 3 mają ten sam rodzaj roślin, ale tylko jedna roślina ma dodany nawóz.

1 i 3 ORAZ Doniczka 1i doniczka 3 mają ten sam typ kwiatów, ale doniczka 1 ma nawóz, a doniczka 3 nie ma i są porównywalne.

1 i 3 ORAZ Ponieważ mają te same kwiaty.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności poprawne numery doniczek bez wzmianki o tym samym gatunku roślin.

Przykłady:

1 i 3 ORAZ *Ponieważ doniczka 1 ma nawóz i Sara chce wiedzieć, więc powinna je porównać.*

1 i 3 ORAZ *Są takie same.*

LUB niepoprawne numery doniczek z nieodpowiednim uzasadnieniem.

Przykłady:

1 i 4 ORAZ *Ponieważ rośliny są takie same i jedna ma nawóz, a druga nie.*

Zadanie 16 Odpowiedź poprawna: A (1 punkt)

Zadanie 17 Odpowiedź poprawna: ABBA (1 punkt)

Zadanie 18 Odpowiedzi poprawne (1 punkt)

Nie ORAZ uzasadnienie z poniższej listy:

- W upieczonym cieście zaszły zmiany chemiczne.
- Ciasto zostało upieczone i nie można tego odwrócić.

Przykłady:

Nie może, bo tam jest przemiana chemiczna.

Bo ciasto zostało upieczone.

Jak raz je upieczesz, nie da się go odpiec.

LUB Nie ORAZ uzasadnienie, że sól się rozpuściła lub że stała się częścią ciasta (bez wyraźnej wzmianki o pieczeniu ciasta).

Przykłady:

Ponieważ sól już się rozpuściła w cieście.

Sól stała się częścią ciasta.

Uwaga: Jeśli żaden kwadracik nie został zaznaczony, należy ocenić odpowiedź na podstawie uzasadnienia.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności Tak niezależnie od uzasadnienia.

Przykłady:

Nie ORAZ *Sól już się zmieszała z ciastem.*

Nie ORAZ *Ciasto jest już zrobione.*

Zadanie 19 Odpowiedź poprawna: C (1 punkt)

Zadanie 20 Odpowiedź poprawna: BBABA (1 punkt)

Zadanie 21 Odpowiedzi poprawne (1 punkt)

Grawitacja (wprost lub domyślnie)

Przykłady:

Grawitacja

Siła ciężkości

Ziemia ściąga ją w dół.

Ziemia przyciąga ją.

Odpowiedzi niepoprawne (0 punktów)

Inne.

Przykłady:

Ziemia

Pochyłość

Zadanie 22 Odpowiedź poprawna: A (1 punkt)

Zadanie 23 Odpowiedź poprawna: D (1 punkt)

Zadanie 24 Odpowiedź poprawna: B (1 punkt)

Zadanie 25 Odpowiedzi poprawne (1 punkt)

8 z resztą 3 LUB $8\frac{3}{5}$, LUB równoważne.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności 8.

Zadanie 26 Odpowiedzi poprawne (2 punkty)

Poprawne obie części.

A. Dowolna liczba między 150 i 199 (włącznie)

B. Dowolna liczba między 450 i 549 (włącznie)

Odpowiedzi częściowo poprawne (1 punkt)

Tylko część A poprawna.

Tylko część B poprawna.

Odpowiedzi niepoprawne (0 punktów)

Inne.

Zadanie 27 Odpowiedź poprawna: C (1 punkt)

Zadanie 28 Odpowiedź poprawna: D (1 punkt)

Zadanie 29 Odpowiedź poprawna: A (1 punkt)

Zadanie 30 Odpowiedź poprawna (1 punkt)

34

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności 7.

Zadanie 31 Odpowiedź poprawna: C (1 punkt)

Zadanie 32A Odpowiedź poprawna (1 punkt)

37

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności 7 lub 47.

Zadanie 32B Odpowiedź poprawna: B (1 punkt)

Zadanie 33 Odpowiedź poprawna: C (1 punkt)

Zadanie 34 Odpowiedzi poprawne (2 punkty)

Poprawne obie części:

A. Zielona anakonda

B. Pyton birmański

Odpowiedzi częściowo poprawne (1 punkt)

Tylko część A poprawna

Tylko część B poprawna

Odpowiedzi niepoprawne (0 punktów)

Inne.

Zadanie 35 Odpowiedź poprawna (1 punkt)

4277

Odpowiedzi niepoprawne (0 punktów)

Inne

Zadanie 36 Odpowiedź poprawna (1 punkt)

4

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności 5 LUB 4 z resztą 2, LUB równoważne.

Zadanie 37 Odpowiedź poprawna (1 punkt)

3 godziny i 14 minut

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności 4 godziny i 46 minut lub 3 godziny i 54 minut.

Zadanie 38 Odpowiedź poprawna (2 punkty)

8 ORAZ 3

Odpowiedzi częściowo poprawne (1 punkt)

Tylko 8 LUB tylko 3

Odpowiedzi niepoprawne (0 punktów)

Inne.

Zadanie 39 Odpowiedź poprawna: B (1 punkt)

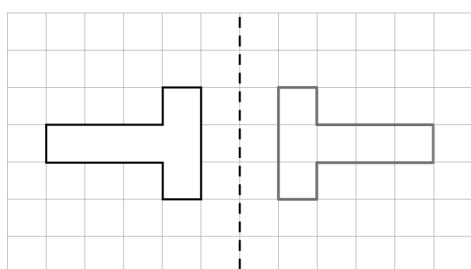
Zadanie 40 Odpowiedź poprawna: C (1 punkt)

Zadanie 41 Odpowiedź poprawna: A (1 punkt)

Zadanie 42 Odpowiedź poprawna: D (1 punkt)

Zadanie 43 Odpowiedzi poprawne (1 punkt)

Figura poprawnie odbita (z tolerancją $\pm \frac{1}{2}$ boku kratki)



Odpowiedzi niepoprawne (0 punktów)

Inna.

Zadanie 44 Odpowiedź poprawna: B (1 punkt)

Zadanie 45 Odpowiedzi poprawne (1 punkt)

Marzec ORAZ kwiecień ORAZ brak innego miesiąca.

Odpowiedzi niepoprawne (0 punktów)

Inne, w szczególności marzec LUB kwiecień ORAZ brak innego miesiąca.

Zadanie 46 Odpowiedź poprawna: C (1 punkt)

Wyniki zbiorów zadań Zeszytu 2

Tabele 6.1 i 6.2 składają się z dwóch części. Górna podaje procenty polskich uczniów wybierających poszczególne opcje w zadaniach wyboru oraz odpowiadających poprawnie i niepoprawnie na pytania krótkiej odpowiedzi. Tabele podają też procent uczniów, którzy opuścili każde pytanie. Podczas obliczania łącznego wyniku testowania pytaniom opuszczonym przypisuje się zero punktów, więc poprawność wchodząca do obliczeń jest zwykle niższa, niż wynikałoby z danych w siedmiu pierwszych kolumnach obu tabel (należy zauważyć, że w każdym wierszu procenty sumują się do 100). Dolna część każdej tabeli podaje takie same dane w odniesieniu do populacji międzynarodowej.

W Tabeli 6.3 znajdują się przedziały wyników surowych, odpowiadające im międzynarodowe oceny w skali od jedynki do piątki oraz procent naszych uczniów znajdujących się w każdym przedziale. Tabela 6.4 podaje podstawowe statystyki krajowego rozkładu wyników surowych: wynik minimalny i maksymalny, średnią wyników oraz odchylenie standardowe.

Komentarz dydaktyczny

Przyroda

Otwierająca rozdział Tabela 3.1 przedstawia osiągnięcia przyrodnicze sprowadzone do jednej liczby. Kolejne tabele w tym rozdziale dostarczają nieco bardziej szczegółowej wiedzy – o osiągnięciach w poszczególnych działach i w poszczególnych funkcjach szkolnej wiedzy przyrodniczej. To uszczegółowienie można kontynuować aż do poziomu zadań, choć tylko tych, które zostały odtajnione.

