



Kontekstowy model oceny efektywności nauczania po pierwszym etapie edukacyjnym



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI



EDUKACYJNA
WARTOŚĆ
DODANA

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOLECZNY



Autorzy:

Roman Dolata, Instytut Badań Edukacyjnych
Anna Hawrot, Instytut Badań Edukacyjnych
Grzegorz Humenny, Instytut Badań Edukacyjnych
Aleksandra Jasińska-Maciążek, Instytut Badań Edukacyjnych
Maciej Koniewski, Instytut Badań Edukacyjnych
Przemysław Majkut, Instytut Badań Edukacyjnych

Redakcja językowa:

Karolina Kwiatosz

Recenzenci:

prof. dr hab. Elżbieta Hornowska
dr hab. Janusz Gęsicki, prof. APS

Wydawca:

Instytut Badań Edukacyjnych
ul. Górczewska 8, 01-180 Warszawa
tel. (22) 241 71 00; www.ibe.edu.pl

Skład:

Andrzej Dziekoński

Druk:

Agencja Reklamowo-Wydawnicza A.Grzegorczyk
ul. Kutrzeby 15, 05-082 Stare Babice

© Copyright by: Instytut Badań Edukacyjnych, Warszawa 2014

ISBN: 978-83-61693-44-4

Publikacja powstała w ramach projektu „Badania dotyczące rozwoju metodologii szacowania wskaźnika edukacyjnej wartości dodanej (EWD)”; współfinansowana została przez Unię Europejską ze środków Europejskiego Funduszu Społecznego w ramach projektu: „Rozwój metody edukacyjnej wartości dodanej na potrzeby wzmocnienia ewaluacyjnej funkcji egzaminów zewnętrznych”.

Egzemplarz bezpłatny

Spis treści

Wprowadzenie i najważniejsze uzyskane wyniki	7
1 Organizacja i metodologia badania podłużnego w szkołach podstawowych	11
Terminarz i zakres badania	11
Próba badawcza	12
Schemat losowania	13
Próba zapasowa	15
Przebieg badania	17
Etap I	17
Etap III, część 1	19
Etap III, część 2	19
Pomiar osiągnięć szkolnych	21
Test umiejętności czytania	21
Test świadomości językowej	22
Test umiejętności matematycznych	22
Reprezentatywność próby	23
Stopy realizacji badania	24
Podejście do braków danych	27
Ograniczenie próby badawczej	27
Analiza reprezentatywności próby	28
Wagi stosowane w analizach	33
2 Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym	35
Cechy ucznia i jego środowiska rodzinnego	35
Płeć	35
Wiek	36
Inteligencja	36
Aspiracje edukacyjne rodziców	37
Status społeczno-ekonomiczny rodziny	38
Struktura rodziny	38
Uczęszczanie do przedszkola i zerówki	39
Czynniki społeczno-gospodarcze na poziomie lokalnym	40

Poziom bezrobocia	40
Struktura gospodarki	41
Powszechność opieki przedszkolnej	41
Wydatki na oświatę	41
Metoda	42
Dane i metody pomiaru	42
Model analizy danych	47
Wyniki	47
Podsumowanie	51
3 Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym	55
Różne modele efektywności nauczania	57
Problem badawczy	60
Metoda	60
Dane i metoda pomiaru	60
Plan analizy danych	60
Model analizy danych	62
Wyniki	62
Zróźnicowanie międzyszkolne cech uczniów mających znaczenie dla osiągnięć szkolnych	62
Kontekstowe modele efektywności nauczania	64
Właściwości kontekstowych wskaźników efektywności nauczania	67
Podsumowanie	83
4 Trafność proponowanego kontekstowego modelu oceny efektywności nauczania	87
Korelaty efektywności nauczania – style zarządzania szkołą	88
Partycypacja w zarządzaniu w ocenie nauczycieli	88
Kultura organizacyjna szkoły	89
Współpraca rodziców ze szkołą	91
Korelaty efektywności nauczania – czynniki związane z nauczycielem	92
Korelaty efektywności nauczania – integracja uczniów ze szkołą	93
Metoda	94
Narzędzia i wskaźniki	94
Model analizy danych	97
Wyniki	97
Style zarządzania szkołą	99
Partycypacja nauczycieli w zarządzaniu szkołą	99
Kultura organizacyjna szkoły	99
Współpraca rodziców ze szkołą	100
Charakterystyki nauczycieli i ich styl pracy z uczniami	100
Staż i wykształcenie	103
Stopień awansu zawodowego	103
Styl pracy nauczyciela z uczniami	103

Poczucie skuteczności w zawodzie nauczyciela	106
Integracja szkolna uczniów	106
Podsumowanie	110
Bibliografia	113
Aneks 1: Raport dla szkoły z badania podłużnego w szkołach podstawowych	123
Aneks 2 (PŁYTA CD)	145
Rozdział 1	145
Rozdział 2	148
Rozdział 4	168

Wprowadzenie i najważniejsze uzyskane wyniki

Prace nad kontekstowymi modelami oceny efektywności nauczania dla pierwszego etapu edukacyjnego to jeden z elementów realizowanego w latach 2007–2013 projektu „Badania dotyczące rozwoju metodologii szacowania wskaźnika edukacyjnej wartości dodanej (EWD)” (www.2013.ewd.edu.pl). Głównym celem tego przedsięwzięcia był rozwój metody edukacyjnej wartości dodanej i propagowanie narzędzi analitycznych pozwalających wykorzystywać ją do analizy wyników egzaminacyjnych. By można było stosować metodę EWD, potrzebne są wyniki pomiaru osiągnięć szkolnych przynajmniej w dwóch punktach czasowych, czyli na początku i na końcu danego cyklu edukacyjnego. Modele wyjście–wejście stosowane są w Polsce do oceny efektywności nauczania w gimnazjach, liceach ogólnokształcących i technikach. Jak jednak oceniać efektywność nauczania, gdy nie dysponujemy danymi o osiągnięciach szkolnych uczniów na progu danego cyklu nauczania? Z taką sytuacją mamy do czynienia na pierwszym etapie edukacyjnym. Choć szkoły coraz częściej dysponują wynikami standaryzowanych pomiarów osiągnięć szkolnych na koniec klasy trzeciej (np. dzięki zorganizowanemu przez Centralną Komisję Egzaminacyjną, a od 2012 r. prowadzonemu przez Instytut Badań Edukacyjnych, *Ogólnopolskiemu badaniu umiejętności trzecioklasistów*), to nie ma naukowo opracowanych metod analiz pozwalających na ocenę na ich podstawie efektywności nauczania. Nieprzetworzone wyniki testów końcowych nie są dobrą miarą efektywności, ponieważ są silnie uzależnione od czynników pozaszkolnych.

W sytuacji braku informacji na temat uprzednich osiągnięć szkolnych, do oceny efektywności nauczania można stosować tzw. modele kontekstowe. Odnoszą one wyniki testu osiągnięć szkolnych przeprowadzonego na koniec danego cyklu nauczania do charakterystyk psychospołecznych uczniów pobierających naukę w danej szkole. Najważniejsze dla trafności modeli są te cechy uczniów, które najsilniej wpływają na osiągnięcia szkolne, a nie są bezpośrednio zależne od działań szkoły.

Wyzwanie stworzenia takiego modelu podjęliśmy w zespole EWD. Zaplanowaliśmy wieloetapowe, podłużne badanie uwarunkowań efektywności wyników nauczania. Rozpoczęło się ono w roku szkolnym 2009/2010 i objęło reprezentatywną, ogólnopolską próbę losową około 300 oddziałów klas pierwszych ze 180 szkół podstawowych. Badanie obejmie pełny cykl szkoły podstawowej i skończy się w 2015 roku. W tej publikacji przedstawimy wyniki opracowane na podstawie danych zabranych w pierwszej fazie badania, czyli w klasach I–III. Dzięki uzyskanym informacjom można testować wiele alternatywnych modeli i analizować ich własności. Ponieważ w badaniu zgromadzono dużo danych pozwalających scharakteryzować środowisko szkolne, została również przeprowadzona ocena trafności proponowanych wskaźników.

Podjęty w naszej książce problem badawczy ma nie tylko wymiar naukowy, ale również praktyczny. Przeprowadzone analizy mają na celu opracowanie takiego kontekstowego

modelu oceny efektywności nauczania dla pierwszego etapu edukacyjnego w Polsce, który będzie można wykorzystywać – po spełnieniu pewnych warunków – do analizy wyników stosowanych w klasach I–III standaryzowanych testów osiągnięć szkolnych. Jeżeli model okaże się dostatecznie dobry, będzie można go wykorzystywać również do analizy standaryzowanych testów diagnostycznych, które – mamy nadzieję – na stałe zagospodszą w szkołach podstawowych po pierwszym etapie edukacyjnym. Należy wyraźnie podkreślić, że nie postulujemy nowego egzaminu. Polskim szkołom potrzebne są dobre, unormowane dla populacji krajowej testy diagnostyczne, z których one same będą korzystać. Jednak by szkoła w ramach ewaluacji wewnętrznej, rozumianej jako zarządzanie procesem jej rozwoju, mogła wyniki tych pomiarów w pełni wykorzystywać, potrzebne są statystyczne modele analizy pozwalające kontrolować wpływ pozaszkolnych czynników determinujących osiągnięcia szkolne.

Warto podkreślić, że w polskim systemie oświaty już w latach 90. ubiegłego wieku podjęto próby stosowania kontekstowych modeli oceny efektywności nauczania. W ramach tzw. eksperymentu kwidzyńskiego (Dolata, Murawska i Putkiewicz, 2001; Dolata, Murawska, Putkiewicz i Żytko, 1997) szkoły biorące udział w programie otrzymywały informacje zwrotne o osiągnięciach szkolnych swoich uczniów nie tylko w postaci rozkładów wyników testów, ale również jako wskaźniki kontekstowe uwzględniające międzyszkolne zróżnicowanie wykształcenia rodziców uczniów. Te pierwsze zastosowania kontekstowych modeli były statystycznie wysoce niedoskonałe (zbyt proste modele statystyczne, nie-szacowanie przedziałów ufności dla wskaźników), ale ich głównym mankamentem był brak dowodów trafności wskaźników. Analizy przedstawione w tym opracowaniu pozwolą zdecydowanie podnieść status metodologiczny tego typu modeli.

Po przedstawieniu celu naszego opracowania pora zaprezentować strukturę raportu i najważniejsze uzyskane wyniki.

W rozdziale 1 przedstawiona została organizacja i metodologia badania podłużnego, które dostarczyło danych niezbędnych do prac nad kontekstowymi modelami oceny efektywności nauczania. Starannie dobrana próba, odpowiednie zaplanowanie pomiarów, dobra realizacja badania w terenie, jak i ex post potwierdzona reprezentatywność próby, wskazują, że zebrane dane są dobrej jakości.

W rozdziale 2 zostały omówione wyniki analiz pozaszkolnych uwarunkowań osiągnięć szkolnych w pierwszym etapie edukacyjnym w Polsce. Przygotowały one grunt pod prace nad modelami oceny efektywności nauczania, które powinny skutecznie kontrolować wpływ tych właśnie pozaszkolnych, kontekstowych czynników. Wyniki przeprowadzonych analiz pozwoliły w większości potwierdzić znaczenie pozaszkolnych czynników wpływających na osiągnięcia szkolne wytypowanych na podstawie dotychczasowych badań. I tak, czynnikiem najsilniej determinującym osiągnięcia szkolne we wszystkich zakresach, jest inteligencja uczniów. Spośród czynników statusowych najsilniejsze efekty stwierdzono dla zasobności gospodarstwa domowego (w tym dla liczby książek w domu ucznia) oraz poziomu wykształcenia rodziców. Znaczenie ma też poziom aspiracji edukacyjnych rodziców. Wskaźniki struktury rodziny okazały się znaczącymi wyznacznikami umiejętności matematycznych i świadomości językowej. Dla osiągnięć szkolnych duże znaczenie ma również opóźniony start szkolny. Z kolei dzieci posłane do szkoły wcześniej – jako sześciolatki – uzyskiwały wyniki porównywalne z siedmiolatkami. Należy jednak pamiętać, że jest to w badanej populacji grupa nieliczna i silnie pozytywnie wyselekcjonowana. Wiek biologiczny w obrębie

głównej kohorty wiekowej wiązał się pozytywnie z wynikami uzyskanymi po trzeciej klasie. Na koniec, okazało się, że wskaźniki z poziomu lokalnego, opisujące społeczno-gospodarcze otoczenie placówki, niewiele wnoszą do wyjaśnienia osiągnięć szkolnych. W przypadku poziomu bezrobocia – jest to wręcz efekt odwrotny od spodziewanego.

Centralną dla tego opracowania częścią jest rozdział 3. Jak brzmią kluczowe ustalenia? Krótkie opisanie uzyskanych wyników już w tym miejscu, pozwoli lepiej zrozumieć cel tego raportu i śledzić kolejne etapy analizy. Badanie wykazało, że szkoły podstawowe już po trzeciej klasie różnią się znacząco poziomem osiągnięć szkolnych swoich uczniów. Podział na szkoły wyjaśnił około 9% wariacji wyników testu umiejętności czytania, 12% wariacji wyników testu świadomości językowej oraz 11% wariacji osiągnięć szkolnych z matematyki. Jednak część tego zróżnicowania daje się wyjaśnić nierównomiernym rozłożeniem pomiędzy szkołami cech uczniów mających znaczenie dla osiągnięć szkolnych. Wykorzystując wnioski z analiz zaprezentowanych w rozdziale 2, zbudowano dla każdego obszaru osiągnięć kontekstowy model efektywności nauczania, który kontrolował wszystkie ważne, potwierdzone w naszych analizach, pozaszkolne uwarunkowania osiągnięć uczniów. Nazwano te modele wyczerpującymi. W modelach tych podział na szkoły wyjaśnił niespełna 4% wariacji testu umiejętności czytania, około 5% dla testu świadomości językowej i poniżej 7% w przypadku testu z matematyki. Oznacza to (przy założeniu trafności modelu analizy), że gdyby szkoły pracowały z uczniami o takich samych charakterystykach psychospołecznych, w takim właśnie stopniu różniłyby się wynikami nauczania.

Wyniki tego modelu stanowiły punkt odniesienia dla modeli prostszych, uwzględniających tylko te zmienne, które można (przy sensownych kosztach) mierzyć i przetwarzać w skali całego systemu oświaty. Na podstawie zbudowanych modeli kontekstowych wyznaczono wskaźniki dla szkół, w których przeprowadzono badanie. Wykonane analizy pokazały, że najlepsze właściwości miały wskaźniki z dwóch modeli. Były one najbliższe pod względem wartości wskaźnikom wyliczonym z modeli wyczerpujących, a także były w najmniejszym stopniu związane z kontekstem pracy szkoły (czyli psychospołecznymi charakterystykami uczniów niezależnymi od szkoły). Dodatkowo wyjaśniły najwięcej wariacji wyników testów. Ponieważ jednak różnice co do oszacowań efektów szkół między tymi modelami były praktycznie pomijalne (korelacje rzędu 0,998-0,999), wybrano model zawierający mniej zmiennych opisujących ucznia i kontekst środowiskowy: płeć, wiek w miesiącach, oznaczenie bycia starszym lub młodszym niż główna kohorta wiekowa, wykształcenie rodziców ucznia w latach nauki (na poziomie indywidualnym oraz jako średnia z poziomu szkoły), aspiracje rodziców co do wykształcenia ich dzieci, informacje o liczbie posiadanych w domu książek. Model ten nazwano złożonym modelem kontekstowym efektywności nauczania. Dodatkowo wybrano do analiz model, w którym uwzględnione zostały tylko te informacje, które są możliwe do pozyskania przez szkoły (płeć i wiek ucznia, oznaczenie bycia starszym lub młodszym niż główna kohorta wiekowa oraz informacja o wykształceniu rodziców uwzględniona na poziomie indywidualnym i jako średnia w szkole). Model ten miał tylko trochę gorsze właściwości niż model złożony, a był znacznie prostszy pod względem możliwości zgromadzenia danych potrzebnych do jego wyliczenia. Dowiedziono, że wskaźniki wyliczone z kontekstowych modeli efektywności nauczania (zarówno złożonych, jak i prostych) są znacznie słabiej powiązane z kontekstem pracy szkoły niż wyniki testów osiągnięć. Średnie wyniki testów osiągnięć szkolnych są skorelowane ze zagregowanymi miarami statusu na poziomie

około 0,7. Dla kontekstowych wskaźników efektywności nauczania nie zaobserwowano natomiast istotnych korelacji. W przypadku średniej inteligencji uczniów w pierwszej klasie, dla wyników testów stwierdzono korelacje rzędu 0,56, natomiast dla kontekstowych wskaźników efektywności nauczania korelacje ponad dwukrotnie niższe. Jest to jeden z dowodów trafności wyliczonych wskaźników jako miary efektywności pracy szkoły. Ich wartość nie zależy bowiem od cech społecznych uczniów, którzy uczęszczają do danej szkoły i w znacznie mniejszym stopniu niż wyniki testów zależy od średniej inteligencji uczniów.

Opracowanie wieńczy rozdział 4, w którym podjęto próbę weryfikacji trafności kontekstowych wskaźników efektywności nauczania dla pierwszego etapu edukacyjnego na podstawie kryteriów zewnętrznych. Podsumowując przedstawione w nim analizy, należy stwierdzić, iż zebrane wyniki nie są rozstrzygające, jeśli chodzi o trafność analizowanych kontekstowych modeli efektywności nauczania. Analiza korelacji ze stylami zarządzania i cechami nauczyciela wskazały na istnienie jedynie części z zakładanych związków. Z drugiej strony, wyniki na pewno nie pozwalają stwierdzić, że szkoły o wysokich wartościach wskaźników efektywności nauczania mają jakieś niepokojące cechy. Natomiast analiza związków z miarami integracji szkolnej uczniów jednoznacznie wskazuje na trafność miar efektywności. Poziom integracji szkolnej okazał się pozytywnie związany z osiągnięciami uczniów w modelach kontekstowych, choć w zależności od obszaru osiągnięć pojawiły się drobne różnice. Efektywność nauczania okazała się korelować ze zagregowanymi na poziomie oddziału miarami integracji społecznej, pozytywnej samooceny kompetencji szkolnych oraz pozytywnego stosunku do szkoły.

Bardzo ważnym uzupełnieniem tekstu są aneksy. Szczególnie lektura pierwszego z nich jest warta polecenia. Zawiera on przykładowy raport dla szkoły, przedstawiający wyniki analizy efektywności nauczania w danej placówce. Raport taki otrzymała każda ze 180 szkół biorących udział w badaniu. Były one omawiane z przedstawicielami szkół na specjalnej konferencji podsumowującej pierwszą fazę badania. Aneks dokumentuje korzyści, jakie wyniosły szkoły z udziału w przedsięwzięciu, ale jest przede wszystkim propozycją analiz, jakie szkoły będą mogły wykonywać, gdy modele oceny efektywności nauczania będą używane w masowej skali. Drugi aneks zawiera dokumentację metodologiczną stosowanych skal i modeli statystycznych analizy danych.

Po przedstawieniu struktury książki, pora na podziękowania. Publikacja nie mogłaby powstać bez wieloletniego wysiłku całego zespołu badawczego zajmującego się rozwojem metody EWD w Polsce. Autorzy pragną podziękować Ewie Stożek, Zofii Lisieckiej, Filipowi Kulonowi, Annie Rappe, Krystynie Szmigel, Cezaremu Kacprzakowi i Danucie Ostrochulskiej. Bez ich zaangażowania nie udałoby się przygotować i przeprowadzić badań. Specjalne podziękowania należą się Arturowi Pokropkowi, który wylosował próbę do badania, a opis doboru próby zawarty w tej książce powstał na podstawie przygotowanej przez niego dokumentacji.

Rozdział 1

Organizacja i metodologia badania podłużnego w szkołach podstawowych

Podstawowe informacje dotyczące organizacji i metodologii będą stanowiły wprowadzenie do badania podłużnego przeprowadzonego w szkołach podstawowych. Omówimy, jak zbierano dane, na podstawie których przeprowadzono analizy i sformułowano wnioski przedstawione w kolejnych rozdziałach książki.

Terminarz i zakres badania

Badanie rozpoczęło się na początku 2010 roku, gdy uczniowie z wylosowanych szkół i oddziałów uczęszczali do klas pierwszych szkoły podstawowej, natomiast zakończyło się jesienią 2012 roku, gdy rozpoczynali naukę w klasie czwartej. Był to pierwszy rocznik objęty reformą programową wdrożoną przez minister edukacji Katarzynę Hall. Zamknięcie pełnego cyklu badawczego zaplanowane jest na moment ukończenia przez kohortę szkoły podstawowej (rok 2015). Prezentowane w książce wyniki są zatem podsumowaniem badania pierwszego etapu edukacyjnego, prowadzonego zgodnie z wytycznymi nowej podstawy programowej¹.

Pierwotnie planowano realizację trzech etapów badania, w tym dwóch obejmujących pomiar osiągnięć szkolnych. Jednakże z powodów niezależnych od zespołu badawczego II etap badania odwołano, wskutek czego modyfikacji uległa koncepcja III etapu². Ostatecznie zrealizowano dwie rundy: etap I oraz dwuczęściowy etap III. W Tabeli 1.1 przedstawiono terminy ich realizacji oraz zbierania danych w terenie (tzw. okienka badawcze).

W I etapie zebrano opinie wychowawców na temat kompetencji uczniów u progu szkoły oraz przeprowadzono badanie inteligencji. Od rodziców (opiekunów) pozyskano informacje o statusie społecznym rodziny (wykształcenie rodziców, ich status na rynku pracy, wykonywany zawód), jej strukturze, kapitale kulturowym i ekonomicznym oraz aspiracjach edukacyjnych wobec dziecka. Nauczycieli i dyrektorów zapytano o opinie na temat egzaminów zewnętrznych oraz sposobów wykorzystania ich wyników w szkole. Ponadto

¹ Badanie kontynuowane jest w ramach projektu systemowego „Rozwój metody edukacyjnej wartości dodanej na potrzeby wzmocnienia ewaluacyjnej funkcji egzaminów zewnętrznych” współfinansowanego ze środków Europejskiego Funduszu Społecznego (POKL 2014–2020, priorytet III, działanie 3.2).

² Informacje o pierwotnych planach badawczych znaleźć można w raporcie Zofii Lisieckiej (2010).

Tabela 1.1. Terminy realizacji kolejnych etapów badania.

Etap	Termin*	Okienko badawcze
I	29.10.2009–07.06.2010	15.02.2010–07.04.2010
III, część 1	18.04.2012–08.08.2012	10.05.2012–26.06.2012
III, część 2	14.08.2012–23.03.2013	17.09.2012–24.10.2012

* Za moment rozpoczęcia badania przyjęto datę podpisania umowy z wykonawcą (I, III etap, cz. 1) lub datę rozpoczęcia kolejnego etapu badania wynikającą z umowy (III etap, cz. 2). Za datę jego zakończenia – przekazanie przez wykonawcę wolnych od wad materiałów badawczych.

od dyrektorów pozyskano dane na temat statystyk rekrutacyjnych oraz stosowanych kryteriów przyjęć do placówki.

W pierwszej części III etapu zebrano dane na temat poziomu inteligencji uczniów, ich integracji ze szkołą i rówieśnikami, motywacji do nauki oraz pozycji socjometrycznej w klasie. Rodziców (opiekunów) pytano o jakość relacji ze szkołą, aspiracje edukacyjne wobec dziecka oraz dodatkowe zajęcia edukacyjne, rekreacyjne i pomoc w nauce poza lekcjami w szkole, z których uczeń korzysta. Nauczycieli ponownie pytano o opinie na temat egzaminów zewnętrznych, stosowane metody analizy wyników egzaminów, doskonalenie zawodowe, wykorzystywane na lekcjach podręczniki oraz metody pracy, kulturę organizacyjną szkoły, partycypację w zarządzaniu szkołą, stosunek do dyrektora, satysfakcję zawodową; mierzono także poczucie skuteczności zawodowej. Do dyrektorów skierowano pytania dotyczące stosowanych w szkole metod analizy wyników egzaminacyjnych, opinii na temat egzaminów zewnętrznych, rekrutacji do szkoły, metod przydziału uczniów do klas i klas do nauczycieli, jak również kultury organizacyjnej placówki, zarządzania nią, partycypacji nauczycieli w tym procesie oraz wizji funkcji dyrektora szkoły.

W toku części drugiej III etapu sprawdzano poziom osiągnięć szkolnych uczniów w zakresie języka polskiego (czytania i świadomości językowej) oraz matematyki. Szczegółowe informacje na temat każdego z etapów znajdują się w dalszych częściach rozdziału.

W aneksie do tego rozdziału zamieszczono listę badanych obszarów wraz z informacją o zastosowanych narzędziach. Należy podkreślić, że do analiz przedstawionych w niniejszej książce wykorzystano tylko część zebranych danych. Na pozostałych oparto się w innych opracowaniach i raportach.

Próba badawcza

Celem badania było jak najdokładniejsze oszacowanie parametrów i wskaźników na poziomach: indywidualnym (uczniów), oddziałów oraz szkół. W zakresie precyzji pomiaru na poziomie uczniów przyjęto, że dokładność oszacowania powinna być nie gorsza niż ta, którą gwarantowałyby dobór prosty co najmniej 400 uczniów. Przyjęcie takiej, efektywnej wielkości próbki pozwala na otrzymanie w przybliżeniu następujących 95-procentowych przedziałów ufności: w przypadku średnich: $M \pm 0,1$ (M – średnia, SE – błąd standardowy na poziomie uczniów); w przypadku odsetków: $p \pm 5$ punktów

procentowych (p – oszacowanie odsetka na poziomie uczniów); w przypadku korelacji: $r \pm 0,1$ (r – współczynnik korelacji na poziomie uczniów).

W związku z tym, że badanie miało na celu również oszacowanie wskaźników na poziomie szkół i oddziałów, potrzebna była także duża próbka na tym poziomie analizy. Dodatkową przesłanką za dużą wielkością próby był fakt, że dobór uczniów do badania miał charakter zespołowy, co wiąże się ze zwiększeniem wartości błędów oszacowań oraz większą homogenicznością różnych charakterystyk uczniów w obrębie zespołu (oddziału). Przyjęto, że odpowiednia wielkość próbki szkół wynosi 180, a w każdej szkole w miarę możliwości do badania losowane będą 2 oddziały klasy pierwszej. Przy tej wielkości próbki szkół i oddziałów, przy stuprocentowym zrealizowaniu próby, oszacowania na poziomie uczniów będą charakteryzowały się większą precyzją, niż założona wyżej. Przewidywano jednak niższy niż stuprocentowy odsetek realizacji próby. Wyniki symulacji przeprowadzonych w celu oszacowania błędów pomiaru, przy założeniu losowego charakteru zjawiska niedostępności szkół, wykazały, że przyjęta liczebność szkół jest optymalna.

Populację docelową w badaniu stanowili uczniowie klas pierwszych szkół podstawowych (z wyłączeniem szkół specjalnych, sportowych, przyszpitalnych, dla dorosłych) w Polsce w roku szkolnym 2009/2010. Operat pierwotny stanowiła lista szkół zawierająca informację o liczbie oddziałów klasy pierwszej sporządzona na podstawie danych z Systemu Informacji Oświatowej (stan na 30 września 2009 r.).

Dodatkowo, by zminimalizować błąd oszacowania parametrów i wskaźników na poziomie szkół, przy narzuconych ograniczeniach kosztowych, postanowiono ograniczyć populację docelową do uczniów szkół publicznych uczących się w szkołach większych niż placówki 10-osobowe na poziomie klasy pierwszej. Populacja ta była rzeczywistą populacją losowania, której odpowiadał dostosowany wtórny operat losowania. Decyzję tę podjęto między innymi po uwzględnieniu wyników symulacji służących do oszacowania błędów próbkowania. Przyjęte ograniczenie populacji wyłączyły z operatu około 2% uczniów. Tabela 1.2 przedstawia strukturę operatu losowania.

Schemat losowania

Losowanie miało charakter: (a) warstwowy (b) wielostopniowy oraz (c) zespołowy. W pierwszym kroku (a) próba została podzielona na 6 warstw według kategorii wielkości miejscowości oraz informacji, czy szkoła prowadzi jeden czy wiele oddziałów na poziomie klasy pierwszej. Następnie dokonano alokacji szkół w warstwach, zbliżonej do procentowego udziału uczniów w danej warstwie, choć odstąpiono nieznacznie od idealnej alokacji proporcjonalnej. Takie rozwiązanie nieznacznie poprawiło wartości oszacowań w stosunku do alokacji wprost proporcjonalnej. Alokacje przedstawiono w Tabeli 1.3.

W kolejnym kroku z każdej warstwy wylosowano zakładaną liczbę szkół (c) w sposób zależny, z prawdopodobieństwami wyboru proporcjonalnymi do liczby uczniów uczących się w klasach pierwszych. Zastosowano klasyczną metodę kumulatywną. Następnie (b) w każdej wylosowanej szkole do udziału w badaniu wylosowano dwa oddziały (c). Jeżeli w danej szkole był jeden lub dwa oddziały, wszystkie były kwalifikowane do

Tabela 1.2. Struktura operatu losowania, podział na warstwy, liczebności dla klasy pierwszej.

Liczba oddziałów (klasa I) w szkole	Struktura operatu	Wieś	Miasto do 20 tys. mieszkańców	Miasto od 20 tys. do 100 tys. mieszkańców	Miasto pow. 100 tys. mieszkańców	Razem
>1 oddział	Liczba uczniów	45 478	51 718	62 413	77 420	237 029
	% uczniów	14,15	16,09	19,42	24,09	73,75
	Liczba szkół	1 139	866	869	1 154	4 028
	Średnia liczebność klas	18,42	21,4	22,8	22,38	2 114
=1 oddział	Liczba uczniów	73 422		10 960		84 382
	% uczniów	22,8		3,3		26,3
	Liczba szkół	4 340		530		487
	Średnia liczebność klas	16,92		20,67		17,33
Razem	Liczba uczniów	118 900	55 137	66 102	81 272	321 411
	% uczniów	37,0	17,2	20,6	25,3	100,0
	Liczba szkół	5 479	1 032	1 049	1 338	8 898
	Średnia liczebność klas	17,23	21,27	22,4	22,18	19,05

badania. Jeżeli liczba oddziałów przekraczała dwa, w sposób prosty zależny z równymi prawdopodobieństwami wyboru (na podstawie dostarczonych liczb losowych) losowano do badania dwa oddziały. Losowanie to przeprowadzał wykonawca badania.

Podziału na warstwy ze względu na lokalizację szkoły dokonano w celu zapewnienia odpowiedniej reprezentacji w próbie szkół z różnych lokalizacji oraz w celu zmniejszenia błędów oszacowania parametrów rozkładu zmiennej zależnej (umiejętności szkolne uczniów). Obserwuje się bowiem dodatnią korelację między wielkością miejscowości a osiągnięciami szkolnymi uczniów. Warstwowanie ze względu na wielkość szkoły (liczbę oddziałów na poziomie klasy pierwszej) przeprowadzono w celu zapewnienia odpowiedniej reprezentacji w próbie szkół małych. Wybór zmiennych warstwujących ograniczony był dostępnością danych w Systemie Informacji Oświatowej.

Decyzja o losowaniu zespołowym podyktowana była przede wszystkim kwestiami merytorycznymi. Takie losowanie umożliwiła późniejsze badanie charakterystyk oddziałów

Tabela 1.3. Alokacja szkół w warstwach.

Warstwa	Lokalizacja szkoły	Liczba klas I w szkole	Liczba szkół wylosowanych do badania	% szkół wylosowanych do badania
1	Wieś		26	14,44
2	Miasto do 20 tys. mieszkańców		29	16,11
2	Miasto od 20 do 100 tys. mieszkańców	> 1 oddział	35	19,44
4	Miasto powyżej 100 tys. mieszkańców		43	23,89
5	Wieś	= 1 oddział	41	22,78
6	Miasto od 20 do 100 tys. mieszkańców		6	3,33

i testowanie dotyczących ich hipotez, co ma kluczowe znaczenie dla badania efektywności nauczania. Oprócz tego włączenie do badania całych oddziałów miało znacznie ułatwić jego realizację. Sprzyja to bowiem zrekutowaniu oraz utrzymaniu próby w kolejnych rundach (np. ze względu na łatwiejsze docieranie do respondentów, mniejszą ingerencję w pracę szkoły podczas prowadzenia badania). Jest to kluczowe w wypadku badania wieloetapowego, choć odbywa się kosztem zwiększenia błędów oszacowań parametrów rozkładu umiejętności uczniów.

Próba zapasowa

Wraz z próbą główną wylosowano próbę zapasową. Najpierw, zgodnie z opisanym powyżej schematem, wylosowano 360 szkół. Następnie spośród nich warstwowo, w sposób prosty wylosowano 180 placówek. Trafiły one do próby głównej. Pozostałe, ułożone w losowej kolejności wewnątrz warstw, trafiły do próby rezerwowej. Dodatkowo została ona podzielona na dwie części, dzięki czemu utworzono dwie próby rezerwowe (każdą szkołę przypisano losowo do jednej z prób rezerwowych).

W razie odmowy udziału w badaniu przez szkołę z próby głównej zastępowano ją pierwszą w kolejności szkołą z tej samej warstwy z próby rezerwowej. Jeżeli odmówiła udziału pierwsza szkoła rezerwowa, wybierana była kolejna szkoła z listy. Uznawano, że szkoła odmówiła udziału w badaniu, jeżeli:

- dyrektor nie udzielił zgody na uczestnictwo szkoły w badaniu lub
- odsetek uczniów, którzy nie zostali objęci badaniem w wylosowanych klasach, przekroczył 30%³.

³ Powodami braku zakwalifikowania był brak zgody rodziców na udział ucznia w badaniu i/lub niepełnosprawność ucznia uniemożliwiająca jego uczestnictwo w badaniu.

Tabela 1.4. Zastępowanie szkół z próby głównej szkołami zapasowymi w warstwach.

Warstwa	Liczba wykluczonych szkół (próba główna)	% szkół w warstwie	W tym zastąpiona szkołą z próby zapasowej		
			Pierwszą	Drugą	Trzecią
1	3	11,54	3	—	—
2	4	13,79	4	—	—
3	4	11,43	2	2	—
4	10	23,26	5	4	1
5	5	12,20	5	—	—
6	3	50,00	3	—	—
Razem	29	—	22	6	1

Z próby głównej z wyżej wskazanych przyczyn wykluczono 29 szkół (16,11%), następnie zastąpiono je szkołami z próby rezerwowej. Zdarzyły się jednak sytuacje, gdy z tych samych przyczyn szkoła z próby rezerwowej musiała być zastąpiona następną w kolejności, zgodnie z opisanymi wcześniej zasadami. Łącznie skontaktowano się z 37 szkołami z próby rezerwowej. Dwadzieścia dwie wykluczone szkoły z próby głównej zastąpiono pierwszymi szkołami z próby rezerwowej. Sześć szkół zastąpiono szkołami drugimi z listy (z powodu wykluczenia szkół pierwszych), natomiast w jednym przypadku – szkołą trzecią. Relatywnie najwięcej wykluczeń szkół wystąpiło w warstwach 4 i 6, czyli zlokalizowanych w miastach o liczbie mieszkańców przekraczającej 100 tysięcy. Dane znajdują się w Tabeli 1.4.

W przypadku dwóch szkół odstąpiono od wykluczenia z badania, pomimo zrealizowania badania poniżej progu 70% w jednym z oddziałów (były to szkoły dwuoddziałowe na poziomie klasy pierwszej). W obu placówkach łączna stopa realizacji w szkole przekroczyła jednak 70%. Decyzja była podyktowana trudnościami w rekrutacji uczniów oraz zbliżającym się terminem zakończenia badania.

Podsumowując, podczas rekrutacji wykluczono z badania 35 szkół (łącznie z próby głównej i rezerwowej). W 17 przypadkach przyczyną był brak zgody dyrektora placówki na uczestnictwo szkoły w badaniu, w 14 przypadkach – niewystarczający odsetek rodziców wyrażających zgodę na udział dziecka w badaniu. Szczegółowe dane zamieszczono w Tabeli 1.5. Istnieje przypuszczenie, iż odmowy te mogły wynikać z niechęci do ujawnienia niskich osiągnięć szkolnych uczniów lub z innych przyczyn z nimi powiązanych. Do problemu tego powrócimy w dalszych częściach rozdziału.

Ostateczna liczebność próby została ustalona podczas sporządzania list uczniów i nauczycieli przez wykonawcę podczas I etapu badania. W następnym etapie listy były weryfikowane. Z próby wykluczano uczniów, którzy zmienili szkołę lub nie mogli uczestniczyć w badaniu z innych przyczyn losowych (długotrwała choroba, niepełnosprawność uniemożliwiająca uczestnictwo w badaniu, śmierć). Przyjęto, że liczebność próby rodziców (opiekunów) jest równa liczebności uczniów. Podkreślić należy, że w każdym etapie badano tych samych uczniów, jednak próba nauczycieli różniła się w kolejnych

Tabela 1.5. Przyczyny wykluczania szkół z próby podczas rekrutacji.

Przyczyna wykluczenia	Liczba szkół
Szkoła wyłącznie z oddziałami 0, I–III	2
Brak zgody dyrektora na realizację badania	17
Odsetek zgód rodziców niższy niż 70% w badanych klasach	14
Wykluczenie z badania oddziału szkolnego (odsetek uczniów niepełnosprawnych w jednej z klas nie pozwala osiągnąć stopy realizacji badania)	2
Razem	35

etapach ze względu na zmiany kadrowe w szkołach oraz przechodzenie nauczycieli na urlopy rodzicielskie, wychowawcze oraz dla poratowania zdrowia. Badaniem objęci zostali wszyscy nauczyciele uczący języka polskiego, matematyki, przyrody oraz nauczyciele nauczania zintegrowanego (z wyjątkiem nauczycieli klas zerowych) uczący w danym momencie w szkołach. Szczegółowe dane na temat liczebności próby w kolejnych etapach przedstawione zostały w dalszych częściach rozdziału.

Przebieg badania

Każdy z etapów poprzedzony został wielomiesięczną pracą zespołu badawczego nad szczegółową koncepcją, narzędziami badawczymi oraz procedurami badania. Realizowano także badania pilotażowe. Ze względu na ograniczoną objętość niniejszego rozdziału, działania te nie zostaną tu opisane.

Realizacja badania w terenie została powierzona firmom badawczym wyłonionym w wyniku przetargów publicznych. Etap I badania przeprowadzony został przez Instytut Badania Rynku i Opinii Publicznej Millward Brown SMG/KRC Poland-Media S.A., natomiast etap III przez ASM Centrum Badań i Analiz Rynku Sp. z o.o. Wykonawca przez cały okres realizacji badania otrzymywał wsparcie zespołu badawczego w zakresie problemów w rekrutacji, prowadzenia badania, interpretacji procedur badawczych oraz postępowania w sytuacjach w nich nieprzewidzianych.

Etap I

Etap ten rozpoczął się na początku stycznia 2010 roku, natomiast zakończył w czerwcu 2010 roku. Podzielony był na trzy fazy.

W fazie pierwszej wykonawca we współpracy z zespołem badawczym rekrutował szkoły wylosowane do badania. W tym celu z Centralnej Komisji Egzaminacyjnej w Warszawie do szkół wysłane zostały listy z prośbą o zgodę na uczestnictwo w badaniu oraz list intencyjny podpisany przez ministra edukacji narodowej. Następnie wykonawca kontaktował się ze szkołami i pozyskiwał zgody dyrektorów szkół. Jeżeli dyrektor nie wyraził zgody na badanie, szkoła zastępowana była placówką z próby rezerwowej według wcześniej

opisanej procedury. Jeżeli dyrektor wyraził zgodę, wykonawca sporządzał imienne listy uczniów i nauczycieli objętych badaniem.

Druga faza polegała na pozyskiwaniu zgód rodziców uczniów oraz realizacji badania w szkołach. W trakcie tej fazy uczniowie, ich rodzice, dyrektor szkoły oraz nauczyciele wybranych przedmiotów wyrażali zgodę na uczestnictwo oraz wypełniali testy i kwestionariusze. W organizacji badania w szkole pomagał koordynator, będący jej pracownikiem (często był to dyrektor lub wicedyrektor szkoły). Był on zobowiązany zapewnić warunki niezbędne do realizacji badania, tzn. przekonać rodziców o znaczeniu i potrzebie jego przeprowadzenia, ułatwić komunikację z nauczycielami, zapewnić odpowiednie warunki lokalowe, pomóc uzyskać dostęp do szkolnej dokumentacji, wesprzeć psychologa realizującego badanie w dystrybucji narzędzi badawczych itp. We współpracy z psychologiem przekazywał on rodzicom (opiekunom) uczniów formularze zgody na udział dziecka w badaniu oraz na przetwarzanie danych osobowych wraz z listem informującym o celach i zakresie badania oraz ankietą dla rodziców. Formularze i wypełnione ankiety odbierał psycholog realizujący badanie w danej szkole. Weryfikował on, czy rodzic wyraził zgodę na badanie dziecka oraz czy ankietą została w całości wypełniona (jeżeli tak nie było, prosił o jej uzupełnienie). Jeżeli w badanych klasach odsetek oświadczeń wyrażających zgodę rodziców nie osiągnął progu 70%, szkoła była wykluczana z badania i zastępowana placówką z próby rezerwowej, zgodnie z opisaną wcześniej procedurą.

Po uzyskaniu zgody rodziców rozpoczynano badanie uczniów za pomocą testu inteligencji. W pierwszym kroku psycholog prowadzący badanie spotykał się z klasą i wyjaśniał, na czym będzie ono polegać. Następnie przystępował do indywidualnego badania każdego ucznia zgodnie z przygotowanym wcześniej harmonogramem. Realizowano je na terenie szkoły podczas lekcji lub pobytu ucznia w świetlicy, jednak nie później niż do godziny 16:00. Dziecko miało zapewnioną ciszę i komfortowe warunki, a w pomieszczeniu, w którym odbywało się badanie, przebywało wyłącznie z psychologiem. Na rozwiązanie testu miało tyle czasu, ile potrzebowało.

Psycholog przeprowadzał także badanie wśród nauczycieli i dyrektorów. Osobiście przekazywał i odbierał od nich ankiety, weryfikował poprawność ich wypełnienia, a w razie potrzeby prosił o uzupełnienie. Wychowawcom klasy przekazywał dodatkowe kwestionariusze opisu umiejętności i szkolnego funkcjonowania uczniów. W celu zapewnienia poufności nie dopuszczono do odbioru ankiet przez osoby trzecie (np. pracowników sekretariatu szkoły lub innych pracowników placówki). Wypełnione przez respondentów kwestionariusze i testy zostały oznaczone odpowiednimi numerami identyfikacyjnymi, by umożliwić połączenie danych zebranych na różnych etapach. Całość tej fazy badania została zrealizowana przez ankietatorów – psychologów mających doświadczenie w przeprowadzaniu badań inteligencji u dzieci; zostali oni przeszkoleni w zakresie procedur badawczych. W fazie trzeciej wykonawca wprowadzał odpowiedzi udzielone przez respondentów do elektronicznej bazy danych.

Realizacja badania w szkołach podlegała kontroli. Sprawdzano, czy badanie przebiega zgodnie z przekazanymi procedurami. Kontrole nie wykazały znaczących uchybień. Informacje o pojawiających się nieznaczących nieprawidłowościach były na bieżąco przekazywane wykonawcy, by mógł uczulić ankietatorów na elementy procedur, których dotyczyły. Również dokumentacja badawcza, kodowanie kwestionariuszy oraz przekazane zbiory danych podlegały szczegółowej kontroli (zob. Lisiecka, 2010).