Rozwiązywanie problemów (zadania 10, 15, 17 i 18). Wśród wymienionych wyróżnia się zadanie 15. Inne sprawdzają, czy uczeń potrafi zastosować wiedzę przyrodniczą do wyjaśnienia danego zjawiska, to natomiast sprawdza, czy uczeń wie, jak zdobywa się tę wiedzę, czyli rozumie istotę eksperymentu przyrodniczego: kanon jedynej różnicy. Zadanie okazało się bardzo trudne. W populacji międzynarodowej w pełni poprawnie wykonało je tylko 22 proc. uczniów. W Polsce nieistotnie więcej: 24 proc. Niższe wyniki niż w Polsce odnotowano m.in. w Portugalii (12 proc.) i we Włoszech (13 proc.). Najlepiej poradzili sobie z zadaniem uczniowie singapurscy (59 proc.) i japońscy (49 proc.). W Europie wyższe wyniki niż

Tabela 6.1. Rozkłady procentowe odpowiedzi na pytania przyrodnicze Zeszytu 2

Numer zadania	Zadania wyboru			Zadania krótkiej odpowiedzi			
	A	B	C	D	Poprawne	Niepoprawne	Opuszczone
Wyniki polskie							
1	71,7	9,5	7,9	9,6			1,4
2					52,7	43,9	3,4
3	78,9	13,4	2,2	5,0			0,6
4	11,7	6,0	71,7	10,0			0,7
5					37,6	54,1	8,3
6					24,8	74,5	0,8
7					49,1	28,0	22,9
8	2,8	3,6	20,1	72,0			1,5
9					32,8	51,7	15,4
10A	4,5	8,2	14,5	71,5			1,3
10B	13,3	7,7	68,9	5,6			4,5
11	24,5	37,9	10,3	24,5			2,8
12					19,9	72,2	7,9
13	6,4	13,4	70,8	4,1			5,3
14					66,3	21,3	6,7
15					24,2	71,7	4,1
16	83,6	9,0	3,3	1,6			2,5
17					41,8	55,7	2,4
18					28,8	69,4	1,8
19	5,7	29,0	53,4	7,9			4,0
20					55,2	42,9	1,9
21					23,2	61,9	14,8
22	86,3	2,3	4,5	4,7			2,1
23	5,4	14,3	2,6	75,3			2,3
24	3,9	73,1	11,3	9,9			1,8
Wyniki międzynarodowe							
1	69,5	8,2	8,0	9,9			4,4
2					49,3	44,3	6,4
3	46,9	13,9	16,4	19,8			2,9
4	11,3	8,2	67,6	10,9			1,9
5					40,8	48,7	10,5
6					29,4	67,8	2,8
7					52,2	34,6	13,1
8	5,7	5,3	23,5	63,1			2,4
9					27,3	57,2	15,5
10A	5,7	7,0	20,8	63,7			2,8
10B	19,3	11,9	53,8	9,6			5,4
11	17,7	42,7	13,1	22,7			3,8
12					29,2	60,4	10,4
13	11,0	14,8	56,8	12,5			4,8
14					50,9	40,1	9
15					22,2	72,9	4,9
16	67,5	6,2	11,7	11,3			3,4
17					36,4	59,9	3,7
18					23,5	72,8	3,7
19	8,8	20,8	50,0	14,6			5,8
20					48,9	47,1	4,1
21					25,6	57,1	17,3
22	70,1	6,7	7,0	11,3			4,8
23	11,6	18,5	5,0	59,0			5,9
24	7,6	62,6	15,2	9,2			5,4

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

W zadaniach wyboru odpowiedź poprawną oznaczono drukiem pogrubionym. Odsetki ważone. Liczba polskich uczniów: 674

Tabela 6.2. Rozkłady procentowe odpowiedzi na pytania matematyczne Zeszytu 2

Numer zadania	Zadania wyboru				Zadania krótkiej odpowiedzi			
	A	B	C	D	Poprawne	Częściowo poprawne	Niepoprawne	Opuszczone
Wyniki polskie								
25					49,4		43,2	7,4
26					28,8	10,4	52,7	8,1
27	10,9	8,7	68,5	10,3				1,5
28	8,6	31,7	8,9	49,4				1,4
29	50,8	35,2	1,3	11,4				1,2
30					71,0		23,0	6,0
31	2,8	41,5	46,8	7,6				1,3
32A					42,1		47,1	10,8
32B	27,2	34,0	9,7	24,3				4,8
33	6,1	7,2	83,8	1,7				1,1
34					88,2		9,2	2,6
35					50,2		46,2	3,7
36					75,3		21,4	3,3
37					35,8		59,2	5,0
38					22,7	15,8	47,2	14,3
39	18,9	40,8	32,6	6,4				1,3
40	23,1	10,2	55,3	9,8				1,7
41	72,9	11,1	5,3	8,9				1,9
42	15,5	21,6	7,9	52,0				3,1
43					88,2		11,0	0,8
44	25,4	45,0	14,9	10,8				3,9
45					31,8		61,9	6,3
46	26,2	5,2	62,6	3,4				2,6
Wyniki międzynarodowe								
25					46,2		48,3	5,5
26					48,4	14,9	31,0	5,6
27	10,2	11,7	65,1	10,3				2,8
28	15,7	32,2	10,7	37,8				3,6
29	44,6	31,2	3,7	17,5				3,0
30					69,4		25,7	4,9
31	6,4	44,3	40,0	6,8				2,6
32A					38,1		53,2	8,7
32B	23,0	35,3	14,5	21,9				5,3
33	15,9	8,9	68,6	2,5				4,0
34					70,2		24,2	5,6
35					56,8		40,2	3,0
36					54,9		41,0	4,1
37					24,9		69,6	5,4
38					16,5	15,9	57,0	10,6
39	27,3	35,5	24,1	8,7				4,4
40	27,4	8,7	49,4	11,0				3,5
41	58,6	10,7	10,7	14,2				5,8
42	15,7	20,8	8,9	47,6				6,9
43					77,0		18,4	4,5
44	25,8	43,3	13,5	12,1				5,3
45					38,8		53,0	8,2
46	31,0	8,3	49,7	4,3				6,7

Źródło: IEA's Trends in International Mathematics and Science Study – TIMSS 2015

W zadaniach wyboru odpowiedź poprawną oznaczono drukiem pogrubionym. Odsetki ważone. Liczba polskich uczniów: 674

Tabela 6.3. Międzynarodowe i krajowe (określone na podstawie Zeszytu 2) poziomy osiągnięć

Nazwa poziomu	Przedział międzynarodowy	Przyroda		Matematyka	
		Przedział krajowy	Procent uczniów	Przedział krajowy	Procent uczniów
Jedynka	Poniżej 400	0–5	2	0–2	3
Dwójka	401–475	6–7	12	3–7	16
Trójka	476–550	8–12	35	8–12	38
Czwórka	551–625	13–18	41	13–19	35
Piątka	626 i więcej	19 i więcej	10	20 i więcej	8

Odsetki ważone. Liczba polskich uczniów: 337. Liczebność uczniów w tej tabeli jest o połowę mniejsza niż w tabelach 6.1 i 6.2. Jest tak dlatego, że przy próbkowaniu macierzowym to samo zadanie wchodzi w skład różnych zeszytów testowych.

polski uczniowie uzyskali m.in. dzieci z krajów skandynawskich: Szwecji (36 proc.), Finlandii (35 proc.) i Norwegii (31 proc.), w których od dawna kładzie się duży nacisk na nauczanie metodą badawczą. Metoda ta zakłada wprowadzanie dzieci w świat przyrody w sposób zgodny zarazem z ich naturalną ciekawością, jak i z procedurami stosowanymi w naukach przyrodniczych. W Polsce pojawia się w nauczaniu dość późno, bo dopiero od klasy czwartej (Poziomek i in., 2015). W krajach zachodnich cieszy się większą popularnością (Minner, Levy i Century, 2009).

W kilku zadaniach wyniki wydają się niespójne. Rozważmy parę 9 i 20. Oba zadania sprawdzają, czy uczeń wie, że wszystkie metale przewodzą prąd elektryczny. W pierwszym uczeń ma wykorzystać tę wiedzę do wyjaśnienia, dlaczego przewody elektryczne wykonuje się z metali, w drugim – wskazać przedmioty przewodzące prąd. Pierwsze okazało się znacznie trudniejsze niż drugie. Różnica, zarówno w Polsce, jak i na świecie, wyniosła 22 punkty procentowe.

Skąd wzięła się ta różnica? Nasuwają się dwie niekonkurencyjne odpowiedzi. Pierwsza zakłada, że czwartoklasistom, a może ludziom w ogóle, łatwiej przywołać z pamięci wiadomość niż posłużyć się

nią do rozwiązania problemu. Potwierdza to trafność klasyfikacji funkcji podmiotowych wprowadzonej w Rozdziale 1. Druga odpowiedź odwołuje się do formy zadania. Zadanie 9 to zadanie krótkiej odpowiedzi, natomiast zadanie 20 to zadanie wyboru. Powszechnie wiadomo, że pierwsza forma jest trudniejsza niż druga. Krótka odpowiedź ma najpierw postać myśli, którą następnie trzeba ubrać się w słowa. Dla czwartoklasisty wyrażenie prostej myśli na piśmie nie stanowi trudności, zwłaszcza że oceniający tolerowali nawet poważne usterki języka i zapisu. Trudność tkwi w samej myśli. Żeby wpaść na właściwą, trzeba dysponować głębszą i lepiej skoordynowaną wiedzą. Nic dziwnego, że diagnostyczność zadania 9 (korelacja z wynikiem sumarycznym $r = 0,43$) jest znacznie większa niż zadania 20 (0,15). Potwierdza to przekonanie, które dominuje wśród polskich badaczy osiągnięć szkolnych: test złożony z samych zadań wyboru jest wprawdzie wielokrotnie tańszy, ale też mniej miarodajny niż test zawierający zadania otwarte.

Umiejętności dziewczynek i chłopców (zadania 7, 9, 20 i 23). Dane przytoczone w Rozdziale 3 świadczą, że osiągnięcia polskich dziewczynek w 2015 r. nie są niższe niż chłopców. Nie oznacza to oczywiście, że jest tak w każdym zadaniu. W opisanym wyżej zadaniu 9 odsetek chłopców, którzy udzielili poprawnej odpowiedzi przewyższył odsetek dziewczynek aż o 18 punktów procentowych, natomiast w zadaniu 23 to dziewczynki przewyższyły o 10 punktów chłopców.

Pewną wskazówkę stanowi fakt, że pierwsze zadanie odnosi się do (paleo)biologii, a drugi do fizyki. Być może zagadnienia życia są bliższe zainteresowaniom dziewczynek, a problematyka prądu elektrycznego bardziej interesuje chłopców. Potwierdzałyby to również 12-punktowa przewaga chłopców w zadaniu 7. Ale nie wszędzie tak bywa. Przykładem jest zadanie 20, w którym w Polsce lepsze wyniki osiągnęli chłopcy, natomiast we wszystkich badanych krajach – dziewczynki.

Tabela 6.4. Podstawowe informacje o zbiorze zadań Zeszytu 2

Statystyka	Przyroda	Matematyka
Liczba pytań	25	23
Liczba badanych	337	337
Minimum	2	2
Maksimum	23	25
Mediana	13	13
Średnia	13,2	13,5
Odchylenie standardowe	4,0	5,2
Korelacja z wynikiem θ całego testu (r)	0,89	0,94

Dane nieważone

Zadania najtrudniejsze dla polskich uczniów (5, 6, 7, 11 i 12). W pięciu zadaniach Zeszytu 2 polscy uczniowie osiągnęli wynik poniżej średniej międzynarodowej. Największą różnicę (o dziewięć punktów procentowych) notujemy w zadaniu 12, które dotyczy widoczności ciał niebieskich na dziennym nieboskłonie. Zasadnicza trudność zadania polega na porównaniu jasności ciał niebieskich; samo stwierdzenie dużej jasności Słońca nie wystarcza. Powodem niskiego wyniku polskich uczniów jest zapewne to, że w polskiej podstawie programowej zagadnienia astronomiczne ograniczają się do elementarnej znajomości Układu Słonecznego i następstw ruchów Ziemi. Ponadto w większości programów nauczania przyrody pojawiają się one dopiero w klasie szóstej.

Podobne wyjaśnienie można zaoferować w odniesieniu do zadań 11, 6 i 7. W pierwszym, które wymaga wybrania opisu działania erozji eolicznej jako czynnika, który wpłynął na kształt zaprezentowanej formy skalnej, polscy uczniowie osiągnęli rezultat niższy od średniej międzynarodowej o pięć punktów procentowych. Nic dziwnego: zagadnienia związane z erozją wiatrową są bardzo rzadko poruszane w polskich programach nauczania. W zadaniu 6 rezultat polskich uczniów był niższy o pięć punktów, a w zadaniu 7 – o trzy punkty procentowe niż w populacji międzynarodowej. Pierwsze zadanie dotyczy magnetyzmu i wymaga wiedzy o oddziaływaniu na siebie biegunów jednoimiennych. Drugie dotyczy obwodów elektrycznych i wymaga zrozumienia, że źródłem energii potrzebnej do świecenia żarówki jest bateria. Oba zadania nawiązują do podstawowych zagadnień obecnych w każdym polskim programie nauczania przyrody, jednak umieszczanych zdecydowanie częściej w klasach piątej i szóstej niż w klasie czwartej.

Na koniec zadanie, którego wyników tak wyjaśnić nie można. Chodzi o zadanie 5. Dotyczy ono dróg szerzenia się chorób i wymaga wiedzy o zakażeniu kropelkowym. Wynik polskich uczniów był o trzy punkty procentowe gorszy od średniej międzynarodowej. Jest to wynik niepokojący, ponieważ wiele programów nauczania już w klasie czwartej wprowadza treść wymaganą w tym zadaniu, a profilaktyka chorób zakaźnych jest również obecna w treści programów pierwszego etapu kształcenia.

Matematyka

Opisane w Rozdziale 2 wyniki testu matematycznego pozwalają na wyciągnięcie wielu wniosków o umiejętnościach matematycznych uczniów z badanych krajów, w tym z Polski. Jeśli jednak chcemy się dowiedzieć, co w nauczaniu matematyki w Polsce należało-

by zmienić, to powinniśmy przyjrzeć się także wynikom poszczególnych zadań. Pokazują one dokładniej, które umiejętności polscy uczniowie opanowali lepiej niż uczniowie z innych krajów, które – gorzej, a które opanowują po prostu inaczej lub w innym tempie. Większość zadań użytych w badaniu pozostaje oczywiście niejawna, ale nawet nieliczne, które możemy pokazać, dają do myślenia.

Matematyka jest dziedziną ścisłą, ale nie sztywną, by trzeba było jej nauczać tylko w jeden sposób. Każdy kraj ma swoją podstawę programową i tradycję nauczania tego przedmiotu. Nawet jeśli programy w dwóch krajach są sformułowane podobnie, nie oznacza to, że matematyki uczy się tam tak samo. Na praktykę nauczania wpływa wiele czynników – ogólnych, takich jak kultura, warunki społeczne i stosunek do edukacji, oraz swoistych, związanych z dydaktyką przedmiotu. Za sprawą odmiennych koncepcji dydaktycznych te same zagadnienia mogą być wprowadzane w różnych klasach i w różnej kolejności. Pojęcie lub własność matematyczna może być intensywnie nauczana w krótkim czasie lub rozciągnięta na kilka klas, problemy i zadania matematyczne typowe w jednym kraju bywają nietypowe w innym, metody nauczania mogą nawiązywać do różnych postaw itd. Wszystko to znajduje odbicie w wynikach poszczególnych zadań. Przyjrzyjmy się im bliżej.

Umiejętności rachunkowe (zadania 35 i 36). Na przykładzie tych kilku zadań rachunkowych wyraźnie widać, jak różnie podchodzi się do nauczania rachunków w różnych krajach. W zadaniu 35 trzeba było wykonać odejmowanie $4\ 809 - 532$. Dla uczniów większości krajów było to zaproszenie do wykorzystania algorytmu odejmowania pisemnego. Polscy uczniowie zaczynają się uczyć tego algorytmu stosunkowo późno, bo dopiero w klasie czwartej, a i w wyższych klasach szkoły podstawowej nie uważa się go za kluczowy warunek sprawności rachunkowej i innych celów nauczania matematyki opisanych w podstawie programowej. Podstawa zakłada, że dużo ważniejsze są umiejętności działań pamięciowych i zrozumienie własności działań.

W zadaniu 35 polscy uczniowie wypadli słabo, ich wynik (50 proc.) był o siedem punktów procentowych gorszy niż średnia wszystkich badanych krajów. Podobnie lub jeszcze słabiej wypadły takie kraje jak Kanada (51 proc.), Holandia (48 proc.) czy Finlandia (42 proc.). Podobnie jak w Polsce, działania pisemne nie odgrywają w tych krajach – słusznie uważanych za wyróżniające się jakością edukacji matematycznej – kluczowej roli w nauczaniu matematyki. Nie znaczy to oczywiście, że im niższy jest wynik w zadaniu 35, tym wyższy poziom nauczania matematyki.

W zadaniu 36 trzeba było nie tylko dobrać odpowiednie działania i wykonać obliczenia, ale też zinterpretować rezultat. Tym razem polscy uczniowie osiągnęli znakomity wynik, o 20 punktów procentowych powyżej średniej międzynarodowej. W Europie nikt nas nie przewyższył, lepsze były tylko kraje azjatyckiej piątki (Hong Kong, Tajwan, Japonia, Korea Południowa i Singapur).

Obliczenia zegarowe (zadania 33 i 37). W zadaniu 33 należało wykazać się znajomością zegara analogowego, a w zadaniu 37 – wykonać obliczenia na cyfrowym zapisie czasu. Oba zadania są dobrym przykładem, jak różnie traktuje się tego typu obliczenia w różnych krajach. W zadaniu 33 polscy uczniowie z wynikiem 84 proc. okazali się najlepsi w Europie, przewyższając o ponad 10 punktów procentowych Niemców czy Francuzów. Na świecie wyprzedzili ich tylko uczniowie z krajów azjatyckich.

W zadaniu 37 należało obliczyć, ile czasu upłynie między godziną 7.52 a 11.06. Polscy uczniowie wypadli nieźle, choć już nie tak dobrze jak w poprzednim. Ich wynik jest prawie o 11 punktów procentowych lepszy od średniej międzynarodowej. Ciekawe, że to zadanie sprawiło dużą trudność uczniom z Japonii i Finlandii – krajów uznawanych za wzorowo nauczające matematyki. Japońscy uczniowie rozwiązali to zadanie nieistotnie gorzej od polskich, co jednak na tle ich wyników w pozostałych zadaniach stanowi istotne zaskoczenie. W Finlandii poprawnej odpowiedzi udzieliło tylko 18 proc. uczniów, niemal dwa razy mniej niż w Polsce. Można przypuszczać, że w obu tych krajach obliczenia zegarowe pojawiają się w wyższych klasach niż w Polsce i że nie ma to większego wpływu na rozwój innych umiejętności matematycznych japońskich i fińskich uczniów.