Etap III, część 1

Ten etap badania, zgodnie z planem oznaczony numerem III, rozpoczął się w kwietniu 2012 roku. On również składał się z trzech faz. W fazie pierwszej wykonawca kontaktował się ze szkołami, żeby wyjaśnić cel i przebieg badania oraz ustalić jego harmonogram. Weryfikował także listy uczniów i nauczycieli objętych badaniem.

Druga faza polegała na realizacji badania w szkołach. W jej trakcie uczniowie, ich rodzice, dyrektor szkoły oraz nauczyciele wybranych przedmiotów wypełniali testy i kwestionariusze. W organizacji badania w szkole pomagał koordynator, będący jej pracownikiem. Badanie wśród uczniów trwało dwa dni powszednie (niekoniecznie następujące po sobie) i miało charakter audytoryjny. Pierwszego dnia przeprowadzono badanie inteligencji. Uczniowie wypełniali je w grupach nie większych niż 15-osobowe, a sesja badawcza nie była ograniczona czasowo (uczniowie mieli na wypełnienie testu tyle czasu, ile potrzebowali). Drugiego dnia, podczas dwóch 35-minutowych sesji przedzielonych przerwą (10-minutową), uczniowie wypełniali kwestionariusze. W czasie badania w klasie przebywali wyłącznie uczniowie nim objęci, ankieter, psycholog-ankieter, nauczyciel i (niekiedy) osoba przeprowadzająca kontrolę zgodności przebiegu badania z procedurami. Uczniom z klas wylosowanych do badania, których rodzice nie wyrazili zgody na uczestnictwo, zapewniono opiekę na terenie szkoły. Wykonawca organizował też sesje uzupełniające, w których uczestniczyli uczniowie nieprzebadani w standardowym terminie.

Kwestionariusz dla rodziców (opiekunów) uczniów objętych badaniem był im przekazywany za pośrednictwem dzieci. W celu zapewnienia poufności znajdował się on w kopercie, by rodzic po wypełnieniu ankiety mógł ją zakleić i bez obaw zwrócić do szkoły za pośrednictwem dziecka (koperty zbierał szkolny koordynator badania). Kwestionariusze dla nauczycieli były im przekazywane (w niezaklejonych kopertach) oraz odbierane od nich osobiście przez psychologa. W celu zapewnienia poufności nie dopuszczono odbioru ankiet nauczycielskich przez osoby trzecie. Dyrektorzy szkół wypełniali skierowany do nich kwestionariusz oraz uczestniczyli w krótkim wywiadzie standaryzowanym. Kwestionariusze i testy wypełniane przez respondentów zostały oznaczone odpowiednimi numerami identyfikacyjnymi, by umożliwić połączenie danych zebranych na różnych etapach. Całość tej fazy badania została zrealizowana przez ankieterów – psychologów mających doświadczenie w przeprowadzaniu badań z udziałem dzieci; zostali oni przeszkoleni w zakresie procedur badawczych. W ostatniej fazie wprowadzano odpowiedzi udzielone przez respondentów do elektronicznej bazy danych.

Podobnie jak w I etapie, realizacja badania w szkołach podlegała kontroli. Nie wykazała ona znaczących uchybień. Informacje o pojawiających się nieznacznych nieprawidłowościach były na bieżąco przekazywane wykonawcy, by mógł uczulić ankieterów na elementy procedur, których dotyczyły. Również dokumentacja badawcza, kodowanie kwestionariuszy oraz przekazane zbiory danych podlegały szczegółowej kontroli (zob. Hawrot, 2013).

Etap III, część 2

Druga część tego etapu rozpoczęła się w sierpniu 2012 roku. Podobnie jak część pierwsza, była także podzielona na fazy. W fazie pierwszej wykonawca kontaktował

się z dyrektorami szkół, przypominał im cel badania, informował o jego przebiegu, weryfikował listy uczniów oraz ustalał harmonogram badania. Podczas drugiej fazy testowano osiągnięcia szkolne uczniów. Ankieterzy przeprowadzali badanie w szkołach, w terminach ustalonych z dyrektorami placówek. W organizacji badania w szkole pomagał szkolny koordynator. Testowanie trwało trzy następujące po sobie dni powszednie i miało charakter audytoryjny. Pierwszego dnia każdy uczeń wypełniał dwa zeszyty testu z języka polskiego, drugiego dnia – dwa zeszyty testu matematycznego, natomiast trzeciego – dwa zeszyty testu z języka polskiego. Każdy test miał dwie wersje (A oraz B). Część zadań w obu wersjach była taka sama, pozostałe występowały tylko w jednej z nich. Każdy uczeń rozwiązywał jedną wersję wszystkich testów (A lub B). Schemat przydziału wersji testów do uczniów ustalany był podczas pierwszego dnia. W czasie badania w klasie przebywali wyłącznie uczniowie nim objęci, psycholog-ankieter, nauczyciel i (niekiedy) osoba przeprowadzająca kontrolę zgodności przebiegu badania z procedurami. Uczniom z oddziałów wylosowanych do badania, których rodzice nie wyrazili zgody na uczestnictwo, zapewniono opiekę na terenie szkoły. Wykonawca, w celu usprawnienia przebiegu badania, mógł poddać testowaniu także uczniów z niego wykluczonych, lecz nie było to wymagane. Dla uczniów, którzy nie mogli wziąć udziału w badaniu w standardowym terminie, wykonawca zorganizował sesje uzupełniające. Testy wypełniane przez respondentów zostały oznaczone odpowiednimi numerami identyfikacyjnymi, by umożliwić połączenie danych zebranych na różnych etapach. Ankieterzy przeprowadzający testowanie mieli doświadczenie w realizacji badań z udziałem dzieci oraz zostali wcześniej przeszkoleni w zakresie procedur badawczych.

Podobnie jak w innych fazach badania, jego realizacja w szkołach podlegała kontroli pod kątem weryfikacji przestrzegania przez ankieterów procedur badawczych. Kontrola ta wykazała pewne nieprawidłowości, m.in. spóźnienia ankieterów (w jednej szkole z tego powodu badanie zostało przesunięte o jeden dzień), jednak w większości przypadków uchybienia nie wpłynęły na jakość zebranych danych. Z powodu nieprzestrzegania procedur badawczych jeden z ankieterów, na żądanie zespołu badawczego, został zastąpiony inną przeszkoloną osobą o wymaganych kompetencjach. Informacje o pojawiających się uchybieniach były na bieżąco przekazywane wykonawcy, by mógł uczulić ankieterów na elementy procedury, których dotyczyły. Dokumentacja ze szkół, w których badania przeprowadzali ankieterzy, wobec których kontrola wykazała uchybienia, została szczególnie dokładnie sprawdzona podczas kontroli materiałów badawczych.

W ostatniej fazie drugiej części badania, rozpoczętej jeszcze podczas realizacji badania w szkołach, zespół koderów kodował wyniki testów osiągnięć szkolnych. Koderzy mieli doświadczenie w zakresie kodowania zdobyte w ogólnokrajowych lub międzynarodowych badaniach osiągnięć szkolnych lub figurowali w ewidencji egzaminatorów sprawdzianu lub egzaminu gimnazjalnego jednej z okręgowych komisji egzaminacyjnych (OKE). Zespół składał się z 38 koderów testów polonistycznych oraz 20 koderów testów matematycznych. Zostali oni przeszkoleni w zakresie zasad kodowania zadań. Kodowanie przebiegło w dwóch krokach. W pierwszym, tzw. kodowaniu próbnym, zakodowane zostały zadania otwarte z 480 zeszytów testowych (po 40 z każdego rodzaju). W jego toku koderzy pod okiem szkolących dokonywali interpretacji klucza kodowego. Kodowanie to zostało zweryfikowane przez zespół badawczy. Po jego zatwierdzeniu rozpoczęło się kodowanie zasadnicze. Zeszyty testowe z każdej badanej klasy były równomiernie rozłożone między wszystkich koderów

kodujących zeszyty danego rodzaju. Koderzy pracowali niezależnie i nie mogli się ze sobą kontaktować. Następnie zespół badawczy wylosował 20% zeszytów, które zostały zakodowane powtórnie, co umożliwiło kontrolę rzetelności kodowania.

Pomiar osiągnięć szkolnych

Do pomiaru osiągnięć szkolnych wykorzystano standaryzowane testy, których właściwości zostały przedstawione w artykule Aleksandry Jasińskiej i Michała Modzelewskiego (2013). W tym miejscu zaprezentujemy więc tylko najważniejsze informacje na ich temat. Najważniejszym elementem zapewniającym ich jakość było duże badanie pilotażowe, które pozwoliło wybrać do ostatecznych wersji testów zadania o najlepszych właściwościach pomiarowych. Podczas konstrukcji narzędzi wykorzystano metodologię *item response theory* (IRT). Powstałe testy mają bardzo dobre właściwości psychometryczne. Ich rzetelność jest wystarczająco wysoka jak na potrzeby badawcze. Współczynniki rzetelności wyliczone na podstawie analizy IRT (Adams, 2005) wyniosły od 0,86 do 0,88.

Przeprowadzone analizy potwierdziły także trafność narzędzi. Zaobserwowano umiarkowanie silne, pozytywne związki między nauczycielskimi ocenami a wynikami testów z danych przedmiotów. Dodatkowym potwierdzeniem była silniejsza korelacja dla odpowiadających sobie przedmiotów (np. między wynikiem testu z matematyki a nauczycielską oceną umiejętności matematycznych ucznia) niż dla różnych przedmiotów. Stwierdzono także związki z innymi miarami tych samych konstruktów, czyli wynikami testów z *Ogólnopolskiego badania umiejętności trzecioklasistów* (OBUT)⁴. Wykorzystane testy osiągnięć szkolnych składały się z trzech skal pomiarowych: skali umiejętności czytania, skali świadomości językowej i skali umiejętności matematycznych. Poniżej scharakteryzowano krótko każdą z nich.

Test umiejętności czytania

Skala tego testu mierzyła poziom rozumienia znaczenia czytanych samodzielnie tekstów różnego typu: literackich (prozatorskich i poetyckich), popularnonaukowych i użytkowych (ogłoszenie, regulamin, ulotka). Niektóre z nich nie miały tradycyjnego charakteru tekstu ciągłego, ale zawierały dodatkowe informacje w ramach i przypisach. Pytania do każdego tekstu sprawdzały różne kompetencje. Najwięcej (35%) wymagało od uczniów wyszukania informacji zawartej w tekście, podjęcia decyzji, które informacje są ważne, a które nie (ze względu na ich związek z tematem lub pytaniem), a także ustalenia kolejności wydarzeń. Inne zadania (33%) mierzyły umiejętność interpretacji tekstu, tj. wydobycia i wyjaśnienia jego sensu, określenia tematu i głównej myśli utworu, porównywania informacji zawartych w tekście, dostrzegania i wyjaśniania przyczyn i skutków zdarzeń, sytuacji, zjawisk w nim opisanych, podania

⁴ Jest to badanie realizowane w ramach projektu systemowego „Badanie umiejętności podstawowych uczniów trzeciej klasy szkoły podstawowej” współfinansowanego przez Unię Europejską z Europejskiego Funduszu Społecznego. Więcej o badaniu na stronie: www.obut.edu.pl

przypuszczalnych motywów działania, zachowania lub postawy bohaterów. Najmniej (32%) odwoływało się do umiejętności dokonania refleksji nad tekstem i jego oceny. Sprawdzały one to, czy uczeń potrafi odnieść tekst do własnego doświadczenia i wiedzy o świecie, czy umie dokonać oceny zdarzeń, postaci i poglądów w kontekście własnego doświadczenia czytelniczego i pozaszkolnego, jak również ocenić kompletność i spójność tekstu.

Test świadomości językowej

Skala testu świadomości językowej składała się z trzech rodzajów zadań: sprawdzających zasób słownikowy uczniów, elementy wiedzy o języku oraz umiejętności związane z pisaniem tekstów.

Wśród zadań mierzących bogactwo słownikowe znalazły się zadania: wymagające utworzenia lub wybrania spośród podanych słów wyrazu o podobnym lub przeciwstawnym znaczeniu; zadania wymagające dobrania poprawnej definicji; zadania wymagające tworzenia, rozpoznawania lub wyjaśniania powszechnie wykorzystywanych porównań i związków frazeologicznych; zadania wymagające rozpoznania niepoprawnego użycia języka.

Sprawdzeniu elementów wiedzy o języku służyły zadania mierzące umiejętność tworzenia i uzupełniania zdań zgodnie z zasadami składni, wyróżniania w tekście i tworzenia zdań oznajmujących, pytających, rozkazujących, a także zadania z zakresu ortografii i interpunkcji oraz poprawności językowej.

Do pomiaru umiejętności związanych z pisaniem tekstów wykorzystano zadania sprawdzające umiejętność redagowania tekstu, dzielenia wypowiedzi na zdania, rozpoznawania i tworzenia czytelnej struktury tekstu, rozpoznawania i nadawania poprawnego stylu wypowiedzi oraz umiejętności argumentowania.

Test umiejętności matematycznych

Każde zadanie testu umiejętności matematycznych można opisać za pomocą dwóch kategorii: umiejętności, którymi należało się posłużyć, by je rozwiązać, oraz treści matematycznych, do których się odwoływało. Zadania mierzyły trzy grupy umiejętności: odtwarzania, powiązania i rozumowania.

Najmniej zadań (21%) odwoływało się do umiejętności odtwarzania wiadomości i dobrze wyćwiczonych schematów, którą zdefiniowano jako umiejętność rozwiązywania zadań typowych, wymagających użycia wyćwiczonych, prostych technik i posłużenia się dobrze znanymi obiektami. Innymi słowy – przywołania z pamięci znanych pojęć lub algorytmów. Najwięcej poleceń (56%) wymagało odwołania się do umiejętności powiązania różnych wiadomości i dobrze wyćwiczonych schematów na potrzeby rozwiązania zadań mniej rutynowych, ale niezbyt odległych od zadań typowych. Uczniowie musieli zwykle wykonać większą liczbę kroków, aby rozwiązać zadanie. Musieli wybrać pojęcia (modele, wzory, procedury) matematyczne odpowiednie dla rozwiązania danego problemu. Od ucznia oczekiwano, że będzie potrafił wykorzystać posiadane wiadomości do rozwiązania zadań, z których nie wynika wprost, jakie pojęcia czy procedury powinno się zastosować, że

będzie umiał porównać ze sobą informacje (zarówno podane w zadaniu, jak i przywołane z pamięci) i zinterpretować je w celu znalezienia rozwiązania.

Umiejętność przeprowadzenia prostego rozumowania matematycznego składającego się z kilku kroków była sprawdzana przez 23% zadań. Od ucznia oczekiwano, że będzie umiał ustalić kolejność czynności prowadzących do rozwiązania problemu (sytuacji nowej, nieoczywistej), że będzie potrafił wyciągnąć wnioski z kilku informacji podanych w różnej postaci.

Zadania mierzące każdą z umiejętności odwoływały się do różnych treści matematycznych, które podzielono na trzy grupy: ilość, przestrzeń i kształt oraz zmiana, związki, zależności.

Obszar „ilość” odnosił się do rozumienia przez uczniów pojęcia liczby, rozumienia i odkrywania relacji między liczbami, umiejętności wykonywania obliczeń oraz rozumienia znaczenia tych operacji, a także umiejętności wykorzystania opisanych kompetencji w sytuacjach praktycznych. W obszarze tym mieściły się także zagadnienia związane z pomiarem właściwości fizycznych przedmiotów. Odwoływały się one do rozumienia problematyki długości, ciężaru, objętości, temperatury i czasu.

Na obszar „przestrzeń i kształt” składały się problemy geometryczne oraz związki przestrzenne między obiektami. Obejmował on umiejętność rozpoznawania i rysowania figur geometrycznych, dostrzegania symetrii i regularności oraz wymagał wykorzystania wyobraźni przestrzennej (zadania te nie wymagały zastosowania wiedzy formalnej).

Zakres obszaru „zmiana, związki, zależności” dobrze oddaje cytat z podstawy programowej kształcenia ogólnego (Podstawa programowa z komentarzami, b.d., s. 34): „czynny udział w zdobywaniu wiedzy matematycznej przybliży dziecko do matematyki, rozwija kreatywność, umożliwi samodzielne odkrywanie związków i zależności; duże możliwości samodzielnych obserwacji i działań stwarza geometria, ale także w arytmetyce można znaleźć obszary, gdzie uczeń może czuć się odkrywcą”. Na obszar ten składa się więc umiejętność dostrzegania przez ucznia związków i zależności, reprezentowanych w sposób graficzny, słowny, tabelaryczny lub symboliczny.

Wyniki każdego z testów zostały wyskalowane jednoparametrycznym modelem IRT (modelem Rascha oraz modelem *partial credit* dla zadań wielopunktowych; patrz np.: De Ayala, 2009) za pomocą programu ConQuest 2.0. Wykorzystano metodę estymacji *marginal maximum likelihood*. W analizach prezentowanych w raporcie wykorzystano punktowe estymatory umiejętności uczniów EAP (*expected a posteriori*). Przeliczono je dla każdego testu na skalę o średniej 100 i odchyleniu standardowym 15 w populacji uczniów. W Tabeli 1.6 przedstawiono podstawowe statystyki opisowe oraz histogramy dla wyników trzech testów, wykorzystanych w analizach prezentowanych w książce.

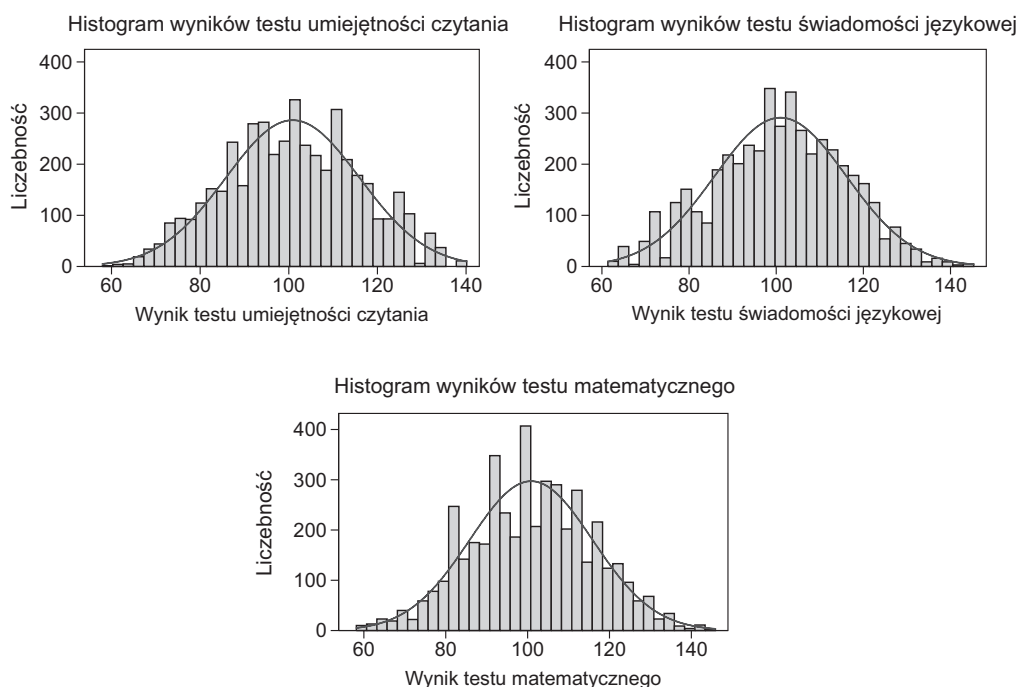
Reprezentatywność próby

Jednym z poważniejszych problemów badań panelowych jest zjawisko polegające na zmniejszaniu się wraz z kolejnymi etapami badania odsetka osób do niego zakwalifikowanych (wylosowanych), które udaje się przebadać. Stanowi to poważne zagrożenie dla reprezentatywności, a tym samym dla możliwości uogólniania wniosków na całą populację.

Tabela 1.6. Statystyki opisowe dla wyników testów.

Test	n	M	SD	Q_1	Me	Q_3	Min	Max	Skośność	Kurtoza
Umiejętności czytania	4 604	100,12	15,08	89,26	100,16	111,01	57,92	140,15	0,046	2,47
Świadomości językowej	4 604	100,13	15,11	89,64	99,96	111,34	61,36	145,33	-0,048	2,51
Umiejętności matematycznych	4 468	100,09	15,05	89,36	100,10	111,16	58,16	145,86	-0,013	2,74

Rysunek 1.1. Histogramy wyników testów osiągnięć.



W celu określenia skali problemu, po każdym etapie przeprowadzono analizę stóp realizacji. Ponadto dokonano weryfikacji reprezentatywności próby wylosowanej i zrealizowanej.

Stopy realizacji badania

W celu osiągnięcia jak najwyższych stóp realizacji badania określono minimalne odsetki przebadanych szkół i respondentów wymagane od wykonawcy, by dany etap (lub jego część) został uznany za zrealizowany należycie. Ponadto wysokość wynagrodzenia wykonawcy uzależniono od osiągniętych stóp realizacji.

Przeanalizowano, jaką liczbę respondentów udało się przebadać w stosunku do liczby respondentów wylosowanych do badania.

1. Organizacja i metodologia badania podłużnego w szkołach podstawowych

Tabela 1.7. Odsetki uczniów i rodziców przebadanych w kolejnych etapach w stosunku do liczby uczniów i rodziców w wylosowanych szkołach i klasach.

Grupa	n*	I etap		III etap, cz. 1		III etap, cz. 2	
		Liczba przebadanych	%**	Liczba przebadanych	%**	Liczba przebadanych	%**
Uczniowie	6 648	5 600	84,24	4 942	74,34	4 894	73,62
Rodzice uczniów	6 648	5 760	86,64	4 713	70,89	—	—

* Liczba respondentów w wylosowanych szkołach i klasach w I etapie badania.

** Odsetek przebadanych w danym etapie w stosunku do liczby respondentów w wylosowanych szkołach i klasach w I etapie badania.

W przypadku uczniów odsetek przebadanych w stosunku do wylosowanych w I etapie badania wyniósł ponad 84%, natomiast w kolejnych etapach był o około 10 punktów procentowych niższy (odpowiednio 74,34% i 73,62%). W I etapie udało się przebadać ponad 86% rodziców uczniów z wylosowanych klas i szkół, natomiast w III etapie, cz. 1 – blisko 71%. Wziąwszy pod uwagę podłużny charakter badania, są to dobre wyniki. Dane znajdują się w Tabeli 1.7.

Z próby wykluczono uczniów, których rodzice nie wyrazili zgody na uczestnictwo w badaniu. Ponadto w każdym etapie z próby wykluczano uczniów, którzy z przyczyn losowych trwale nie mogli uczestniczyć w badaniu (np. z powodu zmiany szkoły, długotrwałej choroby, niepełnosprawności uniemożliwiającej uczestnictwo w badaniu, śmierci). Z tego powodu liczebność próby zmniejszała się – próba uczniów drugiej części III etapu stanowiła 90,36% liczebności próby uczniów, dla których uzyskaliśmy zgodę na uczestnictwo w badaniu w I etapie. Szczegółowe dane zawarto w Tabeli 1.8.

W I etapie badanie zrealizowane zostało wśród 97% uczniów zakwalifikowanych do próby, 100% ich rodziców⁵, 95% nauczycieli oraz 100% dyrektorów szkół, w których się uczyli. W części pierwszej III etapu odsetki te były nieco niższe, lecz wciąż satysfakcjonujące. Udało się przebadać ponad 91% uczniów, którzy w próbie pozostali, 97% rodziców, blisko 90% nauczycieli oraz wszystkich dyrektorów. W stosunku do pierwotnej liczebności próby odsetki przebadanych uczniów i ich rodziców wyniosły odpowiednio 86% i 82%. W drugiej części III etapu badania (objęto nim tylko uczniów) stopa realizacji wyniosła 94% liczebności próby w tym etapie (85% w stosunku do liczebności próby w I etapie). Odsetki te są wysokie i należy je uznać za satysfakcjonujące. Nie dają one jednak pełnego obrazu realizacji próby, bowiem nie pokazują, na ile zebrane dane są kompletne. Tymczasem w analizach wykorzystuje się zwykle dane zebrane w ramach więcej niż jednego etapu, więc wszelkie braki danych etapowych wiążą się z ograniczeniem liczby obserwacji uwzględnianych w analizach. W związku z tym zweryfikowano liczebność uczniów, od których dane pozyskano we wszystkich trzech etapach badania. Wyniosła ona 4648 osób (przy uwzględnieniu jedynie danych uczniów), co stanowi 69,92% pierwotnej liczebności próby uczniów oraz 80,69% liczby uczniów, którzy nie zostali wykluczeni z badania w I etapie. Przy jednoczesnym uwzględnieniu

⁵ Tak wysoki próg realizacji wynika z faktu, że rodzice, udzielając zgody, jednocześnie wypełniali przeznaczony dla nich kwestionariusz.

Tabela 1.8. Stopy realizacji badania wśród respondentów, którzy nie zostali wykluczeni z próby badawczej.

Grupa	I etap		III etap, cz. 1			III etap, cz. 2		
	<i>n</i> *	Stopa realizacji: liczba (%)	<i>n</i> *	Stopa realizacji: liczba (%)	Stopa realizacji w stosunku do <i>N</i> w I etapie	<i>n</i> *	Stopa realizacji: liczba (%)	Stopa realizacji w stosunku do <i>N</i> w I etapie
Uczniowie	5 760	5 600 (97,22)	5 399	4 942** (91,54)	85,80%	5 205	4 894 (94)	84,97%
Rodzice uczniów	5 760	5 760 (100)	5 399	4 713 (87,29)	81,73%	—	—	—
Nauczyciele	2 557	2 432 (95,11)	2 553	2 295 (89,89)	nd.***	—	—	—
Dyrektorzy	180	180 (100)	180	180 (100)	nd.***	—	—	—

* Liczebność próby w danym etapie (liczba respondentów, którzy wyrazili zgodę na badanie i nie zostali wykluczeni z przyczyn losowych).

** Uczniowie, którzy wypełnili wszystkie przewidziane testy i narzędzia.

*** Skład próby zmieniał się z powodu zmian w kadrze pedagogicznej w szkołach.

zarówno danych uczniów, jak i rodziców, liczba kompletnych obserwacji spadła do 4292, co stanowi 64,56% pierwotnej liczebności próby uczniów i 74,51% w stosunku do liczebności uczniów niewykluczonych z badania w I etapie. Odsetki te wskazują na znaczne ubytki w próbie, wciąż jednak są wysokie w porównaniu ze stopami realizacji osiąganymi w innych badaniach podłużnych (np. Karwowski, 2013).

W celu określenia, jak dalece możliwe będzie przeprowadzanie analiz na danych zagregowanych oraz na poziomie ponadindywidualnym, przeprowadzono analizę realizacji badania na poziomie oddziałów i szkół. W I etapie badania stopa realizacji wśród uczniów oraz nauczycieli przekroczyła 80% w każdej szkole, a w przypadku rodziców i dyrektorów wyniosła 100%.

W części pierwszej III etapu badanie zostało przeprowadzone we wszystkich szkołach, jednakże w 9 oddziałach z 7 szkół nie zostały osiągnięte wymagane minimalne progi realizacji badania wśród rodziców (wymagany próg: 75%; najniższa stopa realizacji w oddziale – 47%), w 10 oddziałach z 10 szkół w przypadku uczniów (wymagany próg: 80%; najniższa stopa realizacji w oddziale – 75%) oraz w 1 szkole w przypadku nauczycieli (wymagany próg: 75%; stopa realizacji w szkole – 60%). Łącznie, pomimo przeprowadzenia badania we wszystkich szkołach, nie osiągnięto minimalnych wymaganych progów realizacji w 16 z nich. Przypomnijmy jednak, że udało się osiągnąć wysokie stopy realizacji w całej próbie. W części drugiej III etapu odsetki realizacji we wszystkich 180 szkołach przekroczyły 80%, a w części z nich osiągnęły 100% (22 placówki). Jest to bardzo dobry rezultat.

1. Organizacja i metodologia badania podłużnego w szkołach podstawowych

Podsumowując, stopy realizacji badania na poziomie indywidualnym oraz oddziałów i szkół są satysfakcjonujące. W stosunku do liczby wylosowanych respondentów, odsetki przebadanych nie spadły poniżej progu 70%. Ponadto osiągnięto wysokie odsetki realizacji badania wśród respondentów, którzy pozostali w próbie. Innymi słowy, choć zmniejszenia się liczebności próby nie udało się całkowicie uniknąć, zostało ono znacznie ograniczone.

Podejście do braków danych

Ze względu na bardzo wysokie stopy realizacji badań panelowych oraz niewielkie odsetki braków danych możliwe było usuwanie w poszczególnych analizach wszystkich przypadków z brakami (*listwise deletion*). Podejście takie, w sytuacji niskich stóp realizacji badania, może prowadzić do obciążenia szacowanych parametrów. Jednakże przy wysokiej jakości zgromadzonych danych prowadzi do uzyskiwania w analizach nieobciążonych statystyk (Allison, 2001).

Ograniczenie próby badawczej

Zdecydowano o wykluczeniu z analiz prezentowanych w niniejszej książce 16 szkół ze względu na niewiarygodne wyniki pomiaru inteligencji uczniów uzyskane w I etapie badania. Przemawiały za tym następujące argumenty.

Po pierwsze, analiza wariancji wewnątrzgrupowej (*intra-class correlation, ICC*) wykażowanego pierwszego pomiaru inteligencji w modelu dwupoziomowym (uczniowie–szkoły) dla pełnej próby wskazywała, że niemal 19,5% zmienności inteligencji uczniów w pierwszej klasie wyjaśnia ich przypisanie do szkół. Analogiczne analizy na danych z badania *Szkolnych uwarunkowań efektywności kształcenia* (SUEK) wskazują, że należy się spodziewać zróżnicowania na poziomie niespełna 15%. Tymczasem analiza porównawcza innych danych zebranych w ramach obu badań w szkołach podstawowych wskazuje, że ich rozkłady są porównywalne.

Po drugie, w trzeciej klasie szkoły podstawowej (wśród tych samych uczniów) podział na szkoły wyjaśnia jedynie 8,8% zmienności wyników testu inteligencji.

Zaobserwowano także efekty podważające jakość danych zebranych w poszczególnych szkołach, np. w jednej z placówek średni poziom inteligencji uczniów w klasie trzeciej był niższy niż w klasie pierwszej. Ponadto w siedmiu szkołach średni wynik w pomiarze pierwszym był o ponad dwa odchylenia standardowe wyższy, a w dwóch szkołach o ponad dwa niższy niż średni wynik w całej próbie. W dwóch szkołach wewnętrzne zróżnicowanie wyników stanowiło od $\frac{1}{3}$ do $\frac{1}{2}$ odchylenia standardowego wyników w całej próbie (było więc bardzo małe), zaś w jednej szkole – niemal dwukrotnie większe niż w próbie. Dodatkowo szkoły o nietypowych wartościach pomiaru inteligencji były zlokalizowane w kilku obszarach geograficznych.

W efekcie podjęto decyzję o przeprowadzeniu pogłębionej analizy tak wysokiego międzyszkolnego zróżnicowania wyników pomiaru inteligencji. Przyjęto 6 wskaźników na poziomie szkół mogących ujawniać nieprawidłowości w przeprowadzeniu pomiaru inteligencji w I etapie. Wskaźniki te to:

- wyliczona dla szkoły średnia z reszt z regresji, w której zmienną wyjaśnianą jest poziom inteligencji uczniów w I etapie badania, a zmiennymi wyjaśniającymi są: wynik pomiaru inteligencji w III etapie oraz inne zmienne, zgodnie z teorią skorelowane z inteligencją;
- alfa Cronbacha mniejsza niż 0,9 dla skali zbudowanej z odpowiedzi na pytania testu inteligencji łącznie w pierwszej i trzeciej klasie;
- korelacja pomiaru inteligencji w pierwszej klasie z pomiarem w trzeciej klasie na poziomie szkoły mniejsza niż 0,55;
- wartość odległości Cooka (*Cook's D*) powyżej 0,0222 dla szkół (poziom drugi) na podstawie analizy MLT pakietu Stata 11;
- średni wynik dla szkoły wystandaryzowany w pierwszym pomiarze powyżej $abs(1,96)$; wystandaryzowane odchylenie standardowe w szkole w pierwszym pomiarze powyżej $abs(1,96)$.

Za kryterium usunięcia z analizy przyjęto wyróżnienie szkoły jako nietypowej na przynajmniej dwóch spośród przedstawionych powyżej wskaźników. Ponadto przyjęto ogólną zasadę, że usunięcie części szkół nie powinno istotnie zmienić zróżnicowania międzyszkolnego w pomiarze inteligencji w III etapie badania.

W efekcie usunięto z analiz dane 16 szkół. Współczynnik ICC dla pierwszego pomiaru po usunięciu tych szkół wyniósł 15,49% i był o 4,00% niższy niż w pełnej próbie. Jednocześnie ICC dla pomiaru inteligencji w trzeciej klasie zmniejszyło się do 8,45%, czyli tylko o 0,41%.

Przeprowadzono również porównanie korelacji wyników pomiarów inteligencji w I i III etapie z osiągnięciami uczniów oraz podstawowymi wskaźnikami statusowymi w grupie szkół wykluczonych i pozostawionych w analizach. Dane te znajdują się w Tabeli 1.9. Wszystkie współczynniki okazały się istotne statystycznie.

Współczynniki korelacji wyników Testu Matrycy Ravena w pierwszej klasie ze zmiennymi statusowymi w grupie szkół pozostawionych w analizie są znacznie bardziej zbliżone do analogicznych współczynników korelacji drugiego pomiaru inteligencji niż odpowiednie współczynniki w grupie szkół usuniętych z analizy. Wskazuje to, iż usunięcie części szkół z analizy wprowadziło oczekiwane zmiany. Wspiera także przypuszczenie o nieprawidłowościach w pomiarze inteligencji w I etapie w tych szkołach.

Analiza reprezentatywności próby

Reprezentatywności zrealizowanej próby sprawdzono na podstawie wyników sprawdzianu umiejętności po klasie szóstej w szkołach podstawowych, które znalazły się w próbie. Dzięki temu można było sprawdzić, czy pod względem osiągnięć uczniów przebadane szkoły nie różnią się znacząco od całej populacji szkół. Ten sposób postępowania wydaje się być uzasadniony w świetle celu przeprowadzonego badania, którym było przygotowanie kontekstowych modeli wskaźników efektywności nauczania. Wymagają one bowiem danych na temat osiągnięć szkolnych uczniów reprezentatywnych dla populacji szkół.

Badanie reprezentatywności przeprowadzono z wykorzystaniem wyników sprawdzianu szóstoklasisty w 2013 i 2012 roku. Reprezentatywność sprawdzono na dwa sposoby. Pierwszy z nich polegał na porównaniu rozkładów sprawdzianu w badanych szkołach oraz ich rozkładów w następujących grupach:

1. Organizacja i metodologia badania podłużnego w szkołach podstawowych

Tabela 1.9. Porównanie korelacji obydwu pomiarów inteligencji uczniów z niektórymi zmiennymi wykorzystywanymi w modelach analitycznych w grupie szkół usuniętych i pozostawionych w próbie*.

Korelaty	Szkoły			
	Usunięte		Pozostawione	
	Test Matryc Ravena		Test Matryc Ravena	
	I etap	III etap	I etap	III etap
Test Matryc Ravena w III etapie badania	0,482	1	0,684	1
Wynik testu ze świadomości językowej	0,372	0,535	0,494	0,567
Wynik testu z czytania	0,289	0,487	0,470	0,544
Wynik testu z matematyki	0,381	0,571	0,577	0,661
SES rodziny ucznia	0,163	0,271	0,304	0,306
Wyższy z prestiżów zawodów rodziców	0,161	0,272	0,276	0,279
Wyższa z liczby lat nauki rodziców	0,253	0,377	0,333	0,350
Zasobność gospodarstwa domowego	0,162	0,298	0,301	0,301

* Pogrubieniem zaznaczono korelacje istotne na poziomie 0,01 (dwustronnie).

- cała populacja uczniów;
- populacja uczniów z wyłączeniem uczniów z typów szkół wykluczonych z operatu;
- uczniowie z wylosowanej próby szkół;
- uczniowie ze zrealizowanej próby szkół;
- uczniowie ze zrealizowanej próby szkół, z wyłączeniem uczniów 16 usuniętych szkół;
- uczniowie ze szkół, które odmówiły udziału w badaniu.

W drugim kroku skupiono się na analizach zróżnicowania wariancji międzyszkolnej. Analizy te wykonano dla grup uczniów od 1 do 5.

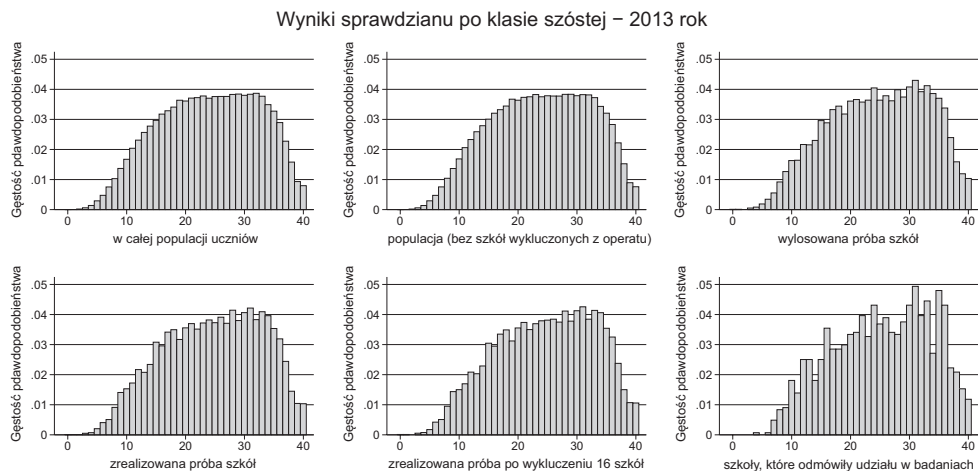
W analizach wykorzystano surowe wyniki sprawdzianu oraz informacje o szkołach pochodzące z bazy wyników egzaminacyjnych przygotowywanej na potrzeby badania *Zrównywanie wyników egzaminów zewnętrznych*⁶ prowadzonego w Instytucie Badań Edukacyjnych. Dane te zostały przygotowane na podstawie danych pozyskiwanych z OKE.

Na Rysunku 1.2 znajdują się histogramy przedstawiające rozkłady wyników uczniów na sprawdzianie w klasie szóstej w 2013 roku. Wykresy znajdujące się na Rysunku 1.1 (zaczynając od lewej strony) sporządzono dla następujących grup szkół:

- cała populacja uczniów;
- populacja uczniów, z wyłączeniem uczniów z typów szkół wykluczonych z operatu;
- uczniowie z wylosowanej próby szkół;
- uczniowie ze zrealizowanej próby szkół;

⁶ Projekt systemowy „Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego” współfinansowany ze środków Europejskiego Funduszu Społecznego (POKL, priorytet III, poddziałanie 3.1.1).

Rysunek 1.2. Rozkłady wyników sprawdzianu w 2013 roku w różnych grupach uczniów.



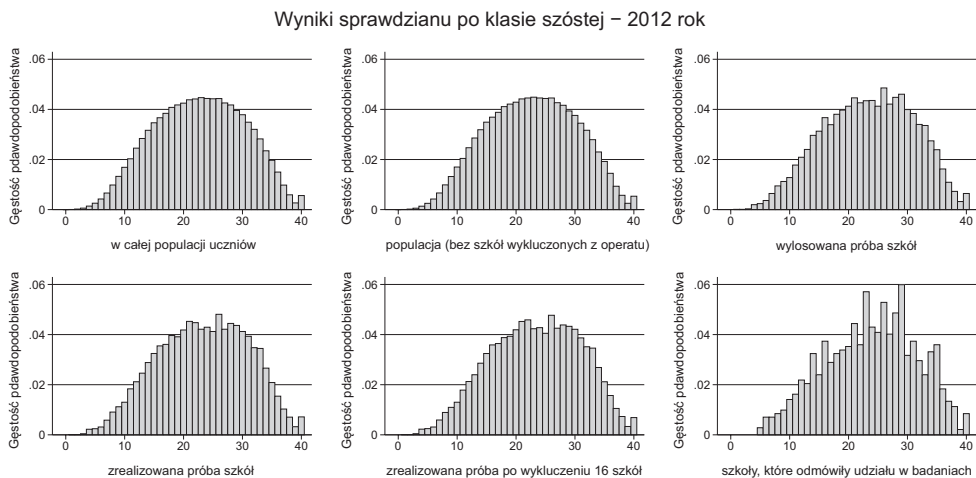
5. uczniowie ze zrealizowanej próby szkół, z wyłączeniem uczniów 16 usuniętych szkół;
6. uczniowie ze szkół, które odmówiły udziału w badaniu.

Punktem odniesienia w analizach jest wykres drugi, opracowany na podstawie wyników całej populacji uczniów, po wyłączeniu szkół nieuwzględnionych w losowaniu. Porównanie go do pozostałych histogramów pozwala stwierdzić, w jakim stopniu próba, na której zostały przeprowadzone analizy, jest reprezentatywna w stosunku do populacji. Rozkłady wyników uczniów w wylosowanej (3), zrealizowanej (4) oraz zrealizowanej próbie po wykluczeniu 16 szkół (5) są w swoim kształcie zbliżone do ustalonego wcześniej punktu odniesienia. Najistotniejsze, iż rozkład wyników w grupie 5 nie odbiega znacząco od histogramu wyników w grupie 2.

Do podobnych wniosków prowadzi analiza statystyk opisowych wyników sprawdzianu we wskazanych grupach szkół w 2013 roku dla wyszczególnionych powyżej grup uczniów (Tabela 1.10). Podstawowe parametry statystyczne są bardzo zbliżone do siebie w grupach 2, 4 i 5. Po pierwsze, oznacza to, że usunięcie 16 szkół z analiz nie wpłynęło znacząco na reprezentatywność pozostałej grupy szkół. Po drugie, widać, że szkoły, w których badania zostały zrealizowane, mają bardzo podobny rozkład wyników sprawdzianu w 2013 roku do populacji. Co prawda średnie wyniki w populacji są nieco niższe niż w grupach 4 i 5, jednak te różnice są bardzo niewielkie, a pozostałe parametry statystyczne są do siebie bardzo zbliżone.

Analogiczne kroki wykonano dla wyników sprawdzianu z 2012 roku. Analiza histogramów (przedstawionych na Rysunku 1.3) wskazuje, że zrealizowana próba uczniów, zarówno pełna (4), jak i po wykluczeniu 16 szkół (5), ma bardzo zbliżony rozkład do populacji (2) pod względem wyników sprawdzianu. Sugeruje to, iż nie ma różnic między analizowanymi grupami szkół pod względem rozkładu tej zmiennej.

Analiza statystyk opisowych (Tabela 1.11) prowadzi nas do takich samych wniosków. Wyniki uczniów w zrealizowanych szkołach są nieco wyższe niż w populacji, ale różnice te

Rysunek 1.3. Rozkłady wyników sprawdzianu w 2012 roku w różnych grupach uczniów.

są znikome. Natomiast między całą grupą zrealizowanych szkół (4) a grupą po usunięciu 16 szkół (5) nie ma praktycznie żadnych różnic. Analiza wyników sprawdzianu z 2012 roku wskazuje, iż analizowana grupa szkół i uczniów jest reprezentatywna dla populacji.

Ostatnim elementem, na który należy zwrócić uwagę podczas analizy wyników sprawdzianu z lat 2013 i 2012, jest kwestia szkół, które odmówiły udziału w badaniu. Istnieje prawdopodobieństwo, iż odmowa mogła być związana z specyficznymi cechami szkół, a te z kolei z bardzo niskimi lub wysokimi wynikami egzaminacyjnymi. Z tego względu wykonano dodatkowe analizy w grupie szkół, które odmówiły udziału w badaniu (grupa 6 w przywołanych powyżej tabelach i rysunkach). Przeglądając się wynikom tej grupy oraz porównując je z pozostałymi grupami, można zauważyć duże podobieństwo osiągnięć uczniów z tych szkół do uczniów z pozostałych grup, zwłaszcza 2, 4 i 5. Choć istnieją pewne różnice w rozkładach oraz w statystykach opisowych sprawdzianu z 2013 roku, wyniki uczniów w szkołach, które odmówiły, są nieco wyższe od uzyskanych

Tabela 1.10. Statystyki opisowe dla wyników sprawdzianu w 2013 roku.

Grupa	N^*/n^{**}	M	SD	Q_1	Me	Q_3	Min	Max	Skośność	Kurtoza
1	350/350*	24,03	8,38	18	24	31	0	40	-0,158	-0,892
2	327/707*	23,95	8,35	17	24	31	0	40	-0,148	-0,890
3	9/561**	24,69	8,31	18	25	32	0	40	-0,214	-0,870
4	9/870**	24,58	8,29	18	25	31	0	40	-0,209	-0,865
5	9/258**	24,64	8,29	18	25	31	0	40	-0,219	-0,859
6	1/438**	25,23	8,28	19	26	32	4	40	-0,227	-0,929

Tabela 1.11. Statystyki opisowe dla wyników sprawdzianu w 2012 roku.

Grupa	N^*/n^{**}	M	SD	Q_1	Me	Q_3	Min	Max	Skośność	Kurtoza
1	355 034*	22,75	7,63	17	23	29	0	40	-0,059	-0,717
2	333 826*	22,69	7,60	17	23	29	0	40	-0,054	-0,713
3	9 619**	23,40	7,64	18	24	29	1	40	-0,140	-0,694
4	9 921**	23,35	7,62	18	24	29	1	40	-0,126	-0,689
5	9 308**	23,37	7,62	18	24	29	1	40	-0,128	-0,686
6	1 419**	23,82	7,75	18	24	29	5	40	-0,183	-0,675

w pozostałych grupach. Na pewno nie można stwierdzić, iż odmowy szkół są znacząco związane z osiągnięciami ich uczniów na egzaminach zewnętrznych.

Podsumowując, zarówno w wylosowanej, jak i zrealizowanej próbie szkół przeciętne wyniki sprawdzianu okazały się minimalnie wyższe niż w populacji szkół uwzględnionych w operacie losowania. Jednak różnice te są na tyle małe, iż należy uznać, że próba pod względem analizowanych wyników dobrze odzwierciedla populację, z której była losowana.

Kolejnym etapem sprawdzenia reprezentatywności próby było porównanie zróżnicowania międzyszkolnego w szkołach, w których zrealizowano badanie, w porównaniu do wyróżnionych grup szkół. W tym celu przeprowadzono serię dwupoziomowych modeli regresji, w których zmienną zależną były wyniki uzyskane na sprawdzianie odpowiednio: w 2012 i 2013 roku (tzw. modele puste).

Modele zbudowano dla:

1. całej populacji uczniów;
2. populacji uczniów, z wyłączeniem uczniów z typów szkół wykluczonych z operatu;
3. uczniów z wylosowanej próby szkół;
4. uczniów ze zrealizowanej próby szkół;
5. uczniów ze zrealizowanej próby szkół, z wyłączeniem uczniów 16 usuniętych szkół.

Wykorzystano oprogramowanie Stata 12, estymacja przeprowadzona została metodą *full maximum likelihood*⁷. W celu sprawdzenia, czy istnieją różnice między wskazanymi powyżej grupami szkół pod względem zróżnicowania międzyszkolnego, otrzymane rezultaty przeliczono do postaci współczynnika ICC. Informuje on, jaki procent zróżnicowania wyników sprawdzianu przypisać można podziałowi na szkoły (współczynniki obliczone zostały na podstawie oszacowań wariancji na poziomie uczniów oraz szkół). Pełne wyniki modelowania zamieszczono w aneksie. Współczynniki z modeli dla lat 2013 i 2012 zostały przedstawione w Tabeli 1.12.

Analiza wartości współczynników korelacji wewnątrzgrupowej wskazuje, że zróżnicowanie międzyszkolne wyników sprawdzianu jest nieco większe w całej populacji (1) niż w badanej populacji (2). Różnice te w latach 2013 i 2012 są na podobnym poziomie – wynoszą około 2%. Przyglądając się zróżnicowaniu w grupach szkół zrealizowanych (4) oraz zrealizowanych bez 16 szkół (5) widać natomiast, że są one do siebie bardzo zbliżone. Różnice między tymi dwoma grupami są niewielkie – wartość różnicy wynosi

⁷ Różnice wersji zastosowanego oprogramowania wynikają z różnic w jego dostępności.

Tabela 1.12. Wyniki dwupoziomowych modeli regresji dla zmiennej zależnej: sprawdzian szóstoklasisty z lat 2013 i 2012 – współczynniki korelacji wewnątrzgrupowej dla poziomu szkół (modele puste).

Grupa	2013 rok (w %)			2012 rok (w %)		
	ICC	Przedział ufności 95%		ICC	Przedział ufności 95%	
1	12,1	11,8	12,4	11,0	10,7	11,3
2	10,1	9,8	10,4	9,1	8,9	9,4
3	10,6	8,6	12,8	11,6	9,5	14,1
4	9,2	7,5	11,3	9,9	8,0	12,1
5	9,4	7,6	11,6	10,4	8,4	12,8

0,2 punktu procentowego w roku 2013 i 0,5 punktu procentowego w roku 2012. Porównanie grup 4 i 5 z badaną populacją (2) nie zmienia tego obrazu. W takim porównaniu różnice są nadal niewielkie. Podsumowując, analiza zróżnicowania wyników sprawdzianu z dwóch kolejnych lat wskazuje, że nie istnieją duże różnice między rozpatrywanymi grupami.

Przeprowadzone analizy reprezentatywności miały odpowiedzieć na pytanie o jakość wylosowanej i zrealizowanej próby. Do tego celu wykorzystano wyniki sprawdzianu szóstoklasisty z lat 2013 i 2012. Wykonane analizy wskazują, że próby: wylosowana, zrealizowana oraz zrealizowana z wykluczeniem 16 szkół są bardzo zbliżone pod względem rozkładu osiągnięć szkolnych do zakładanej populacji szkół. Rozkłady wyników są bardzo podobne, a między parametrami statystycznymi są jedynie minimalne różnice. Analiza zróżnicowania międzyszkolnego prowadzi nas do zbliżonych wniosków. Procent wariacji tłumaczony podziałem na szkoły w porównywanych grupach jest bardzo zbliżony zwłaszcza, gdy weźmiemy pod uwagę przedziały ufności. Możemy zatem przyjąć, iż wykonane analizy wskazują na wysoką jakość zrealizowanej próby. Jest ona reprezentatywna w stosunku do populacji, z której pochodzi, nawet w przypadku, gdy rozpatrujemy ją bez 16 usuniętych szkół.

Wagi stosowane w analizach

Wagi stosowane są w analizach danych społecznych, by lepiej scharakteryzować populację na podstawie danych z próby. Wagi mają szczególne znaczenie przy doborze próby, w którym niektóre podgrupy badanych jednostek są nadreprezentowane lub niedoreprezentowane. W naszych analizach wykorzystaliśmy wagi dla szkół, oddziałów, uczniów i nauczycieli. Odzwierciedlają one sposób doboru próby i uwzględniają prawdopodobieństwo znalezienia się poszczególnych szkół, nauczycieli, oddziałów, uczniów i ich rodziców w próbie.

W analizach trypoziomowych wagi stosowane na poziomie szkół informują, jaką liczbę placówek podobnych (pod względem wielkości miejscowości oraz liczby oddziałów w szkole) reprezentuje wylosowany podmiot. Natomiast wagi na poziomie oddziałów informują, jaką liczbę oddziałów z wylosowanej szkoły reprezentują wylosowane oddziały. Jeśli

w szkole było pięć oddziałów klas pierwszych, to dla każdego z wylosowanych dwóch oddziałów waga ta przyjmie wartość 2,5.

Wagi na poziomie uczniów wprowadzają korektę na poziomie realizacji badania w danej klasie. Informują zatem, jaką liczbę uczniów danego oddziału reprezentują przebadani. Jeśli w klasie, w III etapie badania było 25 osób stanowiących próbę podstawową (osoby niewykluczone z badania, uczące się stale w tym samym oddziale), a ostatecznie kwestionariusze badawcze uzupełnione zostały przez 20 osób, to każdemu uczniowi w tej klasie przyporządkowana została waga 1,25. Wartość tej wagi została ograniczona do wartości 3, podobnie jak to ma miejsce chociażby w badaniach PISA (OECD, 2009). Waga ta jest wykorzystywana w analizach wielopoziomowych na najniższym poziomie analizy. Iloczyn wagi szkoły, oddziału i ucznia to indywidualna waga ostateczna. Na jej podstawie, wykorzystując macierz Hadamarda z korektą Faya 0,5, przygotowano zestaw 76 wag replikacyjnych. Wykorzystanie ich jest niezbędne dla szacowania błędów standardowych oszacowań parametrów statystycznych z użyciem modeli nieuwzględniających pogrupowania uczniów w oddziałach i szkołach.

W analizach, w których nauczyciele traktowani są jako grono pedagogiczne danej szkoły, przyjmuje się dla nich łącznie wagi poszczególnych szkół. W przypadku przyłączania konkretnych nauczycieli do oddziałów, które uczą, stosowane są wagi dla oddziałów lub (w przypadku analiz dwupoziomowych) iloczyn wagi szkoły i wagi klasy.

Usunięcie z analiz danych 16 szkół wskazanych jako niewiarygodne ze względu na pierwszy pomiar inteligencji, nie wpływa pod tym względem na schemat doboru próby, a co za tym idzie, umożliwia stosowanie tych samych wag.

Rozdział 2

Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

W rozdziale tym przedstawimy analizy wpływu czynników pozaszkolnych na osiągnięcia szkolne uczniów w zakresie umiejętności czytania, świadomości językowej oraz umiejętności matematycznych po zakończeniu pierwszego etapu edukacyjnego (klasy I–III). Mają one przynieść odpowiedź na pytanie, jakie czynniki należy uwzględnić w kontekstowych modelach oceny efektywności nauczania. Ponieważ u progu pierwszego etapu edukacyjnego nie prowadzi się pomiaru osiągnięć, logikę klasycznego modelu edukacyjnej wartości dodanej zastąpić musimy statystyczną kontrolą pozaszkolnych uwarunkowań osiągnięć szkolnych.

Cechy ucznia i jego środowiska rodzinnego

Literatura przedmiotu dotycząca indywidualnych i rodzinnych uwarunkowań osiągnięć szkolnych jest bardzo obszerna. W związku z tym jej przegląd będzie opierał się na metaanalizach lub na w miarę aktualnych artykułach odwołujących się do badań prowadzonych w tym zakresie.

Płeć

Różnice międzypłciowe w osiągnięciach uczniów są przedmiotem zainteresowania od wielu lat. Prowadzone analizy wskazują na zmiany różnic międzypłciowych w zakresie osiągnięć szkolnych w skali niemal całego świata. Ich charakter zależy od typu analizowanych osiągnięć. Zmiany te są dwojakiego rodzaju. Po pierwsze, są one międzygeneracyjne. Po drugie, wydaje się, iż różnice występują również na poszczególnych etapach rozwoju. Analizy prowadzone na danych z ostatnich 20 lat wskazują na to, że dziewczynki już na początku szkoły czytają lepiej niż chłopcy, a wraz z wiekiem różnice te powiększają się (Lynn i Mikk, 2009). W przypadku umiejętności matematycznych oraz wiedzy przyrodniczej (*science*) w kolejnych kohortach przewaga chłopców w poziomie osiągnięć maleje. Obecnie, w większości badań międzynarodowych, różnice te

nie są istotne statystycznie (Else-Quest, Hyde i Linn, 2010; Hyde, Lindberg, Linn, Ellis i Williams, 2008; Lindberg, Hyde, Petersen i Linn, 2010; Noble i Bradford, 2000; Steinmayr i Spinath, 2008; Stoet i Geary, 2013). Ze względu na dynamikę zjawisk wiążących płeć i osiągnięcia w różnych dziedzinach istotne jest stałe monitorowanie sytuacji w tym zakresie.

Wiek

Wiele badań pokazuje, że uczniowie starsi (w obrębie tej samej kohorty) przewyższają młodszych w poziomie osiągnięć szkolnych (Lee i Fish, 2010; Morrison, Griffith i Alberts, 1997), ale tempo rozwoju młodszych uczniów w zakresie umiejętności czytania i matematyki nie jest wolniejsze od obserwowanego wśród starszych kolegów z tej samej klasy (Morrison i in., 1997). Niektóre badania pokazują dodatkowo, że związek osiągnięć z wiekiem jest silniejszy wśród dziewczynek niż wśród chłopców, a także wśród uczniów pochodzących z mniej zamożnych środowisk (Smith, 2009). Co ciekawe, interakcje te obserwuje się dopiero w starszych klasach. Wyniki badań nie zawsze są spójne, jeśli chodzi o wielkość i istotność efektu wieku dla osiągnięć szkolnych (por. Martin, 2009), jednak w wielu krajach powtarza się prawidłowość świadcząca o tym, że w kolejnych latach nauki szkolnej siła korelacji osiągnięć z wiekiem dla głównej kohorty wiekowej spada (Lee i Fish, 2010; Smith, 2009). Analizy wykorzystujące dane z polskiego systemu egzaminacyjnego pokazały, że w Polsce również obserwujemy zależność osiągnięć od wieku uczniów (Dolata i Pokropek, 2012).

Analizy prowadzone przez badaczy na całym świecie pokazują, że o ile efekt wieku dla uczniów z głównej kohorty jest raczej niewielki, o tyle różnicowania wynikające z wcześniejszego lub późniejszego posłania dziecka do szkoły oraz braku promocji do następnej klasy są wyraźnie skorelowane z późniejszymi osiągnięciami. Wyniki badań pokazują jednak, że uczniowie, którzy zostali posłani do szkoły później niż ich rocznikowi rówieśnicy, mają niższe (lub rzadziej: porównywalne) osiągnięcia szkolne niż koledzy z tej samej klasy (Martin, 2009; May i Kundert, 1995; Morrison i in., 1997). Na drugim krańcu tego zjawiska są uczniowie o przyspieszonym starcie szkolnym. Ci zazwyczaj mają istotnie wyższe wyniki niż ich de facto starsi koledzy z tej samej klasy (Mayer i Knutson, 1999). Wynika to z tego, że w tej nielicznej grupie wcześniej lub później rozpoczynających naukę szkolną znajdują się uczniowie wyselekcjonowani ze względu na cechy, które są powiązane pozytywnie z późniejszymi sukcesami szkolnymi.

Inteligencja

Inteligencja jest jednym z najważniejszych i najbardziej stabilnych predyktorów osiągnięć szkolnych dziecka – pozwala wyjaśniać od kilkunastu do kilkudziesięciu procent zmienności wyników w nauce (Deary, Strand, Smith i Fernandes, 2007; Duckworth, Quinn i Tsukayama, 2012; Heaven i Ciarrochi, 2012; Hofer, Kuhnle, Kilian i Fries, 2012; Sternberg, Grigorenko i Bundy, 2001; Teo, Carlson, Mathieu, Egeland i Sroufe, 1996).

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

Co istotne, relacja ta ujawnia się systematycznie, pomimo różnorodności stosowanych zarówno ujęć i miar inteligencji, jak i dziedzin i miar osiągnięć szkolnych. Zwykle bywa silniejsza dla przedmiotów ścisłych (np. matematyka, fizyka) w porównaniu do humanistycznych (np. historia, język ojczysty; Ferrer i McArdle, 2004; Heaven i Ciarrochi, 2012; Teo i in., 1996).

Różnice w zakresie operacjonalizacji zarówno osiągnięć, jak i inteligencji, w połączeniu z odmiennymi schematami badawczymi (przekrojowe versus podłużne/z odroczonym pomiarem osiągnięć) oraz różnicami w wieku osób badanych, sprawiają, iż trudno jest precyzyjnie określić siłę zależności między nimi. W badaniach podłużnych wykorzystujących miary inteligencji ogólnej oraz miary inteligencji zawierające zarówno komponenty werbalne, jak i niewerbalne, raportowana w literaturze światowej siła relacji kształtuje się na poziomie 0,6–0,8 (np. Deary i in., 2007; Duckworth i in., 2012; Heaven i Ciarrochi, 2012; Hofer i in., 2012; Sternberg i in., 2001). Słabsze zależności z osiągnięciami notuje się w przypadku inteligencji płynnej – w zależności od wieku osób badanych kształtują się one na poziomie około 0,2 do około 0,4 (Bossaert, Doumen, Buyse i Verschueren, 2011; Hofer i in., 2012). Wyniki polskich badań podsumowane przez Romana Dolatę (2008) wskazują, że na etapie szkoły podstawowej (od około drugiej klasy) do końca gimnazjum testy inteligencji płynnej korelują z wynikami egzaminów zewnętrznych oraz ocenami szkolnymi na poziomie odpowiednio 0,4–0,5 oraz 0,3–0,4.

Aspiracje edukacyjne rodziców

Metaanaliza wpływu czynników związanych z zaangażowaniem rodzicielskim na osiągnięcia szkolne dzieci (Fan i Chen, 2001) wskazuje, że spośród wielu analizowanych aspektów największe znaczenie mają aspiracje i oczekiwania edukacyjne rodziców¹. Analiza ta została wykonana na podstawie prostych współczynników korelacji. Bardziej zaawansowane analizy zostały przeprowadzone z wykorzystaniem modeli równań strukturalnych. W analizach tych (Singh, Bickley, Trivette, Keith i Anderson, 1995) autorzy wskazują również na to, iż spośród kilku wskaźników zaangażowania rodziców w edukację dzieci, aspiracje edukacyjne mają największy wpływ na oceny uzyskiwane przez uczniów. Co więcej, jest to zmienna wpływająca pośrednio także na rozmowy o szkole prowadzone pomiędzy rodzicami oraz między rodzicami a dziećmi, które także wpływają na osiągnięcia edukacyjne. W kontekście głównego tematu naszej książki, wyniki tych analiz wskazują, że aspiracje rodziców są mocno powiązane z wcześniejszymi osiągnięciami szkolnymi uczniów. Układ taki sugeruje istnienie efektu spirali. Wyniki te uzyskane zostały przy kontroli statusu społeczno-ekonomicznego rodziny (SES). Do podobnych wniosków doszedł również Xitao Fan (2001), wykorzystując modele wzrostu (*Latent Growth Curve*, LGC). Aspiracje edukacyjne rodziców mają duże znaczenie dla osiągnięć, niezależnie od dziedziny nauczania, również przy kontroli statusu rodziny.

¹ Co ciekawe, spośród ponad 2000 artykułów i raportów zakwalifikowanych wstępnie do analizy jedynie w 25 z nich znalazły się informacje, które mogły posłużyć do przeprowadzenia analiz. Najważniejszym powodem wykluczenia tekstu z analizy było to, że nie miał on charakteru empirycznego (*sic!*).

Status społeczno-ekonomiczny rodziny

Przegląd badań dotyczących statusowej determinacji osiągnięć szkolnych, również dotyczących polskiej oświaty, można znaleźć w wielu polskich monografiach poświęconych różnym aspektom nierówności edukacyjnych (Dolata, 2008; Dolata, Jakubowski i Pokropek, 2013; Zawistowska, 2012). Jest to czynnik postrzegany od czasów raportu Jamesa Colemana (1966) za kluczowy w kształtowaniu osiągnięć edukacyjnych uczniów.

Intensywne badania w tym zakresie wskazały na to, iż znaczna część wpływu, jaki Coleman przypisał czynnikom statusowym, ma źródła w zmienności genetycznej, co należy uwzględnić w wyjaśnianiu nierówności społecznych (Herrnstein, 1994; Jensen, 1969; Jensen, 1991; Plomin, DeFries, McClearn i McGuffin, 2001). Podsumowując z perspektywy pedagogicznej wyniki badań nad odziedziczalnością, Judith Harris (2000) i Stephen Pinker (2005) doszli do dwóch kluczowych wniosków. Po pierwsze, główne cechy osobowości – w tym inteligencja ogólna – są w znaczącym stopniu determinowane genetycznie. Po drugie, wpływ wychowywania się w tej samej rodzinie na kształtowanie cech osobowości – przy kontroli czynnika przekazu genetycznego – jest słaby lub nie zachodzi. Hans Eysenck i Michael Eysenck (1996, s. 226) piszą:

Ignorując oddziaływanie [czynnika genetycznego - przyp. aut.], Coleman w rzeczywistości dopuścił do tego, by w jego badaniach dziedziczenie oddziaływało na osiągnięcia szkolne w postaci takich czynników, jak status społeczno-ekonomiczny i środowisko rodzinne ucznia, co doprowadziło do sztucznego rozdęcia wpływów środowiska rodzinnego ponad wszelkie rozsądne proporcje. Jeżeli powtórzy analizę jego wyników z należyтым uwzględnieniem oddziaływania czynników dziedzicznych, treść wyciągniętych wniosków ulega radykalnej zmianie i okazuje się, że jakość szkoły w bardzo poważnym stopniu wpływa na poziom edukacyjnych osiągnięć uczniów.

W analizach osiągnięć edukacyjnych należy uwzględnić miary kapitału kulturowego i ekonomicznego rodziny ucznia oraz wskaźniki pozycji rodziny na różnych wymiarach statusu społeczno-ekonomicznego. Natomiast wyniki badań z zakresu genetyki zachowania uczulają nas na problem interpretacji korelacji miar SES z osiągnięciami. Prawdopodobnie za znaczną część tej korelacji odpowiada mechanizm dziedziczenia biologicznego.

Struktura rodziny

Niezależnie od statusu społeczno-ekonomicznego rodziny, jej struktura również jest przedmiotem zainteresowania badaczy. W tym zakresie rozpatruje się najczęściej to, czy rodzina jest pełna (dwoje biologicznych rodziców) czy niepełna, dwu- czy trzypokoleniowa. Innym często analizowanym aspektem struktury rodziny jest liczba rodzeństwa i kolejność urodzenia. Powszechną prawidłowością, jaką zaobserwowano, jest to, iż dzieci z rodzin niepełnych mają niższe osiągnięcia na poszczególnych etapach edukacji. Jednakże analizy prowadzone przy kontroli statusu rodziny prowadzą do sprzecznych wniosków. Część badaczy wskazuje, że uwzględnienie dodatkowych czynników, zwłaszcza statusu ekonomicznego, powoduje, że to, czy rodzina jest pełna czy niepełna, nie ma znaczenia (Carlson i Corcoran, 2001; Ginther i Pollak, 2004). Wskazywałoby to

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

na mechanizm oddziaływania tej cechy rodziny na osiągnięcia szkolne dzieci poprzez SES. Są jednak analizy, w których przy kontroli zmiennych statusowych efekt ten jest słabszy, jednak nadal istotny (Chiu, 2010; Deleire i Kalil, 2002). Może to sugerować, że poza statusem społeczno-ekonomicznym struktura rodziny jednak oddziałuje na osiągnięcia. Podobnie niejednoznaczne wnioski płyną z analiz nad wielopokoleniowością rodzin. Część badań sygnalizuje negatywny wpływ zamieszkiwania z dziadkami (Chiu, 2010). Inne dowodzą z kolei pozytywnych efektów – zwłaszcza w rodzinach niepełnych (Deleire i Kalil, 2002).

Większość badań dotyczących liczby rodzeństwa wskazuje na to, że wraz ze wzrostem liczby rodzeństwa osiągnięcia badanych uczniów obniżają się (Carlson i Corcoran, 2001; Chiu, 2010; Ginther i Pollak, 2004). Z kolei kolejność urodzenia ucznia ma bardziej złożony charakter. Generalnie najstarsze z rodzeństwa ma najwyższe osiągnięcia (Chiu, 2010). Zaobserwowane relacje mogą być tłumaczone koniecznością dystrybucji ograniczonych zasobów rodziny pomiędzy większą liczbę dzieci. Zasoby te to nie tylko dobra ekonomiczne, ale również czas poświęcany dzieciom. W rodzinach, w których jest więcej dzieci, relatywnie najwięcej czasu poświęca się dziecku najstarszemu, które przez pewien czas – siłą rzeczy – jest dzieckiem jedynym.

Uczęszczanie do przedszkola i zerówki

Osobnym problemem wymagającym wnikliwej analizy są efekty opieki przedszkolnej. Istnieje wiele badań w tym zakresie. Przegląd wyników w skali międzynarodowej oparty zostanie na dwóch obszernych metaanalizach badań eksperymentalnych i quasi-eksperymentalnych. Pierwsza z nich ogranicza się do przeglądu badań ze Stanów Zjednoczonych i zawiera zestawienie wyników analiz znajdujących się w 123 raportach lub artykułach spełniających surowe wymogi metodologiczne na poziomie zbierania i analizy danych. Druga z metaanaliz obejmuje wyniki badań z różnych państw. Opisuje ona ponad 400 raportów dokumentujących wyniki ponad 250 badań. Także w tym przypadku założone zostały restrykcyjne kryteria włączenia wyników do analizy, o czym świadczy fakt, iż analizowane teksty zostały wybrane spośród 10 309 raportów (Camilli, Vargas, Ryan i Barnett, 2010; Leak i in., 2010). Przeprowadzone analizy wskazują na istotny, lecz malejący z czasem wpływ edukacji przedszkolnej na osiągnięcia szkolne dzieci. Najmocniejszy wpływ miała edukacja przedszkolna na zakres wiadomości, mniejszy zaś na umiejętności szkolne i kompetencje społeczne. Wpływ ten w okresie 2–4 lat po zakończeniu programu przedszkolnego stawał się zwykle nieistotny statystycznie – dzieci uczęszczające do przedszkoli nie różniły się pod względem osiągnięć od dzieci nieobjętych taką opieką. Przy kontroli innych charakterystyk badanych wpływ ten był wręcz ujemny. Na podstawie danych dla Polski do podobnych wniosków doszedł Krzysztof Konarzewski (2012). W analizach wykorzystał dane z polskiej edycji badania osiągnięć matematycznych i przyrodniczych (TIMSS) oraz osiągnięć w czytaniu (PIRLS). Proste porównanie osiągnięć uczniów uczęszczających wcześniej do przedszkola z rówieśnikami nieobjętymi taką opieką wskazało na wyższy poziom umiejętności dzieci objętych przedszkolnym przygotowaniem. Jednakże przy kontroli płci i SES rodziny dziecka nie wystąpiły różnice w osiągnięciach dziesięciolatków uczęszczających i nieuczęszczających wcześniej do przedszkola.

Czynniki społeczno-gospodarcze na poziomie lokalnym

Szkoła jako instytucja funkcjonuje w określonym otoczeniu społeczno-gospodarczym, które samo w sobie może wpływać na osiągnięcia uczniów. W literaturze zwykle raportowane są wyniki analiz prowadzonych na zagregowanych danych o osiągnięciach uczniów (np. na poziomie gminy) i charakterystyk społeczno-gospodarczych. Brak kontroli cech na poziomie indywidualnym może prowadzić do zawyżenia siły rozpatrywanych związków. Dane zgromadzone w ramach opisywanego badania dają okazję do przeanalizowania najczęściej pojawiających się w literaturze wątków badawczych z wykorzystaniem modeli wielopoziomowych. Omawiane poniżej czynniki społeczno-gospodarcze okazały się istotne w analizach prowadzonych w celu określenia trafności metody edukacyjnej wartości dodanej dla gimnazjów (Dolata i in., 2013).

Poziom bezrobocia

W Polsce wyniki egzaminów zewnętrznych zostały poddane analizie pod kątem oddziaływania lokalnego poziomu bezrobocia na średni wynik na danym terenie przez Jana Herczyńskiego i Mikołaja Herbsta (2002; Herbst, 2004). Dowiedli oni, że wzrost lokalnej stopy bezrobocia (na poziomie powiatu), przy kontroli innych czynników lokalnych, wpływa na obniżenie wyników uzyskiwanych przez gimnazjalistów w testach zewnętrznych. Z kolei na podstawie danych ze Stanów Zjednoczonych Elizabeth Oltmans Ananat, Anna Gassman-Pines, Diana Francis oraz Christina Gibson-Davis (2011) wykazały, że wzrost stopy bezrobocia wpływa na obniżenie osiągnięć szkolnych mierzonych przede wszystkim wynikami zewnętrznych egzaminów z matematyki. Powyższe analizy prowadzone były na danych zagregowanych, co jest istotnym ograniczeniem ze względu na brak kontroli indywidualnej sytuacji dzieci i ich rodzin.

Problem ten rozwiązany został w analizach przeprowadzonych z uwzględnieniem takich informacji. Leon Feinstein i James Symons (1999) stwierdzili, iż bezrobocie na poziomie lokalnym nie wpływa na wyniki uzyskiwane w testach. Z kolei Phillip Levine (2011) na podstawie analizy danych NLSY79 dowiódł, iż negatywne oddziaływanie lokalnej stopy bezrobocia na wyniki w testach osiągnięć dotyczy jedynie dzieci z rodzin o niskim statusie społeczno-ekonomicznym. Do jeszcze innych wniosków doszły Ann Stevens i Jessamyn Schaller (2009), wskazując, że lokalny poziom bezrobocia zwiększa ryzyko powtarzania klasy przez uczniów. W analizie tej kontrolowane było bezrobocie rodziców, które również zwiększało to ryzyko.

Hipotetyczny mechanizm, mogący kryć się za zaobserwowanymi relacjami, wiąże lokalny poziom bezrobocia z osiągnięciami uczniów dwoma rodzajami oddziaływań, także w rodzinach niedotkniętych bezpośrednio bezrobociem. Na obszarach o wyższej stopie bezrobocia średni poziom płac jest niższy (Blanchflower i Oswald, 2005), co obniża zamożność rodzin i może wpływać na niższe osiągnięcia szkolne dzieci. Na obszarach tych w wyższym stopniu odczuwane jest zagrożenie utratą pracy, co również może

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

przekładać się na psychiczny stan rodziców, ich relacje z dziećmi (Fenwick i Tausig, 1994), a w konsekwencji na ich edukację.

Struktura gospodarki

W analizach lokalnych uwarunkowań osiągnięć szkolnych wskazuje się również na wpływ struktury lokalnej gospodarki. Jest to jednak temat relatywnie rzadko eksplorowany. W badaniach (Broomhall i Johnson, 1994; James, 2001) pozytywnie zweryfikowano hipotezę, iż na obszarach, na których jest niewiele podmiotów gospodarczych zgłaszających zapotrzebowanie na wysoko wykwalifikowanych pracowników, młodzi ludzie są słabiej motywowani do nauki. Brak w lokalnym otoczeniu przykładów transmisji wyższego wykształcenia na wyższe dochody prowadzi do przykładania mniejszej wagi do osiągnięć szkolnych. W efekcie uczniowie ci szybciej niż ich rówieśnicy z innych rejonów rezygnują z dalszej edukacji.

Powszechność opieki przedszkolnej

Poziom upowszechnienia opieki przedszkolnej może również wpływać na osiągnięcia szkolne. Na podstawie porównań międzynarodowych wykazali to pośrednio Gabriela Schütz, Heinrich Ursprung i Ludger Woessmann (2005). Dowiedli, że istnieje zależność pomiędzy upowszechnieniem nauczania przedszkolnego a siłą wpływu „zaplecza rodzinnego” (*family background*) na osiągnięcia edukacyjne. Związek tych czynników jest nieliniowy. Wzrost udziału dzieci w nauczaniu przedszkolnym do poziomu około 60% powoduje wzrost siły związku zaplecza rodzinnego i osiągnięć. Powyżej tego odsetka siła rodzinnej determinacji spada. Bezpośrednio związek poziomu upowszechnienia opieki przedszkolnej z późniejszymi osiągnięciami uczniów analizowany był w badaniach trafności wskaźników EWD dla gimnazjów (Dolata i in., 2013). Po dziewięciu latach nauki w szkole osiągnięcia uczniów nie były związane z tym czynnikiem. W literaturze brak jednak analiz rozpatrujących ten związek wśród młodszych dzieci, z tego powodu prezentowane poniżej analizy będą miały przede wszystkim charakter eksploracyjny.

Wydatki na oświatę

Analizy uwzględniające wpływ wydatków na oświatę prowadzone były zarówno w obrębie krajów, jak i na poziomie międzynarodowym. Wykazano na przykład brak związku pomiędzy wydatkami na edukację a osiągnięciami polskich uczniów (Herbst, 2004). Do podobnych wniosków prowadzą również porównania międzynarodowe. Eric Hanushek i Ludger Woessmann (2011), na podstawie przeglądu literatury w tym zakresie, stwierdzili, że wzrost wydatków na edukację w przeliczeniu na ucznia nie przekłada się na podniesienie poziomu jego osiągnięć. Z tego względu w literaturze przedmiotu (Costrell, Hanushek i Loeb, 2008; Hanushek, 2001; 2013) wskazuje się na konieczność uwzględniania

w analizach wydatków dodatkowych czynników kontrolnych, zwłaszcza w zakresie jakości i organizacji nauczania.

Metoda

Dane i metody pomiaru

Analizowane czynniki można podzielić na niezmiennie, ulegające niewielkim wahaniom w czasie oraz mogące w szerszym zakresie zmieniać swoje wartości. Do pierwszej grupy należą zmienne o podłożu biologicznym oraz dane dotyczące przeszłości. Dane biologiczne to płeć i względny wiek (w stosunku do rówieśników). Termin pomiaru tego typu czynników nie ma znaczenia dla możliwości ich wykorzystywania w analizach. Z kolei zmienne dotyczące przeszłości dotyczą na przykład tego, czy i jak długo dziecko uczęszczało do przedszkola czy zerówki. W przypadku tego typu danych należy przyjąć, że informacje są bardziej wiarygodne, jeśli zbierane są w jak najkrótszym okresie od wystąpienia analizowanego zdarzenia.

Dla pozostałych zmiennych – zwłaszcza tych, których wartości mogą ulegać znacznym zmianom w czasie – ważny jest termin wykonania pomiaru. Informacje dotyczące dziecka i jego rodziny opisują jego sytuację na początku nauki w pierwszej klasie. Podejście takie jest uzasadnione tym, iż poziom części czynników może być uzależniony od efektywności pracy szkoły. Przykładowo, aspiracje rodziców kształtowane są częściowo przez osiągnięcia dziecka, te zaś zależą od efektywności szkoły. Jeśli w analizach wykorzystywane byłyby aspiracje rodziców mierzone pod koniec trzeciej klasy, to obciążone byłyby one częściowo tym, jak pracuje szkoła. Taka zmienna nie spełniłaby odpowiedniej roli kontrolnej w modelach efektywności pracy szkół.

Podobna sytuacja ma miejsce w przypadku inteligencji uczniów. W przypadku uczniów gimnazjów (Dolata i in., 2013) wykazano, że zmiany w poziomie inteligencji powiązane są z efektywnością pracy szkoły. Przyjęcie pomiaru inteligencji w trzeciej klasie utrudniłoby oszacowanie tej efektywności. Z tego powodu do szacowania kontekstowych modeli EWD wykorzystywany będzie pomiar inteligencji przeprowadzony na początku nauki w klasie pierwszej.

Również zmienne opisujące status rodziny – takie jak wykształcenie rodziców, ich status ekonomiczny, prestiż wykonywanego przez nich zawodu czy też zamożność – powinny opisywać sytuację na początku ścieżki edukacyjnej dziecka. Takie podejście – po wprowadzeniu tych zmiennych do modeli regresyjnych – pozwala „zrównać” rodzinne zaplecze dzieci w czasie ich pierwszego kontaktu ze szkołą, a przez to lepiej oszacować efektywność pracy szkoły.

Ostatnia grupa to czynniki o relatywnie niewielkiej zmienności w czasie. Zaliczyć do nich można przede wszystkim informacje opisujące lokalny kontekst funkcjonowania szkoły. Dla czternastu wykorzystywanych w analizach wskaźników społeczno-gospodarczych charakteryzujących wszystkie gminy w Polsce, współczynnik korelacji pomiędzy poszczególnymi latami był bardzo wysoki. Dla dziesięciu wskaźników wartości były wyższe niż 0,95, zaś dla pozostałych wahały się od 0,64 do 0,90. Ponadto znacznym ograniczeniem przy korzystaniu z informacji pochodzących z Głównego Urzędu Statystycznego jest opóźnienie,

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

Tabela 2.1. Poziom wykształcenia, do którego aspirują rodzice ucznia, liczony w latach nauki.

Wykształcenie	Lata nauki
Zasadnicze zawodowe	10
Średnie zawodowe	13
Średnie ogólnokształcące	12
Policealne lub pomaturalne	13
Wyższe licencjackie lub inżynierskie	15
Wyższe studia magisterskie	17
Stopień doktora, doktora habilitowanego lub profesora	20

z jakim są one udostępniane. W trakcie przygotowywania zbioru danych, tj. w 2012 roku, w okresie realizacji III etapu badania, najświeższe dostępne dane obejmowały 2010 rok. Z tych dwóch powodów zdecydowano o przyjęciu do analiz najświeższych dostępnych danych. Opisują one sytuację w gminach w pierwszym roku kalendarzowym, w którym dzieci uczęszczały do szkoły podstawowej.

Informacja o płci była zbierana w I i III etapie badania. W przypadku niezgodności kodowania prowadzona była procedura ponownego wprowadzania informacji na podstawie danych znajdujących się na papierowych wersjach narzędzi badawczych.

Informacja o wieku uczniów zbierana była w I etapie poprzez kwestionariusze dla rodziców, których proszono o podanie dziennej daty urodzenia ucznia. Na tej podstawie oraz na podstawie zakodowanej daty przeprowadzenia pierwszego z testów osiągnięć, obliczony został wiek ucznia w miesiącach, w dniu udziału w teście. Zmienna ta została wycentrowana tak, aby uczeń urodzony 1 stycznia 2002 roku miał dla tej zmiennej wartość zero. Dodatkowo zakodowano także to, czy uczniowie należą do kohorty głównej (uczniowie, którzy poszli do szkoły w roku kalendarzowym, w którym skończyli 7 lat), do kohort starszych niż główna (uczniowie z odroczonym obowiązkiem szkolnym lub powtarzający klasę), czy do kohorty młodszej niż główna (uczniowie, którzy poszli do szkoły w wieku 6 lat).

Na podstawie deklaracji rodziców zbieranych podczas I etapu badania ustalono, ile lat dziecko łącznie uczęszczało do przedszkola przed rozpoczęciem nauki w szkole podstawowej (nie licząc zerówki). Zero oznacza, że dziecko nie uczęszczało do przedszkola. Opierając się na danych rodzicielskich z tego samego etapu badania, ustalono, czy w ogóle uczeń uczęszczał do zerówki, a jeśli tak, to czy była to grupa w przedszkolu, czy też oddział zerowy w szkole podstawowej.

Aspiracje edukacyjne rodziców ustalone zostały w I etapie badania na podstawie odpowiedzi na pytanie: „Jaki poziom wykształcenia chciałoby Państwo, żeby osiągnęło Wasze dziecko?”. Odpowiedzi zostały przekodowane na liczbę lat nauki, które w „regularnym” trybie pozwalają na osiągnięcie danego poziomu. Przedstawiono je w Tabeli 2.1.

Wyróżnionym w ankiecie skierowanej do rodziców poziomom wykształcenia przypisano wartości skalowe w latach nauki. Pewne decyzje podjęto arbitralnie. Na przykład przypisanie wykształceniu niepełnemu podstawowemu 6 lat nauki jest uproszczeniem,

choć faktycznie przez długi czas utrzymywała się w polskiej szkole praktyka, że dopiero po ukończeniu szóstej klasy można było kierować „przerośniętych” uczniów do przysposobienia zawodowego. Bardziej arbitralne jest przypisanie do kategorii „stopień doktora, doktora habilitowanego lub profesora” wartości skalowej 20.

Wskaźnik używany w analizach to poziom wykształcenia dla tego rodzica, dla którego przybiera on wyższą wartość. W wypadku braku danych dla jednego z rodziców, przypisywano wartość wykształcenia drugiego.

W badaniach socjologicznych procesów ruchliwości zwykle stosuje się miary SES, które wskazują na pozycję jednostki w strukturze społecznej na podstawie wykonywanego zawodu. Zawody rodziców uczniów zostały zakodowane zgodnie z Międzynarodową Klasyfikacją Zawodów ISCO-08 (International Labour Organization, 2007). Klasyfikacja ta jest powszechnie przyjętą podstawą do prowadzenia porównań międzynarodowych w zakresie rynków pracy².

Na podstawie ISCO-08 oraz odpowiednich kluczy przejścia każdemu z zawodów przyporządkowane zostały dwie miary społecznej klasyfikacji zawodów. Pierwsza z nich to międzynarodowa skala statusu społeczno-ekonomicznego – *International Socio-Economic Index* – (ISEI; Ganzeboom, De Graaf i Treiman, 1992), druga to międzynarodowa skala prestiżu zawodów – *Standard International Occupational Prestige Scale* (SIOPS; Treiman, 1977)³.

Do analizy poziomu zamożności gospodarstwa domowego ucznia zaadaptowana została pula pytań dotycząca wyposażenia gospodarstwa domowego, wykorzystywana w badaniach PISA OECD. W wersji użytej w badaniu składała się ona z 17 dychotomicznych pytań dotyczących następujących elementów wyposażenia otoczenia dziecka: biurka do nauki; własnego pokoju; spokojnego miejsca do nauki; komputera, którego dziecko może używać do nauki; edukacyjnych programów komputerowych; dostępu do internetu; kalkulatora; literatury pięknej (np. Sienkiewicz); zbiorów poezji; dzieł sztuki (np. obrazy); encyklopedii; słowników; zmywarki do naczyń; DVD albo wideo; anteny satelitarnej lub telewizji kablowej; aparatu cyfrowego; mikroskopu lub lunety. Do konstrukcji skali użyto też predykcyjnego pytania o liczbę książek w domu.

Ze względu na problemy związane z estymacją modelu pomiarowego zgodnie z zasadami przyjętymi w badaniach PISA oraz ze względu na chęć zachowania porównywalności miar zasobności gospodarstwa domowego z wynikami badań nad trafnością EWD w gimnazjach (Dolata i in., 2013), wykorzystano taki sam sposób przygotowania wskaźników w tym zakresie. Zastosowano model BIFACTOR w ramach podejścia ESEM (RMSEA = 0,025; CFI = 0,988; TLI = 0,977; WRMR = 1,19). Poza czynnikiem głównym, mierzącym zasobność gospodarstwa domowego rodziny ucznia, wyróżnione zostały trzy czynniki lokalne: miejsce do nauki, sprzęt elektroniczny oraz dobra kulturowe. W analizach użyto miary ogólnej zasobności gospodarstwa domowego.

Informacje potrzebne do opracowania wskaźników opisujących strukturę rodziny zebrane zostały w I etapie badania za pomocą ankiety rodzicielskiej. Pytano w niej między innymi o to, czy dziecko ma rodzeństwo, o rok urodzenia każdego z rodzeństwa,

² Dobry przewodnik do kodowania poszczególnych zawodów wraz z wyjaśnieniami znajduje się na stronie Harry’ego Ganzebooma: home.fsw.vu.nl/hbg.ganzeboom/isco08/

³ W analizach używamy wartości wskaźników ISEI i SIOPS tego z rodziców, dla którego przybiera on wyższą wartość. W wypadku braku danych dla jednego z rodziców, przypisywano wartość wskaźnika drugiego. Zmienne te będziemy skrótowo nazywać HISEI i HSIOPS.