Rozwiązywanie problemów (zadania 28 i 38). W zadaniach wymagających rozumowania polscy uczniowie wypadli bardzo dobrze. Pokazują to ogólne wyniki opisane w Rozdziale 2. Z jakiego typu problemami matematycznymi dobrze sobie poradzili, możemy zobaczyć na przykładzie dwóch zadań.

W zadaniu 28 trzeba było przeprowadzić rozumowanie związane z obliczeniami. Istotne przy tym było wyłuskanie najistotniejszych informacji podanych w tekście zadania. Poprawną odpowiedź podało 49 proc. polskich uczniów – o 12 punktów procentowych więcej niż w populacji międzynarodowej. W Europie lepiej z tym zadaniem poradzili sobie tylko uczniowie duńscy i litewscy.

Zadanie 38 jest najtrudniejsze w całym zeszycie. Trzeba w nim przeprowadzić rozumowanie niedające się ująć w jednym zapisie działań, podobne do metod stosowanych później w algebrze. Średni wynik

wszystkich krajów to 17 proc. W Polsce to 22 proc. W Europie więcej poprawnych odpowiedzi udzieliли uczniowie tylko w czterech krajach: Szwecji (30 proc.), Finlandii (29 proc.), na Cyprze (24 proc.) i w Norwegii (23 proc.).

Umiejętności dziewczynki i chłopców (zadania 42 i 46). Wyniki polskich dziewczynki i chłopców w całym teście są statystycznie takie same. Nie oznacza to, że jest tak w każdym zadaniu. Są zadania, w których znacznie lepiej wypadają dziewczynki, i zadania, w których lepsi okazali się chłopcy.

Na przykład dziewczynki lepiej niż chłopcy rozwiązały zadanie 42. Odsetek poprawnych odpowiedzi wśród dziewczynki był wyższy o 8 punktów procentowych. W zadaniu trzeba było wskazać własność narysowanych trójkątów prostokątnych, przy czym istotną była umiejętność uważnego przeczytania polecenia. Ale w zadaniu 46, w którym należało zinterpretować i połączyć informacje podane w tekście zadania i na diagramie kołowym, to chłopcy przewyższyli dziewczynki o 6 punktów procentowych.

Zadań, które dziewczynki w Polsce i w innych krajach wykonywały inaczej niż chłopcy, było znacznie więcej. Często zdarzało się, że to, co w jednym kraju było łatwiejsze dla dziewczynki, w innym okazywało się łatwiejsze dla chłopca. Na podstawie badania TIMSS nie sposób dociec, jakie są ogólne czynniki tej statystycznej interakcji.

Zadania najtrudniejsze dla polskich uczniów (32B, 35, 26 i 45). Tylko w czterech zadaniach z Zeszytu 2 polscy uczniowie osiągnęli wynik poniżej średniej międzynarodowej, w tym w jednym (32B) nieistotnie. Zadanie 35 zostało już omówione. Zadanie 26 wymagało głębszego zrozumienia własności zaokrąglanych liczb. Wynik niższy od średniej międzynarodowej o 20 punktów procentowych świadczy, że polscy czwartoklasiści jeszcze tego nie potrafią. Nic dziwnego: w typowym programie nauczania matematyki zaokrąglanie liczb występuje we wszystkich klasach drugiego etapu kształcenia, ale w klasie czwartej uczniowie zdobywają tylko podstawowe umiejętności z tego zakresu.

W zadaniu 45 wymagającym kojarzenia danych w tabeli i na diagramie polscy czwartoklasiści uzyskali wynik o 7 punktów procentowych niższy od średniej międzynarodowej i od wszystkich krajów europejskich. W Polsce umiejętności związane ze statystyką opisową kształtuje się już od klasy pierwszej, ale w okresie edukacji wczesnoszkolnej poświęca się im mniej uwagi niż w większości innych krajów. Nauczanie statystyki opisowej ciągnie się u nas aż do klasy szóstej, a nawet do początkowych klas gimnazjum. Czwartoklasiści będą mieli wiele okazji do nadrobienia braków.

Nota prawna

Materiał zamieszczony na następnych stronach jest chroniony prawem autorskim. Copyright © 2015 by International Association for the Evaluation of Educational Achievement (IEA). Wydawca: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. Komercyjne wykorzystanie, rozpowszechnianie, reprodukcja lub przekazywanie w jakiegokolwiek formie i jakiegokolwiek środkami, włącznie z elektronicznymi lub mechanicznymi, całości lub części udostępnionych zadań lub ich przekładów na inne języki jest dozwolone wyłącznie za pisemną zgodą IEA.

Wskazówki

W tym sprawdzianie będziesz odpowiadać na pytania z matematyki i przyrody. Niektóre pytania mogą być łatwe, inne trudniejsze. Staraj się odpowiedzieć na wszystkie pytania – i na łatwe, i na trudniejsze.

Przy niektórych pytaniach są podane różne odpowiedzi. Musisz wybrać jedną, którą uważasz za poprawną, i zamalować kółko obok niej. Pierwszy przykład pokazuje takie pytanie z zamalowanym kółkiem obok poprawnej odpowiedzi.

Z ilu minut składa się godzina?

- (A) 12
- (B) 24
- (C) 60
- (D) 120

Przykład 1

Kółko z literą „C” jest zamalowane, ponieważ godzina składa się z 60 minut. Jeśli nie wiesz, jak odpowiedzieć na pytanie, zakreśl kółko obok odpowiedzi, która wydaje Ci się najlepsza, i przejdź do następnego pytania.

Jeśli zechcesz zmienić swoją odpowiedź, przekreśl ją tak: ~~⊗~~. Potem zamaluj kółko przy nowej odpowiedzi. Drugi przykład pokazuje, jak to zrobić.

Z ilu minut składa się godzina?

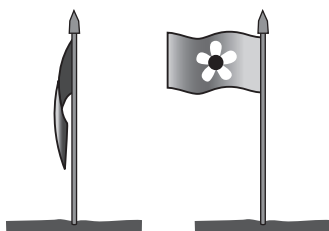
- ~~⊗~~ (A) 12
- (B) 24
- (C) 60
- (D) 120

Przykład 2

Wskazówki (ciąg dalszy)

Na inne pytania musisz napisać odpowiedź pod pytaniem. Mogą to być słowa, rysunki lub liczby. Trzeci przykład pokazuje takie pytanie.

W ogrodzie Janka znajduje się flaga. Czasem flaga wisi na maszcie, a czasem powiewa, jak widać na rysunku.



Co sprawia, że flaga powiewa?

To wiatr porusza flagę.

Trzeci przykład

Żeby dostać maksymalną liczbę punktów za niektóre pytania, musisz podać odpowiedź i uzasadnić ją albo przedstawić rysunek lub obliczenia. Pisz starannie. Na pytania z matematyki odpowiadaj w najprostszym sposobie. Przy pytaniach, w których mówi się o pieniądzach, wyobraź sobie, że jesteś w Zedlandii – kraju, w którym ludzie płacą „zedami”, tak jak w Polsce złotymi.

Gdy napiszesz odpowiedź, sprawdź, czy Twoje pismo da się przeczytać. Dobrze się zastanów nad każdym pytaniem i podawaj pełne odpowiedzi. Jeśli nie wiesz, jak odpowiedzieć na pytanie, napisz odpowiedź, która wydaje Ci się najlepsza, i przejdź do następnego pytania.

Będziesz mieć 36 minut na odpowiedzi na pierwszą część pytań. Potem będzie krótka przerwa. Po przerwie będziesz pracować przez następne 36 minut nad drugą częścią pytań.

Podczas sprawdzianu nie wolno używać kalkulatora.

Instrukcja do części 1

Przeczytaj uważnie każde pytanie i odpowiedz najlepiej, jak potrafisz.
Jeśli dokładnie nie wiesz, jak odpowiedzieć na pytanie, wybierz lub napisz
odpowiedź, która wydaje Ci się najlepsza, i przejdź do następnego pytania.

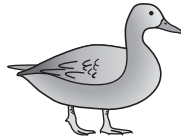
Masz 36 minut na rozwiązanie tej części sprawdzianu.

Nie zaczynaj pracy, dopóki Ci nie powiemy.

1

W której ramce są dwa zwierzęta znoszące jaja?

Ⓐ



kaczka



żaba

Ⓑ

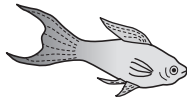


struś



królik

Ⓒ



złota rybka

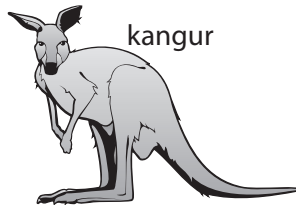


nietoperz

Ⓓ



wróbel



kangur

2

Rośliny i zwierzęta to organizmy.