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

kto wychowuje dziecko, z ilu osób składa się gospodarstwo domowe ucznia oraz czy zamieszkuje w nim przynajmniej jeden z dziadków lub jedna z babć. Na podstawie tych informacji przygotowano wskaźniki:

- struktury rodziny – pełna dwupokoleniowa, pełna trzypokoleniowa, niepełna dwupokoleniowa, niepełna trzypokoleniowa, inna (rodziny zastępcze, domy dziecka itp.);
- kolejności urodzenia ucznia w grupie rodzeństwa – jedynak, najstarszy, środkowy, najmłodszy;
- wielkości gospodarstwa domowego, który to wskaźnik, przy kontroli struktury rodziny, pośrednio informuje o liczbie rodzeństwa. Wykorzystanie zmiennej opisującej liczbę rodzeństwa w analizach było niemożliwe ze względu na współliniowość z kolejnością urodzenia dziecka.

Pomiar inteligencji płynnej uczniów wykonano w I i III etapie z wykorzystaniem Testu Matryc Ravena. Do oszacowania poziomu inteligencji w I etapie badania wykorzystano metodę kalibracji łącznej, zgodnie z którą w jednym kroku szacowano model IRT dla pomiaru pierwszego i drugiego. Zdecydowano się na model dwuparametryczny ze względu na lepsze dopasowanie do danych. Analizy wykonano w programie Mplus 6 (Muthén i Muthén, 1998–2012), definiując strukturalny model czynnikowy dla zmiennych porządkowych. Wykorzystano estymator WLSMV. Model taki jest równoważny dwuparametrycznemu modelowi IRT (zob. np.: Bartholomew, Knott i Moustaki, 2011). Przyjmował założenie, że poziom inteligencji z I etapu wpływa na poziom inteligencji z III etapu. W modelu zdefiniowano, że te same zadania mają w obu pomiarach takie same parametry trudności i dyskryminacji. Model dopuszczał różne wartości średnich i odchyłeń standardowych dla obu pomiarów. Dla pierwszego pomiaru zostały one zafiksowane na wartościach odpowiednio 0 i 1, natomiast dla drugiego zostały uwolnione i wyrażone na skali pierwszego. Dodatkowo uwzględniono pogrupowanie uczniów w klasach. Model uzyskał dobre dopasowanie do danych (RMSEA = 0,012; CFI = 0,936; TLI = 0,936). Korelacja latentna między poziomami inteligencji uczniów w I etapie i III etapie badania jako cechami ukrytymi wyniosła 0,7. Na podstawie tego modelu wyznaczono wyniki dla uczniów dla obu pomiarów jako wartości czynnikowe dla zmiennych ukrytych. Uczniom, którzy mieli dane tylko z jednego pomiaru, przypisano najbardziej wiarygodny wynik na podstawie rekordu odpowiedzi z drugiego pomiaru i zależności między pomiarami obserwowanymi w danych. Imputacji tej dokonano dla 2,5% wyników dla pierwszego pomiaru. Uzyskane wyniki przeliczono na skalę o średniej 0 i odchyleniu standardowym 1 w próbie uczniów z I etapu badania.

Informacje o czynnikach lokalnych pobrane zostały z Banku Danych Lokalnych. Jest on „największym w Polsce uporządkowanym zbiorem informacji o sytuacji społeczno-gospodarczej, demograficznej, społecznej oraz o stanie środowiska, opisującym województwa, powiaty oraz gminy jako podmioty systemu organizacji społecznej i administracyjnej państwa” (GUS, 2013). Dane do analiz zostały pobrane dla wszystkich gmin w Polsce, łącznie z kodem jednostki samorządu terytorialnego. Do analiz wykorzystano najświeższe dostępne w chwili kwerendy bazy danych (za 2010 rok). Informacje, po odpowiednim przygotowaniu, zostały dołączone do zbioru danych indywidualnych uczniów.

Główny Urząd Statystyczny publikuje informację na temat stopy bezrobocia jedynie na poziomie powiatów lub wyższym. Na poziomie gmin dostępna jest jedynie statystyka

udziału zarejestrowanych bezrobotnych w liczbie osób w wieku produkcyjnym. Ta ostatnia przyjmuje znacznie niższe wartości. Na przykład stopa bezrobocia w Polsce w 2011 roku wyniosła 12,5%, natomiast odsetek bezrobotnych wśród osób w wieku produkcyjnym – 8,0%. W skali kraju relacja wskaźnika udziału bezrobotnych do stopy bezrobocia na przestrzeni lat 2004–2011 była stabilna, w granicach 0,64 ($\pm 0,01$). Ze względu na możliwość skorzystania z informacji na niższym poziomie agregacji – a więc precyzyjniej opisującej lokalne uwarunkowania funkcjonowania szkół – w analizach jako wskaźnik sytuacji na gminnym rynku pracy wykorzystana została informacja na temat udziału zarejestrowanych bezrobotnych wśród ludności w wieku produkcyjnym.

Do skonstruowania wskaźników opisujących strukturę lokalnej gospodarki wykorzystane zostały informacje na temat liczby podmiotów ujętych w rejestrze REGON według sekcji i działów Polskiej Klasyfikacji Działalności z 2007 roku (PKD; GUS, 2007) (Sekcje od A do U, działy od 1 do 98) na poziomie gmin. Są to najdokładniejsze dane dostępne na tym poziomie podziału administracyjnego. Należy mieć na uwadze, iż rejestr REGON nie jest idealnym odzwierciedleniem rzeczywistości. Obarczony jest błędem wynikającym z tego, że funkcjonują w nim również podmioty, które zaprzestały już prowadzenia działalności. Na potrzeby prowadzonych analiz przyjęto założenie, iż wielkość tego błędu jest niezależna od gminy. Na podstawie wyżej wymienionych danych skonstruowano wskaźnik opisujący liczbę podmiotów wysokich technologii (*high tech*, HT) na 100 mieszkańców.

Przedsiębiorstwa wysokich technologii charakteryzują się relatywnie wysokimi nakładami na badania i rozwój, wysokim poziomem zatrudnienia pracowników naukowo-technicznych, intensywną współpracą z ośrodkami naukowo-badawczymi, szybkim procesem starzenia się opracowanych produktów i technologii oraz efektami działalności innowacyjnej w postaci patentów i licencji. Na tej podstawie wyróżnione zostały sekcje PKD, będące składową przemysłu wysokich technologii w Polsce (Świdurska, 2009). Opierając się na klasyfikacji PKD 2007, do podmiotów wysokich technologii zaliczono w Polsce te, które były ujęte w następujących sekcjach i działach klasyfikacji: Sekcja C dział 20, Sekcja C dział 21, Sekcja C dział 26, Sekcja C dział 27, Sekcja C dział 28, Sekcja C dział 29, Sekcja C dział 30, Sekcja J, Sekcja M, Sekcja P. Na ich podstawie obliczono wskaźnik opisujący liczbę podmiotów wysokich technologii na 100 mieszkańców.

Wskaźnikiem upowszechnienia nieobowiązkowej opieki przedszkolnej był odsetek dzieci w wieku 3–5 lat objętych taką opieką w 2007 roku (główna kohorta miała wtedy 5 lat). W analizach wykorzystano również informacje o wydatkach samorządu gminnego na wychowanie i oświatę, w przeliczeniu na mieszkańca gminy. Dane te dotyczą łącznie finansowania przez gminę wszystkich rodzajów szkół, a nie – jak byśmy chcieli – tylko szkół podstawowych. Wskazują jednak na zróżnicowanie finansowego zaangażowania samorządu w realizację zadań oświatowych.

W analizach uwarunkowań lokalnych ważnym czynnikiem jest również typ gminy. Często stosowany jest podział administracyjny na gminy miejskie, wiejskie oraz miejsko-wiejskie. Jego zasadniczą wadą jest to, iż niezbyt dobrze odzwierciedla zróżnicowanie gmin w zakresie składu społecznego mieszkańców. W klasyfikacji tej jako gminy wiejskie traktowane są zarówno gminy ościenne wielkich aglomeracji miejskich, jak i gminy peryferyjne bez ośrodka miejskiego. W tym kontekście lepszym wskaźnikiem zróżnicowania wydaje się wielkość gminy, zwłaszcza ze względu na prowadzenie analiz, łącznie z omówionymi powyżej

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

czynnikami lokalnymi. Ze względu na to, iż zmienna ta nie ma rozkładu normalnego, w analizach został wykorzystany logarytm naturalny wielkości gminy⁴.

Tabela 1 w aneksie do niniejszego rozdziału przedstawia statystyki opisowe dla wskaźników wykorzystywanych w prezentowanych poniżej modelach.

Model analizy danych

W celu odpowiedzi na pytanie badawcze użyto hierarchicznych modeli regresji. Zastosowano modele trzypoziomowe z losową stałą, uwzględniające poziom ucznia, oddziału i szkoły. Analizy przeprowadzono oddzielnie dla umiejętności matematycznych, umiejętności czytania oraz dla świadomości językowej. Wyniki testów umiejętności w tych zakresach są przedstawiane na skali 100;15 w próbie. Usunięto z modeli te czynniki, których związek z osiągnięciami był nieistotny statystycznie na poziomie 0,05.

Analizy wykonano za pomocą programu MPlus 7, wykorzystując metodę estymacji *full maximum likelihood*. Do oszacowania istotności efektów stałych wykorzystano odporne (*robust*) błędy standardowe. Wartości istotne na poziomie istotności $p < 0,05$ zaznaczono w tabelach pogrubionym drukiem. Analizy wykonano z uwzględnieniem wag warunkowych (zob. rozdział 1) na wszystkich trzech poziomach, odpowiednio przeskalowanych za pomocą zastosowanego programu.

Wyniki

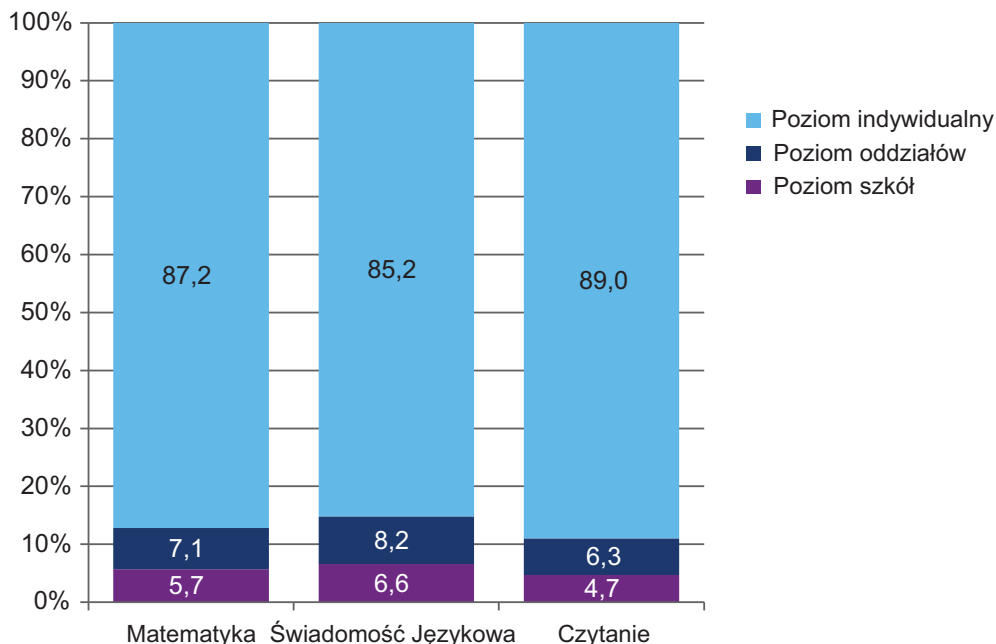
Na wstępie przedstawimy wyniki analiz opisujących międzyszkolne i międzyoddziałowe zróżnicowanie wyników testów osiągnięć uczniów w trzech analizowanych zakresach, tj. umiejętności czytania, świadomości językowej oraz umiejętności matematycznych. Przeprowadzono je, wykorzystując trzypoziomowe modele puste. Oszacowanie efektów losowych na poziomie szkół, oddziałów i uczniów pozwoliło wyliczyć współczynniki korelacji wewnątrzgrupowej. Ich wartość mówi, jaką część zmienności miar poszczególnych umiejętności można przypisać podziałowi uczniów na szkoły i oddziały.

Przeprowadzona analiza wykazała istotne międzyszkolne i międzyoddziałowe zróżnicowanie ze względu na badane umiejętności. Podział uczniów na szkoły wyjaśnił 4,7% wariancji umiejętności czytania, 5,7% wariancji umiejętności matematycznych i 6,6% zmienności miary świadomości językowej (Rysunek 2.1). W tych zakresach podział na oddziały w obrębie szkoły wyjaśnił odpowiednio 6,3%, 7,1% oraz 8,2% zmienności. W każdym z analizowanych zakresów zróżnicowanie międzyszkolne było mniejsze niż zróżnicowanie międzyoddziałowe. Jednocześnie skala tego zróżnicowania daje podstawy do prowadzenia analiz z wykorzystaniem modeli trzypoziomowych i poszukiwania czynników warunkujących osiągnięcia uczniów klas I–III.

Analizy przeprowadzone dla poszczególnych umiejętności mierzonych w badaniu dają nieco inny obraz.

⁴ W efekcie interpretacja współczynnika regresji przy tej zmiennej (po podzieleniu przez 100) będzie mówiła, o ile zmieni się wartość zmiennej zależnej, jeśli wielkość gminy zmieni się o 1%.

Rysunek 2.1. Wariancja międzyszkolna, międzyoddziałowa i wewnątrzoddziałowa jako odsetek wariancji całkowitej.



Na osiągnięcia z matematyki nie oddziałuje płeć, nie ma znaczenia, czy dziecko posłane zostało do szkoły jako sześciolatek. Nie ma również znaczenia czas uczęszczania do przedszkola ani to, czy uczeń uczęszczał do zerówki i czy była to zerówka przy szkole, czy prowadzona w ramach przedszkola. Osiągnięcia z matematyki nie są powiązane ani z prestiżem zawodu wykonywanego przez rodziców, ani z ich statusem społeczno-ekonomicznym. Nie ma znaczenia zasobność gospodarstwa domowego, jego wielkość ani kolejność urodzenia dziecka.

Na podstawie wartości standaryzowanych można wskazać, że najsilniejszy wpływ na osiągnięcia z matematyki ma przynależność do starszej kohorty (opóźnienie w starcie szkolnym) – uczniowie z tej grupy osiągają przeciętnie wynik o 12,8 punktu niższy niż uczniowie z kohorty głównej (skala o średniej 100 i odchyleniu standardowym 15 dla wszystkich uczniów wypełniających test z matematyki). Na podstawie przeprowadzonych analiz nie da się jednak stwierdzić, czy gorsze wyniki są efektem późniejszego posłania dziecka do szkoły, czy silnej selekcji negatywnej ze względu na niższy potencjał ucznia. Nie można również zweryfikować, czy lepszą strategią jest opóźnianie startu szkolnego, czy też posyłanie dzieci o niższym potencjale z kohortą podstawową. Jest to temat do dalszych badań.

Drugim pod względem siły wpływu czynnikiem jest poziom inteligencji płynnej. Uczniowie, których inteligencja jest o jedno odchylenie standardowe wyższa niż przeciętna w próbie, uzyskali średnio o 7,59 punktu więcej w teście. Kolejnym czynnikiem pod względem siły oddziaływania jest zamieszkiwanie w niepełnej rodzinie trzypokoleniowej – obniża ono osiągnięcia o 2,81 punktu na skali umiejętności matematycznych. Następny

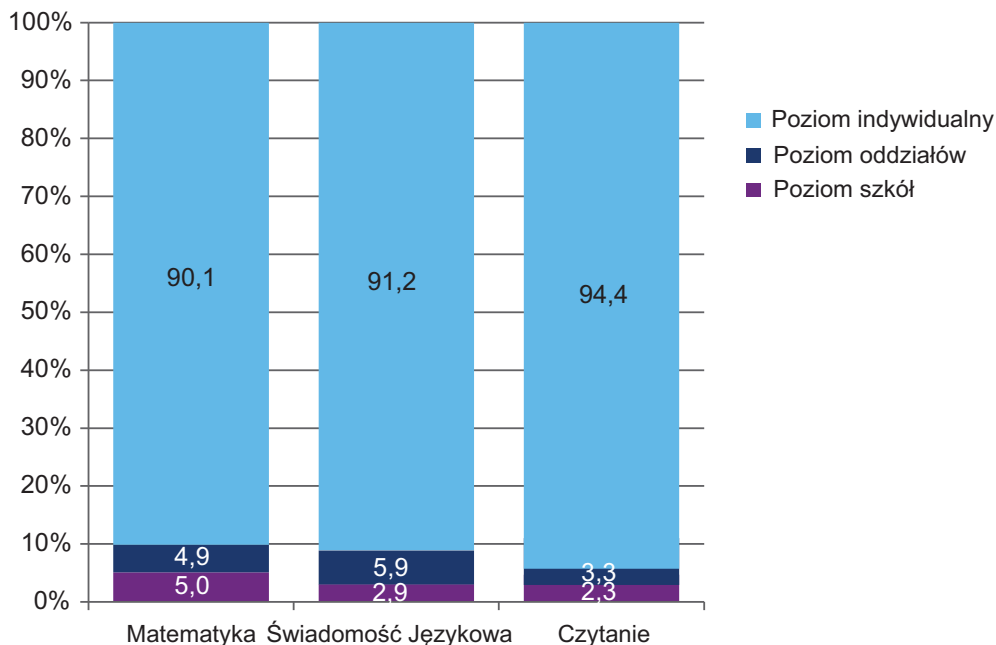
2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

czynnikiem to wykształcenie rodziców. Wzrost liczby lat nauki o rok zwiększa osiągnięcia z matematyki o 0,66 punktu. Z kolei uczniowie, których przynajmniej jedno z rodziców jest bezrobotne, uzyskują średnio o 1,70 punktu mniej niż uczniowie, których żadne z rodziców nie jest bezrobotne. Kolejnym pod względem siły oddziaływania czynnikiem jest poziom aspiracji edukacyjnych rodziców wobec dzieci na początku pierwszej klasy. Wyrażając poziom tych aspiracji liczbą lat nauki potrzebnych do osiągnięcia danego wykształcenia, można powiedzieć, że wzrost o rok prowadzi do podniesienia wyników w teście osiągnięć o 0,51 punktu. Pozostałe dwa czynniki mają porównywalną siłę oddziaływania: wzrost wieku o miesiąc podnosi uzyskiwane wyniki o 0,22 punktu, a wzrost zasobności gospodarstwa o jedno odchylenie standardowe prowadzi do wzrostu uzyskiwanych wyników o 1,15 punktu. Z czynników na poziomie lokalnym znaczenie ma odsetek dzieci objętych w gminie opieką przedszkolną (choć, jak pamiętamy, uczęszczanie do przedszkola rozpatrywane na poziomie indywidualnym nie miało znaczenia). Szczegóły przedstawione zostały w Tabeli 2.2, w kolumnie „Matematyka”.

Dla wyników testu świadomości językowej nie ma znaczenia, czy dziecko zostało posłane do szkoły jako sześciolatek. Nie ma również znaczenia długość uczęszczania do przedszkola ani to, czy uczeń uczęszczał do zerówki i czy była to zerówka przy szkole, czy prowadzona w ramach przedszkola. Świadomość językowa nie jest powiązana ani z bezrobociem rodziców, ani z ich statusem społeczno-ekonomicznym. Nie ma znaczenia również wielkość gospodarstwa domowego, w którym dziecko się wychowuje.

Poniżej omówione zostaną czynniki indywidualne, oddziałujące na świadomość językową uczniów w kolejności od najmocniej do najslabiej powiązanego (na podstawie wartości standaryzowanych). Podobnie jak w przypadku umiejętności matematycznych, najsilniejszy wpływ na wyniki ma przynależność do starszej niż główna kohorty wiekowej. Uczniowie rozpoczynający naukę w pierwszej klasie jako rocznikowe ósmiolatki uzyskują średnio o 12,82 punktu mniej. Drugim pod względem siły wpływu czynnikiem jest poziom inteligencji uczniów na początku pierwszej klasy. Wzrost poziomu o jedno odchylenie standardowe powoduje wzrost wyników w teście świadomości językowej o 5,89. Trzecim czynnikiem jest płeć. Dziewczynki uzyskały średnio o 4,42 punktu więcej. Na kolejnym miejscu jest wychowywanie się w niepełnej rodzinie trzypokoleniowej – obniża ono osiągnięcia o 3,32 w stosunku do wychowania w rodzinie pełniej dwupokoleniowej. Wychowanie w innym typie rodziny oznacza spadek osiągnięć o 3,88 punktu. Zajmowanie w strukturze rodzeństwa środkowej pozycji pod względem kolejności urodzenia, oznacza w stosunku do bycia jedynakiem spadek o 3,43 punktu. Również bycie najmłodszym przekłada się na ubytek 1,82 punktu. Wzrost o jedno odchylenie standardowe wskaźnika zasobności gospodarstwa domowego podnosi osiągnięcia o 2,14 punktu. Z kolei wzrost wykształcenia rodzica o rok podnosi osiągnięcia o 0,54 punktu. Również wzrost o jedno odchylenie wartości wskaźnika aspiracji edukacyjnych rodziców w stosunku do dzieci podnosi osiągnięcia o 0,47 punktu. Ważnym czynnikiem jest również wiek uczniów. Każdy miesiąc więcej wpływa na podniesienie poziomu świadomości językowej o 0,23 punktu. Wzrost prestiżu zawodu wykonywanego przez rodziców o jeden punkt na skali podnosi osiągnięcia ucznia o 0,08 punktu. Trudny do zinterpretowania, na szczęście bardzo słaby efekt odnotowano dla czynnika posiadania przez dziecko miejsca do nauki. Na poziomie lokalnym znaczącym wyznacznikiem osiągnięć jest gminny poziom bezrobocia. Podobnie jak w wypadku miejsca do nauki,

Rysunek 2.2. Wariancja międzyszkolna, międzyoddziałowa i wewnątrzoddziałowa jako odsetek wariancji całkowitej po uwzględnieniu czynników pozaszkolnych analizowanych w badaniu.



jest to efekt trudny do interpretacji. Wzrost udziału osób bezrobotnych wśród osób w wieku produkcyjnym o 1 punkt procentowy wpływa na podniesienie świadomości językowej uczniów o 0,30 punktu. Jest to relacja o niespodziewanym kierunku – odmiennym od wyników uzyskanych w badaniu gimnazjalistów (Dolata i in., 2013), z tego względu wymaga ona pogłębionych analiz. Kolejny czynnik lokalny to wydatki gminy na oświatę i wychowanie w przeliczeniu na mieszkańca. Wzrost nakładów o 1000 zł prowadzi do podniesienia osiągnięć o 4,11 punktu. Szczegóły przedstawiono w Tabeli 2.2, w kolumnie „Świadomość językowa”.

Zestaw czynników powiązanych istotnie statystycznie z umiejętnością czytania jest najkrótszy. Nie mają znaczenia żadne czynniki z poziomu lokalnego oraz ani jeden wskaźnik struktury rodziny. Podobnie jak w wypadku umiejętności omawianych powyżej, nie ma znaczenia uczęszczanie do przedszkola i zerówki, a także nie ma znaczenia, czy dziecko należy do głównej kohorty, czy też posłane zostało do szkoły jako sześciolatek. Z osiągnięciami tymi nie jest powiązany również status społeczno-ekonomiczny rodziny ucznia, prestiż wykonywanego zawodu, ani bycie bezrobotnym przynajmniej przez jednego z rodziców.

Szeregując – jak powyżej – czynniki malejąco pod względem wartości standaryzowanych współczynników regresji, można stwierdzić, że największe znaczenie ma późniejsze niż główna kohorta rozpoczęcie nauki w klasie pierwszej. Wyniki uzyskiwane przez takich uczniów są średnio o 10,94 punktu niższe niż uczniów z głównej kohorty. Drugim pod względem ważności czynnikiem jest inteligencja. Wzrost jej poziomu o jedno odchylenie

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

standardowe przekłada się na umiejętność czytania wyższą o 5,74 punktu. Na trzecim miejscu znalazła się płeć ucznia. Dziewczynki wypadają w teście czytania przeciętnie o 3,33 punktu lepiej. Kolejnym czynnikiem jest wykształcenie rodziców. Wzrost jego poziomu o rok przekłada się na umiejętności wyższe o 0,90 punktu. Piąty czynnik stanowi zamożność gospodarstwa domowego. Jej wzrost o jedno odchylenie standardowe oznacza wzrost umiejętności czytania o 1,58 punktu. Znaczenie mają również aspiracje edukacyjne rodziców wyrażone liczbą lat nauki potrzebnych do ich zrealizowania. Wzrost o jeden rok przekłada się na wzrost umiejętności czytania o 0,48 punktu. Ostatnim z czynników jest wiek uczniów. Uczniowie starsi o miesiąc uzyskiwali średnio o 0,17 punktu więcej. Szczegóły przedstawiono w Tabeli 2.2, w kolumnie „Czytanie”.

Wprowadzenie do modeli zmiennych z poziomu indywidualnego oraz lokalnego (tam, gdzie mają one znaczenie) pozwoliło ograniczyć różnicowanie międzyszkolne oraz międzyoddziałowe osiągnięć szkolnych. Współczynnik warunkowej korelacji wewnątrzgrupowej – którą można interpretować jako różnicowanie po zrównaniu poziomów pomiędzy szkołami i oddziałami pod kątem uwzględnionych w modelu zmiennych niezależnych – zmniejszył się zarówno dla poziomu szkół jak, i oddziałów.

Podział uczniów na szkoły wyjaśnił 2,3% wariacji umiejętności czytania (w stosunku do 4,7% bez zmiennych niezależnych), 5,0% wariacji umiejętności matematycznych (w stosunku do 5,7%) i 2,9% zmienności miary świadomości językowej (w stosunku do 6,6%). W tych zakresach podział na oddziały w obrębie szkoły wyjaśnia odpowiednio 3,3%, 4,9% i 5,9% zmienności (w stosunku do 6,3%, 7,1% oraz 8,2% zmienności w modelach pustych). Stopień, w jakim szkoły mogą różnić się efektywnością nauczania, pokazano na Rysunku 2.2.

Podsumowanie

Wyniki przeprowadzonych analiz pozwalają w większości potwierdzić znaczenie wytypowanych na podstawie dotychczasowych badań pozaszkolnych czynników wpływających na osiągnięcia szkolne.

Analizy wskazują, że dla osiągnięć w badanych obszarach bardzo duże znaczenie ma opóźniony start szkolny. Potwierdza to dotychczasową wiedzę w tym zakresie. W kontekście wieku należy zwrócić uwagę, że dzieci posłane do szkoły wcześniej – sześciolatki – uzyskiwały wyniki porównywalne z siedmiolatkami. Jednakże każdy miesiąc więcej – od urodzenia do pomiaru umiejętności – dawał badanemu uczniowi dodatkowe punkty w teście. Zgodnie z oczekiwaniami najistotniejszym czynnikiem kształtującym wyniki testów we wszystkich zakresach jest inteligencja uczniów na starcie szkolnym. Tak jak w innych badaniach, okazało się, że ma ona nieco większe znaczenie dla osiągnięć z matematyki niż ze świadomości językowej czy umiejętności czytania.

Spośród czynników statusowych niezmiennie oddziałuje zasobność gospodarstwa domowego oraz poziom wykształcenia rodziców. Nieistotnym predyktorem z tej grupy jest międzynarodowy wskaźnik statusu społeczno-ekonomicznego (ISEI) rodziny. To wynik niespodziewany. Może mieć to źródło w sposobie jego konstrukcji jako uśrednionego wyniku statusu z licznych badań narodowych. W przyszłych analizach można byłoby podjąć próbę wykorzystania wskaźników statusu społeczno-ekonomicznego stworzonych jedynie dla

Polski. Pozostałe dwa wskaźniki (prestż zawodu i bezrobocie) sprawdziły się częściowo, a ich znaczenie należy pogłębić w kolejnych badaniach i/lub analizach. Wskaźniki struktury rodziny okazały się niezłymi predyktorami umiejętności matematycznych i świadomości językowej.

Wskaźniki z poziomu lokalnego, opisujące społeczno-gospodarcze otoczenie szkoły, nie wydają się na etapie szkoły podstawowej oddziaływać w ten sam sposób na poziom poszczególnych umiejętności. W przypadku poziomu bezrobocia jest to wręcz efekt odwrotny od spodziewanego. Z tego względu nie ma powodów, aby uwzględniać te wskaźniki w modelach efektywności nauczania.

Tabela 2.2. Uwarunkowania osiągnięć szkolnych uczniów po pierwszym etapie edukacyjnym w zakresie umiejętności matematycznych, świadomości językowej oraz umiejętności czytania. Niestandaryzowane efekty stałe (bez czynników nieistotnych dla żadnej ze skal)*.

	Matematyka		Świadomość językowa		Czytanie	
	Wsp. niestand.	SE	Wsp. niestand.	SE	Wsp. niestand.	SE
Oszacowanie efektów stałych						
Poziom ucznia						
Płeć	-0,472	0,432	4,421	0,388	3,325	0,427
Starszy niż główna kohorta	-12,800	2,155	-12,823	1,671	-10,943	1,935
Młodszy niż główna kohorta	0,522	1,355	1,092	1,283	-2,201	1,154
Wiek ucznia	0,219	0,062	0,230	0,067	0,173	0,071
Aspiracje edukacyjne rodziców	0,506	0,125	0,473	0,117	0,479	0,132
Wykształcenie rodziców	0,661	0,109	0,537	0,122	0,904	0,108
Prestż zawodu			0,078	0,024		
Bezrobocie jednego z rodziców	-1,698	0,796				
Zamożność gosp. domowego	1,148	0,298	2,141	0,320	1,579	0,345
Miejsce do pracy dla ucznia			-0,944	0,374		
Najmłodszy (ref. jedynak)			-0,159	0,639		
Średni (ref. jedynak)			-3,426	0,823		
Najstarszy (ref. jedynak)			-1,822	0,573		
Rodzina pełna, trzypokoleniowa**	-0,011	0,480	-0,034	0,472		
Rodzina niepełna, dwupokoleniowa**	-1,109	0,937	-1,715	0,998		
Rodzina niepełna, trzypokoleniowa**	-2,806	1,133	-3,319	1,055		
Inna struktura rodziny**	-3,078	1,998	-3,880	1,522		
Test Matrycy Ravena (I pomiar)	7,590	0,239	5,889	0,263	5,738	0,297
Stała	100,324	1,065	94,170	2,567	100,010	0,564

2. Pozaszkolne uwarunkowania osiągnięć szkolnych na pierwszym etapie edukacyjnym

	Matematyka		Świadomość językowa		Czytanie	
	Wsp. niestand.	SE	Wsp. niestand.	SE	Wsp. niestand.	SE
Poziom szkoły						
Poziom bezrobocia w gminie			0,298	0,125		
% dzieci w przedszkolu w gminie	0,041	0,019				
Wydatki na oświatę i wychowanie w tys. zł na osobę w gminie			4,109	1,751		
Oszacowanie efektów losowych						
Wariancja efektów szkół	6,735	2,627	4,125	2,038	3,644	1,765
Wariancja efektów oddziałów	6,651	2,244	8,312	2,103	5,183	1,969
Wariancja na poziomie ucznia	122,581	3,484	129,203	3,362	146,568	4,429
Współczynniki korelacji wewnątrzgrupowej						
Poziom szkół		0,050		0,029		0,023
Poziom oddziałów		0,049		0,059		0,033
Poziom szkół i oddziałów łącznie		0,099		0,088		0,056
Podsumowanie						
Deviance		26 231,48		26 893,68		27 636,93
Liczba szacowanych parametrów		18		23		12
Liczba uczniów		4 364		4 436		4 503
Liczba klas		286		286		286
Liczba szkół		164		164		164

* Pogrubieniem zaznaczono wartości współczynników istotne statystycznie na poziomie $p < 0,05$.

** Grupą odniesienia jest rodzina pełna, dwupokoleniowa.

Rozdział 3

Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Szkoły różnią się średnimi wynikami uzyskiwanymi przez ich uczniów w testach osiągnięć. Zróżnicowanie to obserwujemy już po pierwszym etapie edukacyjnym. Czy oznacza to, że szkoły, które osiągnęły średnio wyższe wyniki, uczyły efektywniej niż te, które uzyskały słabsze rezultaty w testach? Niekoniecznie. Wyniki w testach osiągnięć nie mówią nam całej prawdy o jakości nauczania w szkole. Zależą one bowiem również od takich cech uczniów czy środowiska lokalnego, na które szkoła nie ma wpływu. Placówce zlokalizowanej w bardziej sprzyjającym środowisku, do której uczęszczają uczniowie o wyższym poziomie ogólnych zdolności poznawczych, pochodzący z rodzin o wyższym statusie społeczno-ekonomicznym, łatwiej będzie wypracować wysokie osiągnięcia niż takiej, która pracuje w mniej sprzyjających warunkach. Charakterystyki uczniów i ich rodzin oraz cechy środowiska lokalnego, mające znaczenie dla osiągnięć szkolnych po pierwszym etapie edukacyjnym, a także skala różnic między nimi, zostały szczegółowo przeanalizowane w poprzednim rozdziale. Wyniki pokazują, że nie możemy obarczać szkoły odpowiedzialnością za poziom osiągnięć szkolnych uczniów. Jeśli chcemy ewaluować pracę szkół (lub dostarczyć narzędzi do ewaluacji wewnętrznej), musimy posłużyć się takimi wskaźnikami, które w sytuacji idealnej obejmują tylko to, na co szkoła ma wpływ.

Wyniki analiz omówione w rozdziale 2 pokazały, że poziom osiągnięć szkolnych zależy od wielu cech uczniów, takich jak inteligencja, wiek, różne wymiary statusu społeczno-ekonomicznego rodziny pochodzenia, aspiracje rodziców związane z wykształceniem ich dzieci czy struktura rodziny. Jednak, gdyby polski system edukacyjny był zorganizowany w taki sposób, że uczniowie byłiby przydzielani do szkół równomiernie pod względem tych cech, zagregowane miary osiągnięć szkolnych (takie jak np. średni wynik testu w szkole) nie byłyby obciążone istnieniem tych zależności. Każda szkoła miałaby takie same szanse na wysoki wynik, ponieważ w konsekwencji takiego przydziału uczniów do szkół, do każdej placówki uczęszczaliby podobni uczniowie. Tak jednak nie jest. W przypadku szkół podstawowych wybór placówki jest przede wszystkim konsekwencją rejonizacji. Powodować to może istnienie znaczących różnic międzyszkolnych w charakterystykach uczniów mających znaczenie dla ich osiągnięć. Im większe będą te różnice dla zmiennych, które są silniej powiązane z wynikami uczenia się, tym bardziej średni wynik szkoły w testach osiągnięć będzie nieadekwatną miarą jakości jej pracy. Skala tych różnic zostanie pokazana w tym rozdziale.

Wypracowanie miary, która będzie pokazywała rzeczywisty wkład szkoły w osiągnięcia uczniów, wymaga uwzględnienia tego, z jakimi uczniami (o jakich cechach) pracowała dana placówka. Zanim jednak przejdziemy do omówienia modeli, które uwzględniają przy wyznaczaniu wyników różne czynniki niezależne od szkoły, zwróćmy uwagę na jeszcze jeden aspekt tych wskaźników. Wszystkie z nich koncentrują się na poznawczych wynikach kształcenia, mierzonych testami osiągnięć szkolnych. Oczywiście nie wynika to z przeświadczenia, że szkoła nie realizuje innych celów i że osiągnięcia intelektualne są jedynymi wartymi uwagi wynikami kształcenia. Niemniej inne obszary pracy szkoły, jak wyniki wychowania czy jakości opieki, nie są łatwo mierzalne i trudno sobie wyobrazić ich pomiar na szeroką skalę. Dlatego rozważania tego rozdziału skupiają się na wskaźnikach efektywności nauczania, w zakresie mierzonym testami osiągnięć szkolnych.

Miara taka wydaje się niezwykle potrzebna, a nie ma obecnie systemowych rozwiązań mających na celu jej dostarczenie. Obserwuje się jednak duże zainteresowanie szkół podstawowych oceną własnej pracy oraz diagnozą umiejętności uczniów już po pierwszym etapie edukacyjnym. Do dobrowolnego badania OBUT w 2013 roku przystąpiło ponad 70% szkół podstawowych w Polsce (Pregler, 2013). Oferty firm komercyjnych także cieszą się zainteresowaniem. W 2013 roku do testu kompetencji trzecioklasisty, przygotowanego przez jedno z najpopularniejszych wydawnictw oferujących taki produkt, przystąpili uczniowie od 43 do 61% szkół, w zależności od województwa¹. Część szkół i nauczycieli przeprowadza zapewne także własne testy diagnostyczne. Jednak, aby wyniki testów osiągnięć mogły być wykorzystywane na potrzeby ewaluacji, niezbędna jest możliwość porównania rezultatów danej placówki z wynikami innych szkół. Podstawowym punktem odniesienia są wszystkie szkoły w kraju lub mniejszej jednostce terytorialnej (np. województwie). Jednak bardziej wartościowa dla szkoły będzie możliwość zobaczenia swoich wyników na tle reprezentatywnej grupy szkół pracujących z podobnymi uczniami i w podobnych warunkach. Szkoły, analizując swoje wyniki pod takim kątem, uzyskająby informację o tym, czy udało im się wypracować więcej czy mniej, niż inne szkoły pracujące z uczniami o podobnych cechach. Aby dobrze odpowiedzieć na potrzeby placówek, należy spełnić dwa warunki. Po pierwsze, wspomniany punkt odniesienia do porównań międzyszkolnych musi zostać ustalony na podstawie reprezentatywnej dla kraju próby szkół. Zaznaczyć należy, że nie jest konieczne, by wszystkie szkoły w kraju zostały objęte pomiarami osiągnięć oraz dodatkowych cech kontekstowych. Wystarczy, że zostanie przeprowadzone badanie na reprezentatywnej próbie szkół, a uzyskane w ten sposób dane posłużą do wyliczenia odpowiednich modeli statystycznych. Następnie każda szkoła zainteresowana takimi wskaźnikami mogłaby przeprowadzić u siebie analogiczne pomiary i porównać je z wynikami badania reprezentatywnego. Po drugie, co już zasugerowane zostało wcześniej, należałoby zebrać informację o tym, z jakimi uczniami i w jakich warunkach szkoła pracuje.

Najpopularniejsze modele biorące pod uwagę cechy uczniów rozpoczynających naukę w danej szkole to modele edukacyjnej wartości dodanej. Modele te uwzględniają przede wszystkim poziom osiągnięć uczniów na progu danej szkoły. W rozdziale 2 wspomniano, że w przypadku pierwszego etapu edukacyjnego pomiar poziomu umiejętności mających znaczenie dla osiągnięć szkolnych na progu szkoły jest utrudniony. Do niedawna brakowało

¹ Według danych wydawnictwa: <http://www.operon.pl/oskt/wyniki>

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

dobrych narzędzi diagnostycznych, które można by wykorzystywać do tego celu. Dzięki pracom Instytutu Badań Edukacyjnych powstało narzędzie mierzące umiejętności na starcie szkolnym (Test Umiejętności na Starcie Szkolnym, TUNSS) na trzech skalach: umiejętności matematycznych, pisania i czytania (Karwowski i Dziedziewicz, 2012). Jeśli narzędzie to będzie rozwijane, a bank zadań systematycznie uzupełniany, możliwe stanie się wykorzystywanie jego na potrzeby wyliczania wskaźników dla pierwszego pomiaru edukacyjnego. Niemniej problem, który nadal by pozostał, to koszt takiego badania (zarówno finansowy, jak i czasowy). Zarówno TUNSS, jak i inne narzędzia, które mogą być wykorzystane przy badaniu dzieci sześć-, siedmioletnich, wymagają indywidualnie przeprowadzanych badań, z uwagi choćby na to, że uczeń w tym wieku zwykle nie umie samodzielnie przeczytać poleceń do zadań. Dlatego, niezależnie od dostępności narzędzi do pomiaru umiejętności uczniów na progu szkoły podstawowej, zbieranie tych danych jest co najmniej kłopotliwe.

Ale czy to oznacza, że w przypadku pierwszego etapu edukacyjnego jesteśmy skazani na posługiwanie się tylko średnimi wynikami testów w ewaluacji pracy szkół? Nie, z pomocą przychodzą inne metody, które wywodzą się z tej samej grupy co modele edukacyjnej wartości dodanej. W modelach tych nie uwzględnia się jednak uprzednich osiągnięć, w związku z czym interpretacja ich wyników jest trochę odmienna. Celem tego rozdziału jest przedstawienie modeli efektywności nauczania wykorzystujących informacje o cechach uczniów warunkujących osiągnięcia szkolne, ocena możliwości wypracowania i użyteczności takich modeli dla pierwszego etapu edukacyjnego w polskim systemie edukacji oraz pokazanie ograniczeń i problemów związanych z ich zastosowaniem.

Różne modele efektywności nauczania

Najpopularniejszymi modelami efektywności nauczania, stosowanymi na szeroką skalę, są modele wartości dodanej. Jest to zbiór modeli statystycznych, których celem jest szacowanie wkładu szkoły w postęp, jaki zrobili uczniowie w zakresie realizacji określonych celów edukacyjnych (np. wyników nauczania). Ponieważ modele te koncentrują się na postępie, ich podstawą jest co najmniej dwukrotny pomiar stopnia realizacji celów kształcenia (OECD, 2008). Modele EWD uwzględniające dwa pomiary osiągnięć szkolnych: „na wejściu” (na progu danego etapu kształcenia) i „na wyjściu” (na zakończenie edukacji na danym etapie) są stosowane w Polsce dla gimnazjów, liceów ogólnokształcących i techników². Mogą one jednak uwzględniać większą liczbę pomiarów osiągnięć szkolnych (jeśli tylko są one dostępne), zwiększając w ten sposób rzetelność oszacowań (Raudenbush i Bryk, 2002).

Modele EWD mogą się różnić także dodatkowymi zmiennymi (poza pomiarami osiągnięć szkolnych), które są w nich uwzględniane. Potrzeba włączenia do modelu dodatkowych zmiennych kontekstowych wynika z przeświadczenia, że na wzrost poziomu umiejętności uczniów mogą wpływać także inne, niezależne od szkoły czynniki. Uważa się, że przy wyliczaniu efektów szkół należy je wziąć pod uwagę, by jak najlepiej

² Więcej informacji na ten temat można znaleźć na stronie projektu: www.ewd.edu.pl oraz w publikacjach zespołu, które można pobrać ze strony.

oczyścić wskaźnik z tego, co jest niezależne od szkoły. Najprostsze pod względem tego kryterium modele EWD nie uwzględniają żadnych zmiennych kontekstowych, wykorzystują jedynie wyniki pomiaru osiągnięć szkolnych. Zakłada się w nich, że wpływ większości ważnych czynników niezależnych od szkoły na osiągnięte wyniki jest kontrolowany przez uwzględnienie w modelu pomiaru uprzednich osiągnięć ucznia. Przykładem modeli EWD, w których nie bierze się pod uwagę zmiennych kontekstowych, jest system stosowany obecnie w Anglii, w którym w ostatnich latach zrezygnowano z uwzględniania w modelu dodatkowych czynników na rzecz modeli prostszych w interpretacji³, oraz amerykański system EVAAS (*Education Value-Added Assessment System*). Jednak w tym drugim przypadku spełnionych jest kilka dodatkowych warunków, które pozwalają lepiej wytrącić ze wskaźników EWD wpływ dodatkowych charakterystyk uczniów (McCaffrey, Lockwood, Koretz, Hamilton i RAND Education, 2003). Modele te są liczone z wykorzystaniem corocznych pomiarów osiągnięć szkolnych, wskaźniki są wyliczane w odniesieniu do nauczycieli (a nie szkół), a przydział nauczycieli do uczniów zmienia się co roku.

Natomiast modele EWD, które poza pomiarami osiągnięć szkolnych uwzględniają różne zmienne kontekstowe, najwygodniej jest podzielić według często wykorzystywanej w literaturze, a wprowadzonej przez Stephena Raudenbusha i J. Douglasa Willmsa (1995), klasyfikacji na modele typu A i B. Podział ten wyrasta z różnych celów, jakim mają służyć wskaźniki. Wskaźniki typu A odpowiadają potrzebom rodziców chcących wybrać najlepszą szkołę dla swojego dziecka. Uwzględniają one bowiem tylko cechy indywidualne uczniów, takie jak np. płeć, wiek, pochodzenie społeczne, pomijając charakterystykę środowiska, w którym pracuje szkoła, jak np. średni poziom osiągnięć uczniów w szkole, poziom bezrobocia w otoczeniu szkoły, jej lokalizację. Pokazują więc one, w której szkole uczeń ma największe szanse na wysoki wynik, niezależnie od tego, w jakim stopniu jest on efektem działania szkoły, a w jakim korzystnego środowiska, w którym żyje. Efekty typu B są interesujące z kolei z perspektywy ewaluatora, który chce ocenić efektywność pracy szkoły. W tym wypadku ważne jest jak najlepsze oczyszczenie wskaźników także z wpływu czynników związanych ze środowiskiem, w którym funkcjonuje placówka. Szkoły mogą się bowiem różnić wynikami nie tylko dlatego, że pracują z uczniami o różnych cechach, ale także dlatego, że pracują w różnych środowiskach (np. w otoczeniu, w którym obserwujemy większy lub mniejszy poziom bezrobocia, który może przekładać się na aspiracje edukacyjne i motywację do nauki młodzieży, a te z kolei na wyniki). Dlatego w modelach typu B uwzględnia się dodatkowo zmienne kontekstowe, charakteryzujące otoczenie szkoły. Modele, które oprócz cech indywidualnych ucznia uwzględniają także dodatkowe zmienne, nazywamy kontekstowymi modelami EWD.

Wskaźniki wyliczone z różnych modeli będą ze sobą dość silnie skorelowane, jednak dla niektórych szkół mogą dawać znacząco odmienne rezultaty (Lenkeit, 2013; Timmermans, Doolaard i de Wolf, 2011). Dlatego dobór zmiennych kontekstowych nie może być przypadkowy. Powinien być konsekwencją przyjętego podstawowego celu, jakiemu mają służyć wskaźniki. Dodatkowo zasadność i sposób uwzględnienia każdej zmiennej w modelu powinien być poddany szczegółowym analizom.

³ <https://www.raiseonline.org/News.aspx?NewsID=303>

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Wszystkie modele EWD są oparte na założeniu, że na osiągnięcia szkolne wpływają różne czynniki, takie jak cechy uczniów i ich rodzin oraz cechy środowiska lokalnego szkoły, a ponieważ są to czynniki niezależne od placówki, ich znaczenie powinno być wytracone ze wskaźnika opisującego jakość pracy szkoły. Analogiczne założenia stoją za inną klasą modeli, które tym różnią się od modeli edukacyjnej wartości dodanej, że nie wykorzystują pomiaru uprzednich osiągnięć, lecz jedynie różne zmienne kontekstowe charakteryzujące uczniów i środowisko pracy szkoły. Modele te wywodzą się z potrzeby badań przekrojowych, w których nie ma możliwości zebrania danych w kilku momentach czasowych dla tych samych uczniów (Lenkeit, 2013). W takiej sytuacji dysponuje się więc wynikami jednego pomiaru osiągnięć oraz danymi o zmiennych kontekstowych, które często zostały zebrane także w tym samym czasie. Modele takie przyjęło się nazywać *contextualised attainment models* (OECD, 2008), co zostało przez nas przetłumaczone jako *kontekstowe modele efektywności nauczania*. Najczęściej modele takie uwzględniają informacje o cechach rodziny pochodzenia ucznia, takich jak status społeczno-ekonomiczny. Efekty szkół wyliczone z takiego modelu są interpretowane jako różnica w wynikach, w porównaniu do szkół, które pracowały z uczniami o takich samych cechach rodziny pochodzenia. Wiemy jednak, że status rodziny pochodzenia nie jest jedynym predyktorem sukcesu szkolnego dziecka. Należy jednak zdawać sobie sprawę z tego, że wybór określonych zmiennych kontekstowych do modelu kształtuje rozumienie efektywności nauczania, więc powinien on być teoretycznie uzasadniony (Coe i Fitz-Gibbon, 1998).

Jakie są więc różnice między kontekstowymi modelami efektywności nauczania a modelami EWD (poza techniczną kwestą uwzględnienia w nich lub nieuwzględnienia poziomu umiejętności na progu szkoły)? Uważa się, że dzięki uwzględnieniu w modelach EWD uprzednich osiągnięć, pozwalają one lepiej oszacować efektywność nauczania niż modele kontekstowe (OECD, 2008), ponieważ lepiej odzwierciedlają wkład szkoły w postępy, jakie robią uczniowie. Argumentując, zwraca się uwagę na kumulatywny charakter osiągnięć szkolnych i podkreśla fakt, że pomiar uprzednich osiągnięć zawiera informację nie tylko o cechach uczniów, ale także o ich uprzednich doświadczeniach edukacyjnych (Lenkeit, 2013; Raudenbush, 2004). Poziom umiejętności na progu danej szkoły jest także nośnikiem informacji o innych cechach uczniów, które mają znaczenie dla osiągnięć szkolnych, a są trudne do uchwycenia czy zbadania i rzadko są brane pod uwagę w modelach kontekstowych (Raudenbush, 2004). Dlatego o modelach kontekstowych zwykło się myśleć raczej jako o bardziej „sprawiedliwym” sposobie szacowania wyników szkół w porównaniu do surowego wyniku testu osiągnięć, niż o sposobie mierzenia EWD szkoły. Termin „edukacyjna wartość dodana” sugeruje bowiem, że mówimy o czymś, co w całości możemy przypisać działaniom szkoły (Fitz-Gibbon i Koch, 2000). Niemniej należy sobie zdawać sprawę z tego, że tak jak modele kontekstowe mają swoje słabe strony (nie uwzględniają być może wszystkich ważnych zmiennych, tak i modele edukacyjnej wartości dodanej przyjmują pewne uproszczone założenia).

Celem, jaki sobie stawiamy, jest zaproponowanie możliwych do zastosowania na szeroką skalę modeli efektywności nauczania, na podstawie danych zabranych w badaniu, a następnie ocena ich jakości i użyteczności. Postulat dotyczący ich wprowadzenia do systemu edukacji zostanie sformułowany tylko w sytuacji, jeżeli okaże się, że modele takie mogą dawać użyteczną i wartościową informację o pracy szkoły.

Problem badawczy

Podjęty tutaj problem badawczy ma wymiar praktyczny. Celem przedstawionych analiz jest opracowanie kontekstowego modelu efektywności nauczania dla pierwszego etapu edukacyjnego w Polsce oraz ocena jego użyteczności dla ewaluacji pracy szkół. Omówione zostaną także warunki, jakie musiałyby być spełnione, by wskaźniki dla zaproponowanych modeli mogły być wyliczane dla szkół podstawowych.

Metoda

Dane i metoda pomiaru

W analizach wykorzystano dane z badania podłużnego w szkołach podstawowych realizowanego w ramach projektu EWD. Najważniejsze informacje o organizacji i metodologii tego badania zawarte są w rozdziale 1. W tym miejscu przypomnimy tylko, że głównym elementem badania są trzy testy osiągnięć szkolnych (umiejętności czytania, świadomości językowej i umiejętności matematycznych), których wyniki zostały wyrażone na skali o średniej 100 i odchyleniu standardowym 15 w populacji uczniów rozpoczynających naukę w czwartej klasie szkoły podstawowej.

W analizach wykorzystano wiele wskaźników charakteryzujących uczniów i ich rodziny. Większość z nich została opisana wcześniej, stąd nie będą tu charakteryzowane. Opis ten uzupełnimy tylko o te wskaźniki, które pierwszy raz pojawiają się w analizach opisanych w niniejszym rozdziale.

Liczba książek. Wskaźnik ten przyjmuje cztery wartości: (a) do 25 książek, (b) od 26 do 100 książek, (c) od 101 do 200 książek, (d) powyżej 200 książek. Informacja ta była zbierana podczas I etapu badania za pomocą pytania zamkniętego w ankiecie rodzicielskiej.

Prosty indeks zasobności domu ucznia. Ponieważ dla wskaźników, które miałyby być wyliczane na szeroką skalę, ważne jest to, by dane potrzebne do ich wyznaczenia mogły być łatwo zebrane i przetworzone, skonstruowano prosty indeks zasobności na podstawie pytań o posiadane w gospodarstwie domowym dobra, z wybranych pozycji skali znajdującej się w ankiecie rodzicielskiej wykorzystanej w badaniu na I etapie. W pytaniu tym proszono rodziców o udzielenie informacji, czy w domu, w którym mieszka dziecko, znajdują się dobra takie jak komputer, którego dziecko może używać do nauki, literatura piękna, encyklopedie (łącznie 17 pozycji). Ze wszystkich pozycji wykluczono trzy, które były najsłabiej powiązane z ogólnym wymiarem zasobności. Z pozostałych utworzono prosty wskaźnik będący sumą posiadanych dóbr spośród wymienionych.

Plan analizy danych

Proces budowania i oceny kontekstowych modeli efektywności nauczania składał się z kilku etapów. Pierwszym była ocena konieczności i użyteczności wprowadzenia takich miar dla pierwszego etapu edukacyjnego na podstawie analizy zróżnicowania międzyszkolnego cech uczniów niezależnych od szkoły, a mających znaczenie dla osiągnięć szkolnych.

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Gdyby się bowiem okazało, że szkoły nie różnią się między sobą pod względem cech uczniów, z którymi pracują, średni wynik testu osiągnięć byłby wystarczającą miarą jakości nauczania. Każda szkoła miałaby bowiem taką samą szansę na wysoki wynik.

Następnie, korzystając z wniosków płynących z analiz pozaszkolnych uwarunkowań wyników nauczania (rozdział 2), wypracowano modele efektywności nauczania, które można uznać za wyczerpujące w kontekście posiadanych danych oraz wiedzy na temat uwarunkowań osiągnięć szkolnych. Uzupełniając analizy przedstawione w rozdziale 2, zbadano także znaczenie wybranych charakterystyk środowiska, w którym pracuje szkoła. Badania bowiem pokazują, że efekty kompozycyjne, takie jak średni status społeczno-ekonomiczny uczniów w szkole, mają znaczenie dla wyników nauczania niezależnie od wpływu danego czynnika na poziomie indywidualnym (Teddlie, Stringfield i Reynolds, 2000). Wyniki takiego modelu, który możemy roboczo nazywać wyczerpującym modelem kontekstowym efektywności nauczania, potraktowano jako punkt odniesienia dla modeli prostszych, mających jednak szansę na praktyczne zastosowanie. Model wyczerpujący jest bowiem trudny do powszechnego wykorzystania ze względu na to, że w jego skład wchodzi wyniki pomiarów, które wymagałyby zbyt dużych nakładów finansowych (jak np. indywidualny pomiar inteligencji uczniów wykonywany przez psychologa) i byłyby problematyczne ze względu na standaryzację. Nie bez znaczenia są też problemy etyczne związane z powszechnym pomiarem inteligencji.

W kolejnych krokach testowano więc modele zawierające tylko takie informacje, które szkoły mogłyby z powodzeniem samodzielnie zbierać i przetwarzać. Modele testowano, włączając krokowo kolejne zmienne mające znaczenie dla osiągnięć szkolnych. Każdy bardziej złożony model konfrontowano z prostszym, oceniając, czy włączenie następnej zmiennej do modelu znacząco poprawia jego dopasowanie i przyczynia się do redukcji niewyjaśnionej wariancji na poziomie szkół lub indywidualnym. W ten sposób uzyskano kilka modeli konkurencyjnych, różniących się stopniem złożoności.

W celu wyboru najlepszych modeli, wyniki porównywano z modelem wyczerpującym, analizując różnice wskaźników efektywności nauczania. Kryterium jakości modelu była jak najwyższa korelacja między efektami szkół z modelu wyczerpującego i prostszego. Przyjęto założenie, że w sytuacji, w której korelacja jest wysoka, a różnice w oszacowaniach dla poszczególnych szkół są niewielkie, model prostszy jest lepszy ze względu na to, że dostarcza niemal tę samą informację, natomiast koszt zebrania danych jest znacznie niższy. Szukano więc najprostszego modelu, który równocześnie da wyniki najbliższe modelowi wyczerpującemu.

Dodatkowo porównywano inne właściwości dla wskaźników wyliczonych z różnych modeli prostszych. Od wskaźników efektywności nauczania (wskaźniki typu B) oczekuje się, że będą w jak najmniejszym stopniu związane z cechami niezależnymi od szkoły, takimi jak średni poziom inteligencji czy średni status społeczno-ekonomiczny uczniów. Dodatkowym kryterium dla wyboru wskaźników była więc analiza związków ze zagregowanymi zmiennymi opisującymi kontekst pracy szkoły. Im niższe zaobserwowano korelacje, tym lepiej świadczyło to o jakości wskaźnika.

Zbadano także użyteczność wyliczonych wskaźników, analizując, w jakim stopniu za ich pomocą można odróżnić od siebie szkoły pracujące z efektywnością poniżej i powyżej przeciętnej. Oceniono również, czy wyliczone wskaźniki dają znacząco inną informację

niż nieprzetworzone wyniki w testach osiągnięć, a także, czy wskaźniki są specyficzne dla badanych obszarów nauczania.

Model analizy danych

W analizach posłużono się hierarchicznymi modelami liniowymi z losową stałą (Raudenbush i Bryk, 2002). Wskaźniki efektywności nauczania wyznaczono z modeli dwupoziomowych jako Bayesowskie predykcje a posteriori efektów losowych (*best linear unbiased predictions*). Czyli wyliczono je w podobny sposób, jaki wykorzystuje się przy szacowaniu wskaźników EWD dla gimnazjów, liceów i techników w Polsce (Pokropek i Żóltak, 2012). By móc wprost oszacować efekty dla szkół, w modelach uwzględniono dwa poziomy analizy, mimo że schemat doboru próby zdefiniował trzypoziomą strukturę danych. Wadą zastosowanych modeli jest więc to, że w analizach nie zawsze wykorzystuje się informacje o wszystkich uczniach na danym poziomie nauczania (w szkołach o większej liczbie oddziałów niż dwa wykorzystuje się losową próbę dwóch oddziałów). Do wyznaczenia efektów szkół wykorzystano oprogramowanie Stata 12 (procedurę XTMIXED). Analizy wykonano metodą największej wiarygodności, wykorzystując wagi warunkowe na obu poziomach analizy (zob. rozdział 1).

Wyniki

Zróznicowanie międzyszkolne cech uczniów mających znaczenie dla osiągnięć szkolnych

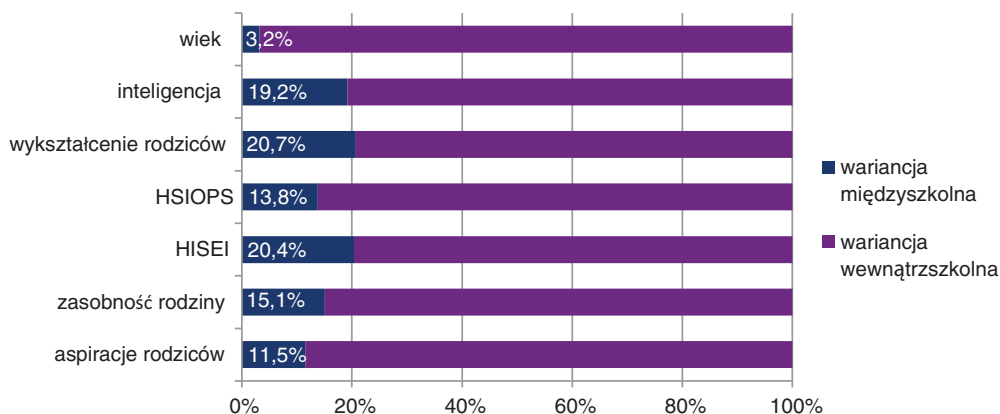
Prezentację wyników analiz zaczniemy od odpowiedzi na pytanie, czy szkoły podstawowe różnią się pod względem cech uczniów mających znaczenie dla ich osiągnięć szkolnych. Gdyby okazało się, że zróznicowanie międzyszkolne tych cech jest praktycznie pomijalne, nie byłoby potrzeby konstruowania wskaźników efektywności nauczania uwzględniających kontekst pracy szkoły. Średnie wyniki testów osiągnięć byłyby wystarczającą miarą jakości nauczania.

Na poniższym wykresie przedstawiono, jaką część wariancji poszczególnych zmiennych można przypisać podziałowi uczniów na szkoły. Wskaźnik ten pokazuje zatem, jak bardzo szkoły różnią się między sobą pod względem tych cech. Im większa wartość wskaźnika, tym większe zróznicowanie międzyszkolne. Udział wariancji międzyszkolnej w wariancji całkowitej wyliczono na podstawie oszacowań efektów losowych z dwupoziomowych modeli pustych ze stałą losową. Dla każdej zmiennej wyliczono osobny model.

Przeprowadzona analiza wykazała istnienie istotnego zróznicowania międzyszkolnego ze względu na wszystkie cechy uczniów wymienione na Rysunku 3.1. Szkoły w najmniejszym stopniu różnią się pod względem wieku uczniów (wyrażonego w miesiącach). Podział na szkoły wyjaśnia jedynie 3% zróznicowania tej zmiennej. Należy jednak zwrócić uwagę na to, że odsetek ten jest trochę wyższy niż w przypadku kohorty o rok starszej, w której wyniósł on około 1% (Jasińska i Modzelewski, 2013). Różnice te mogą być konsekwencją

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Rysunek 3.1. Zróźnicowanie międzyszkolne cech uczniów mających znaczenie dla osiągnięć szkolnych*.



* Udział wariancji międzyszkolnej w stosunku do wariancji całkowitej, oszacowania na podstawie modeli pustych z losową stałą.

okresu przejściowego reformy obniżającej wiek obowiązku szkolnego. Szkoły w dość dużym stopniu różnią się poziomem inteligencji uczniów. Wariancja międzyszkolna wyników testu inteligencji przeprowadzanego w pierwszej klasie wyniosła około 19%. Odsetek ten jest także trochę wyższy niż zaobserwowany dla kohorty o rok starszej w badaniu uczniów z trzeciej klasy szkoły podstawowej (Jasińska i Modzelewski, 2013). Jego oszacowanie może być jednak trochę zawyżone ze względu na zaobserwowane problemy z wiarygodnością danych z pomiaru inteligencji na pierwszym etapie badania (więcej na ten temat w rozdziale 1). Dla czynników statusowych także zaobserwowano istotne i dość wysokie zróźnicowanie międzyszkolne, co znajduje potwierdzenie w literaturze przedmiotu (Dolata, 2008). Podział na szkoły wyjaśnił ponad 20% zmienności wykształcenia rodziców uczniów, podobnie jak w przypadku wskaźnika statusu społeczno-ekonomicznego HISEI. Oszacowania te są wyższe niż w przypadku wspomnianych już danych z badania przeprowadzonego na starszej o rok kohorcie uczniów i znajdują się na poziomie obserwowanego zróźnicowania międzyszkolnego tych wskaźników dla gimnazjów (Dolata i in., 2013). Nieco niższe niż dla wykształcenia rodziców i HISEI zróźnicowanie międzyszkolne obserwuje się dla wskaźników prestiżu wykonywanego zawodu (HSIOPS), zasobności rodziny ucznia oraz aspiracji rodziców wobec dziecka (wyrażonych w latach nauki) mierzonych na początku formalnej nauki szkolnej.

Wyniki analiz potwierdziły istnienie znaczącego zróźnicowania międzyszkolnego cech uczniów wpływających na ich osiągnięcia szkolne. Szkoły w największym stopniu różnią się pod względem inteligencji uczniów oraz rodzinnych czynników statusowych. Warto zauważyć, że informacja o wykształceniu rodziców, która jest względnie prosta do pozyskania w systemie oświaty, dobrze wyjaśnia międzyszkolne zróźnicowanie statusu rodzin uczniów. Dodatkowo należy zastrzec, że porównania zróźnicowania międzyszkolnego dla różnych wskaźników trzeba traktować jako przybliżenie tego obrazu. Analizowane wskaźniki

prawdopodobnie różnią się bowiem rzetelnością, a obserwowane różnice w zakresie udziału wariancji międzyszkolnej w wariancji całkowitej mogą być częściowo konsekwencją realnych różnic, częściowo mogą też wynikać z wyższej lub niższej rzetelności pomiaru poszczególnych zmiennych.

Kontekstowe modele efektywności nauczania

W poprzedniej części rozdziału wykazano, że miara efektywności nauczania wymaga uwzględnienia tego, że szkoły różnią się między sobą pod względem cech uczniów mających znaczenie dla osiągnięć szkolnych. W tej części zaprezentowane zostaną modele efektywności nauczania uwzględniające różne zmienne kontekstowe. Tłem do porównań będą wyniki modelu pustego, nieuwzględniającego żadnych zmiennych kontrolnych oraz wyniki modelu zawierającego wszystkie ważne zmienne kontekstowe nazwanego przez nas modelem wyczerpującym.

W Tabelach 3.1–3.3 zaprezentowano wyniki konkurencyjnych modeli efektywności nauczania dla poszczególnych testów osiągnięć szkolnych: testu umiejętności czytania, świadomości językowej i umiejętności matematycznych. Modele (0) to tzw. modele puste, uwzględniające tylko podział uczniów na szkoły. Modele (1) to wyczerpujące modele efektywności nauczania, w których zostały uwzględnione te spośród zmiennych niezależnych, co do których stwierdzono, że mają znaczenie dla osiągnięć szkolnych (na podstawie wyników analiz opisanych w rozdziale 2) oraz zmienne będące dodatkową informacją o kontekście pracy szkoły (do modeli włączono średni poziom inteligencji uczniów oraz średni poziom wykształcenia rodziców uczniów). Uwzględnienie w modelu tylko jednego zagregowanego wskaźnika statusu jest wystarczające, a zarazem wskazane ze względu na to, że wskaźniki te (średnie wykształcenie rodziców, średnia wartość wskaźników HISEI, HSIOPS, zasobność rodziny) są ze sobą wysoko skorelowane (korelacje: 0,85–0,94). Dodatkowo średnie wykształcenie rodziców jest najwyżej skorelowane ze wszystkimi pozostałymi. Kolejne modele to konkurencyjne modele efektywności nauczania, w których uwzględniono tylko te zmienne opisujące uczniów i kontekst pracy szkoły, które mają szansę być mierzone na szeroką skalę. Modele (0) oraz (2)–(6) zostały policzone na próbie tych samych uczniów (wyłączono z analizy wszystkich uczniów, dla których nie zgromadzono kompletu danych), by można było poprawnie porównywać jakość ich dopasowania. Modele (1) zostały policzone z uwzględnieniem mniejszej liczby obserwacji, z uwagi na braki danych dla dodatkowych zmiennych kontrolnych. W tabelach pogrubionym drukiem zaznaczono wartości współczynników istotne statystycznie na poziomie $p < 0,05$.

Wyniki modeli pustych (0) pokazują, jak bardzo szkoły różnią się między sobą wynikami nauczania po pierwszym etapie kształcenia. W przyjętym dwupoziomowym modelu analizy uczniowie–szkoły podział na placówki wyjaśnia 8,6% wariancji wyników testu umiejętności czytania, 11,8% wariancji wyników testu świadomości językowej oraz 10,5% wariancji osiągnięć szkolnych z matematyki. Oszacowania te są spójne z wynikami innego badania prowadzonego w Polsce na uczniach będących w podobnym wieku (Jasińska i Modzelewski, 2013). Do porównywalnych wniosków prowadzą także analizy krajowych danych zebranych w międzynarodowych badaniach osiągnięć uczniów PIRLS i TIMSS. Jednak należy zwrócić uwagę, że oszacowania udziału wariancji międzyszkolnej w wariancji

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

całkowitej na podstawie przyjętego w poprzednim rozdziale modelu trzypoziomowego uczniowie–oddziały–szkoły dają znacząco inne wyniki. Różnice wynikają z odmienności algorytmów dekompozycji wariancji w modelu dwu- i trzypoziomym oraz – przede wszystkim – z tego, że znaczna część szkół podstawowych to placówki jednooddziałowe, czyli oddział jest dla nich tożsamy ze szkołą. Choć dla oddania hierarchicznej struktury próby model trzypoziomowy uwzględniający pogrupowanie uczniów w oddziały jest bardziej adekwatny (z tego powodu wykorzystaliśmy go do zbadania wyznaczników osiągnięć szkolnych), to dla analiz efektywności pracy szkoły z oczywistych powodów bardziej adekwatny będzie model dwupoziomowy. Daje on też bardziej wiarygodne oszacowania wariancji międzyszkolnej.

Wyniki te pokazują, że międzyszkolne zróżnicowanie osiągnięć uczniów jest znaczące i istotne statystycznie, choć nie możemy powiedzieć, że jest bardzo duże. Dla porównania zróżnicowanie międzyszkolne wyników egzaminu gimnazjalnego w części humanistycznej w 2013 roku było na poziomie około 24% (Dolata, 2013)⁴. Zatem po pierwszym etapie kształcenia obserwujemy mniejsze zróżnicowanie międzyszkolne osiągnięć uczniów, niż ma to miejsce pod koniec nauki w gimnazjach, jednak stwierdzone zróżnicowanie należy uznać za znaczące. Jego część związana jest z tym, że szkoły różnią się pod względem tego, jacy uczniowie do nich uczęszczają. Dowodzi tego spadek wariancji międzyszkolnej efektów losowych w modelach (1). Uwzględniono w nich te zmienne niezależne, które są związane z osiągnięciami szkolnymi uczniów. Nie będziemy tu dyskutować zaobserwowanych efektów, bo nie jest to celem tego rozdziału, związku te zostały omówione w rozdziale 2.

Komentarza wymagają jednak efekty dla zmiennych z poziomu szkół. Przeprowadzone analizy pokazały, że w szkole średni poziom wykształcenia rodziców uczniów jest pozytywnie związany z wynikami trzech testów, dodajmy, że po wytrąceniu znaczenia tej zmiennej na poziomie indywidualnym. Wzrost średniego wykształcenia rodziców o jeden rok nauki potrzebny do osiągnięcia danego poziomu wykształcenia przekłada się na wzrost wyników osiągnięć uczniów w szkole o około 1 punkt. Zaobserwowany efekt jest spójny z wynikami innych badań (Teddlie i in., 2000). Zaskoczeniem może być natomiast ujemny efekt dla średniego poziomu inteligencji uczniów w szkołach. Jednak w kontekście zaobserwowanych problemów z systematycznym błędem pomiaru inteligencji w pierwszej klasie (które opisano w rozdziale 1), nie powinniśmy interpretować stwierdzonego efektu jako opisującego kontekstowe uwarunkowania osiągnięć. Może on bowiem być tylko konsekwencją tego, że nie udało się wystarczająco dobrze oczyścić danych. Jeśli jednak tak jest, to włączenie tej zmiennej do modelu pozwoli zniwelować wpływ ewentualnego systematycznego błędu pomiaru inteligencji na oszacowane efekty szkół.

Uwzględnienie w modelach (1) wszystkich ważnych dla osiągnięć szkolnych, a niezależnych od szkoły zmiennych, pozwoliło w znacznym stopniu zredukować niewyjaśnioną wariancję zarówno na poziomie indywidualnym, jak i na poziomie szkół. W modelach tych podział na szkoły wyjaśnił znacznie mniejszy odsetek wariancji zmiennej zależnej: 3,5% dla testu umiejętności czytania, 5,4% dla testu świadomości językowej i 6,5% w przypadku testu z matematyki. Oznacza to, że gdyby szkoły pracowały z uczniami o takich samych

⁴ Wyniki te zostały policzone z uwzględnieniem dodatkowego, wyższego poziomu analizy (podziału na gminy), stąd przytaczany odsetek dotyczy średniego zróżnicowania międzyszkolnego wyników egzaminu wewnątrz gmin. Gdyby zróżnicowanie międzyszkolne zostało policzone bez podziału na gminy, uzyskane oszacowania byłyby nieco wyższe.

cechach, w znacznie mniejszym stopniu różniłyby się wynikami nauczania. Niemniej nadal istniałoby to zróżnicowanie i byłoby ono związane z tym, jak pracują szkoły.

Modele (2)–(6) to wybrane spośród wszystkich testowanych modele kontekstowe, pokazujące kolejne etapy konstrukcji modelu optymalnego. Model (2) jest najprostszy i zawiera podstawowe oraz łatwo dostępne dane o uczniu: jego płeć i wiek oraz wskaźnik statusu społecznego rodziny pochodzenia, jakim jest informacja o wykształceniu rodziców (uwzględniona jako zmienna na poziomie indywidualnym). Wskaźnik ten wybrano dlatego, że jest dobrym predyktorem sukcesu szkolnego dziecka, a także jest to informacja najłatwiejsza do pozyskania w systemie oświaty. Dodatkowo, jak pokazano wcześniej, dość dobrze odzwierciedla zróżnicowanie międzyszkolne ze względu na status rodzin uczniów. Zestaw zmiennych uwzględnionych w modelu (2) pozwala na wyjaśnienie od 11 do 15% wariacji na poziomie uczniów (por. pseudo R^2 (p1)) oraz od 57 do 65% wariacji z poziomu szkół (pseudo R^2 (p2))⁵. Jednak dodanie do modelu średniego wykształcenia rodziców uczniów danej szkoły (jako zmienną z poziomu szkoły) pozwala w jeszcze większym stopniu zredukować niewyjaśnioną wariację międzyszkolną. Pokazują to wyniki modelu (3). Zmienne w nim uwzględnione wyjaśniają 66–72% wariacji z poziomu szkół. Pod tym względem modele (3) są porównywalne z modelami wyczerpującymi (1). Jednak modele (3) w znacznie mniejszym stopniu pozwalają na redukcję niewyjaśnionej wariacji na poziomie indywidualnym. Dlatego kolejne modele zawierają dodatkowe zmienne z poziomu indywidualnego, pozwalające przewidywać osiągnięcia szkolne. Zmierzonymi tymi są dodatkowe wskaźniki powiązane ze statusem społeczno-ekonomicznym rodziny: aspiracje edukacyjne rodziców, liczba książek, prosty indeks zasobności rodziny, bezrobocie co najmniej jednego rodzica. Wyniki modelu (4) pokazują, że dodanie do modelu aspiracji rodziców pozwala zredukować niewyjaśnioną wariację nie tylko na poziomie indywidualnym, ale także na poziomie międzyszkolnym. Znacząco poprawia także dopasowanie modeli do danych (dla testu czytania Δ deviance(1) = 2814,36; dla testu świadomości językowej Δ deviance(1) = 3725,88; dla testu matematycznego Δ deviance(1) = 3279,98; $p = 0,05$). Liczba książek również jest dobrym predyktorem osiągnięć szkolnych. Istotnie związana jest z wynikami każdego z trzech testów (modele (5)). Dodanie jej do modelu zmniejsza niewyjaśnioną wariację na poziomie indywidualnym oraz – w mniejszym stopniu – również na poziomie szkół. Znacząco poprawia się także dopasowanie modeli do danych (dla testu czytania Δ deviance(3) = 3001,32; dla testu świadomości językowej Δ deviance(3) = 4082,50; dla testu matematycznego Δ deviance(3) = 3911,22; $p = 0,05$). Pozostałe dwie zmienne: prosty indeks zasobności rodziny (suma posiadanych dóbr z wymienionych, Min = 0, Max = 14) oraz dwuwartościowa zmienna kategoryjna, zdająca sprawę z tego, czy co najmniej jeden z rodziców dziecka jest bezrobotny, nie przynoszą już tak dużej poprawy w zakresie dopasowania modelu i redukcji wariacji (por. modele (6)). Trochę większe znaczenie mają dla wyjaśniania wariacji wyników testu umiejętności czytania niż pozostałych. Prostą indeks zasobności rodziny nie jest istotnie powiązany z wynikami testów z matematyki.

Z opisanych powyżej modeli oszacowano efekty dla szkół, będące wskaźnikami efektywności nauczania. W kolejnej części rozdziału porównamy właściwości tych wskaźników.

⁵ Interpretację oraz sposób wyznaczania wartości pseudo R^2 można znaleźć np. w: Domański i Pokropek, 2011.

Właściwości kontekstowych wskaźników efektywności nauczania

Przypomnijmy przyjęte kryteria oceny kontekstowych wskaźników efektywności nauczania. Po pierwsze, założyliśmy, że ich oszacowania powinny być jak najbliższe wyliczonym z modelu wyczerpującego. Po drugie, powinny być one w jak najmniejszym stopniu powiązane z czynnikami od szkoły niezależnymi (takimi jak średni poziom inteligencji uczniów czy średni status społeczno-ekonomiczny). Po trzecie, informacje potrzebne do ich zbudowania powinny być stosunkowo łatwe do pozyskania na szeroką skalę. W świetle tych kryteriów ocenimy wskaźniki wyliczone ze wcześniej opisywanych modeli. Posłużymy się analizą korelacji między wskaźnikami wyliczonymi dla szkół. W Tabelach 3.4–3.6 przedstawiono wartości współczynników korelacji liniowej Pearsona dla wskaźników odpowiadających trzem testom osiągnięć. Współczynniki istotne statystycznie na poziomie $p < 0,05$ zaznaczono pogrubionym drukiem.

Wskaźniki wyliczone na podstawie modeli wyczerpujących (1) będą traktowane jako punkt odniesienia dla pozostałych. Dzięki porównaniu wskaźników z modeli prostszych z wyznaczonymi z modeli (1) możliwe będzie udzielenie odpowiedzi na pytanie, czy modele prostsze są wystarczającymi modelami efektywności nauczania, czyli czy niosą za sobą tę samą informację.

Spośród wszystkich zaprezentowanych modeli najsilniejsze korelacje ze wskaźnikami z modeli (1) obserwujemy dla modeli (6) oraz (5). Dla modeli (5) współczynniki korelacji są niższe tylko o około 0,001. Zaobserwowane korelacje są wysokie (od 0,95 do 0,97), co pokazuje, że wskaźniki wyliczone z tych modeli są niemal identyczne jak wyliczone z modeli wyczerpujących. Zdecydowanie najslabiej związane ze wskaźnikami z modeli wyczerpujących są wskaźniki z modeli (2), które kontrolowały jedynie znaczenie płci, wieku i wykształcenia rodziców uczniów uwzględnionego tylko na poziomie indywidualnym. Korelacje te wyniosły 0,87–0,89. Wskaźniki z modelu (3) zawierającego informację o średniej wykształcenia rodziców uczniów prezentują się pod tym względem lepiej w porównaniu do wskaźników z modelu (2). Natomiast te, które zostały wyliczone z modelu (4) także trochę lepiej niż wyznaczone z modelu (3). Oczywiście zaobserwowane zależności nie dziwią. Im model bardziej złożony, tym jego wyniki bardziej podobne do wyników modelu wyczerpującego. Uwagę jednak należy zwrócić na to, w jakim stopniu określona komplikacja modelu poprawia właściwości wskaźników. W tym przypadku widzimy, że model (5) daje praktycznie tak samo dobre oszacowania jak model (6).

Przejdźmy do kolejnego kryterium. Wskaźniki efektywności nauczania nie powinny być skorelowane z charakterystykami szkoły, na które – przynajmniej bezpośrednio – nie może ona wpływać. I tak, wszystkie wskaźniki, poza wyliczonymi z modelu (2), nie są istotnie skorelowane ze zagregowanymi wskaźnikami statusu społeczno-ekonomicznego rodziny uczniów, takimi jak średnia wykształcenia rodziców, średni prestiż wykonywanego zawodu, średnia zasobność rodziny ucznia. Oczywiście należało się tego spodziewać, bowiem przy ich obliczaniu uwzględniono średni poziom wykształcenia rodziców. Skutecznie pozwoliło to kontrolować nie tylko znaczenie tej zmiennej, ale także innych wymiarów statusu. Jednak wskaźniki z modelu (2) są najsilniej spośród wszystkich konkurencyjnych powiązane ze średnim poziomem inteligencji uczniów w pierwszej klasie. Dla pozostałych wskaźników obserwujemy około dwukrotnie słabszy związek ze średnią inteligencją uczniów w pierwszej klasie, jest on najslabszy dla wskaźników z modeli (6) i (5).

Tabela 3.1. Kontekstowe modele efektywności nauczania dla umiejętności czytania.

Zmienna zależna: wynik testu umiejętności czytania	(0)	(SE)	(1)	(SE)
Oszacowanie efektów stałych				
Poziom ucznia				
Płeć (efekt dla dziewcząt)			3,31	0,433
Wiek w miesiącach			0,176	0,072
Starszy niż główna kohorta			-11,13	1,908
Młodszy niż główna kohorta			-2,06	1,133
Wykształcenie rodziców			0,677	0,140
Aspiracje rodziców			0,483	0,136
Liczba książek – kat. 2				
Liczba książek – kat. 3				
Liczba książek – kat. 4				
Prosty indeks zasobności rodziny				
Co najmniej jeden rodzic bezrobotny				
Test Matryc Ravena			6,10	0,318
HSIOPS			0,06	0,026
Zasobność rodziny			1,50	0,348
Stała	98,70	0,497	68,12	4,759
Poziom szkół				
Średnie wykształcenie rodziców			0,94	0,365
Średnia inteligencja uczniów			-3,43	1,061
Oszacowanie efektów losowych				
Wariancja efektów szkół	19,29	3,70	5,31	1,29
Wariancja na poziomie ucznia	205,67	5,75	148,45	4,38
Podsumowanie				
Deviance	1 420 249,3		1 341 923,3	
Log pseudolikelihood	-710 124,7		-670 961,6	
Pseudo R^2			0,32	
Pseudo R^2 (p2)			0,72	
Pseudo R^2 (p1)			0,28	
Liczba uczniów	4 508		4 444	
Liczba szkół	164		164	

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

(2)	(SE)	(3)	(SE)	(4)	(SE)	(5)	(SE)	(6)	(SE)
4,00	0,474	3,98	0,478	3,88	0,464	3,87	0,454	3,91	0,453
0,342	0,074	0,336	0,07	0,341	0,075	0,344	0,073	0,345	0,074
-15,83	1,779	-16,08	1,77	-15,02	1,790	-14,86	1,781	-14,72	1,797
-0,408	1,232	-0,394	1,25	-0,655	1,247	-0,881	1,228	-0,91	1,216
1,913	0,099	1,753	0,109	1,476	0,112	1,155	0,117	1,077	0,122
				0,885	0,144	0,767	0,149	0,720	0,152
						2,72	0,548	2,28	0,586
						4,02	0,85	3,43	0,83
						6,55	0,977	5,79	1,023
								0,272	0,13
								-2,17	0,945
74,63	1,58	63,07	3,439	54,31	3,47	60,378	3,74	60,52	3,74
		1,06	0,272	0,93	0,267	0,73	0,278	0,69	0,270
6,83	1,77	5,32	1,41	5,06	1,29	4,95	1,36	4,77	1,33
179,97	5,23	179,81	5,25	177,01	4,97	173,97	5,12	173,41	5,13
1 393 294,6		1 392 339,8		1 389 525,5		1 386 524,1		1 385 869,4	
-696 647,3		-696 169,9		-694 762,7		-693 262,1		-692 934,7	
0,17		0,18		0,19		0,20		0,21	
0,65		0,72		0,74		0,74		0,75	
0,12		0,13		0,14		0,15		0,16	
4 508		4 508		4 508		4 508		4 508	
164		164		164		164		164	

Tabela 3.2. Kontekstowe modele efektywności nauczania dla świadomości językowej.

Zmienna zależna: wynik testu świadomości językowej	(0)	(SE)	(1)	(SE)
Oszacowanie efektów stałych				
Poziom ucznia				
Płeć (efekt dla dziewcząt)			4,44	0,389
Wiek w miesiącach			0,234	0,067
Starszy niż główna kohorta			-13,06	1,72
Młodszy niż główna kohorta			1,35	1,27
Wykształcenie rodziców			0,446	0,127
Aspiracje rodziców			0,507	0,118
Liczba książek – kat. 2				
Liczba książek – kat. 3				
Liczba książek – kat. 4				
Prosty indeks zasobności rodziny				
Co najmniej jeden rodzic bezrobotny				
Test Matryc Ravena			6,16	0,267
HSIOPS			0,076	0,025
Zasobność rodziny			2,09	0,326
Miejsce			-0,979	0,372
Rodzina pełna, trzypokoleniowa*			0,092	0,464
Rodzina niepełna, dwupokoleniowa*			-1,89	1,03
Rodzina niepełna, trzypokoleniowa*			-3,30	0,992
Inna struktura rodziny*			-3,85	1,56
Najstarszy z rodzeństwa			0,100	0,636
W środku z rodzeństwa			-3,28	0,850
Najmłodszy z rodzeństwa			-1,68	0,587
Stała	98,41	0,565	65,49	5,139
Poziom szkół				
Średnie wykształcenie rodziców			1,34	0,386
Średnia inteligencja uczniów			-4,18	1,32
Oszacowanie efektów losowych				
Wariancja efektów szkół	26,41	4,58	7,58	1,95
Wariancja na poziomie ucznia	197,26	4,59	131,96	3,41
Podsumowanie				
Deviance	1 414 956,1		1 320 083,2	
Log pseudolikelihood	-707 478,0		-660 041,6	
Pseudo R^2			0,38	
Pseudo R^2 (p2)			0,71	
Pseudo R^2 (p1)			0,33	
Liczba uczniów	4 508		4 430	
Liczba szkół	164		164	

* Grupą odniesienia jest rodzina pełna, dwupokoleniowa.

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

(2)	(SE)	(3)	(SE)	(4)	(SE)	(5)	(SE)	(6)	(SE)
5,15	0,462	5,14	0,465	5,03	0,450	5,02	0,442	5,06	0,439
0,398	0,074	0,393	0,075	0,399	0,075	0,406	0,074	0,406	0,074
-18,25	1,681	-18,54	1,70	-17,37	1,67	-17,22	1,61	-17,05	1,618
2,80	1,377	2,85	1,38	2,56	1,41	2,38	1,34	2,35	1,353
1,91	0,096	1,734	0,102	1,42	0,106	1,07	0,108	0,983	0,113
				0,990	0,120	0,846	0,125	0,791	0,132
						3,37	0,567	2,86	0,622
						5,12	0,75	4,43	0,76
						7,04	0,841	6,16	0,905
								0,317	0,14
								-2,01	0,876
74,15	1,554	58,11	3,196	48,33	3,16	55,07	3,38	55,06	3,378
		1,43	0,25	1,27	0,242	1,06	0,25	1,02	0,247
10,75	2,78	7,85	2,33	6,63	1,87	6,46	1,97	6,59	2,03
168,39	4,14	168,18	4,11	165,12	4,07	161,27	4,10	160,48	4,08
1 383 826,9		1 382 320,7		1 378 594,8		1 374 512,3		1 373 748,9	
-691 913,4		-691 160,3		-689 297,4		-687 256,2		-686 874,5	
0,20		0,21		0,23		0,25		0,25	
0,59		0,70		0,75		0,76		0,75	
0,15		0,15		0,16		0,18		0,19	
4 508		4 508		4 508		4 508		4 508	
164		164		164		164		164	

Tabela 3.3. Kontekstowe modele efektywności nauczania dla umiejętności matematycznych.