Pomyśl, czego potrzebują do życia **zarówno** rośliny, jak i zwierzęta. Podaj dwie takie rzeczy.

1.

2.

S061010

3

Który z niżej wymienionych pokarmów zawiera dużo białka?

- (A) Ryba.
- (B) Chleb.
- (C) Świeże owoce.
- (D) Warzywa zielone.

S061028

4

Który z poniższych środków transportu uczniów do szkoły powoduje **najmniejsze** zanieczyszczenie powietrza?

- Ⓐ Autobus na benzynę.
- Ⓑ Samochód na benzynę.
- Ⓒ Samochód elektryczny.
- Ⓓ Pociąg na ropę.

S061065

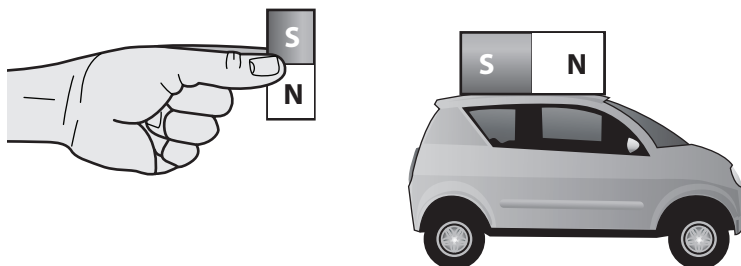
5

W jaki sposób możesz zarazić się chorobą (taką jak grypa) od kogoś, kto kaszle w Twojej obecności, chociaż nawet go nie dotkniesz?

S061130

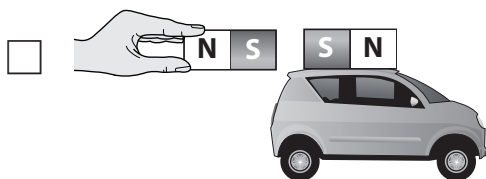
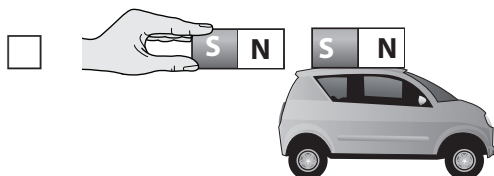
6

Magnes jest przyklejony do dachu plastikowego samochodziku. Ola chce odepchnąć samochodzik za pomocą innego magnesu.



W jaki sposób Ola powinna trzymać magnes, żeby odepchnąć samochodzik?

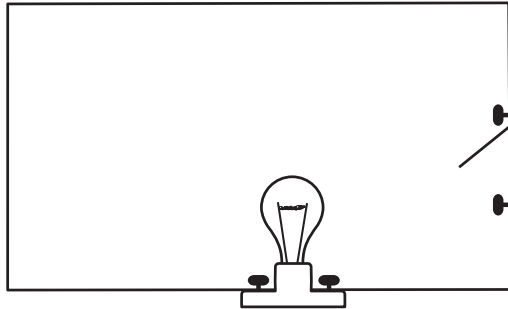
(Zaznacz jeden kwadracik)



Uzasadnij odpowiedź.

7

Bartek zbudował obwód z żarówką i przełącznikiem, jak widać na rysunku.



Kiedy Bartek zamyka obwód za pomocą przełącznika, żarówka nie świeci. Potem Bartek wstawia do obwodu baterię i żarówka świeci.

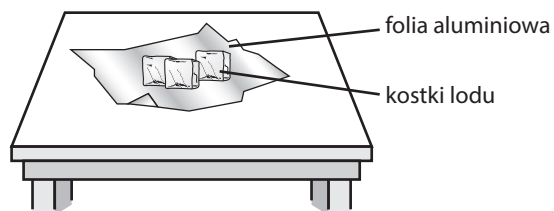
Wyjaśnij, dlaczego włączenie baterii do obwodu pozwala żarówce świecić.

S02_07



8

Stefan kładzie kostki lodu na folii aluminiowej rozłożonej na stole. Po jakimś czasie kostki lodu się topią.



Jaki jest powód tej przemiany?

- Ⓐ Stół pobiera ciepło z kostek lodu.
- Ⓑ Powietrze pobiera ciepło ze stołu.
- Ⓒ Folia aluminiowa pobiera ciepło z kostek lodu.
- Ⓓ Kostki lodu pobierają ciepło z powietrza.

S061075

9

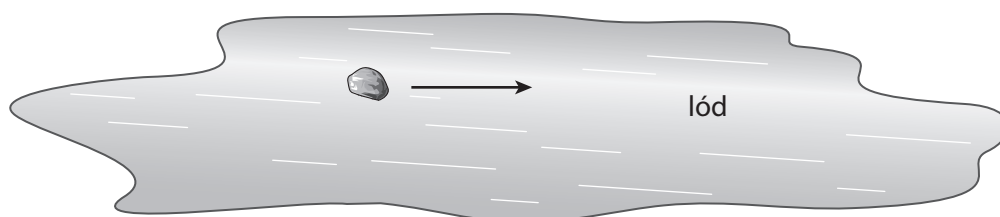
Dlaczego przewody elektryczne robi się z metalu?



S061031

10

Rysunek poniżej ukazuje gładki kamień sunący po lodzie w kierunku wskazanym przez strzałkę.



A. Jeśli Łukasz chce, żeby kamień wrócił tą samą drogą, to w jakim kierunku powinien przyłożyć siłę?

(A) 

(B) 

(C) 

(D) 

B. Jeśli kamień sunie w kierunku wskazanym przez strzałkę, a Łukasz chce, żeby sunął po skosie w prawo w dół, to w jakim kierunku powinien przyłożyć siłę?

(A) 

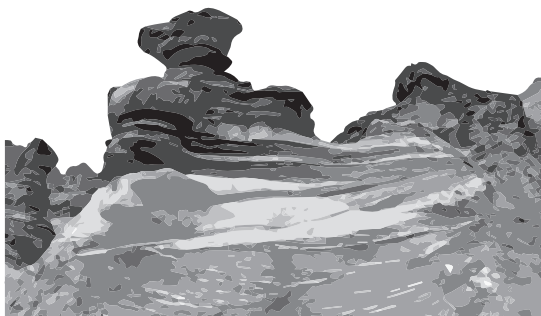
(B) 

(C) 

(D) 

11

To jest naturalna formacja skalna występująca na pustyni.



W jaki sposób wiatr wpłynął na uformowanie tych skał?

- (A) Wiatr ochładzał skały, które następnie pękały.
- (B) Wiatr niósł ze sobą ziarenka piasku, które tarły o skałę.
- (C) Wiatr na pustyni był gorący i topił skały.
- (D) Wiatr niósł ze sobą wilgoć, która wymywała skały.

S061098

12

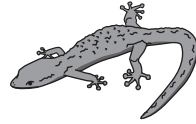
Dlaczego w ciągu dnia gwiazdy są niewidoczne?

S061172

13

Które z tych zwierząt jest ssakiem?

Ⓐ



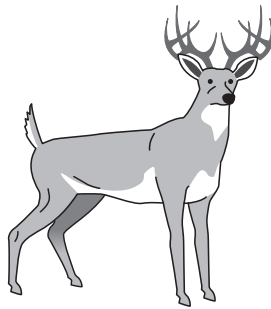
jaszczurka

Ⓑ



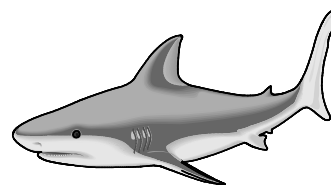
pingwin

Ⓒ



jeleń

Ⓓ



rekin

S051041

14

Gdy Paulina biegnie, to czuje, że w jej ciele zachodzą różne zmiany. Jedną zmianą jest to, że Paulinie robi się gorąco.

Podaj dwie inne zmiany, które może odczuć Paulina.

1.

2.





S051037



15

Sara chce się przekonać, czy nawożenie ma wpływ na wzrost roślin.

Sara napełniła cztery doniczki taką samą ziemią i posadziła w nich rośliny. Do dwóch doniczek dodała nawóz, jak widać na rysunku.

			
Doniczka 1	Doniczka 2	Doniczka 3	Doniczka 4
z nawozem	z nawozem	bez nawozu	bez nawozu

Które dwie doniczki Sara powinna porównać, żeby się przekonać, czy nawożenie ma wpływ na wzrost roślin?

Doniczkę _____ i doniczkę _____.

Uzasadnij odpowiedź.

S051008

16

Upierzenie niektórych ptaków ma kolor podobny do koloru otoczenia.

W jaki sposób pomaga to przetrwać tym ptakom?

- (A) Ukrywa je przed niebezpieczeństwem.
- (B) Pozwala łatwiej wypatrzeć pożywienie.
- (C) Daje osłonę przed niepogodą.
- (D) Pozwala łatwiej się spotkać.