Zmienna zależna: wynik testu umiejętności matematycznych	(0)	(SE)	(1)	(SE)
Oszacowanie efektów stałych				
Poziom ucznia				
Płeć (efekt dla dziewcząt)			-0,556	0,435
Wiek w miesiącach			0,219	0,061
Starszy niż główna kohorta			-12,59	1,97
Młodszy niż główna kohorta			0,491	1,29
Wykształcenie rodziców			0,605	0,116
Aspiracje rodziców			0,523	0,127
Liczba książek – kat. 2				
Liczba książek – kat. 3				
Liczba książek – kat. 4				
Prosty indeks zasobności rodziny				
Co najmniej jeden rodzic bezrobotny			-1,868	0,778
Test Matrycy Ravena			7,99	0,258
HSIOPS				
Zasobność rodziny			1,22	0,309
Rodzina pełna, trzypokoleniowa*			0,027	0,484
Rodzina niepełna, dwupokoleniowa*			-0,937	0,952
Rodzina niepełna, trzypokoleniowa*			-2,90	1,078
Inna struktura rodziny*			-3,32	1,932
Stała	98,47	0,550	69,03	5,21
Poziom szkół				
Średnie wykształcenie rodz.			1,26	0,390
Średnia inteligencja uczniów			-5,10	1,21
Oszacowanie efektów losowych				
Wariancja efektów szkół	23,24	3,93	8,65	1,90
Wariancja na poziomie ucznia	198,92	5,36	124,39	3,45
Podsumowanie				
Deviance	1 371 109,5		1 286 640,1	
Log pseudolikelihood	-685 554,8		-643 320,0	
pseudo R^2			0,40	
pseudo R^2 (p2)			0,63	
pseudo R^2 (p1)			0,37	
Liczba uczniów	4 368		4 358	
Liczba szkół	164		164	

* Grupą odniesienia jest rodzina pełna, dwupokoleniowa.

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

(2)	(SE)	(3)	(SE)	(4)	(SE)	(5)	(SE)	(6)	(SE)
0,360	0,497	0,336	0,501	0,249	0,487	0,242	0,479	0,293	0,481
0,442	0,077	0,436	0,077	0,439	0,075	0,447	0,072	0,450	0,071
-18,67	2,21	-18,86	2,183	-17,77	2,17	-17,68	2,137	-17,60	2,111
2,31	1,45	2,30	1,456	2,01	1,46	1,82	1,428	1,79	1,427
1,86	0,105	1,70	0,109	1,40	0,115	1,03	0,124	0,969	0,128
				0,965	0,147	0,834	0,149	0,805	0,156
						3,11	0,543	2,84	0,583
						5,23	0,77	4,91	0,841
						7,26	0,765	6,81	0,783
								0,165	0,136
								-2,65	0,933
77,37	1,84	63,63	3,75	54,10	3,786	61,10	3,854	61,69	3,85
		1,234	0,273	1,08	0,268	0,850	0,27	0,81	0,267
10,00	2,49	7,86	2,13	7,76	1,94	7,43	1,89	7,14	1,88
176,91	5,04	176,74	5,00	173,29	4,78	169,37	4,80	168,89	4,79
1 347 999,0		1 346 906,1		1 343 626,1		1 339 714,9		1 339 111,1	
-673 999,5		-673 453,0		-671 813,1		-669 857,4		-669 555,6	
0,16		0,17		0,19		0,20		0,21	
0,57		0,66		0,67		0,68		0,69	
0,11		0,11		0,13		0,15		0,15	
4 368		4 368		4 368		4 368		4 368	
164		164		164		164		164	

Tabela 3.4. Macierz korelacji między wskaźnikami dla szkół – obszar czytania.

Zmienna	IQ – I klasa	HEDU	HSIOPS	Zasob. rodz	Wynik testu	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Średnie IQ – I klasa	1,000										
Średni HEDU	0,552	1,000									
Średni HSIOPS	0,547	0,933	1,000								
Średnia zasobność rodziny	0,579	0,883	0,850	1,000							
Średni wynik testu czytania	0,563	0,731	0,697	0,704	1,000						
Wskaźnik kontekstowy – model (1)	0,005	0,008	-0,006	0,031	0,624	1,000					
Wskaźnik kontekstowy – model (2)	0,409	0,347	0,338	0,386	0,859	0,892	1,000				
Wskaźnik kontekstowy – model (3)	0,237	0,014	0,022	0,097	0,655	0,950	0,940	1,000			
Wskaźnik kontekstowy – model (4)	0,227	0,016	0,026	0,091	0,650	0,953	0,934	0,994	1,000		
Wskaźnik kontekstowy – model (5)	0,208	0,014	0,015	0,049	0,647	0,959	0,928	0,987	0,992	1,000	
Wskaźnik kontekstowy – model (6)	0,205	0,015	0,015	0,037	0,644	0,960	0,927	0,985	0,990	0,998	1,000

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Tabela 3.5. Macierz korelacji między wskaźnikami dla szkół – obszar świadomości językowej.

Zmienna	IQ – I klasa	HEDU	HSIOPS	Zasob. rodz.	Wynik testu	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Średnie IQ – I klasa	1,000										
Średni HEDU	0,552	1,000									
Średni HSIOPS	0,547	0,933	1,000								
Średnia zasobność rodziny	0,579	0,883	0,850	1,000							
Średni wynik testu św. językowej	0,547	0,737	0,694	0,699	1,000						
Wskaźnik kontekstowy – model (1)	0,025	0,002	-0,026	0,016	0,627	1,000					
Wskaźnik kontekstowy – model (2)	0,416	0,420	0,398	0,440	0,897	0,871	1,000				
Wskaźnik kontekstowy – model (3)	0,205	0,008	0,007	0,083	0,651	0,962	0,908	1,000			
Wskaźnik kontekstowy – model (4)	0,197	0,010	0,010	0,079	0,645	0,963	0,902	0,993	1,000		
Wskaźnik kontekstowy – model (5)	0,175	0,010	0,002	0,035	0,642	0,972	0,896	0,986	0,991	1,000	
Wskaźnik kontekstowy – model (6)	0,171	0,011	0,001	0,021	0,642	0,973	0,895	0,984	0,989	0,998	1,000

Tabela 3.6. Macierz korelacji między wskaźnikami dla szkół – dla obszaru umiejętności matematycznych.

Zmienna	IQ – I klasa	HEDU	HSIOPS	Zasob. rodz.	Wynik testu	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Średnie IQ – I klasa	1,000										
Średni HEDU	0,552	1,000									
Średni HSIOPS	0,547	0,933	1,000								
Średnia zasobność rodziny	0,579	0,883	0,850	1,000							
Średni wynik testu matematycznego	0,569	0,720	0,704	0,704	1,000						
Wskaźnik kontekstowy – model (1)	-0,008	-0,012	0,020	0,031	0,627	1,000					
Wskaźnik kontekstowy – model (2)	0,424	0,370	0,388	0,413	0,883	0,875	1,000				
Wskaźnik kontekstowy – model (3)	0,231	-0,007	0,034	0,086	0,657	0,947	0,924	1,000			
Wskaźnik kontekstowy – model (4)	0,220	-0,004	0,038	0,081	0,652	0,951	0,918	0,993	1,000		
Wskaźnik kontekstowy – model (5)	0,197	-0,005	0,027	0,034	0,645	0,954	0,910	0,985	0,990	1,000	
Wskaźnik kontekstowy – model (6)	0,195	-0,004	0,026	0,026	0,643	0,956	0,909	0,983	0,989	0,999	1,000

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

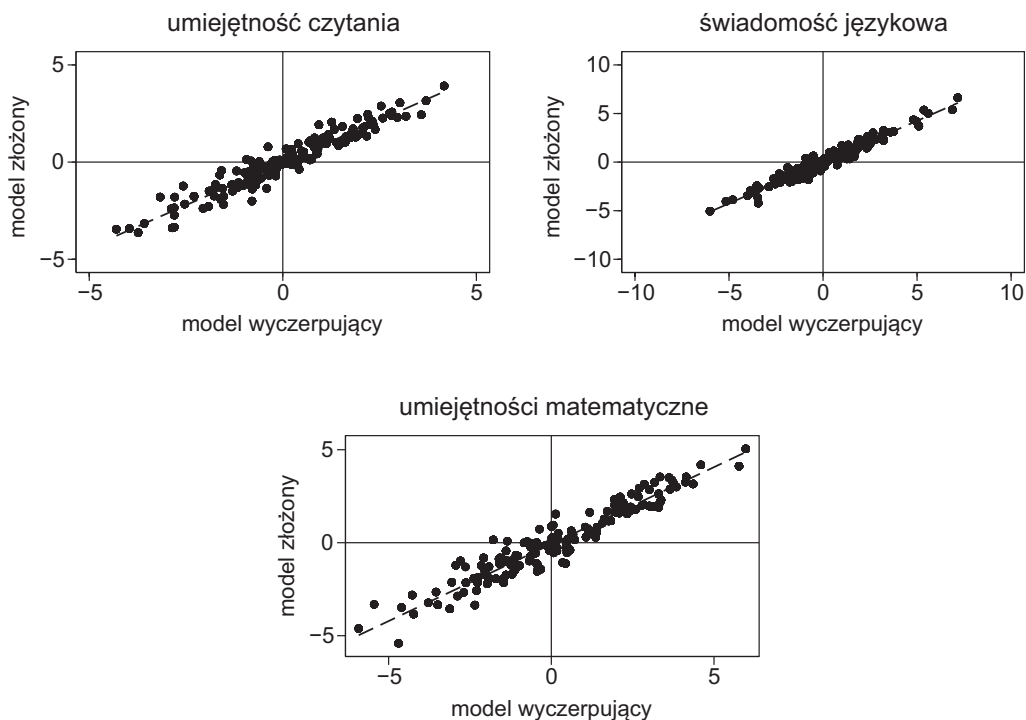
Przeprowadzone analizy pokazały, że najlepsze właściwości mają wskaźniki z modeli (6) i (5). Dają one wyniki najbliższe wyliczonym z modeli wyczerpujących, a także w najmniejszym stopniu są związane z kontekstem społecznym szkoły. Dodatkowo modele (6) i (5) wyjaśniają najwięcej wariacji wyników każdego z trzech testów. Model (6) ma nieco lepsze właściwości od modelu (5), jednak różnice są niewielkie. Oprócz tego korelacja między wskaźnikami z modeli (5) i (6) wynosi 0,998–0,999, co oznacza, że wyliczone wskaźniki są niemal identyczne. Model (6) zawiera jednak dodatkowe zmienne, których pozyskanie może być utrudnione. Pytanie o bezrobocie oraz o posiadane dobra materialne mogą być odbierane przez rodziców uczniów jako zagrażające, a ponieważ uwzględnienie tych informacji w modelu przynosi niewielkie korzyści, model (5) uznajemy za najlepszy z powyżej prezentowanych. Warto jednak zauważyć, że model (3), w którym uwzględnione są tylko zmienne względnie najprostsze do pozyskania w systemie oświaty (płeć i wiek ucznia oraz informacja o wykształceniu rodziców – na poziomie indywidualnym i jako średnia w szkole), ma tylko trochę gorsze właściwości niż model (5). Dlatego też zbadamy dokładniej, czy taki prosty model nie byłby wystarczający do oceny efektywności nauczania.

W dalszych analizach prezentowanych zarówno w tym, jak i w następnym rozdziale, wykorzystane zostaną modele (3) oraz (5). By uprościć przekaz, model (3) będziemy nazywać prostym modelem kontekstowym efektywności nauczania, natomiast model (5) – modelem złożonym.

Wspominaliśmy, że korelacje między wskaźnikami wyliczonymi ze złożonych modeli kontekstowych a wskaźnikami z modeli wyczerpujących są wysokie i wynoszą 0,95–0,97 (w zależności od testu). Należy jednak zweryfikować, czy niezależnie od głównej tendencji, dla konkretnych szkół nie obserwuje się wyjątkowo dużych różnic prowadzących do odmiennych wniosków w zależności od zastosowanego modelu. Na poniższych wykresach zamieszczonych na Rysunku 3.2 pokazano wartości wskaźników wyliczonych z modelu wyczerpującego (oś pozioma) i złożonego modelu kontekstowego (oś pionowa) dla każdego z trzech testów. Widzimy, że wartości wskaźników szkół wyliczone z tych dwóch modeli niewiele różnią się od siebie (największe różnice obserwujemy w przypadku matematyki). Jednak najbardziej niepokojącymi różnicami byłyby takie, które dostarczyłyby odmienną informację o efektywności nauczania w zależności od zastosowanego modelu. Czyli na przykład sytuacja, w której na podstawie wyników jednego modelu stwierdzilibyśmy ponadprzeciętną efektywność pracy szkoły, a na podstawie wyników drugiego modelu – poniżej przeciętną (lub vice versa). Zatem najbardziej niepokoją nas oszacowania dla szkół znajdujące się w II i IV ćwiartce wyznaczonej przez układ współrzędnych przebiegający przez punkt (0, 0). Widzimy jednak, że przypadków takich jest niewiele. W przypadku umiejętności czytania dotyczy to 7,3% szkół, świadomości językowej – 4,9%, a w przypadku matematyki dla 11% szkół znak wartości wskaźnika z obu modeli jest różny. Należy jednak zwrócić uwagę na to, że dotyczy to sytuacji, w których wartość wskaźnika z obu modeli jest bliska zeru. Dlatego, gdyby rozważyć szacowanie efektywności z przedziałem ufności, na podstawie obu modeli szkoły uzyskałyby informację, że nauczają z przeciętną efektywnością.

W przypadku modelu prostego dla trochę większej liczby szkół zauważamy następujące różnice: 9,6% dla umiejętności czytania, 7,9% dla świadomości językowej i 15,2% dla matematyki (Rysunek 3.3). Jednak znów są one relatywnie niewielkie.

Rysunek 3.2. Porównanie kontekstowych wskaźników efektywności nauczania – model złożony (5) a model wyczerpujący (1).



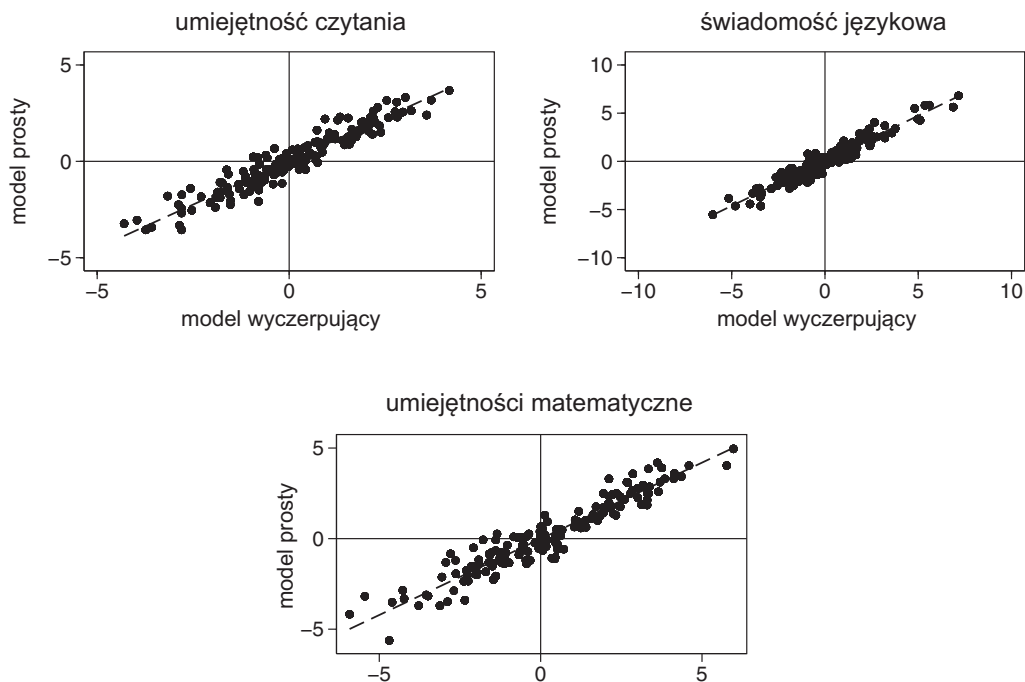
Kolejną ważną kwestią związaną z użytecznością zaproponowanych modeli jest pytanie o to, na ile są one pomocne w wyłanianiu szkół ponad i poniżej przeciętnych, jeśli chodzi o efektywność nauczania. Oczywiście, na to składa się z jednej strony faktycznie obserwowane zróżnicowanie międzyszkolne w tym zakresie, ale z drugiej strony, pewność oszacowań, która jest reprezentowana przedziałem ufności. Gdyby okazało się, że nie można odróżnić od siebie szkół ze względu na efektywność ich pracy, coroczne liczenie takiego wskaźnika nie miałyby żadnej wartości praktycznej.

Pewność oszacowań jest reprezentowana wielkością przedziałów ufności. Im większa pewność, tym krótsze przedziały ufności, a tym samym o większym odsetku szkół będzie można powiedzieć, że uczą z ponad lub poniżej przeciętną efektywnością. Błąd standardowy pomiaru (na podstawie którego wyznaczane są przedziały ufności) zależy od dwóch czynników: niewyjaśnionej przez model wariancji wewnątrzszkolnej oraz liczby obserwacji wewnątrz grupy (w tym przypadku liczby uczniów w szkole uwzględnionych w analizie; por. Pokropek, 2009). Im mniejsza wariancja na poziomie indywidualnym (czyli im większy procent wariancji udało nam się w modelu wyjaśnić) oraz im większa liczba przypadków uwzględniona w analizie w ramach danej grupy, tym błąd standardowy oszacowania efektu dla szkoły mniejszy.

W przeprowadzonym badaniu nie zbieraliśmy jednak informacji o wszystkich uczniach w szkole. W przypadku szkół, w których było więcej niż dwa oddziały na danym poziomie,

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Rysunek 3.3. Porównanie kontekstowych wskaźników efektywności nauczania – model prosty (3) a model wyczerpujący (1).



ograniczono liczbę badanych oddziałów do dwóch. Dodatkowo, w trakcie badania nie udało się pozyskać wszystkich danych dla wszystkich uczniów z badanych klas. Wyznaczone przedziały ufności dla tej próbki danych są więc trochę szersze niż byłyby, gdyby w badaniu uczestniczyli wszyscy uczniowie w wylosowanych szkołach. Gdyby wskaźniki efektywności nauczania były wyliczane systemowo, z dużą pewnością pomiarami zostaliby objęci wszyscy lub prawie wszyscy uczniowie w szkole. Dlatego, aby oszacować wielkości przedziałów ufności na potrzeby odpowiedzi na pytanie, jaki odsetek szkół miałby istotnie odróżniały od zera wskaźnik efektywności nauczania, wzięto pod uwagę informację o faktycznej wielkości szkoły, czyli o liczbie uczniów we wszystkich klasach na danym poziomie nauczania. Do wyliczenia przedziałów ufności wykorzystano zatem wyniki dekompozycji wariancji z modelu policzonego na danych z próby oraz przypisanej każdej szkole liczby uczniów na danym poziomie nauczania objętym badaniem testowym. Następnie dla każdej szkoły zweryfikowano, czy wyznaczony przedział ufności nie zawiera wartości 0. W ten sposób wskazano w próbie szkoły, które mają efektywność nauczania istotnie różną od przeciętnej. W kolejnym kroku, uwzględniając wagi próbkowania, obliczono, jaki odsetek szkół w populacji miałby wskaźnik efektywności nauczania istotnie różny od zera. Wyniki te zaprezentowano w Tabeli 3.7.

W zależności od obszaru badanych osiągnięć i zastosowanego modelu, odsetek szkół z istotnie różną od przeciętnej efektywnością nauczania waha się od 4,5 do prawie 14% dla 95-procentowych przedziałów ufności i od prawie 7 do 21% dla 90-procentowych

Tabela 3.7. Odsetek szkół, dla których wskaźnik efektywności nauczania jest istotnie różny od zera*.

Test	Przedział ufności			
	Model złożony		Model prosty	
	95	90	95	90
Umiejętności czytania	4,52	6,68	11,67	14,44
Świadomości językowej	6,44	7,71	12,22	18,33
Umiejętności matematycznych	7,54	12,52	13,89	21,11

* Dane na podstawie modeli wyliczonych dla wszystkich 180 szkół biorących udział w badaniu.

przedziałów ufności. Wydaje się, że to niewiele, szczególnie w przypadku umiejętności czytania. Jeśli jednak przełożymy te odsetki na liczbę szkół, okaże się, że od ponad 400 do prawie 2000 szkół miałyby szansę dowiedzieć się, że podejmowane przez nich działania w zakresie nauczania danych umiejętności są nieskuteczne lub szczególnie efektywne. Jeśli choć część tych szkół potrafiłaby dobrze wykorzystać tę informację, wielu uczniów mogłoby na tym zyskać.

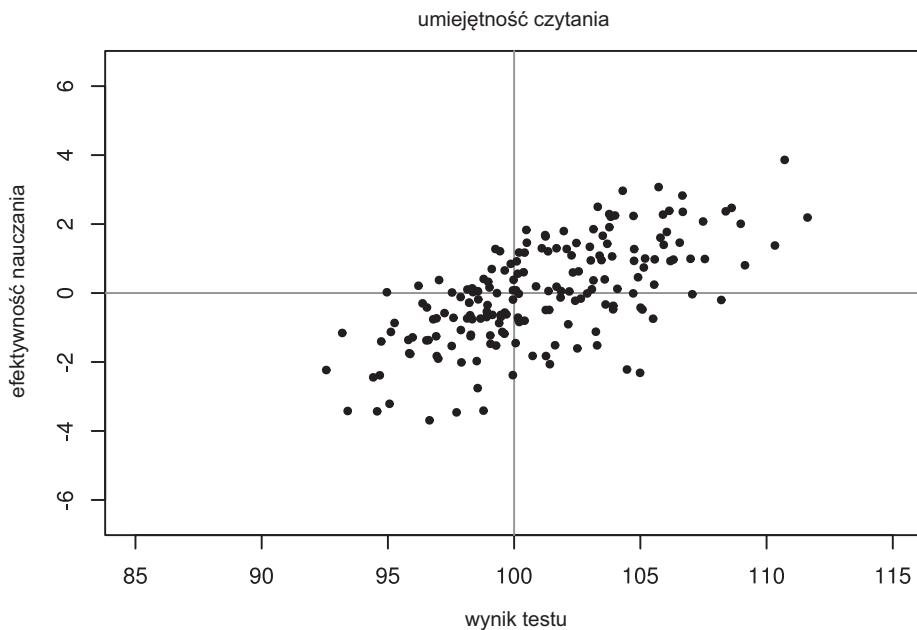
Na dane te należy spojrzeć jeszcze od innej strony. Skoro odsetek szkół z odróżnialną od przeciętnej efektywnością nauczania jest relatywnie nieduży, oznacza to, że w większości szkół na pierwszym etapie nauczania uczy się z podobną efektywnością. Jest to taki obraz, który zapewne chcielibyśmy zaobserwować. Gdyby było inaczej, okazałoby się, że jakość edukacji na tym poziomie jest bardzo zróżnicowana, a szansa ucznia na sukces szkolny w dużej mierze byłaby determinowana przez to, w jakiej szkole spędził pierwsze lata swojej nauki.

Kolejna warta poruszenia kwestia dotyczy tego, na ile informacja płynąca ze wskaźników efektywności jest wartościowym uzupełnieniem wyników testów, czyli czy dostarcza swoistej wiedzy, różnej od nieprzetworzonych wyników testów osiągnięć. Korelacje między wynikami testów a kontekstowymi wskaźnikami efektywności nauczania wynoszą około 0,64 dla modelu złożonego oraz około 0,65 dla modelu prostego, czyli mniej niż 50% zmienności tych miar to wariancja wspólna. Nawet biorąc pod uwagę rzetelność wskaźników, i tak rozpatrywane miary mają duże obszary wariancji swoistej. Pozytywny związek jest naturalny, bowiem wyższa lub niższa efektywność nauczania przekłada się na wyższe lub niższe wyniki. Związek między tymi zmiennymi zobrazujemy dodatkowo na Rysunkach 3.4–3.6 (tylko dla wskaźników z modelu złożonego). Naniesiono na nie pozycje szkół ze względu na wartość średniego wyniku w teście (oś pozioma) i wskaźnika efektywności nauczania (oś pionowa). Dodatkowo na wykresie zaznaczono liniami średni wynik w teście w populacji uczniów oraz średnią efektywność nauczania.

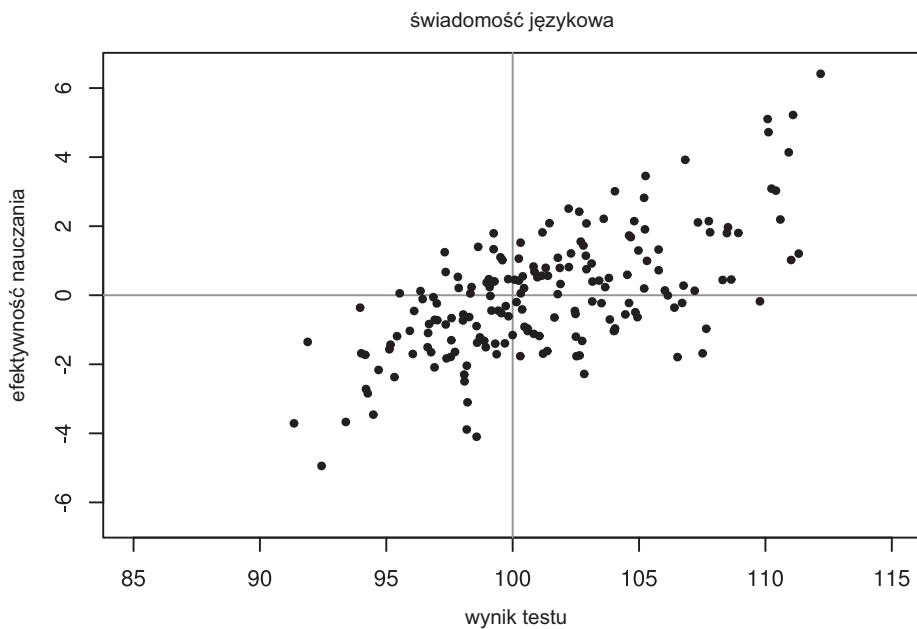
Widzimy wyraźnie, że informacje z obu wskaźników: miary efektywności i wyniku w teście osiągnięć wzajemnie się uzupełniają. Dzięki takiemu obrazowi możemy na przykład wyłonić szkoły, które mają ten sam wynik w teście, ale pracują z różną efektywnością, a także takie, które różnią się wynikiem w teście, mimo iż pracują z podobną efektywnością. Każdy z obu wskaźników jest miarą czego innego, ale razem dają pełniejszy obraz pracy szkoły.

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Rysunek 3.4. Wykres rozrzutu wyniku testowego i efektywności nauczania – umiejętność czytania.



Rysunek 3.5. Wykres rozrzutu wyniku testowego i efektywności nauczania – świadomość językowa.



Rysunek 3.6. Wykres rozrzutu wyniku testowego i efektywności nauczania – umiejętności matematyczne.

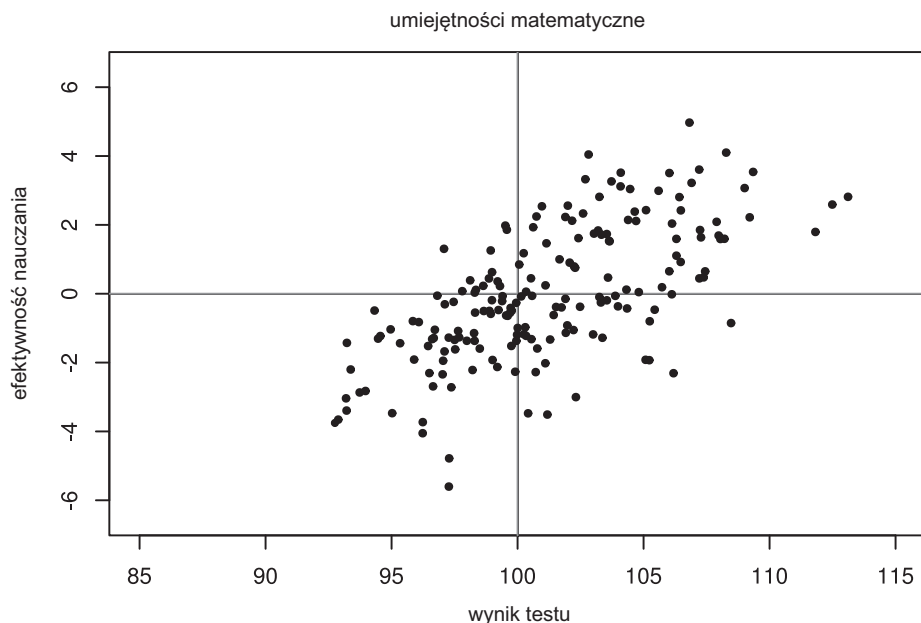


Tabela 3.8. Korelacje między wskaźnikami wyznaczonymi za pomocą różnych modeli efektywności nauczania.

Model efektywności nauczania	Model złożony		Model prosty	
	Umiejętność czytania	Świadomość językowa	Umiejętność czytania	Świadomość językowa
Model złożony – umiejętność czytania	1			
Model złożony – świadomość językowa	0,749	1		
Model złożony – umiejętności matematyczne	0,737	0,674		
Model prosty – umiejętność czytania			1	
Model prosty – świadomość językowa			0,766	1
Model prosty – umiejętności matematyczne			0,752	0,698

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

Ostatnia kwestia, którą należy poruszyć, dotyczy tego, czy każdy z trzech wskaźników efektywności (obliczony dla każdego testu osiągnięć) wnosi swoistą informację o pracy szkoły, czy można byłoby ograniczyć się do jednego wskaźnika. W Tabeli 3.8 zaprezentowano macierz korelacji dla kontekstowych wskaźników efektywności nauczania. Wszystkie korelacje są istotne statystycznie na poziomie istotności $p < 0,05$. Efektywność nauczania w trzech badanych zakresach jest dość silnie i pozytywnie ze sobą powiązana. Oczywiście nie jest to zaskoczeniem. Tym bardziej że w klasach I–III wszystkich tych przedmiotów naucza jeden nauczyciel. Nie mniej korelacje te nie są na tyle wysokie, by każdy ze wskaźników dawał dokładnie tę samą informację. Dlatego wydaje się uzasadnione posługiwanie się trzema niezależnymi wskaźnikami efektywności nauczania.

Podsumowanie

Głównym celem tego rozdziału było przedstawienie możliwych do zastosowania modeli oceny efektywności nauczania dla pierwszego etapu kształcenia w Polsce. Modele takie są niezwykle potrzebne, bowiem surowe wyniki z testów osiągnięć nie informują wystarczająco o efektywności nauczania w szkole. Zależą one również od takich cech uczniów, czy środowiska lokalnego, na które szkoła nie ma wpływu. Szkoły podstawowe różnią się pod względem poziomu osiągnięć szkolnych swoich uczniów już po trzeciej klasie. Jednak część tego zjawiska można wyjaśnić nierównomiernym rozłożeniem pomiędzy szkołami cech uczniów mających znaczenie dla efektów nauczania. Wyniki przeprowadzonych analiz potwierdziły istnienie znaczącego zróżnicowania międzyszkolnego charakterystyk uczniów wpływających na ich osiągnięcia szkolne. Szkoły w największym stopniu różnią się pod względem inteligencji uczniów oraz czynników statusowych. Dlatego nieprzetworzone wyniki z testów osiągnięć nie są adekwatną, a tym samym sprawiedliwą miarą pracy szkoły. Placówki pracujące w bardziej korzystnym środowisku mają większą szansę na uzyskanie lepszych rezultatów.

W rozdziale omówiono takie modele efektywności nauczania, które mogłyby być używane w polskim systemie oświaty. Dopóki nie powstanie realna możliwość pomiaru umiejętności uczniów na progu szkoły podstawowej na szeroką skalę, dopóty wyliczenie klasycznych modeli edukacyjnej wartości dodanej będzie niemożliwe. Ponieważ obecnie pomiar taki wydaje się co najmniej problematyczny, zaproponowano modele niewymagające pomiaru osiągnięć szkolnych na początku danego etapu kształcenia. Modele te nazywaliśmy kontekstowymi modelami efektywności nauczania.

Wykorzystując wnioski z analiz omówionych w rozdziale 2, zbudowano dla każdego obszaru osiągnięć wyczerpujący kontekstowy model efektywności nauczania, który kontrolował wszystkie ważne dla osiągnięć szkolnych zmienne niezależne od szkoły. W modelach tych podział na szkoły wyjaśnił 3,5% wariacji zmiennej zależnej dla testu umiejętności czytania, 5,4% dla testu świadomości językowej i 6,5% w przypadku testu z matematyki. Jeżeli założenia przyjętego modelu byłyby spełnione, oznaczałoby to, że gdyby szkoły pracowały z uczniami o takich samych cechach, w takim właśnie stopniu różniłyby się między sobą wynikami nauczania. Tę właśnie część zróżnicowania wyników nauczania można zatem przypisywać pracy szkoły. Wyniki modelu wyczerpującego stanowiły punkt odniesienia dla modeli prostszych, w których uwzględniono tylko te charakterystyki,

które można mierzyć i przetwarzać systemowo. Ze wszystkich zbudowanych modeli kontekstowych wyznaczono wskaźniki dla 180 szkół z próby badawczej.

Przeprowadzone analizy pokazały, że najlepsze właściwości miały wskaźniki z dwóch modeli: (5) i (6). Były one najbliższe (pod względem wartości) wskaźnikom wyliczonym z modeli wyczerpujących, a także były w najmniejszym stopniu związane z kontekstem pracy szkoły. Dodatkowo wyjaśniały najwięcej wariacji wyników testów. Ponieważ jednak różnice co do oszacowań efektów szkół między modelem (5) a (6) były praktycznie pomijalne (korelacje rzędu 0,998–0,999), wybrano model prostszy, czyli model (5), zawierający następujące zmienne kontekstowe: płeć, wiek w miesiącach, oznaczenie bycia starszym lub młodszym niż główna kohorta wiekowa, wykształcenie rodziców ucznia w latach nauki (na poziomie indywidualnym oraz jako średnia obliczona dla szkoły), aspiracje rodziców wobec wykształcenia ich dzieci oraz informacja o liczbie posiadanych w domu książek. Model ten nazwano złożonym modelem kontekstowym efektywności nauczania. Dodatkowo wybrano model (3), w którym uwzględnione zostały tylko zmienne względnie najprostsze do pozyskania w systemie oświaty (płeć i wiek ucznia, oznaczenie bycia starszym lub młodszym niż główna kohorta wiekowa oraz informacja o wykształceniu rodziców – uwzględnione na poziomie indywidualnym i jako średnia w szkole). Model ten miał nieznacznie gorsze właściwości niż model (5), był za to znacznie prostszy pod względem możliwości pozyskania danych potrzebnych do jego wyliczenia. Nazwano go prostym modelem kontekstowym efektywności nauczania.

Korelacje między wskaźnikami wyliczonymi ze złożonych modeli kontekstowych a wskaźnikami z modeli wyczerpujących były wysokie i wyniosły 0,95–0,97 (w zależności od testu). Dla poszczególnych szkół zaobserwowano pewne różnice w oszacowanych wartościach wskaźników wyliczonych z obu modeli, jednak były one niewielkie (relatywnie największe w przypadku matematyki) i w oszacowaniach przedziałowych nie prowadziły do odmiennych wniosków na temat poniżej lub powyżej przeciętnej efektywności nauczania. Dla modeli prostych zaobserwowane korelacje były nieznacznie niższe (0,95–0,96) i prowadziły do podobnych wniosków.

Dowodzono, że wskaźniki wyliczone z kontekstowych modeli efektywności nauczania (zarówno z modeli złożonych, jak i prostych) są znacznie słabiej powiązane z kontekstem pracy szkoły niż wyniki testów osiągnięć. Średnie wyniki testów osiągnięć szkolnych są skorelowane ze zagregowanymi miarami statusu na poziomie około 0,7. Dla kontekstowych wskaźników efektywności nauczania nie zaobserwowano natomiast istotnych korelacji. W przypadku średniej inteligencji uczniów w pierwszej klasie, dla wyników testów osiągnięć stwierdzono korelacje rzędu 0,56, natomiast dla kontekstowych wskaźników efektywności nauczania korelacje ponad dwukrotnie niższe. Jest to jeden z dowodów trafności wyliczonych wskaźników jako miary efektywności pracy szkoły. Ich wartość nie zależy bowiem od składu społecznego danej placówki, i w znacznie mniejszym stopniu niż wyniki testów zależy od średniej inteligencji uczniów.

W analizach podjęto też kwestię użyteczności wypracowanych wskaźników. Jednym z kryteriów było pytanie o to, na ile są one pomocne w wyłanianiu szkół ponad i poniżej przeciętnych, jeśli chodzi o efektywność nauczania. Przeprowadzone analizy pokazały, że dla 4,5 do 21% szkół (w zależności od testu i zastosowanego modelu) można byłoby powiedzieć, że nauczają z poniżej lub ponad przeciętną efektywnością. Oznacza to, że od ponad 400 do prawie 2000 placówek miałyby szansę dowiedzieć się, że podejmowane

3. Kontekstowe modele oceny efektywności nauczania po pierwszym etapie edukacyjnym

przez nie działania w ramach nauczania poszczególnych umiejętności są wyjątkowo skuteczne lub mało efektywne.

Kolejnym kryterium użyteczności było to, na ile informacja płynąca ze wskaźników efektywności jest wartościowym uzupełnieniem wyników testów, czyli czy dostarcza innej wiedzy ponad to, co można wyczytać z nieprzetworzonych wyników. Korelacje między wynikami testów osiągnięć a kontekstowymi wskaźnikami efektywności nauczania wyniosły około 0,64 dla modelu złożonego oraz około 0,65 dla modelu prostego. Przeprowadzona analiza pokazała, że informacje z obu wskaźników: miary efektywności i wyniku w teście osiągnięć wzajemnie się uzupełniają.

Ostatni przeanalizowany problem dotyczył tego, czy każdy z trzech wskaźników efektywności (dla każdego testu osiągnięć) wnosi swoistą informację o pracy szkoły, czy też można byłoby się ograniczyć do jednego wskaźnika. Przeprowadzone analizy pokazały, że mimo, iż wskaźniki te są ze sobą wysoko skorelowane (korelacje 0,67–0,76, w zależności od modelu i wskaźnika), to jednak nie dają tożsamej informacji, dlatego uzasadnione byłoby posługiwanie się wszystkimi trzema miarami dla uzyskania pełniejszego obrazu efektywności nauczania w szkole.

Ostatnia część podsumowania zostanie poświęcona krótkiemu omówieniu warunków, które musiałyby zostać spełnione, by zaproponowane wskaźniki mogły być obliczane i wykorzystywane w systemie edukacji. Po pierwsze, niezbędny byłby dobry pomiar osiągnięć szkolnych po trzeciej klasie szkoły podstawowej. Nie jest jednak konieczne, by zostały nim objęte wszystkie szkoły w kraju. Wystarczające byłoby badanie na reprezentatywnej próbie szkół, którego wyniki posłużyłyby do wyliczenia modeli i ustalenia punktu odniesienia do porównań dla innych szkół. Pozostałe placówki, zainteresowane takimi wskaźnikami, mogłyby dobrowolnie przeprowadzić u siebie takie pomiary i porównać ich rezultaty z wynikami badania reprezentatywnego. W podobnej formule jest realizowane badanie OBUT, co pokazuje, że potencjalnie mogłoby stanowić ono (lub inne realizowane w podobny sposób) podstawę do wyliczania takich wskaźników. Oczywiście, konieczna byłaby krytyczna analiza jakości testów, które byłyby wykorzystane do wyliczenia wskaźników, bowiem od ich trafności i rzetelności zależałaby jakość wyliczonych na ich podstawie miar.

Poza pomiarem osiągnięć, potrzebne byłyby dane o uczniach, które musiałyby być zbierane i przechowywane przez szkoły, a następnie łączone z wynikami testów osiągnięć. Dla wskaźników z prostych modeli efektywności nauczania należałoby zebrać informację o wieku, płci dziecka i wykształceniu jego rodziców. Informacje te mogłyby być zbierane w trzeciej klasie za pomocą krótkiej ankiety skierowanej do rodziców, w momencie, w którym szkoła zdecyduje się przystąpić do pomiaru osiągnięć. W przypadku wskaźników wyliczonych z modeli złożonych, sytuacja jest trochę bardziej skomplikowana. Poza wymienionymi już danymi oraz informacją o liczbie książek, którą także można pozyskać w podobny sposób, potrzeba dodatkowej informacji na temat aspiracji rodziców wobec własnych dzieci. W przypadku tego pytania, ważne jest, by było ono skierowane do rodziców na początku edukacji ich dzieci w klasie pierwszej. Aspiracje rodziców mogą się bowiem zmienić pod wpływem osiągnięć szkolnych dzieci oraz działań szkoły. Dlatego, jeśli chcemy w modelu uwzględnić czynnik niezależny od pracy szkoły, ważne jest, by pomiar aspiracji przeprowadzać już w pierwszej klasie. To jednak rodzi dodatkową komplikację w postaci konieczności zbierania danych kontekstowych dla badanego rocznika już na początku nauki oraz przechowywania ich i ochrony przez szkoły przez dwa kolejne lata. Oczywiście

nie jest to nierealne, a ankieta kierowana do rodziców na początku drogi edukacyjnej ich dzieci mogłaby stanowić dodatkową okazję dla szkół do zebrania dodatkowych danych o uczniu, pomagając w ten sposób dostosować ofertę edukacyjną do jego potrzeb.

I na koniec kwestia wykorzystywania zebranych informacji. Po przeprowadzeniu pomiaru osiągnięć szkolnych na koniec pierwszego etapu edukacyjnego zebrane przez placówkę dane mogłyby być wprowadzane do specjalnie przygotowanej aplikacji (np. podobnej do funkcjonującego już Kalkulatora EWD 100⁶), w której uprzednio zaimplementowano by modele efektywności nauczania, wyliczone na podstawie danych z badania na reprezentatywnej próbie. W ten sposób osoby zainteresowane mogłyby analizować efektywność nauczania w danej szkole czy w podgrupach uczniów w tej szkole, bez konieczności wysyłania danych do zarządzającej tym procesem jednostki centralnej.

⁶ <http://ewd.edu.pl/kewd100/>

Rozdział 4

Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

Studium metody oceny efektywności pracy szkół podstawowych na pierwszym etapie edukacyjnym, oprócz przedstawienia propozycji modelu statystycznego szacowania tej efektywności, musi uwzględniać kwestię jej trafności. W tym rozdziale przedstawiamy analizy służące weryfikacji trafności kontekstowego modelu oceny efektywności nauczania.

Nie chcemy tu rozstrzygać obecnych w literaturze sporów na temat tego, co kryje się pod poszczególnymi typami trafności, ani tym bardziej, z jakim typem trafności mamy do czynienia w sytuacji badania trafności kontekstowych wskaźników efektywności nauczania. Typy te wydają się dość nieostre i nierozłączne. Istnieją jednak dobre opracowania, które w przystępny sposób starają się porządkować te kwestie (Carmines i Zeller, 1979). Użyteczną perspektywę wyznacza rozumienie trafności w kategoriach interpretacji wyników badania (Cronbach, 1971). Także standardy dla testów stosowanych w psychologii i edukacji definiują trafność jako stopień, w jakim dowody empiryczne i teoria wspierają interpretację wyników badania (AERA, APA i NCME, 1999). W podobny sposób charakteryzują również podejście argumentacyjne (*an argument-based approach to validity*; Kane, 1992). Wyznacza ono strategie budowania argumentacji wspierającej interpretację wyników. Polega na poszukiwaniu dowodów, które będą popierały dokonane na podstawie wyników badania interpretacje, sformułowane wnioski oraz podjęte decyzje i jednocześnie będą odpierały kontrargumenty odwołujące się do konkurencyjnych interpretacji wyników badania.

Przywołane już *Standardy...* (AERA, APA i NCME, 1999), mimo że opracowane z myślą o testach psychologicznych i edukacyjnych, wskazują na źródła, w których można szukać danych dotyczących trafności kontekstowych wskaźników efektywności nauczania dla szkół podstawowych. Jednym z nich jest analiza związków badanego konstruktów z innymi zmiennymi zidentyfikowanymi jako ważne korelaty efektywności nauczania.