S051004

17

W łańcuchu pokarmowym **drapieżnik** to zwierzę, które zjada inne zwierzę. Zwierzę, które jest zjadane, nazywamy **ofiarą**.

Które zdania o drapieżnikach i ofiarach są prawdziwe, a które nieprawdziwe?

Zamaluj jedno kółko przy każdym zdaniu.

Prawda Nieprawda

Zwierzę, które ma ostre zęby, jest zwykle drapieżnikiem. ----- (A) ----- (B)

Drapieżniki są zawsze większe od swoich ofiar. ----- (A) ----- (B)

Duże zwierzę nie może być ofiarą. ----- (A) ----- (B)

Niektóre zwierzęta mogą być zarówno drapieżnikami, jak i ofiarami. ----- (A) ----- (B)

S051026

18

Ania i Jacek pomagali w przygotowaniu przyjęcia urodzinowego. Upiekli ciasto, ale przez pomyłkę zamiast cukru użyli soli. Tuż przed przyjęciem Jacek zjadł kawałek ciasta i stwierdził, że jest słone.

Czy Jacek może usunąć sól z ciasta i zastąpić ją cukrem?

(Zaznacz jeden kwadracik)

Tak

Nie

Uzasadnij odpowiedź.

S051130



19

Cztery przedmioty podzielono na dwie grupy.

Grupa 1	 szklanka	 plastikowe okulary
Grupa 2	 metalowa łyżka	 drewniany talerz

Jaką cechę wzięto pod uwagę, dzieląc przedmioty na te grupy?

- (A) Uginanie się.
- (B) Unoszenie się na wodzie.
- (C) Przepuszczanie światła.
- (D) Przyciąganie przez magnes.

20

Które przedmioty z pokazanych poniżej przewodzą prąd elektryczny?
Zamaluj jedno kółko przy każdym przedmiocie.

Czy przewodzi prąd?

Tak

Nie



Drewniana łyżka ----- (A) ----- (B)



Plastikowy grzebień ----- (A) ----- (B)



Srebrny łańcuszek ----- (A) ----- (B)



Gumowa piłka ----- (A) ----- (B)

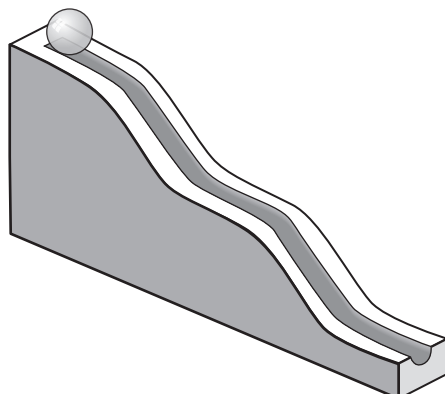


Żelazny klucz ----- (A) ----- (B)



21

Marek położył szklaną kulkę na górze stromego toru, jak widać na rysunku.



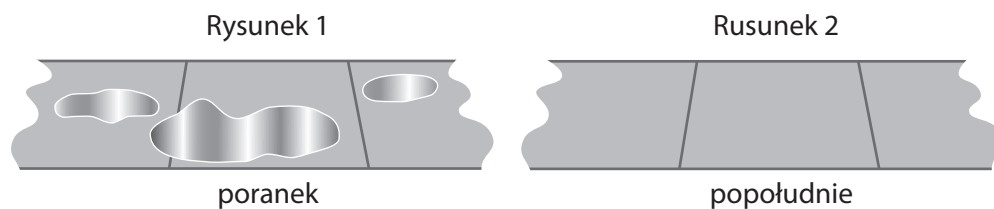
Kulka stacza się po torze.

Podaj nazwę siły, która porusza kulkę.

S051147

22

Rysunek 1 przedstawia kałuże na betonowym chodniku wczesnym rankiem. Po południu ten chodnik jest suchy, jak widać na rysunku 2.



Co stało się z wodą?

- (A) Uniosła się w powietrze.
- (B) Zmieniła się w kurz.
- (C) Zużyły ją drzewa.
- (D) Rozlała się na jezdnię.

S051105

23

Uczony znajduje skamieliny skorupiaków w skałach leżących na ziemi.
Co mówi uczonemu to znalezisko?

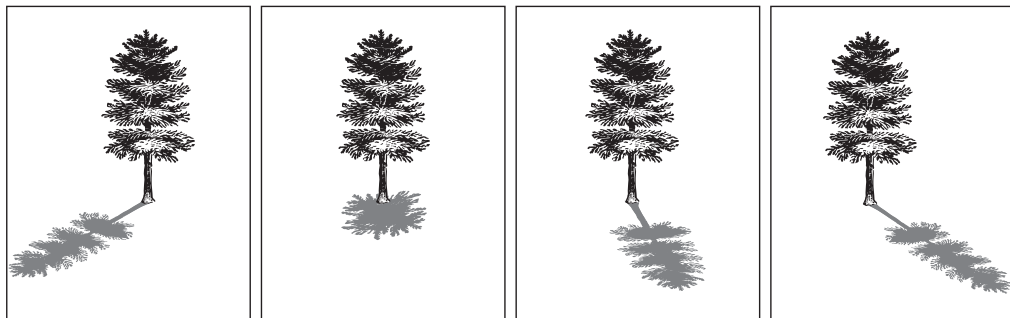


- (A) W tej okolicy żyły różne gatunki roślin.
- (B) Skorupiaki były stare, gdy umierały.
- (C) Skorupiaki zjadały inne skorupiaki.
- (D) Ta okolica była kiedyś pod wodą.

S051110

24

Dzieci obserwują cień drzewa w różnych porach dnia.
Który cień widzą w południe?



(A)

(B)

(C)

(D)

S051111



**Nie zaczynaj rozwiązywać części 2,
dopóki Ci nie powiemy.**

Jeśli skończyłaś/skończyłeś część 1 sprawdzianu przed czasem, przejrzyj i sprawdź swoje odpowiedzi.

Instrukcja do części 2

W drugiej części sprawdzianu będziesz odpowiadać na inne pytania z matematyki lub przyrody. Masz 36 minut na rozwiązanie tej części sprawdzianu.

Czytaj uważnie każde pytanie i odpowiadaj najlepiej, jak potrafisz. Jeśli nie wiesz, jak odpowiedzieć na pytanie, wybierz lub napisz odpowiedź, która wydaje Ci się najlepsza, i przejdź do następnego pytania.

Nie zaczynaj pracy, dopóki Ci nie powiemy.

25

$43 : 5 =$

Odpowiedź: _____

M02_01

M061272

26

Jurek zaokrągla liczby do setek.

A. Napisz Jurkowi liczbę mniejszą od 200, która po zaokrągleniu będzie równa 200.

Odpowiedź: _____

B. Napisz liczbę inną niż 500 i większą od 200, która po zaokrągleniu będzie równa 500.

Odpowiedź: _____

M02_02

M061243

27

Basia podała następujące informacje o pewnej 4-cyfrowej liczbie:

Cyfra setek to 7.

Cyfra tysięcy jest większa od cyfry setek.

Cyfra jedności jest mniejsza niż cyfra setek.

Jaka to liczba?

- Ⓐ 2708
- Ⓑ 4733
- Ⓒ 8726
- Ⓓ 9718

28

Sylwia ma 12 drucików, 40 okrągłych koralików i 48 płaskich koralików.

Na zrobienie jednej bransoletki Sylwia potrzebuje 1 drucik, 10 okrągłych koralików i 8 płaskich koralików.

Wszystkie bransoletki, które robi Sylwia, są takie same. Ile bransoletek może zrobić?

- (A) 40
- (B) 12
- (C) 5
- (D) 4

M061031

29

Jaką liczbą należy zastąpić \triangle , żeby równość była prawdziwa?

$$6 + 15 = \triangle + 10$$

- (A) 11
- (B) 21
- (C) 25
- (D) 31

M061050



30

Hania zaczyna pisać liczby według pewnej reguły:

6, 13, 20, 27, ...

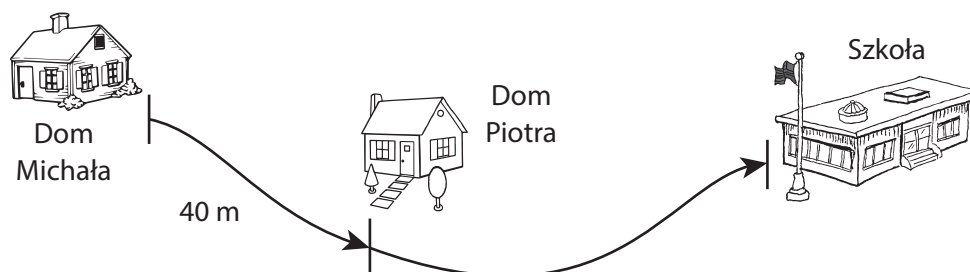
Żeby otrzymać kolejną liczbę, do ostatniej zapisanej liczby dodaje pewną liczbę, zawsze tę samą.

Jaka jest następna liczba, którą Hania powinna zapisać zgodnie z tą regułą?

Odpowiedź: _____

M061167

31



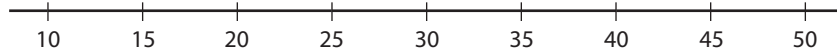
Ścieżka z domu Michała do domu Piotra ma 40 metrów. Ścieżka prowadzi dalej do szkoły.

Jak długa jest ścieżka z domu Michała do szkoły?

- (A) 40 m
- (B) 80 m
- (C) 100 m
- (D) 130 m

M061206

Ela i Kuba grają w pewną grę na osi liczbowej. W tej grze w jednym ruchu można się przesunąć tylko w prawo lub tylko w lewo.



- A. Ela zaczęła od 27, następnie przesunęła się o 10 jednostek. Skończyła na 17. W jakim innym punkcie mogłaby skończyć?

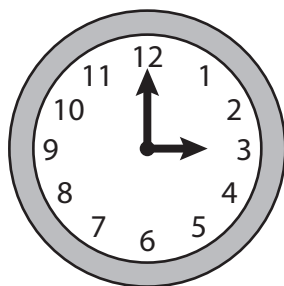
Odpowiedź: _____

- B. Kuba zaczął od 35 i przesunął się o 13 jednostek w lewo. W następnym ruchu przesunął się jeszcze o 2 jednostki. W którym punkcie mógł skończyć?

- (A) 22
- (B) 24
- (C) 48
- (D) 50



33



O godzinie 3:00 wskazówki zegara tworzą kąt prosty. Wskaż inną godzinę, o której wskazówki tworzą kąt prosty.

- (A) 3:15
- (B) 3:45
- (C) 9:00
- (D) 9:45

M061185

34

W tabeli poniżej podano wagę i długość wielkich węży.

Gatunek węża	Waga (w kilogramach)	Długość (w metrach)
Boa dusiciel	27	4
Pyton birmański	90	5 do 7
Anakonda zielona	227	6 do 9
Kobra królewska	9	4

A. Kuba zobaczył węża, który miał 8 metrów długości. Jaki to mógł być gatunek węża?

Odpowiedź: _____

B. Naima zobaczyła węża, który miał 6 metrów długości i ważył około 80 kilogramów. Jaki to mógł być gatunek węża?

Odpowiedź: _____



M061239

35

$4809 - 532 =$

Odpowiedź: _____

M051205

M03_01

36

Marysia ma 50 zedów.

Każda książka kosztuje 12 zedów.

Jaka jest największa liczba książek, które Marysia może kupić?

Odpowiedź: _____

M051039

M03_02

37

Pociąg wyjeżdża z Krakowa o 7.52 rano i przyjeżdża do Warszawy o 11.06 rano tego samego dnia.

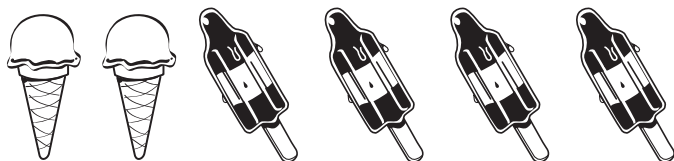
Ile czasu trwa podróż?

Odpowiedź: _____ godz. i _____ min.

M051055

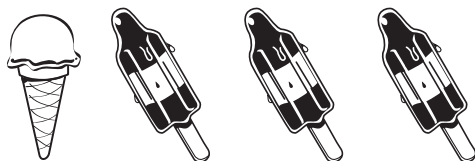
M03_03

Bartek kupił:



Zapłacił
22 zedy

Jola kupiła:



Zapłaciła
14 zedów

Ile zedów kosztuje  i  razem?

Odpowiedź: _____

Ile zedów kosztuje  ?

Odpowiedź: _____



39

Karol miał 20 zedów.

Karol kupił książkę za 3,65 zedów.

Kupił też czasopismo za 2,70 zedów.

Ile pieniędzy mu zostało?

- (A) 6,35 zedów
- (B) 13,65 zedów
- (C) 14,65 zedów
- (D) 16,35 zedów





M051070

40

Obrazki w rolce naklejek powtarzają się co cztery naklejki, jak widać na rysunku.



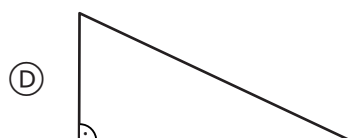
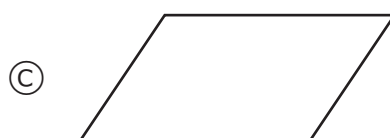
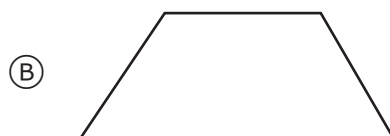
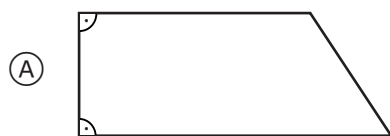
Który obrazek pojawi się na 39 naklejce w rolce?

- (A) 
- (B) 
- (C) 
- (D) 

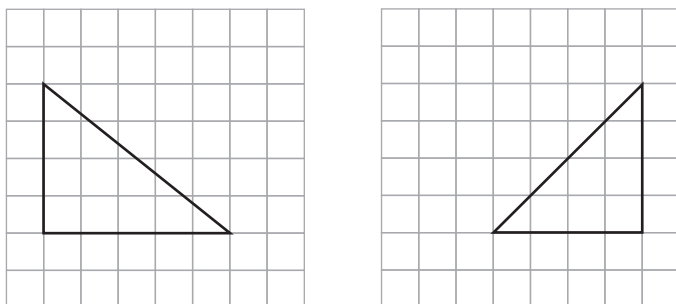
M051018

41

Która z tych figur jest zbudowana z jednego trójkąta i jednego prostokąta położonych jeden przy drugim?



42



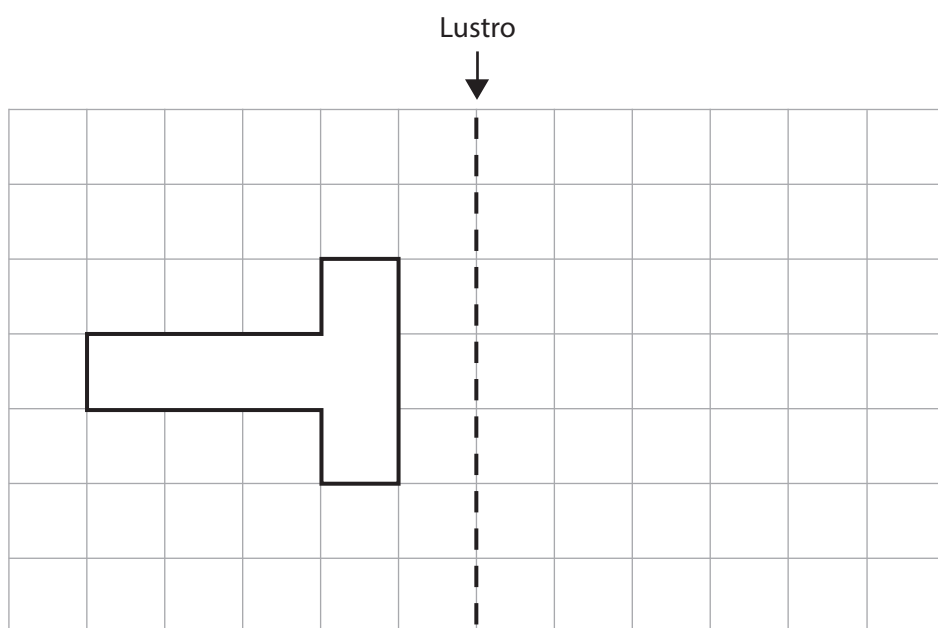
Które zdanie na temat tych trójkątów jest prawdziwe?

- (A) Każdy trójkąt ma 2 boki równe.
- (B) Każdy trójkąt ma 3 boki o różnych długościach.
- (C) Każdy trójkąt ma kąt większy od kąta prostego.
- (D) Każdy trójkąt ma kąt prosty.

M051410

43

Narysuj odbicie poniższej figury.



M051059

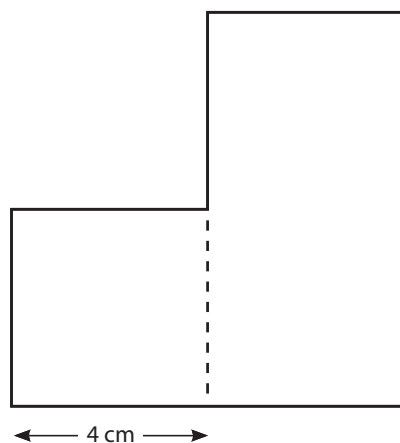
44

Ta figura składa się z kwadratu i prostokąta.

Szerokość prostokąta jest taka sama jak szerokość kwadratu.

Wysokość prostokąta jest dwa razy większa niż jego szerokość.

Oblicz obwód tej figury.