Dane zebrane podczas badania podłużnego w szkołach podstawowych pozwalają na przeprowadzenie takich analiz. W tym celu wybrano kilka grup zmiennych, potencjalnie użytecznych w analizach trafności. Kryterium wyboru stanowiła ich obecność w dyskursie naukowym dotyczącym kryteriów opisu efektywności pracy szkół w kontekście szkolnych osiągnięć uczniów oraz powiązanie z realizacją pozadydaktycznych zadań stawianych szkołom¹. Grupy zmiennych uwzględnionych w analizach to:

¹ Zadania te stawia ustawa o systemie oświaty (Dz.U. z 1991 r., Nr 95 poz. 425 z późn. zm).

- w obszarze funkcjonowania szkół: wymiary stylu zarządzania szkołą – partycypacyjność, kultura organizacyjna, współpraca z rodzicami;
- w obszarze dotyczącym nauczycieli: charakterystyki nauczycieli oraz ich pracy z uczniami;
- w obszarze dotyczącym uczniów: integracja szkolna.

Korelaty efektywności nauczania – style zarządzania szkołą

Jednym z kryteriów, który rozważano w odniesieniu do trafności kontekstowych modeli efektywności kształcenia dla szkół podstawowych, były style zarządzania placówką edukacyjną. Ponieważ zakładamy, że kontekstowe wskaźniki efektywności nauczania dla pierwszego etapu kształcenia odzwierciedlają wkład szkoły w proces uczenia się, to zależności między stylem zarządzania a miarą efektywności nauczania będą argumentem na rzecz jej trafności. Szkoły o stylach zarządzania, które zgodnie z teorią są korzystne dla efektywnego nauczania, powinny mieć wysokie wyniki kontekstowych miar efektywności nauczania. Należy zauważyć, że analizy dotyczące tego zjawiska są szczególnie istotne, ponieważ weryfikują powiązania modeli kontekstowych efektywności nauczania z cechami szkoły, które są od niej zależne.

Definicje stylów zarządzania oraz ich konceptualizacja mogą bardzo się różnić w zależności od przyjętej orientacji teoretycznej (Bush, 2003). W tym rozdziale skupiono się na trzech zagadnieniach: postrzeganiu przez nauczycieli partycypacji w decyzjach podejmowanych w szkole (Miller i Rowan, 2006), kulturze organizacyjnej szkoły w opiniach nauczycieli (Maslowski, 2001; Hartnell, Ou i Kinicki, 2011) oraz na dokonanej przez rodziców ocenie jakości swoich relacji ze szkołą (Comer i Haynes, 1991; Driessen, Smit i Slegers, 2005; Fan i Chen, 2001).

W literaturze wskazuje się, że wpływ stylów zarządzania – w tym nauczycielskiej partycypacji i typów kultury organizacyjnej – na osiągnięcia uczniów ma charakter pośredni (Miller i Rowan, 2006). Wpływają one na praktyki nauczania, które już bezpośrednio przekładają się na osiągnięcia uczniów (Raudenbush, Rowan i Kang, 1991). W wypadku zaangażowania rodziców w proces edukacyjny swoich dzieci badacze tematu wskazują na pozytywny związek między ich zaangażowaniem w życie szkoły a efektywnością procesu nauczania (Comer i Haynes, 1991). Opis konstruktów teoretycznych leżących u podstawy przyjętych koncepcji badawczych opisano w dalszej części rozdziału.

Partycypacja w zarządzaniu w ocenie nauczycieli

Partycypacja jest pojęciem znaczeniowo zróżnicowanym. Można ją rozumieć jako stopień, w jakim na funkcjonowanie szkoły wpływają nauczyciele (Conley, 1991). Oczywiście należy zauważyć, że poziom partycypacji zależy od uwarunkowań prawnych. W silnie zhierarchizowanym systemie szkolnym, którego przykładem może być system polski, włączanie nauczycieli do podejmowania decyzji dotyczących funkcjonowania szkoły jest

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

trudne, jednak możliwe (Tołwińska, 2011). Z drugiej strony, niektórzy badacze podkreślają, że istotne może być także postrzeganie przez nauczycieli stopnia ich wpływu na proces decyzyjny w szkole, a nie faktyczny wpływ (Miller i Rowan, 2006).

W badaniach dotyczących partycypacji nauczycieli można wyróżnić dwa główne podejścia. Pierwsze z nich skupia się na efektach partycypacji odczuwanych przez nauczycieli. Wskazuje się tutaj na jej związek z satysfakcją z pracy, stresem, poczuciem alienacji, współpracą z innymi nauczycielami, rodzicami i dyrektorem (Conley, 1991). Drugie podejście skupia się na badaniu wpływu partycypacji na funkcjonowanie szkoły, a w szczególności na jej znaczenie dla osiągnięć szkolnych uczniów. Wyniki badań prowadzonych w tym nurcie nie są jednoznaczne – część z nich wskazuje na istnienie pozytywnego związku partycypacji nauczycieli w decyzjach podejmowanych w szkole na osiągnięcia uczniów (Rowan, Chiang i Miller, 1997), inne natomiast nie dają w tym względzie rozstrzygających wyników (Miller i Rowan, 2006; Raudenbush i in., 1991). Sugeruje się, że wyniki badań wskazują na pośredni wpływ partycypacji nauczycieli na osiągnięcia szkolne uczniów. Wskazuje się przede wszystkim na większe możliwości podejmowania przez nauczycieli decyzji dotyczących przebiegu lekcji w szkołach nastawionych na partycypację. Dzięki temu nauczyciel może bardziej elastycznie kształtować proces nauczania (Marks i Louis, 1997; Smylie, Lazarus i Brownlee-Conyers, 1996).

Rozważając zagadnienie partycypacji nauczycieli w kontekście decyzji podejmowanych w szkole, należy podkreślić rolę dodatkowych czynników, które bezpośrednio wpływają na możliwość zaistnienia pozytywnego wpływu tej partycypacji na funkcjonowanie szkoły (Heller, Pusic, Strauss i Wilpert, 1998). Są to: zaufanie do współpracowników, zaufanie do przełożonego, możliwość uzyskania informacji zwrotnej o swojej pracy, możliwość samorozwoju, identyfikacja z instytucją. Związki między wymienionymi czynnikami a samą partycypacją są opisywane następująco: warunkiem koniecznym partycypacji jest zaufanie zarówno do współpracowników (czyli innych nauczycieli), jak i do przełożonego (czyli dyrektora i jego zastępców). W sytuacji braku zaufania rośnie ryzyko pojawienia się konfliktów, które wynikają z samego dopuszczenia pracowników do procesu podejmowania decyzji. Brak zaufania może sprawić, że partycypacja stanie się bezproduktywna (Heller i in., 1998). W partycypacyjnym stylu zarządzania istotna staje się także możliwość uzyskania informacji zwrotnej o jakości swojej pracy. Bez niej osoby zaangażowane w proces decyzyjny mogą nie mieć wiedzy o skuteczności swoich działań (Heller i in., 1998). Możliwość samorozwoju oraz identyfikację z instytucją należy traktować jako elementy stymulujące pojawienie się w szkole partycypacji (Glew, O'Leary-Kelly, Griffin i Van Fleet, 1995).

Kultura organizacyjna szkoły

W badaniu podłużnym prowadzonym w szkołach podstawowych przyjęto, że kulturę organizacyjną rozumie się jako pewien nieformalny zestaw przekonań, wartości i podstawowych założeń dotyczących funkcjonowania placówki (Masłowski, 2001; Schein, 2010). Należy jednak zauważyć, że badania dotyczące kultury organizacyjnej opierają się na nieco innych definicjach tego zjawiska – zwykle przyjmują, że jest to konglomerat różnego rodzaju aspektów zarządzania szkołą (Tsereteli, Martskvishvili i Aptarashvili, 2011; Zhu, Devos i Li, 2011).

W naszym badaniu użyliśmy narzędzia do badania kultury organizacyjnej zbudowanego na podstawie koncepcji Kima Camerona i Roberta Quinna (2006). W jej obrębie wyróżnia się dwa główne wymiary kultury organizacji: orientację na otoczenie instytucji versus orientację na wnętrze instytucji oraz elastyczność versus stabilność. Rozpatrywanie tych dwóch wymiarów wspólnie pozwala wyróżnić cztery typy kultury organizacji: klanu, adhokracji, rynku i hierarchii.

Kultura klanu charakteryzuje się naciskiem na stabilność wewnętrzną oraz elastyczność. Przyjazna pracownikom, premiuje współpracę. Więzy między pracownikami mogą nabrać charakteru osobistego, prawie rodzinnego. Lider organizacji często jest mentorem. Instytucja nastawiona jest na długoterminowy rozwój, głównie dzięki rozwojowi osobistemu pracowników (nauczycieli). Widoczny jest nacisk na spójność grupową. Wyznacznikami sukcesu organizacji są: lojalność, zaangażowanie, praca zespołowa oraz konsensus. Kultura adhokracji jest zorientowana na otoczenie oraz elastyczna w działaniu. Premiowane wśród pracowników cechy to kreatywność, dynamika, przedsiębiorczość. Dyrektor szkoły pełni rolę innowatora wytyczającego długoterminowe cele. Organizacja jest nastawiona na ciągły rozwój, także przez podejmowanie wyzwań i pozyskiwanie nowych możliwości rozwoju. Pracownicy są zachęcani do wykazywania się inicjatywą, samodzielnego działania. Unikalna oferta, niestandardowe sposoby postępowania stają się kryterium sukcesu w tego rodzaju organizacji.

Kultura rynku jest nastawiona na stabilność i kontrolę, przy jednoczesnym nakierowaniu na otoczenie. W instytucji tego rodzaju istotne są rezultaty realizowanych działań. Nauczyciele są ambitni i skupieni na osiąganiu celów. Lider jest nastawiony na rezultaty, konkutowanie z innymi. Spójność organizacji jest możliwa dzięki dążeniu do osiągnięcia sukcesu, bycia lepszą od innych, podobnych instytucji. Ważne dla szkoły są: reputacja, rezultaty i sukces w osiąganiu celów. Wskaźnikiem realizacji celów jest pozycja względem innych.

Kultura hierarchii jest nastawiona na stabilność, kontrolę oraz integrację wewnętrzną. Obowiązują w niej ścisłe procedury oraz sztywna hierarchia. Przywódca to sprawny organizator i koordynator działań pracowników. Celem instytucji jest sprawne funkcjonowanie, które ma być zapewnione przez formalne przepisy i procedury postępowania, zapewniające instytucji spójność. Dla pracowników ważne jest bezpieczeństwo zatrudnienia, stałość i przewidywalność działań. Sukces w tej organizacji to przewidywalność rezultatów, dotrzymywanie formalnych zobowiązań.

Analizy oparte na danych zebranych w gimnazjach wskazują na związek między kulturą organizacyjną a osiągnięciami uczniów i efektywnością pracy szkoły (Dolata i in., 2013). Zaobserwowano dodatnie efekty dla rynkowego i klanowego typu kultury. Na tej podstawie można zatem przewidywać, że kultura organizacyjna będzie także w związku z osiągnięciami uczniów na pierwszym etapie kształcenia. Nie można jednak założyć, że wzorzec zależności zaobserwowany w gimnazjach powtórzy się w szkołach podstawowych. Ze względu na przypuszczalne niewielkie znaczenie współpracy między nauczycielami nauczania zintegrowanego dla osiągnięć uczniów (każda klasa ma jedynie nauczyciela-wychowawcę, który prowadzi większość zajęć z uczniami), można spodziewać się braku efektu dla kultury klanu. Przypuszczać można także, że znaczenia nie będzie miała kultura rynku, wymaga ona bowiem dokonywania porównań z innymi szkołami i konkutowania z nimi. Tymczasem egzamin zewnętrzny, który by na to

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

pozwolił, odbywa się dopiero w klasie szóstej, czyli po drugim etapie edukacyjnym, jest więc zbyt odległy.

Spodziewać się można negatywnych zależności z kulturą hierarchii. Ta bowiem skutkować może zbyt dużym usztywnieniem pracy nauczycieli i w efekcie – niedostosowaniem działań opiekuńczo-wychowawczych i edukacyjnych do zróżnicowanych potrzeb dzieci. Z drugiej jednak strony elastyczność nauczania początkowego i związana z tym duża swoboda nauczyciela w organizacji procesu nauczania, nawet w przypadku obowiązywania w szkole sztywnych zasad i procedur, może chronić przed możliwym negatywnym znaczeniem kultury klanu dla osiągnięć uczniów.

Trudno formułować przypuszczenia dotyczące wpływu kultury adhocracji na osiągnięcia uczniów w pierwszym etapie edukacyjnym. Charakterystyczne dla tego typu kultury jest nastawienie na ciągły rozwój, elastyczność, doskonalenie zawodowe i poszukiwanie nowych metod nauczania. Te cechy powinny sprzyjać efektywniejszej pracy nauczycieli, przekładającej się na umiejętności ich podopiecznych. Jednak w analizach przeprowadzonych dla gimnazjów nie wykazano takiego związku (Dolata i in., 2013). Możliwe, że inne cechy charakterystyczne tej kultury organizacyjnej – głównie jej tymczasowość oraz chęć podejmowania ryzyka w działaniu – sprawiają, że nie przekłada się ona na wyższe osiągnięcia. W przypadku pierwszego etapu kształcenia istotne jest, by miało ono charakter stabilny, uporządkowany, zaplanowany i w ten sposób przewidywalny dla ucznia, sprzyjający poczuciu bezpieczeństwa. Z tego względu można spodziewać się negatywnych związków między tym typem kultury organizacyjnej a miarami efektywności nauczania.

Współpraca rodziców ze szkołą

Ostatnim zagadnieniem rozpatrywanym w związku ze stylami zarządzania jest współpraca rodziców ze szkołą. W literaturze tematu zjawisko to jest definiowane na wiele sposobów, jednak panuje zgoda, że na najbardziej ogólnym poziomie współpracę rodziców ze szkołą można traktować jako składową zaangażowania rodziców w edukację dzieci (Fantuzzo, Tighe i Child, 2000). Obecnie powszechnie w analizie tego zjawiska używa się typologii zaproponowanej przez Joyce Epstein i Susan Dauber (1991). Jej najbardziej aktualna wersja wyróżnia następujące elementy zaangażowania:

- wychowanie (zapewnienie opieki, bezpieczeństwa, warunki domowe wspierające naukę, zapewnienie szkole informacji pozwalającej na lepsze poznanie ucznia, umiejętności rodziców w nawiązywaniu interakcji z dzieckiem);
- komunikowanie na linii szkoła–dom oraz dom–szkoła;
- wolontariat (angażowanie się w organizację i współuczestnictwo w wycieczkach klasowych lub inną formę pomocy szkole);
- wspomaganie nauki w domu (pomoc w pracach domowych, w dokonywaniu wyborów edukacyjnych);
- współuczestniczenie w podejmowaniu decyzji w szkole (uczestnictwo w radach rodziców);
- współpraca ze wspólnotą szkolną (działania wykraczające poza zwyczajne ramy współpracy ze szkołą).

Typologię tę często redukuje się do dwóch podstawowych wymiarów – zaangażowania w życie szkolne oraz zaangażowania w pomoc edukacyjną dziecku w domu (Mattingly, Prislin, McKenzie, Rodriguez i Kayzar, 2002).

Niektóre badania opisujące to zjawisko dotyczą wpływu zaangażowania rodziców w proces edukacyjny na osiągnięcia szkolne uczniów (Fan i Chen, 2001; González i Jackson, 2013; Jeynes, 2007). Postulowana zależność jest stosunkowo prosta – większe zaangażowanie rodziców w proces edukacyjny swoich dzieci prowadzi do lepszych osiągnięć edukacyjnych. Formy tego zaangażowania mogą być bardzo zróżnicowane, począwszy od pomocy w pracach domowych, komunikacji ze szkołą czy też współuczestnictwa w decyzjach podejmowanych w szkole. Metaanalizy dotyczące tego tematu wskazują, że istnieje pozytywny, istotny związek między zaangażowaniem rodziców a osiągnięciami uczniów, szczególnie jeśli bierzemy pod uwagę wyniki uczniów na egzaminach (Fan i Chen, 2001; Jeynes, 2007).

Korelaty efektywności nauczania – czynniki związane z nauczycielem

Zgodnie ze stanem wiedzy na temat uwarunkowań wyników nauczania, należy spodziewać się istotnych efektów nauczyciela. Jeśli na podstawie teorii i dowodów empirycznych wiemy, że konkretne charakterystyki nauczycieli i ich pracy istotnie pomagają wyjaśniać zróżnicowanie wyników uczniów, to chcemy, aby w ramach testowanych modeli efektywności nauczania ich korelacje z wynikami uczniów były istotne i przebiegały w zgodnym z oczekiwaniami kierunku. Dodając do modelu efektywności nauczania kolejne czynniki nauczycielskie, należy się spodziewać poprawy dopasowania modelu do danych, czyli bardziej dokładnego przewidywania wyników uczniów na wyjściu.

Niestety gromadzone w rozmaitych badaniach dowody empiryczne nie dostarczają jednoznacznych wskazań, za pomocą jakich zmiennych skutecznie można wyjaśnić zróżnicowanie efektywności pracy nauczycieli. Największe autorytety w zakresie badań edukacyjnych nad czynnikami nauczycielskimi w osobach Barbary Nye, Larry’ego Hedgesa, Erica Hanushka, Richarda Laine’a, Roberta Greenwalda, Stevena Rivkina, Johny’ego Kaina są zgodne, że efektywny nauczyciel ma większy wpływ na procesy uczenia niż jakiegokolwiek inne czynniki kontrolowane przez system szkolny. Jednocześnie przyznają, że mierzalne charakterystyki nauczyciela wyjaśniają niewiele wariacji efektywności nauczycieli. W literaturze przedmiotu można jednak znaleźć wskazówki, jakie charakterystyki nauczycieli mogą być powiązane z wynikami szkolnymi uczniów. Rezultaty badań pokazują dodatni, jednak nie zawsze istotny, związek stażu zawodowego z wynikami uczniów (Klitgaard i Hall, 1975; Murnane i Phillips, 1981; Rosenholtz, 1986). Barbara Nye, Spyros Konstantopoulos i Larry Hedges (2004) zaobserwowali istnienie liniowego, pozytywnego związku stażu nauczania z wynikami uczniów z matematyki oraz kompetencji językowych. Jednakże w badaniach nad trafnością gimnazjalnych wskaźników EWD (Dolata i in., 2013; Koniewski, 2013) nie zaobserwowano istotnego związku stażu pracy z wynikami uczniów.

Badania poświęcone czynnikowi wykształcenia nauczycieli w większości dotyczą zależności między wykształceniem przedmiotowym (w zakresie nauczanego przedmiotu)

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

nauczycieli a osiągnięciami szkolnymi ich uczniów. Generalnie obserwuje się dodatni, ale nie zawsze istotny, związek wiedzy przedmiotowej z wynikami uczniów (Ashton i Crocker, 1987; Byrne, 1983). Kompetencje językowe nauczycieli wydają się słabiej korelować z wynikami uczniów z testów językowych (Bowles i Levin, 1968; Ehrenberg i Brewer, 1995; Ferguson i Ladd, 1996; Hanushek, 1971; 1992) niż kompetencje matematyczne nauczycieli z wynikami uczniów z testów matematycznych (Ferguson i Ladd, 1996; Rowan, Chiang i Miller, 1997). Z punktu widzenia nauczania w klasach I–III szkoły podstawowej większe znaczenie niż przygotowanie przedmiotowe ma formalne wykształcenie pedagogiczne. Carolyn Evertson, Willis Hawley i Marylin Zlotnik (1985), a także Colin Byrne (1983) wykazali pozytywny wpływ formalnego wykształcenia pedagogicznego nauczycieli na umiejętności szkolne uczniów. Spośród 13 badań, 11 raportowało wyższą efektywność nauczycieli z formalnym wykształceniem pedagogicznym w stosunku do nauczycieli bez takiego przygotowania. W badaniu trafności gimnazjalnych wskaźników EWD zaobserwowano istotny związek poziomu wykształcenia nauczycieli (mierzonego posiadaniem dyplomu studiów wyższych) z wynikami uczniów tylko z części matematycznej egzaminu (Dolata i in., 2013). W tych samym badaniu stopień awansu zawodowego nauczycieli nie pozwalał na przewidywanie wyników egzaminu gimnazjalnego.

W przypadku badań prowadzonych w Polsce zaobserwowano, że stosowanie przez nauczyciela aktywizujących metod nauczania istotnie wyjaśnia różnicowanie efektywności nauczania matematyki, choć efekt nie jest zbyt silny (każde jedno odchylenie standardowe na skali metod aktywizujących przekładało się na wyższe średnio o 0,1 odchylenia standardowego wyniki z części matematycznej egzaminu gimnazjalnego). Podobną zależność zaobserwowano dla umiejętności utrzymania przez nauczyciela dyscypliny² (Koniewski, 2013).

Do badania poczucia skuteczności wykorzystano adaptację skali zaproponowanej przez Megan Tschannen-Moran i Anitę Woolfolk Hoy (2001). Poczucie własnej skuteczności (*self-efficacy*) jest przekonaniem nauczycieli o posiadaniu wiedzy i umiejętności wystarczających do osiągnięcia planowanych efektów w pracy z uczniami, nawet tymi o niskiej motywacji do nauki i niezdyscyplinowanymi (Armor i in., 1976; Bussey i Bandura, 1999). Przekonanie nauczycieli o własnej skuteczności jest pozytywnie związane z wynikami uczniów (Armor i in., 1976), ich motywacją (Midgley, Feldlaufer i Eccles, 1989) oraz poczuciem skuteczności samych uczniów (Anderson, Greene i Loewen, 1988).

Korelaty efektywności nauczania – integracja uczniów ze szkołą

W publicznym dyskursie szkołom skoncentrowanym na osiągnięciach niejednokrotnie zarzuca się, że promują wśród uczniów niezdrową rywalizację i atmosferę wrogości. Zwolennicy tej tezy wskazują, że za wysokie wyniki egzaminacyjne, czy szerzej – wyniki

² Przy kontroli wyników sprawdzianu, płci ucznia, orzeczenia o trwałej dysleksji, inteligencji mierzonej Testem Matryc Ravena, wykształcenia rodziców, zamożności gospodarstwa domowego, liczby osób w gospodarstwie domowym, średniego wyniku sprawdzianu w oddziale, lokalizacji gimnazjum, średniej dynamiki bezrobocia (2006–2010) w gminie, w której znajdowało się gimnazjum.

nauczania – wysoką cenę płać uczniowie. Funkcjonują bowiem w ciągłym stresie, zaburzeniu ulega ich rozwój społeczny, nie mają szans na rozwijanie ważnych umiejętności, np. pracy w zespole (Katulska, 2010; Radwańska, 2008).

Jeżeli opinia ta trafnie opisuje szkolną rzeczywistość, to wysoki poziom osiągnięć uczniów powinien współwystępować z niechęcią wobec rówieśników i wobec samej szkoły, wynikającą z presji osiągnięć i podsycecia rywalizacji między uczniami. Zależność ta powinna ujawnić się także w kontekstowych modelach efektywności nauczania, podważając ich trafność. Jeżeli bowiem wskaźniki obliczane na ich podstawie mają dostarczać informacji o tym, jak dobrze szkoła pracuje z uczniami, tj. realizuje stawiane jej cele ogólne, również pozadydaktyczne (np. zapewnia warunki do rozwoju uczniów w różnych sferach³), to zaobserwowane negatywne zależności będą sygnałem, że interpretowanie ich w takich kategoriach jest nieuprawnione. Odzwierciedlając bowiem realizację celów dydaktycznych, wskaźniki będą ignorować realizację przez szkoły celów pozadydaktycznych.

Stąd kolejnym konstruktem uwzględnionym w badaniu trafności modeli kontekstowych jest integracja uczniów ze szkołą, w rozumieniu zespołu Ursa Haeblerlina (Haeblerlin, Bless, Moser i Klaghofer, 1991). Ujmuje ona funkcjonowanie uczniów w szkołach na trzech wymiarach:

- integracji emocjonalnej, wyrażającej pozytywne nastawienie do szkoły;
- integracji społecznej, związanej z satysfakcją z kontaktów z klasowymi kolegami i
- integracji motywacyjnej, która odnosi się do obrazu siebie jako ucznia i znajduje wyraz w samoocenie szkolnych kompetencji.

Tak postrzegana integracja szkolna wydaje się ważną cechą uczniów, na którą wpływ wywiera placówka. Na etapie edukacji wczesnoszkolnej szczególną rolę odgrywają nauczyciele-wychowawcy. Pracują bowiem z tylko jednym oddziałem, jednocześnie każdy oddział ma zwykle tylko jednego nauczyciela (z wyjątkiem szczególnych przedmiotów, np. religii).

Metoda

Narzędzia i wskaźniki

Do badania partycypacji w zarządzaniu wykorzystano zmodyfikowaną wersję Indeksu Partycypacji przygotowanego i udostępnionego przez Interdyscyplinarne Centrum Badań i Rozwoju Organizacji przy Zakładzie Psychologii Społecznej Instytutu Psychologii Uniwersytetu Jagiellońskiego. Narzędzie składało się z 43 stwierdzeń dotyczących różnych aspektów partycypacji nauczycieli w zarządzaniu szkołą, do których respondent ustosunkowywał się na 5-stopniowej skali. W celu przygotowania wskaźnika wyspecyfikowano model konfirmacyjnej analizy czynnikowej dla zmiennych porządkowych za pomocą oprogramowania MPlus 7 (Muthén i Muthén, 1998–2012). W toku analiz wyłączono 7 pytań ze względu na ich niejednoznaczność przynależność czynnikową. Model okazał się dobrze dopasowany do danych (RMSEA = 0,039;

³ Obowiązek ten nakłada na szkoły ustawa o systemie oświaty (Dz.U. z 1991 r., Nr 95 poz. 425 z późn. zm).

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

CFI = 0,976). Na jego podstawie wygenerowano wartości czynnikowe dla poszczególnych obserwacji, które następnie wyrażono w jednostkach odchylenia standardowego. Szczegóły znajdują się w aneksie. W analizach wykorzystano wyniki zagregowane do średniej szkolnej.

Do badania kultury organizacyjnej wykorzystano adaptację⁴ kwestionariusza *Organizational Culture Assessment Instrument* (OCAI; Hartnell i in., 2011). Test składał się z 24 pozycji zebranych w 6 bloków dotyczących różnych właściwości kultury organizacyjnej szkoły spostrzeganych przez respondenta obecnie, oraz stanu, jaki będzie preferował za 3 lata. W każdym bloku respondent rozdzielał 100 punktów między 4 stwierdzenia, z których każde charakteryzowało jeden z typów kultury organizacyjnej. W analizach wykorzystano wyniki tylko dla stanu obserwowanego przez respondentów w chwili wypełniania narzędzia. W celu utworzenia wskaźników przydzielone punkty zostały zsumowane i podzielone przez liczbę pytań wskaźnikujących dany typ kultury. W rezultacie dla każdego nauczyciela otrzymano 4 wskaźniki, każdy z nich przyjął wartości od 0 do 100. Wyższy wynik oznacza, że nauczyciel w większym stopniu spostrzega placówkę, w której pracuje, jako charakteryzującą się danym typem kultury organizacyjnej. W analizach wykorzystano wyniki zagregowane do średniej szkolnej.

Do badania współpracy rodziców ze szkołą wykorzystano kwestionariusz składający się z 15 stwierdzeń dotyczących tego obszaru. Do każdej pozycji respondent ustosunkowywał się na 5-stopniowej skali. W celu opracowania wskaźników przeprowadzono eksploracyjną analizę czynnikową (EFA) dla zmiennych porządkowych za pomocą oprogramowania Mplus 6.1⁵. W jej toku wyodrębniono dwa czynniki – jeden dotyczący jakości współpracy z dyrektorem, drugi – jakości relacji z nauczycielami (usunięto jedną pozycję ze względu na jej niejasną przynależność czynnikową). Następnie w eksploracyjnym modelu równań strukturalnych (ESEM) wygenerowano oszacowania czynnikowe dla poszczególnych obserwacji. Model ESEM okazał się w akceptowalnym stopniu dopasowany do danych (RMSEA = 0,092; CFI = 0,957). Szczegóły analiz znajdują się w aneksie. Uzyskane oszacowania czynnikowe wyrażono w jednostkach odchylenia standardowego, a następnie zagregowano do średniej szkolnej oraz odwrócono ich kierunek tak, by wzrost pozycji na skalach oznaczał zwiększenie natężenia czynnika. Dane dotyczące charakterystyk nauczycieli zebrano za pomocą kwestionariusza wypełnianego przez nauczycieli. Utworzono następujące wskaźniki: staż pracy (deklarowana liczba lat pracy w zawodzie nauczyciela); wykształcenie (deklarowane posiadanie dyplomu: licencjata, magistra); stopień awansu zawodowego (deklarowany stopień awansu: dyplomowany, mianowany, kontraktowy, stażysta).

Dane dotyczące charakterystyk pracy nauczyciela z uczniami zebrano również za pomocą kwestionariuszy kierowanych do nauczycieli, a dotyczących ich pracy na lekcji. Każda pozycja miała postać stwierdzenia, do której nauczyciel ustosunkowywał się na 4-stopniowej skali. Analiza struktury odpowiedzi pozwoliła stworzyć skalę „Aktywny styl nauczania”. Jej ostateczna wersja składała się z 14 pozycji zgrupowanych w 5 podskal: „Indywidualizacja nauczania”, „Multimedia/gry”, „Internet”, „Podsumowania przerebionego materiału”, „Stymulowanie pracy w grupach”. W celu utworzenia łącznego

⁴ Adaptację przygotowało i udostępniło Interdyscyplinarne Centrum Badań i Rozwoju Organizacji przy Zakładzie Psychologii Społecznej Instytutu Psychologii UJ.

⁵ Różnice wersji zastosowanego oprogramowania wynikają z różnic w jego dostępności.

wskaźnika stosowania metod aktywizujących wyspecyfikowano model konfirmacyjnej analizy czynnikowej dla zmiennych porządkowych. Model okazał się w akceptowalnym stopniu dopasowany do danych (RMSEA = 0,067; CFI = 0,921). Na jego podstawie wygenerowano wartości czynnikowe dla poszczególnych obserwacji, które następnie poddano standaryzacji.

Do pomiaru stopnia, w jakim nauczyciel utrzymuje dyscyplinę w klasie, stworzono na podstawie pozycji kwestionariusza skalę „Autorytet nauczyciela/utrzymanie dyscypliny”. Jej ostateczna wersja składała się z 3 pozycji. W celu utworzenia wskaźnika wyspecyfikowano model konfirmacyjnej analizy czynnikowej dla zmiennych dyskretnych. Okazał się on dobrze dopasowany do danych (RMSEA = 0,031; CFI = 0,995). Na jego podstawie wygenerowano wartości czynnikowe dla poszczególnych obserwacji, które następnie poddano standaryzacji.

Do pomiaru stopnia, w jakim nauczyciel wspiera uczniów w nauce podczas lekcji, skonstruowano na podstawie deklaracji ankietowych skalę „Pomocny nauczyciel”. Jej ostateczna wersja składała się z 5 pozycji. W celu utworzenia wskaźnika wyspecyfikowano model konfirmacyjnej analizy czynnikowej dla zmiennych dyskretnych. Okazał się on dobrze dopasowany do danych (RMSEA = 0,033; CFI = 0,972). Na jego podstawie wygenerowano wartości czynnikowe dla poszczególnych obserwacji, które następnie poddano standaryzacji.

Pomiar poczucia skuteczności w zawodzie nauczyciela został przeprowadzony za pomocą wstępnej adaptacji narzędzia *Teachers' Sense of Efficacy Scale* (TSES, Tschannen-Moran i Woolfolk Hoy, 2001). Składało się ono z 24 pytań, na które nauczyciele odpowiadali, zaznaczając jedną z 9 odpowiedzi w kafeterii. W celu utworzenia wskaźnika przeprowadzono konfirmacyjną analizę czynnikową dla zmiennych porządkowych. Ostateczna wersja modelu osiągnęła akceptowalne dopasowanie do danych (RMSEA = 0,080; CFI = 0,903) i składała się z 2 czynników: „Nauczanie” oraz „Zarządzanie”. Na podstawie modelu wygenerowano wartości czynnikowe dla poszczególnych obserwacji, które następnie poddano standaryzacji.

Pomiaru stopnia integracji szkolnej dokonano za pomocą Kwestionariusza Integracji Uczniowskiej (KIU) w adaptacji Grzegorza Szumskiego (np. Szumski i Karwowski, 2012). Narzędzie składało się z 45 pozycji podzielonych na 3 podskale, do których dziecko ustosunkowywało się, wybierając jedną z 4 możliwych odpowiedzi. W celu weryfikacji właściwości narzędzia przeprowadzono konfirmacyjną analizę czynnikową dla zmiennych porządkowych. Model osiągnął akceptowalne dopasowanie do danych (RMSEA = 0,05; CFI = 0,918). Na jego podstawie wygenerowano wartości czynnikowe dla poszczególnych obserwacji, które następnie wystandaryzowano. W analizach wykorzystano dane zagregowane do średniej oddziałowej w każdej podskali, przyjmując je za wskaźniki procesów zachodzących między uczniami oraz ich stosunku do szkoły i siebie samych w roli ucznia. Założono również, że te relacje kształtowane są przez nauczyciela w toku nauczania. Zwiększenie wartości wskaźnika oznacza wyższe nasilenie danego aspektu integracji szkolnej. Korelacje między podskalami integracji emocjonalnej i motywacyjnej wyniosły 0,414, emocjonalnej i społecznej: 0,533, natomiast społecznej i motywacyjnej: 0,590.

Narzędzia do pomiaru pozostałych zmiennych oraz utworzone wskaźniki zostały opisane we wcześniejszych rozdziałach, w związku z tym nie będą ponownie omawiane.

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

Przypomnimy jedynie, że osiągnięcia szkolne mierzono na skali o średniej 100 i odchyleniu standardowym 15.

Model analizy danych

Celem weryfikacji trafności dwóch wariantów kontekstowych modeli osiągnięć (prostych i złożonych) wyspecyfikowano serie trzypoziomowych modeli regresji z losową stałą, z uwzględnieniem wag próbkowania (opisanych w rozdziale 1). Zmienną zależną stanowiły odpowiednio: poziom osiągnięć w zakresie matematyki, świadomości językowej oraz czytania. W pierwszej kolejności obliczono tzw. modele puste, tj. nie zawierające żadnych predyktorów, lecz uwzględniające grupowanie uczniów w szkołach i oddziały wewnątrz szkół. Służyły one za punkt odniesienia dla innych modeli w zakresie redukcji wariancji na poszczególnych poziomach analizy, pozwalały też ocenić konsekwencję ograniczenia liczby obserwacji w analizach. W drugim kroku do modeli wprowadzono (na poziomie szkoły lub oddziału) zmienne niezależne będące przedmiotem zainteresowania, tj. służące weryfikacji trafności. W kroku trzecim specyfikowano proste kontekstowe modele osiągnięć uczniów; uwzględniały one płeć, wykształcenie rodziców (w latach), wiek w miesiącach, bycie starszym lub młodszym niż główna kohorta uczniów oraz średnie wykształcenie rodziców liczone w latach nauki (na poziomie szkoły). Następnie do prostego modelu kontekstowego dodawano zmienne z kroku drugiego. Krok piąty polegał na specyfikacji złożonego kontekstowego modelu osiągnięć uczniów, natomiast szósty – na dodaniu jako predyktorów zmiennej/zmiennych z kroku drugiego. Przyjęto, że jeżeli kierunek zależności w modelach z kroku czwartego i szóstego był zgodny z oczekiwaniami, świadczy to na korzyść tezy o trafności wskaźników (zob. rozdział 3).

Dla uproszczenia narracji opisu wyników, model z kroku pierwszego nazywaliśmy modelem pustym, model z kroku trzeciego – modelem prostym (jako skrót od prostego modelu kontekstowego osiągnięć uczniów), natomiast model z kroku piątego – modelem złożonym (jako skrót od złożonego modelu kontekstowego osiągnięć uczniów). Ponadto w przypadku dodania do tych modeli dodatkowych zmiennych (kroki drugi, czwarty i szósty), dodaliśmy do jego nazwy znak „+”.

Przedstawiony schemat był wspólny dla wszystkich analiz przedstawionych w niniejszym rozdziale. Analizy przeprowadzono za pomocą oprogramowania HLM 6.06 (Scientific Software International, 2012) oraz Mplus 7.

Wyniki

Wyniki analiz przedstawione zostaną w blokach podzielonych na wskazane wcześniej obszary tematyczne: problematykę stylów zarządzania szkołą, charakterystyki nauczycieli i ich pracy z uczniami oraz integrację szkolną uczniów w oddziałach. Zaznaczyć należy, że zaprezentowane i opisane zostaną jedynie wybrane rezultaty analiz. Pozostałe zamieszczono w odpowiednich tabelach w aneksie, Czytelnik może je więc samodzielnie prześledzić.

Tabela 4.1. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających dane z obszaru „partycypacja”.

Zmienna niezależna	Modele								
	Umiejętność czytania		Świadomość językowa		Umiejętności matematyczne				
	Pusty+	Prosty+	Złożony+	Pusty+	Prosty+	Złożony+			
Partycypacja operacyjna	-3,537 (1,244)	-3,032 (0,894)	-2,787 (0,884)	-2,901 (1,785)	-2,440 (1,308)	-2,178 (1,210)	-3,097 (1,400)	-2,744 (0,970)	-2,479 (0,976)
Partycypacyjny styl przywództwa	2,865 (2,020)	2,924 (1,435)	2,488 (1,487)	2,518 (2,570)	2,483 (1,785)	2,018 (1,707)	1,968 (2,377)	2,409 (1,569)	1,913 (1,649)
Zaufanie do przełożonego	-1,154 (1,454)	-1,102 (1,107)	-0,939 (1,120)	-1,491 (1,764)	-1,113 (1,255)	-0,939 (1,191)	-0,023 (1,505)	-0,137 (1,019)	0,060 (1,061)
Zaufanie do współpracowników	-0,712 (0,893)	-0,074 (0,524)	-0,261 (0,521)	-0,340 (0,969)	0,325 (0,552)	0,127 (0,519)	-0,252 (0,976)	0,373 (0,587)	0,156 (0,567)
Identyfikacja z instytucją	-1,099 (1,218)	0,506 (0,812)	0,502 (0,822)	-1,716 (1,315)	0,126 (0,859)	0,154 (0,807)	-1,413 (1,337)	0,407 (0,810)	0,441 (0,771)
Możliwość samorozwoju	1,826 (1,203)	-0,681 (0,851)	-0,619 (0,873)	3,339 (1,292)	0,327 (0,923)	0,343 (0,853)	2,633 (1,218)	-0,259 (0,770)	-0,219 (0,749)
Informacja zwrotna	1,481 (0,832)	0,89 (0,579)	0,95 (0,608)	0,705 (1,033)	0,10 (0,660)	0,17 (0,601)	-0,290 (0,973)	-0,71 (0,641)	-0,65 (0,631)

Style zarządzania szkołą

Partycypacja nauczycieli w zarządzaniu szkołą

W przypadku partycypacji otrzymane wyniki są niejednoznaczne, ale wskazują na powiązania niektórych czynników związanych z partycypacją z kontekstowymi modelami efektywności nauczania. Część z tych związków jest jednak niezgodna z oczekiwaniami. W przypadku umiejętności czytania rezultaty wskazują na negatywne związki partycypacji operacyjnej z osiągnięciami uczniów. Efekt ten jest obecny w każdym z analizowanych modeli, choć należy uznać go za raczej niewielki. Najsilniej partycypacja operacyjna oddziałuje na osiągnięcia uczniów w modelu pustym+ (około 24% odchylenia standardowego). W przypadku kontekstowego modelu prostego (prostego+) negatywny wpływ partycypacji na wyniki jest mniejszy i wynosi około 20% odchylenia standardowego wyniku. Bardzo podobne rezultaty otrzymano w przypadku złożonego modelu kontekstowego (złożony+) – istotny, negatywny wpływ partycypacji na poziomie około 19% odchylenia standardowego wyniku uczniów z testu umiejętności czytania. Należy zauważyć, iż w modelu prostym+ istotny pozytywny wpływ uzyskano także w przypadku partycypacyjnego stylu przywództwa (na poziomie około 20% odchylenia standardowego). Ten rezultat wskazuje na dodatni związek poziomu osiągnięć z czytania uczniów (przy kontroli zmiennych kontekstowych) z postrzeganiem dyrektora szkoły przez nauczycieli jako sprzyjającego partycypacji.

Analizy wykonane dla modeli, w których zmienną zależną stanowiły osiągnięcia szkolne w zakresie świadomości językowej, dały odmienne rezultaty. Jedynie w modelu prostym+ uzyskano pozytywny związek postrzeganej przez nauczycieli możliwości samorozwoju z osiągnięciami uczniów w teście świadomości językowej (na poziomie 22% odchylenia standardowego wyniku uczniów). W przypadku pozostałych modeli (pusty+, złożony+) zmienne związane z partycypacją okazały się nieistotne.

Rezultaty uzyskane dla modeli uwzględniających osiągnięcia matematyczne są bardzo podobne do rezultatów dla modeli osiągnięć w zakresie czytania. Im większy spostrzegany przez nauczycieli wpływ na decyzje podejmowane w szkole, tym niższe osiągnięcia uczniów. Najsilniejsze zależności zanotowano w modelu pustym+ (na poziomie 21% odchylenia standardowego). W kontekstowych modelach osiągnięć, zarówno prostym+, jak i złożonym+, związek ten także okazał się istotny, przy sile odpowiednio 18% i 16% odchylenia standardowego. Dodatkowo w modelu pustym+, spostrzegana przez nauczycieli możliwość samorozwoju okazała się istotnie dodatnio związana z wynikami (na poziomie 17% odchylenia standardowego). Wybrane wyniki analiz znajdują się w Tabeli 4.1. W tabelach efekty istotne statystycznie na poziomie $p < 0,05$ zaznaczono pogrubieniem. W komórkach podano wartości współczynników regresji oraz błędów standardowych (w nawiasach).

Kultura organizacyjna szkoły

Analizy modeli dla osiągnięć uczniów w zakresie czytania nie wykazały istnienia zależności z kulturą organizacyjną szkoły. W każdym z analizowanych modeli, w tym

Tabela 4.2. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających dane z obszaru „kultura organizacyjna szkoły”.

Zmienna niezależna: typ kultury	Świadomość językowa – modele		
	Pusty+	Prosty+	Złożony+
Klanowy	-0,262 (0,950)	0,415 (0,477)	0,461 (0,450)
Adhokracyjny	-0,781 (0,594)	-0,877 (0,337)	-0,785 (0,321)
Rynkowy	0,103 (0,934)	0,772 (0,495)	0,741 (0,473)
Hierarchiczny	-0,009 (0,635)	-0,04 (0,367)	-0,05 (0,343)

w kontekstowych modelach efektywności nauczania, zmienne określające ten wymiar zarządzania szkołą okazały się nieistotne.

W przypadku modeli dla świadomości językowej zaobserwowano związki między adhokracyjnym typem kultury a osiągnięciami uczniów w modelach prostym+ i złożonym+. Uzyskane rezultaty są ujemne – wskazuje to na obniżanie się osiągnięć uczniów wraz ze wzrostem pozycji szkoły na skali „Kultura adhokracyjna”. Zaobserwowane efekty są jednak bardzo słabe – wynoszą one odpowiednio 6% i 5% odchylenia standardowego (Tabela 4.2).

W modelach dla osiągnięć uczniów z matematyki nie zaobserwowano istotnych statystycznie związków z kulturą organizacyjną szkoły w każdej z przeprowadzonych analiz (modele: pusty+, prosty+, złożony+). Kultura organizacyjna szkoły nie jest związana z osiągnięciami matematycznymi uczniów.