- (A) 28 cm
- (B) 32 cm
- (C) 36 cm
- (D) 40 cm

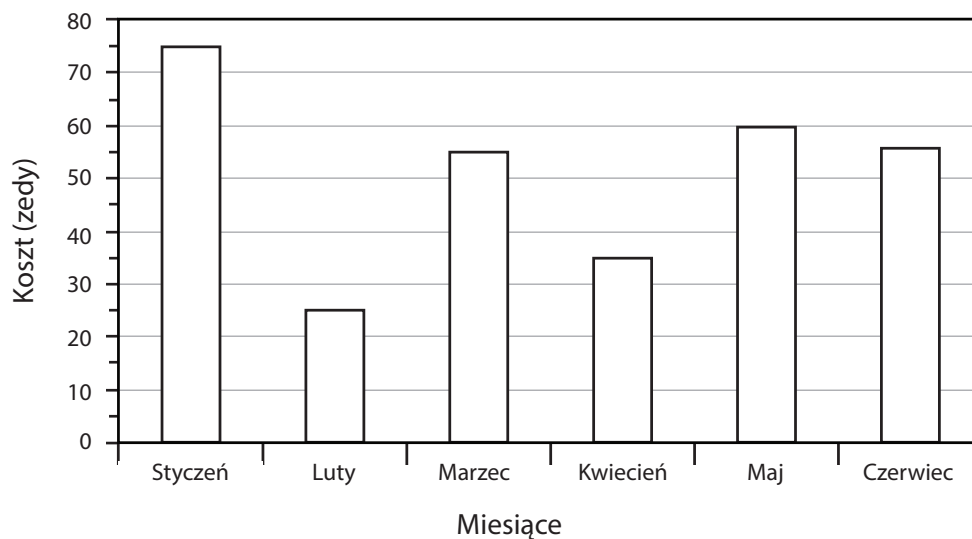
Tyle wyniosły rachunki za telefon Piotra w pierwszych sześciu miesiącach 2008 roku.

Rachunki za telefon Piotra w 2008 roku

Miesiąc	Styczeń	Luty	Marzec	Kwiecień	Maj	Czerwiec
Koszt (zedy)	65	20	60	40	60	45

Tyle wyniosły rachunki za telefon Piotra w pierwszych sześciu miesiącach 2009 roku.

Rachunki za telefon Piotra w 2009 roku



W których miesiącach Piotr zapłacił **mniej** za telefon w 2009 roku niż w 2008 roku?

Odpowiedź: _____

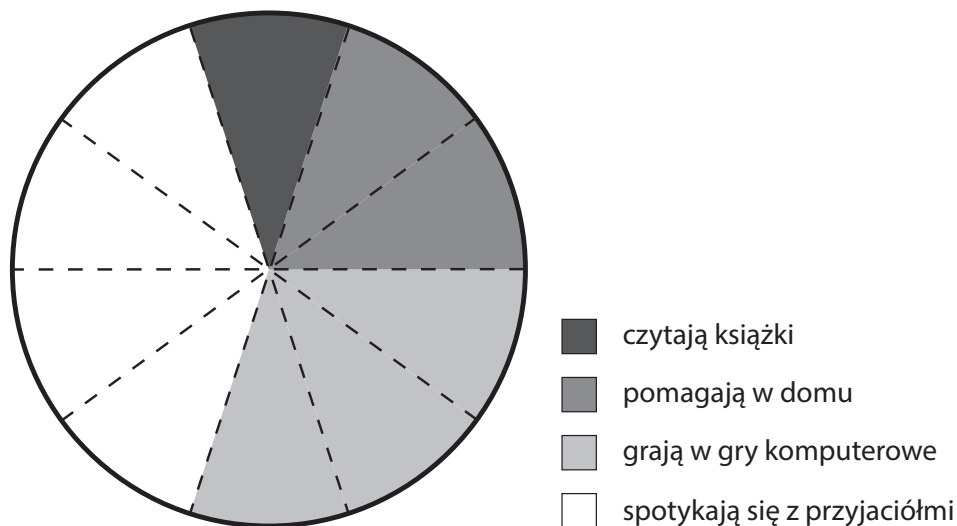


46

Rysunek pokazuje, co uczniowie robią po lekcjach.

Rysunek jest podzielony na 10 równych części.

Po lekcjach



20 uczniów czyta książki. Ilu uczniów spotyka się z przyjaciółmi?

- (A) 40
- (B) 60
- (C) 80
- (D) 100

M051077



Nie ma więcej pytań w tym sprawdzianie.

Jeśli skończyłaś/skończyłeś część 2 sprawdzianu przed czasem, przejrzyj i sprawdź swoje odpowiedzi w części 2.

Dziękujemy Ci za poświęcenie czasu na staranne odpowiedzi na pytania.

Literatura

- Bourdieu, P. (2005). *Dystynkcja. Społeczna krytyka władzy sądzona*. Przełożył P. Biłos. Warszawa: Wydawnictwo Naukowe „Scholar”.
- Davies von, M., Gonzales, E. i Mislevy, R. J. (2009). What are plausible values and why they are useful? *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, 2, 9–36.
- Dolata, R. (2008). *Szkoła – segregacje – nierówności*. Warszawa: Wydawnictwa UW.
- Dolata, R. i Pokropek, A. (2012). Czy warto urodzić się w styczniu? Wiek biologiczny a wyniki egzaminacyjne. W: B. Niemierko i M. K. Szmigel (red.), *Regionalne i lokalne diagnozy edukacyjne. Materiały XVIII Konferencji Diagnostyki Edukacyjnej* (s. 52–58). Warszawa: Polskie Towarzystwo Diagnostyki Edukacyjnej.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth K. i Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, 43(6), 1428–1446.
- Główny Urząd Statystyczny (2009). *Oświata i wychowanie w roku szkolnym 2008–2009*. Warszawa: Autor.
- Główny Urząd Statystyczny (2010). *Oświata i wychowanie w roku szkolnym 2009–2010*. Warszawa: Autor.
- Główny Urząd Statystyczny (2011). *Oświata i wychowanie w roku szkolnym 2010–2011*. Warszawa: Autor.
- Główny Urząd Statystyczny (2012). *Oświata i wychowanie w roku szkolnym 2011–2012*. Warszawa: Autor.
- Główny Urząd Statystyczny (2013). *Oświata i wychowanie w roku szkolnym 2012–2013*. Warszawa: Autor.
- Główny Urząd Statystyczny (2014). *Oświata i wychowanie w roku szkolnym 2013–2014*. Warszawa: Autor.
- Główny Urząd Statystyczny (2015). *Oświata i wychowanie w roku szkolnym 2014–2015*. Warszawa: Autor.
- Główny Urząd Statystyczny (2016). *Oświata i wychowanie w roku szkolnym 2015–2016*. Warszawa: Autor.
- Grygiel, P. (2016). Dynamika poczucia integracji z rówieśnikami w klasach 3–6 szkoły podstawowej. Model latentnych krzywych rozwojowych. *Edukacja*, 138(3), w druku.
- Hambleton, R. K., Swaminathan, H. i Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park: Sage.
- Jakubowski, M. i Sakowski, P. (2006). Quasi-experimental estimates of class size effect in primary schools in Poland. *International Journal of Educational Research*, 45(3), 202–215.
- Konarzewski, K. (2004). *Podstawa programowa i warunki kształcenia*. Warszawa: Instytut Spraw Publicznych.
- Konarzewski, K. (2007). PIRLS 2006. *Jak czytają dzieci w Polsce i na świecie*. Warszawa: Centralna Komisja Egzaminacyjna.
- Konarzewski, K. (2012). TIMSS i PIRLS 2011. *Osiągnięcia szkolne polskich trzecioklasistów w perspektywie międzynarodowej*. Warszawa: Centralna Komisja Egzaminacyjna.
- Konarzewski, K. (2013). Wiek startu szkolnego a osiągnięcia w nauce w okresie wczesnoszkolnym. *Edukacja*, 124(4), 5–19.
- Minner, D. D., Levy, A. J. i Century, J. (2009). Inquiry-based science instruction –What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching*, 47(4), 474–496.
- Mullis I. V. S., Martin, M. O., Minnich, Ch. A., Stanco G. M., Arora, A., Centurino, V. A. S. i Castle, C. E. (eds.). (2016).

- TIMSS 2015 Encyclopedia. Education Policy and Curriculum in Mathematics and Science.* Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Oshima, T. C. i Domaleski, Ch. S. (2006). Academic performance gap between summer-birthday and fall-birthday children in grades K–8. *Journal of Educational Research*, 99(3), 212–217.
- Pokropek, A. (red.) (2015). *Modele cech ukrytych w badaniach edukacyjnych, psychologii i socjologii. Teoria i zastosowania.* Warszawa: IBE.
- Poziomek, U., Tarłowski, A., Marszał, D. i Ostrowska, E. B. (2015). Edukacja przyrodnicza małych dzieci – opis przyrody czy jej badanie? *Edukacja Biologiczna i Środowiskowa*, 4, 48–61.
- Raudenbush, S. W. i Bryk, A. S. (2002). *Hierarchical linear models. Applications and data analysis methods* (wyd. 2). Thousand Oaks: Sage.
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31, 114–128.