Współpraca rodziców ze szkołą

W przypadku współpracy rodziców ze szkołą znaleziono istotne statystycznie zależności jedynie w przypadku osiągnięć uczniów w zakresie świadomości językowej. Okazała się ona istotnie statystycznie związana w każdym z analizowanych modeli. W modelu, w którym współpraca rodziców ze szkołą była jedynym predyktorem, zanotowano jej negatywną zależność z osiągnięciami uczniów. Siła efektu wyniosła około 21% odchylenia standardowego. W modelach prostym+ i złożonym+ związek jakości współpracy rodziców z nauczycielami okazał się dodatni. Siła efektu wyniosła odpowiednio 14% i 13% odchylenia standardowego (Tabela 4.3).

W modelach dotyczących umiejętności czytania i matematycznych nie zaobserwowano istotnych statystycznie związków.

Charakterystyki nauczycieli i ich styl pracy z uczniami

Przed prezentacją wyników dla zmiennych nauczycielskich wskazać należy na specyficzne elementy przeprowadzonych analiz. Po pierwsze, uwzględniono w nich dane tylko tych nauczycieli-wychowawców klas I–III, którzy uczestniczyli w obu pomiarach (tj. w etapie I oraz III, cz. 1), oraz co do których nie było wątpliwości, że uczyli dany oddział przez

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

Tabela 4.3. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających dane z obszaru „kultura organizacyjna szkoły”.

Zmienna niezależna	Świadomość językowa – modele		
	Pusty+	Prosty+	Złożony+
Jakość relacji z nauczycielami	-3,114 (0,913)	2,074 (0,959)	1,879 (0,928)

3 lata. W próbie ostatecznie pozostało 3783 uczniów w 247 oddziałach, w 156 szkołach oraz 247 nauczycieli. Dzięki takiemu przygotowaniu danych możliwe było szacowanie efektu nauczyciela, każdej klasie został bowiem przypisany tylko jeden nauczyciel.

Ze względu na znaczne ograniczenie liczby obserwacji (w tym liczby szkół i oddziałów) uwzględnionych w analizach, w pierwszej kolejności przedstawimy wyniki dekompozycji wariancji odpowiednio dla poziomu szkół, oddziałów i poziomu ucznia, a także inne statystyki podsumowujące analizowane modele. Znajdują się one w Tabeli 4.4.

W modelu pustym, uwzględniającym jedynie grupowanie uczniów w oddziałach i szkołach, grupowanie uczniów w oddziałach wyjaśniło 4,7% wariancji wyników z testu mierzącego umiejętność czytania, 7,4% wariancji wyników z testu dotyczącego świadomości językowej i 5,7% wariancji wyników z testu umiejętności matematycznych. Z kolei grupowanie uczniów w szkołach wyjaśniło odpowiednio: 6,1%, 7,3% i 6,3% wariancji wyników. Niektórzy badacze (np. Scheerens i Bosker, 1997) interpretują wariancję międzyoddziałową wyników uczniów w modelu pustym jako efekt nauczyciela. Podejście to ma swoje uzasadnienie, ponieważ efektywność pracy nauczycieli jest najsilniejszym predyktorem osiągnięć uczniów na poziomie oddziału. Dlatego też podane oszacowania można traktować jako efekt nauczyciela w edukacji wczesnoszkolnej.

Na podstawie interpretacji współczynników korelacji wewnątrzgrupowej można stwierdzić, że uczniowie w nauczaniu początkowym najbardziej różnią się między szkołami ze względu na poziom świadomości językowej uczniów, najmniej zaś ze względu na poziom umiejętności czytania. W przypadku modeli z predyktorami, tj. prostego i złożonego, dla współczynników korelacji wewnątrzgrupowej można stosować taką samą interpretację jak dla modelu pustego. Wartości te to procent wyjaśnionej wariancji wyników uczniów z wariancji pozostałej do wyjaśnienia na poziomie oddziałów i szkół, po uwzględnieniu danego zestawu zmiennych niezależnych. Zgodnie z oczekiwaniami, wariancji pozostałej do wyjaśnienia jest coraz mniej wraz ze wzrostem złożoności modelu. Stąd też współczynniki korelacji wewnątrzgrupowej są niższe dla bardziej złożonych modeli.

Mimo że analizy wpływu charakterystyk nauczycieli na wyniki uczniów są prowadzone na próbie o zredukowanej liczebności (3783 uczniów w stosunku do 4254 dla wskaźników kultury organizacyjnej i partycypacji rodziców oraz 4126 dla KIU), różnice wyników dekompozycji wariancji oraz współczynników kontekstowych modeli efektywności są niewielkie. Mniejsza próba w przypadku analiz nauczycielskich jest wynikiem procedur łączenia oddziałów z nauczycielami i przyjętych kryteriów włączenia obserwacji do analiz.

W dalszej części rozdziału przedstawimy rezultaty analiz dotyczących związku poszczególnych charakterystyk nauczycieli i ich pracy z osiągnięciami uczniów.

Tabela 4.4. Podsumowanie analizowanych modeli.

	Modele									
	Umiejętność czytania			Świadomość językowa			Umiejętności matematyczne			
	Pusty	Prosty	Złożony	Pusty	Prosty	Złożony	Pusty	Prosty	Złożony	
Oszacowania efektów losowych										
Wariancja efektów szkół	13,78	2,36	2,58	16,37	0,25	0,57	13,87	0,79	0,71	
Wariancja efektów oddziałań	10,54	5,47	4,47	16,56	12,43	10,1	12,52	8,78	7,94	
Wariancja na poziomie ucznia	202	179,99	173,72	192,24	167,39	160,68	193,95	174,11	165,86	
Współczynniki korelacji wewnątrzgrupowej										
Poziom szkół	6,10%	1,30%	1,40%	7,30%	0,10%	0,30%	6,30%	0,40%	0,40%	
Poziom oddziałań	4,70%	2,90%	2,50%	7,40%	6,90%	5,90%	5,70%	4,80%	4,60%	
Łącznie	10,70%	4,20%	3,90%	14,60%	7,00%	6,20%	12,00%	5,20%	5,00%	
Podsumowanie										
Deviance	31 079,75	30 515,03	30 374,29	30 953,07	30 303,63	30 132,19	30 946,69	30 413,15	30 223,74	
Parametry	4	10	14	4	10	14	4	10	14	

Staż i wykształcenie

Staż pracy nauczyciela okazał się nieistotnym korelatem wyników uczniów. Jeśli chodzi o wykształcenie, to w szkołach podstawowych uczniowie nauczycieli z wykształceniem wyższym uzyskują w zależności od szacowanego modelu wyniki wyższe średnio od 17,4 do 18,9% odchylenia standardowego wyników testu z czytania, niż ich rówieśnicy nauczani przez nauczycieli z dyplomem studiów licencjackich. W przypadku świadomości językowej jest to odpowiednio od 11,8 do 16,2%. Z kolei wpływ wykształcenia nauczyciela na wyniki uczniów z testu sprawdzającego umiejętności matematyczne okazał się nieistotny (Tabela 4.5).

Zgodnie z oczekiwaniami wraz ze wzrostem złożoności modelu współczynniki dla efektu wykształcenia powinny być niższe. Należy zauważyć, że w przypadku wyników uczniów w zakresie czytania i świadomości językowej tak nie jest. Może to być spowodowane tym, że wykształcenie nauczyciela wchodzi w interakcję z którąś ze zmiennych z poziomu indywidualnego wprowadzonych do modelu złożonego. Dodatkowe analizy potwierdzają to przypuszczenie. Obserwujemy istotny efekt interakcji poziomu aspiracji edukacyjnych rodziców uczniów w oddziale z poziomem wykształcenia nauczyciela. W przypadku osiągnięć szkolnych w zakresie czytania jest to efekt wielkości 1,2% odchylenia standardowego wyników. Dla osiągnięć w zakresie świadomości językowej to 0,7%. Efekt nie jest istotny w wyjaśnianiu wyników testu sprawdzającego umiejętności matematyczne.

Stopień awansu zawodowego

Wyniki analiz uwzględniających dane na temat stopnia awansu zawodowego nauczycieli znajdują się w Tabeli 4.6. Okazał się on stosunkowo silnym i istotnym predyktorem osiągnięć uczniów w zakresie czytania i świadomości językowej. Jednak wraz z dodaniem korelatów osiągnięć szkolnych uczniów na poziomie indywidualnym pozytywny efekt awansu zawodowego zanika.

Styl pracy nauczyciela z uczniami

W przypadku skali „Aktywny styl nauczania” nie zaobserwowano istotnych statystycznie efektów. Deklaracje dotyczące stosowania przez nauczycieli aktywnego stylu nauczania nie korelują z osiągnięciami szkolnymi uczniów. Podobnie też nie zaobserwowano istotnych statystycznie efektów dla skali „Pomocny nauczyciel”. Istotne efekty otrzymano dla skali „Autorytet nauczyciela/utrzymanie dyscypliny” (Tabela 4.7). W nauczaniu początkowym skala ta istotnie koreluje z osiągnięciami uczniów w zakresie świadomości językowej i matematyki w modelach prostym oraz złożonym. Umiejętności uczniów z oddziałów, w których nauczyciel jest w stanie utrzymać dyscyplinę, są wyższe średnio o 7,4–8,2% odchylenia standardowego w przypadku świadomości językowej, oraz o 6,2–7,1% odchylenia standardowego w przypadku matematyki, w porównaniu do umiejętności rówieśników nauczanych przez nauczycieli słabiej radzących sobie z utrzymaniem dyscypliny.

Tabela 4.5. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających informacje o poziomie wykształcenia nauczyciela.

Zmienna niezależna	Modele								
	Umiejętność czytania		Świadomość językowa		Umiejętności matematyczne				
	Pusty+	Prosty+	Złożony+	Pusty+	Prosty+	Złożony+			
Magister*	2,84 (1,20)	2,61 (0,606)	2,64 (0,599)	2,43 (1,25)	1,76 (0,906)	1,83 (0,780)	0,981 (1,03)	0,985 (1,02)	0,926 (0,944)

* Kategoria odniesienia: licencjat ($n = 113$).**Tabela 4.6.** Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających informacje o stopniu awansu zawodowego nauczyciela.

Zmienna niezależna	Modele								
	Umiejętność czytania		Świadomość językowa		Umiejętności matematyczne				
	Pusty+	Prosty+	Złożony+	Pusty+	Prosty+	Złożony+			
Dyplomowany*	3,05 (1,03)	-0,118 (2,44)	-0,118 (2,44)	2,18 (0,987)	-1,21 (2,66)	-1,07 (2,14)	1,07 (0,878)	-2,27 (2,06)	-2,13 (1,54)
Mianowany*	3,42 (1,41)	0,235 (2,52)	0,235 (2,52)	3,69 (1,25)	0,161 (2,68)	0,260 (2,17)	1,47 (1,24)	-1,70 (2,16)	-1,64 (1,68)
Kontraktowy*	4,35 (2,13)	1,03 (2,93)	1,02 (2,93)	3,55 (2,65)	-0,307 (3,50)	-0,037 (2,97)	-0,090 (2,21)	-3,56 (2,80)	-3,31 (2,26)

* Kategoria odniesienia: nauczyciel-stażysta ($n = 76$).

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

Tabela 4.7. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających wyniki skali „Autorytet nauczyciela/utrzymanie dyscypliny”.

Zmienna niezależna	Modele								
	Umiejętność czytania		Świadomość językowa		Umiejętności matematyczne				
	Pusty+	Prosty+	Złożony+	Pusty+	Prosty+	Złożony+			
Autorytet nauczyciela/ utrzymanie dyscypliny	0,345 (0,609)	0,858 (0,465)	0,745 (0,444)	0,710 (0,639)	1,23 (0,482)	1,11 (0,455)	0,503 (0,590)	1,06 (0,490)	0,936 (0,461)

Tabela 4.8. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających wyniki dla czynnika „Nauczanie” z kwestionariusza TSES.

Zmienna niezależna	Modele								
	Umiejętność czytania		Świadomość językowa		Umiejętności matematyczne				
	Pusty+	Prosty+	Złożony+	Pusty+	Prosty+	Złożony+			
Czynnik „Nauczanie”	0,976 (0,380)	0,176 (0,295)	0,165 (0,280)	1,04 (0,418)	0,079 (0,299)	0,070 (0,287)	1,10 (0,387)	0,293 (0,281)	0,286 (0,266)

Tabela 4.9. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających wyniki dla czynnika „Zarządzanie” z kwestionariusza TSES.

Zmienna niezależna	Modele								
	Umiejętność czytania		Świadomość językowa		Umiejętności matematyczne				
	Pusty+	Prosty+	Złożony+	Pusty+	Prosty+	Złożony+			
Czynnik „Zarządzanie”	0,584 (0,367)	0,056 (0,282)	0,046 (0,280)	0,682 (0,435)	0,058 (0,290)	0,040 (0,275)	0,772 (0,384)	0,244 (0,262)	0,239 (0,247)

Tabela 4.10. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą i ze zmiennymi niezależnymi na poziomie oddziałów: integracja emocjonalna, społeczna, motywacyjna.

Zmienna niezależna	Model pusty+		
	Umiejętności matematyczne	Świadomość językowa	Umiejętność czytania
Integracja emocjonalna	-2,79 (1,19)	-1,56 (1,06)	-1,64 (1,00)
Integracja społeczna	3,18 (1,68)	3,09 (1,44)	3,03 (1,51)
Integracja motywacyjna	6,83 (1,37)	7,66 (1,31)	7,10 (1,33)

Poczucie skuteczności w zawodzie nauczyciela

W Tabelach 4.8 i 4.9 znajdują się wyniki analiz uwzględniających jako zmienną niezależną poczucie własnej skuteczności nauczyciela. Przyjęto hipotezę, że wysokie wyniki w obu podskalach narzędzia, tj. poczucie wysokiej skuteczności w zakresie stosowanych strategii nauczania („Nauczanie”) i wysokiej skuteczności w zakresie panowania nad zachowaniem klasy („Zarządzanie”), dodatnio korelują z osiągnięciami szkolnymi uczniów. Jedynie część tej hipotezy została zweryfikowana pozytywnie. Dodatni i istotny efekt zaobserwowano dla poczucia skuteczności w zakresie nauczania – pozytywnie korelowało ono z wynikami we wszystkich trzech obszarach (umiejętności czytania, świadomości językowej i umiejętności matematycznych). Jednak po uwzględnieniu korelatów osiągnięć szkolnych uczniów z poziomu indywidualnego istotność czynnika „Nauczanie” zanika.

W podsumowaniu analiz czynników nauczycielskich możemy stwierdzić, co następuje:

- Zaobserwowano istotny statystycznie związek wykształcenia nauczyciela z świadomością językową i umiejętnością czytania uczniów. Skala „Autorytet nauczyciela/utrzymanie dyscypliny” to także ważna zmienna w wyjaśnianiu zróżnicowania wyników z testów mierzących świadomość językową i kompetencje matematyczne.
- Dla skali „Aktywny styl nauczania”, skali „Pomocny nauczyciel” oraz czynnika „Utrzymanie dyscypliny” skali „Efektywność nauczyciela”, mimo ich dobrych właściwości statystycznych, nie zaobserwowano związku z wynikami uczniów.
- Z kolei stopień awansu zawodowego oraz czynnik „strategie nauczania” skali „Efektywność nauczyciela”, mimo że istotnie związany z osiąganymi wynikami, po uwzględnieniu indywidualnych charakterystyk uczniów traci swe znaczenie predykcyjne.

Integracja szkolna uczniów

W Tabeli 4.10 przedstawiono podsumowanie modeli, w których zmienną zależną stanowiły kolejno: osiągnięcia szkolne z matematyki, świadomości językowej i czytania. Modele zawierały (jako zmienne niezależne) wskaźniki integracji szkolnej. Współczynniki regresji z ich błędami standardowymi podano w nawiasach, a wartości istotne statystycznie pogrubiono.

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

Przeciętny stopień integracji motywacyjnej w oddziale okazał się najsilniejszym korelatem osiągnięć szkolnych. Wraz ze wzrostem o jedno odchylenie standardowe osiągnięcia uczniów w każdym z trzech obszarów okazały się rosnać o około 7 punktów (niemal pół odchylenia standardowego) przy kontroli pozostałych wymiarów integracji szkolnej. Przeciętny stopień integracji społecznej w oddziale korelował pozytywnie z osiągnięciami uczniów w zakresie świadomości językowej i czytania – wzrost ich poziomów o jedno odchylenie standardowe związany jest ze wzrostem wyników uczniów o około 3 punkty (około 20% odchylenia standardowego). Nieoczekiwany rezultat uzyskano dla integracji emocjonalnej – negatywnie koreluje ona z osiągnięciami, jednak zależność istotna statystycznie ujawniła się tylko dla matematyki.

W kolejnych tabelach (4.11 i 4.12) przedstawiono podsumowanie wyników analiz uwzględniających zmienne znajdujące się w prostych i złożonych modelach kontekstowych osiągnięć uczniów oraz dane na temat przeciętnego stopnia integracji szkolnej w oddziałach (modele proste+ i złożone+). Zmienną zależną stanowiły w nich kolejno: osiągnięcia szkolne z czytania, świadomości językowej i matematyki. W tabelach podano wartości współczynników regresji (wraz z ich błędami standardowymi w nawiasach) tylko dla integracji szkolnej na poziomie oddziałów. Każda kolumna z wartościami prezentuje wyniki innego modelu. Wartości istotne statystycznie pogrubiono. Ze względu na skorelowanie wyników w podskalach, gdy dla któregoś podskali uzyskiwano wynik nieistotny statystycznie, usuwano ją z modelu. Pełne wyniki modelowania zamieszczono w aneksie. Czytelnik może je odnaleźć, wykorzystując podane numery modeli.

W przypadku osiągnięć szkolnych w zakresie czytania każdy z aspektów integracji szkolnej na poziomie oddziału okazał się istotny statystycznie, jednak w przypadku integracji emocjonalnej tylko w sytuacji, gdy z modelu wyłączono integrację społeczną. Wzorzec ten ujawnił się zarówno w modelach prostych+ (Tabela 4.11), jak i złożonych+ (Tabela 4.12).

Siła efektu integracji motywacyjnej w obu wariantach modeli wahała się (w zależności od tego, jakie aspekty integracji szkolnej uwzględniono w modelu) w granicach 16,5–25% odchylenia standardowego. Dla integracji społecznej wyniosła około 17–20% odchylenia standardowego, dla integracji emocjonalnej – 10–12% (tylko w przypadku, gdy z modeli usunięto integrację społeczną). Siły zależności nieznacznie różniły się między modelami prostymi+ i złożonymi+, jednak biorąc pod uwagę błędy standardowe oszacowań współczynników, różnice te można uznać za pomijalne.

Wyniki modeli dla umiejętności uczniów w zakresie świadomości językowej okazały się bardzo zbliżone do wyników w zakresie czytania. Podobnie, zarówno w modelach prostych+ (Tabela 4.13), jak i złożonych+ (Tabela 4.14) istotne statystycznie okazały się wszystkie trzy aspekty integracji szkolnej, przy czym w przypadku integracji emocjonalnej – tylko w sytuacji, gdy z modelu wyłączono integrację społeczną.

Siła efektu integracji motywacyjnej w modelach dla obu obszarów osiągnięć wahała się (w zależności od uwzględnionych w modelu aspektów integracji szkolnej) w granicach od około 20–26% odchylenia standardowego w modelach prostych+ do 18–24% w modelach złożonych+. Dla integracji społecznej wyniosła około 16–20% odchylenia standardowego w obu wariantach, natomiast dla integracji emocjonalnej – około 12–13% (tylko w przypadku, gdy z modelu usunięto integrację społeczną). Różnice zaobserwowane między modelami prostymi+ a złożonymi+ można jednak uznać za pomijalne, wzięwszy pod uwagę błędy standardowe oszacowań.

Tabela 4.11. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających zmienne z prostego kontekstowego modelu osiągnięć z zakresu umiejętności czytania i dane na temat stopnia integracji szkolnej na poziomie oddziałów.

Zmienna zależna	Zmienne niezależne – poziom oddziałów	Współczynniki regresji: model prosty+		
		Model 23	Model 24	Model 25
Umiejętność czytania	Integracja emocjonalna	0,454 (0,867)		1,57 (0,774)
	Integracja społeczna	2,54 (1,17)	2,82 (1,01)	
	Integracja motywacyjna	2,76 (1,08)	2,85 (1,08)	3,75 (1,01)

Tabela 4.12. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających zmienne ze złożonego kontekstowego modelu osiągnięć z zakresu umiejętności czytania i dane na temat stopnia integracji szkolnej na poziomie oddziałów.

Zmienna zależna	Zmienne niezależne – poziom oddziałów	Współczynniki regresji: model złożony+		
		Model 26	Model 27	Model 28
Umiejętność czytania	Integracja emocjonalna	0,664 (0,809)		1,79 (0,738)
	Integracja społeczna	2,55 (1,12)	2,97 (0,984)	
	Integracja motywacyjna	2,48 (1,08)	2,61 (1,07)	3,48 (1,01)

Tabela 4.13. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających zmienne z prostego kontekstowego modelu osiągnięć w zakresie świadomości językowej i dane na temat stopnia integracji szkolnej na poziomie oddziałów.

Zmienna zależna	Zmienne niezależne – poziom oddziałów	Współczynniki regresji: model prosty+		
		Model 14	Model 15	Model 16
Świadomość językowa	Integracja emocjonalna	0,693 (0,882)		1,80 (0,765)
	Integracja społeczna	2,42 (1,04)	2,86 (0,886)	
	Integracja motywacyjna	3,01 (1,06)	3,15 (1,05)	3,86 (1,03)

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

Tabela 4.14. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających zmienne ze złożonego kontekstowego modelu osiągnięć w zakresie świadomości językowej i dane na temat stopnia integracji szkolnej na poziomie oddziałów.

Zmienna zależna	Zmienne niezależne – poziom oddziałów	Współczynniki regresji: model złożony+				
		Model 17	Model 18	Model 19		
Świadomość językowa	Integracja emocjonalna	0,991	(0,794)	2,03	(0,687)	
	Integracja społeczna	2,35	(0,986)	2,99	(0,836)	
	Integracja motywacyjna	2,66	(1,03)	2,82	(1,03)	3,58

Tabela 4.15. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających zmienne z prostego kontekstowego modelu osiągnięć z matematyki i dane na temat stopnia integracji szkolnej na poziomie oddziałów.

Zmienna zależna	Zmienne niezależne – poziom oddziałów	Współczynniki regresji: model prosty+				
		Model 4	Model 5	Model 6		
Umiejętności matematyczne	Integracja emocjonalna	-0,821	(1,03)	0,476	(0,993)	
	Integracja społeczna	2,94	(1,29)	2,43	(1,20)	
	Integracja motywacyjna	2,65	(1,18)	2,49	(1,17)	3,78

Tabela 4.16. Wybrane wyniki analiz wykonanych za pomocą trzypoziomowych modeli regresji z losową stałą, zawierających zmienne ze złożonego kontekstowego modelu osiągnięć z matematyki i dane na temat stopnia integracji szkolnej na poziomie oddziałów.

Zmienna zależna	Zmienne niezależne – poziom oddziałów	Współczynniki regresji: model prosty+				
		Model 7	Model 8	Model 9	Model 10	
Umiejętności matematyczne	Integracja emocjonalna	-0,539				(0,996)
	Integracja społeczna	2,85	2,51	3,67		(1,29) (1,21) (1,01)
	Integracja motywacyjna	2,35	2,24		3,78	(1,18) (1,16) (0,978)

W przypadku modelu prostego dla umiejętności matematycznych (Tabela 4.15), zmiennymi istotnymi statystycznie okazały się: przeciętny stopień integracji motywacyjnej oraz przeciętny stopień integracji społecznej w oddziale. Wzrost każdego z nich o jedno odchylenie standardowe związany był ze wzrostem poziomu osiągnięć uczniów o nieco ponad 16% odchylenia standardowego. Nie zanotowano istotnego efektu dla integracji emocjonalnej.

W przypadku modelu złożonego+ wyniki były zbliżone (Tabela 4.16) – statystycznie istotne zależności uzyskano dla stopnia integracji społecznej motywacyjnej w oddziale. Jednak w przypadku tego ostatniego wzorzec wyników był niestabilny – zmieniał się w zależności od uwzględnianych aspektów integracji szkolnej. Wzrost przeciętnego poziomu integracji społecznej w oddziale o jedno odchylenie standardowe związany był ze wzrostem poziomu osiągnięć z matematyki (w zależności od modelu) o około 17–24% odchylenia standardowego. W przypadku integracji motywacyjnej wahał się w podobnych granicach. Nie zanotowano istotnego efektu dla integracji emocjonalnej.

Podsumowanie

Celem analiz zaprezentowanych w rozdziale było sprawdzenie trafności kontekstowych wskaźników efektywności nauczania. Uzyskane wyniki są niejednoznaczne, ale kilka ważnych teoretycznie efektów przemawia za ich trafnością. Natomiast poza negatywnym efektem dla partycypacji nie uzyskano wyników wskazujących na to, że w szkołach o wysokich wartościach wskaźników efektywności nauczania mogą występować niepokojące zjawiska.

Analizy związane ze stylami zarządzania jedynie w części potwierdzają trafność kontekstowych modeli efektywności nauczania. Wykazano istnienie negatywnych związków partycypacji w przypadku modeli dla wyników testów z zakresu umiejętności czytania i matematycznych, co jest niepokojącym rezultatem. Ten czynnik jest bowiem w literaturze wskazywany jako pozytywnie związany z osiągnięciami uczniów. Należy jednak zauważyć, iż istnieją wyjaśnienia teoretyczne, które wskazują, że brak odpowiednich czynników moderujących może w taki sposób wpłynąć na partycypacyjny sposób zarządzania. Podkreśla się w tym kontekście rolę zaufania do współpracowników i przełożonego, możliwość samorozwoju, otrzymania informacji zwrotnej o swojej pracy oraz identyfikacji ze szkołą. W naszych analizach, w zdecydowanej większości rozpatrywanych modeli wyjaśniających osiągnięcia szkolne, zmienne te okazywały się być nieistotne statystycznie. Może to oznaczać, że partycypacja nie jest związana pozytywnie z efektywnością pracy szkoły z powodu niezastnienia pozostałych czynników, które warunkują jej właściwą ekspresję.

W przypadku kultury organizacyjnej istotne statystycznie rezultaty osiągnięto jedynie dla wyników testu z zakresu świadomości językowej – kultura adhokratyczna okazała się negatywnie związana z osiągnięciami uczniów w kontekstowych modelach osiągnięć. Ze względu na charakterystykę tego rodzaju kultury organizacji (tymczasowość, elastyczność w działaniu oraz podejmowanie ryzyka) możemy uznać, iż ten rezultat wskazuje na trafność kontekstowych wskaźników efektywności. Wyniki uzyskane dla czynnika współpracy rodziców ze szkołą także wskazują na trafność kontekstowych wskaźników efektywności. Istotne statystycznie rezultaty uzyskano dla skali określającej

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

jakość relacji rodziców z nauczycielami w modelu dla osiągnięć szkolnych uczniów w zakresie świadomości językowej.

W przypadku analiz związanych z charakterystyką nauczyciela i skal opisujących styl jego pracy część wyników potwierdza tezę o trafności kontekstowych modeli osiągnięć. Istotnie pozytywnie związane z wskaźnikami efektywności okazało się wykształcenie (w przeciwieństwie do stażu pracy). Nie uzyskano natomiast istotnych zależności w modelach uwzględniających skale „Aktywny nauczyciel” oraz „Pomocny nauczyciel” (opisujące styl pracy nauczyciela z uczniami podczas lekcji). Istotne związki w kontekstowych modelach efektywności nauczania uzyskano dla skali „Autorytet nauczyciela/utrzymanie dyscypliny”. Okazała się ona istotna w przypadku modeli dla wyników testów mierzących świadomość językową i umiejętności matematyczne. Nie stwierdzono związków poczucia skuteczności nauczycieli z miarami efektywności nauczania.

Analizy uwzględniające integrację szkolną dały najbardziej jednoznaczne rezultaty. Zaobserwowano pozytywne zależności kontekstowych wskaźników efektywności ze średnimi dla oddziału ocenami jakości relacji uczniów między sobą, samooceną ich kompetencji szkolnych oraz stosunkiem do szkoły. Można zatem uznać, iż wyniki te wskazują na trafność analizowanych wskaźników efektywności nauczania.

Bibliografia

- Adams, R. J. (2005). Reliability as a measurement design effect. *Studies in Educational Evaluation*, 31(2), 162–172.
- Anderson, R., Greene, M. i Loewen, P. (1988). Relationships among teachers' and students' thinking skills, sense of efficacy, and student achievement. *Alberta Journal of Educational Research*, 34(2), 148–165.
- AERA, APA i NCME (1999). Standards for educational and psychological testing. Washington, D.C.: Author.
- Allison, P. D. (2002). *Missing data*. Thousand Oaks, CA: Sage Publications.
- Ananat, E. O., Gassman-Pines, A., Francis, D. V. i Gibson-Davis, C. M. (2011). *Children left behind: the effects of statewide job loss on student achievement*. Cambridge, MA: National Bureau of Economic Research, Inc. Pobrano z <http://ideas.repec.org/p/nbr/nberwo/17104.html>.
- Armor, D., Conroy-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E. i Zellman, G. (1976). Analysis of the school preferred reading programs in selected Los Angeles minority schools, REPORT NO. R-2007-LAUSD. Santa Monica, CA: Rand Corporation (ERIC Document Reproduction Service No. 130 243).
- Ashton, P. i Crocker, L. (1987). Systematic study of planned variations-the essential focus of teacher-education reform. *Journal of Teacher Education*, 38(3), 2–8.
- Bartholomew, D. J., Knott, M. i Moustaki, I. (2011). *Latent variable models and factor analysis: a unified approach*. New York, NY: John Wiley & Sons.
- Blanchflower, D. G. i Oswald, A. J. (2005). The wage curve reloaded. *National Bureau of Economic Research Working Paper Series, No. 11338*. Pobrano z <http://www.nber.org/papers/w11338>
- Bossaert, G., Doumen, S., Buyse, E. i Verschueren, K. (2011). Predicting children's academic achievement after the transition to first grade: a two-year longitudinal study. *Journal of Applied Developmental Psychology*, 32(2), 47–57. doi:10.1016/j.appdev.2010.12.002
- Bowles, S. S. i Levin, H. M. (1968). More on multicollinearity and the effectiveness of schools. *The Journal of Human Resources*, 3(3), 393–400.
- Broomhall, D. E. i Johnson, T. G. (1994). economic factors that influence educational performance in rural schools. *American Journal of Agricultural Economics*, 76(3), 557–567.
- Bush, T. (2003). *Theories of educational leadership and management* (wyd. 3). London: Sage Publications.
- Bussey, K. i Bandura, A. (1999). Social cognitive theory of gender development and differentiation, *Psychological Review*, 106(4), 676–713.
- Byrne, C. (1983). Teacher knowledge and teacher effectiveness: a literature review, theoretical analysis and discussion of research strategy. Referat wygłoszony na konferencji Northwestern Educational Research Association, Ellenville, NY.

- Cameron, K. S. i Quinn, R. E. (2006). *Diagnosing and changing organizational culture: based on the competing values framework* (wyd. uzup.). San Francisco, CA: Jossey-Bass.
- Camilli, G., Vargas, S., Ryan, S. i Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, 112(3), 579–620.
- Carlson, M. J. i Corcoran, M. E. (2001). Family structure and children's behavioral and cognitive outcomes. *Journal of Marriage and Family*, 63(3), 779–792. doi:10.1111/j.1741-3737.2001.00779.x
- Carmines, E. G. i Zeller, R. A. (1979). Reliability and validity assessment. Newbury Park, CA: Sage Publications.
- Chiu, M. M. (2010). Effects of inequality, family and school on mathematics achievement: country and student differences. *Social Forces*, 88(4), 1645–1676.
- Coe, R. i Fitz-Gibbon, C. T. (1998). School effectiveness research: criticisms and recommendations. *Oxford Review of Education*, 24(4), 421–38.
- Coleman, J. S. (1966). *Equality of Educational Opportunity* (No. FS 5.238:38001). Washington, D.C.: The National Center For Educational Statistics.
- Comer, J. P. i Haynes, N. M. (1991). Parent involvement in schools: an ecological approach. *Elementary School Journal*, 91(3), 271–277.
- Conley, S. (1991). Review of research on teacher participation in school decision making. *Review of Research in Education*, 17(1), 225–266.
- Costrell, R., Hanushek, E. i Loeb, S. (2008). What do cost functions tell us about the cost of an adequate education? *Peabody Journal of Education*, 83(2), 198–223. doi:10.1080/01619560801996988
- Cronbach, L. J. (1971). Test validation. W: R. L. Thorndike (red.), *Educational measurement* (wyd. 2, s. 443–507). Washington, D.C.: American Council on Education.
- Deary, I. J., Strand, S., Smith, P. i Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence*, 35(1), 13–21. doi:10.1016/j.intell.2006.02.001
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. New York, NY: Guilford Press.
- Deleire, T. i Kalil, A. (2002). Good things come in threes: single-parent multigenerational family structure and adolescent adjustment. *Demography*, 39(2), 393–413. doi:10.1353/dem.2002.0016
- Dolata, R. (2008). *Szkola, segregacje, nierówności*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Dolata, R. (2013). *Międzyszkolne zróżnicowanie wyników nauczania na poziomie szkoły podstawowej i gimnazjum. Raport podsumowujący*. Warszawa: Instytut Badań Edukacyjnych.
- Dolata, R., Hawrot, A., Humenny, G., Jasińska, A., Koniewski, M., Majkut, P. i Żóltak, T. (2013). Trafność metody edukacyjnej wartości dodanej dla gimnazjów. Warszawa: Instytut Badań Edukacyjnych.
- Dolata, R., Jakubowski, M. i Pokropek, A. (2013). *Polska oświata w międzynarodowych badaniach umiejętności uczniów PISA OECD*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

- Dolata, R., Murawska, B. i Putkiewicz, E. (2001). *Monitorowanie osiągnięć szkolnych jako metoda wspierania lokalnego środowiska edukacyjnego*. Warszawa: Żak.
- Dolata, R., Murawska, B., Putkiewicz, E. i Żytko, M. (1997). *Monitorowanie osiągnięć szkolnych jako metoda doskonalenia edukacji. Zarys metody oraz przykłady zastosowań w edukacji początkowej*. Warszawa: Żak.
- 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 z XVIII Krajowej Konferencji Diagnostyki Edukacyjnej*. Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.
- Domański, H. i Pokropek, A. (2011). *Podziały terytorialne, globalizacja a nierówności społeczne: wprowadzenie do modeli wielopoziomowych*. Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii PAN.
- Driessen, G., Smit, F. i Slegers, P. (2005). Parental involvement and educational achievement. *British Educational Research Journal*, 31(4), 509–532. doi:10.1080/01411920500148713
- Duckworth, A. L., Quinn, P. D. i Tsukayama, E. (2012). What «No Child Left Behind» leaves behind: the roles of IQ and self-control in predicting standardized achievement test scores and report card grades. *Journal of Educational Psychology*, 104(2), 439–451. doi:10.1037/a0026280
- Ehrenberg, R. G. i Brewer, D. J. (1995). Did teachers' verbal ability and race matter in the 1960's? Coleman Revisted. *Economics of Education Review*, 14(1), 1–21.
- Else-Quest, N. M., Hyde, J. S. i Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological Bulletin*, 136(1), 103–127. doi:10.1037/a0018053
- Epstein, J. i Dauber, S. (1991). School programs and teacher practices of parent involvement in inner-city elementary and middle schools. *The Elementary School Journal*, 91(3), 289–305.
- Evertson, C. M., Hawley, W. D. i Zlotnik, M. (1985). Making a difference in educational quality through teacher-education. *Journal of Teacher Education*, 36(3), 2–12.
- Eysenck, H. J. i Eysenck, M. (1996). *Podpatrywanie umysłu: dlaczego ludzie zachowują się tak, jak się zachowują?* Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Fan, X. (2001). Parental involvement and students' academic achievement: a growth modeling analysis. *Journal of Experimental Education*, 70(1), 27– 61.
- Fan, X. i Chen, M. (2001). Parental involvement and students' academic achievement: a meta-analysis. *Educational Psychology Review*, 13(1), 1–22. doi:10.1023/A:1009048817385
- Fantuzzo, J., Tighe, E. i Child, S. (2000). Family involvement questionnaire: a multivariate assessment of family participation in early childhood education. *Journal of Educational Psychology*, 92(2), 367– 376.
- Feinstein, L. i Symons, J. (1999). Attainment in secondary school. *Oxford Economic Papers*, 51(2), 300–321. doi:10.1093/oepp/51.2.300
- Fenwick, R. i Tausig, M. (1994). The macroeconomic context of job stress. *Journal of Health and Social Behavior*, 35(3), 266–282.

- Ferguson, R. F. i Ladd, H. L. (1996). *How and why money matters: an analysis of Alabama schools*. W: H. F. Ladd (red.), *Holdine schools accountable*. Washington, D.C.: The Brookings Institution.
- Ferrer, E. i McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, 40(6), 935–952. doi:10.1037/0012-1649.40.6.935
- Fitz-Gibbon, C. T. i Koch, S. (2000). School effectiveness and education indicators. W: Ch. Teddlie i D. Reynolds (red.), *The international handbook of school effectiveness research* (s. 257–282). London–New York: Falmer Press.
- Ganzeboom, H. B. G. (bdw). How was new ISEI-08 constructed. Pobrano z http://home.fsw.vu.nl/hbg_ganzeboom/isco08/qa-isei-08.htm
- Ganzeboom, H. B. G. (2010). Tools for deriving occupational status measures from ISCO-08 with interpretative notes to ISCO-08. Pobrano z http://home.fsw.vu.nl/hbg_ganzeboom/isco08/
- Ganzeboom, H. B. G., De Graaf, P. M. i Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21(1), 1–56. doi:10.1016/0049-089X(92)90017-B
- Ganzeboom, H. B. G. i Treiman, D. J. (1996). Internationally comparable measures of occupational status for the 1988 International Standard Classification of Occupations. *Social Science Research*, 25(3), 201–239. doi:10.1006/ssre.1996.0010
- Ganzeboom, H. G. i Treiman, D. (2003). Three internationally standardised measures for comparative research on occupational status. W: J. P. Hoffmeyer-Zlotnik i C. Wolf (red.), *Advances in cross-national comparison* (s. 159–193). Pobrano z http://dx.doi.org/10.1007/978-1-4419-9186-7_9
- Ginther, D. i Pollak, R. (2004). Family structure and children's educational outcomes: blended families, stylized facts, and descriptive regressions. *Demography*, 41(4), 671–696. doi:10.1353/dem.2004.0031
- Glew, D. J., O'Leary-Kelly, A. M., Griffin, R. W. i Van Fleet, D. D. (1995). Participation in organizations: a preview of the issues and proposed framework for future analysis. *Journal of Management*, 21(3), 395–421.
- Główny Urząd Statystyczny (2007). Polska Klasyfikacja Działalności (PKD 2007). Pobrano z http://www.stat.gov.pl/klasyfikacje/pkd_07/pkd_07.htm
- Główny Urząd Statystyczny (2013). BDL: Bank Danych Lokalnych. O Banku. Pobrano z http://www.stat.gov.pl/bdl/app/strona.html?p_name=obanku
- González, R. L. i Jackson, C. L. (2013). Engaging with parents: the relationship between school engagement efforts, social class, and learning. *School Effectiveness and School Improvement*, 24(3), 316–335.
- Haeberlin, U., Bless, G., Moser, U. i Klaghofer, R. (1991). *Die Integration von Lernbehinderten*. Bern–Stuttgart: Verlag Paul Haupt.
- Haeberlin, U., Moser, U. i Bless, G. (1989). Fragebogen zur Erfassung von Dimensionen der Integration von Schülern (FDI 4–6) – Questionnaire for recording dimensions of integration of pupils. *Psychologie in Erziehung und Unterricht: Zeitschrift für Forschung und Praxis*, 36(H. 1), 17–26.
- Hanushek, E. A. (1971). Teacher characteristics and gains in student achievement – estimation using micro data. *American Economic Review*, 61(2), 280–288.

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

- Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100(1), 81–117.
- Hanushek, E. A. (2001). Spending on schools. W: T. Moe (red.), *A primer on American education* (s. 69–88). Stanford, CA: Hoover Institution Press. Pobrano z <http://hanushek.stanford.edu/publications/spending-schools>
- Hanushek, E. A. (2013). Financing schools. W: J. Hattie i E. M. Anderman (red.), *International guide to student achievement* (s. 134–136). New York, NY: Routledge. Pobrano z <http://hanushek.stanford.edu/publications/financing-schools>
- Hanushek, E. A. i Woessmann, F. (2011). The economics of international differences in educational achievement. W: E. A. Hanushek, S. Machin i L. Woessmann (red.), *Handbook of the economics of education* (t. 3). Amsterdam–London: North-Holland. Pobrano z <http://www.sciencedirect.com/science/book/9780444534293>
- Hartnell, CH. A., Ou, A. Y. i Kinicki, A. (2011). Organizational culture and organizational effectiveness: a meta-analytic investigation of the competing values framework's theoretical suppositions. *The Journal of Applied Psychology*, 96(4), 677–694.
- Harris, J. R. (2000). *Geny czy wychowanie? Co wyrosnie z naszych dzieci i dlaczego*. Warszawa: Wydawnictwo Jacek Santorski.
- Hartnell, C. A., Ou, A. Y. i Kinicki, A. (2011). Organizational culture and organizational effectiveness: a meta-analytic investigation of the competing values framework's theoretical suppositions. *The Journal of Applied Psychology*, 96(4), 677–694.
- Hawrot, A. (2013). *Raport techniczny z III etapu badania podłużnego w szkołach podstawowych*. Warszawa: Instytut Badań Edukacyjnych. Pobrano z http://2013.ewd.edu.pl//badania-szkoly-podstawowe/raport_techniczny_sp.pdf
- Heaven, P. C. L. i Ciarrochi, J. (2012). When IQ is not everything: intelligence, personality and academic performance at school. *Personality and Individual Differences*, 53(4), 518–522. doi:10.1016/j.paid.2012.04.024
- Heller, F. A., Pusic, E., Strauss, G. i Wilpert, B. (1998). *Organizational participation: myth and reality*. Oxford–New York: Oxford University Press.
- Herczyński, J. i Herbst, M. (2002). *Pierwsza odłona: społeczne i terytorialne zróżnicowanie wyników sprawdzianu szóstoklasistów i egzaminu gimnazjalnego przeprowadzonych wiosną 2002 roku*. Warszawa: Fundacja Klub Obywatelski.
- Herbst, M. (2004). Zróżnicowanie jakości kapitału ludzkiego w Polsce. Od czego zależą wyniki edukacyjne? *Studia Regionalne i Lokalne*, 17(3), 89–104.
- Herrnstein, R. J. (1994). *The bell curve: intelligence and class structure in American life*. New York, NY: Free Press.
- Hofer, M., Kuhnle, C., Kilian, B. i Fries, S. (2012). Cognitive ability and personality variables as predictors of school grades and test scores in adolescents. *Learning and Instruction*, 22(5), 368–375. doi:10.1016/j.learninstruc.2012.02.003
- Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B. i Williams, C. C. (2008). Diversity: gender similarities characterize math performance. *Science*, 321(5888), 494–495. doi:10.1126/science.1160364
- International Labour Organization (2007). Bureau of Statistics, work unit of the Policy Integration Department. Pobrano z <http://www.ilo.org/public/english/bureau/stat/isco/intro.htm>

- James, R. (2001). Participation disadvantage in Australian higher education: an analysis of some effects of geographical location and socioeconomic status. *Higher Education*, 42(4), 455–472.
- Jasińska, A. i Modzelewski, M. (2013). Międzyszkolne zróżnicowanie wyników nauczania po pierwszym etapie kształcenia. W: B. Niemierko i M. K. Szmigel (red.), *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych* (s. 165–178). Materiały z XIX Krajowej Konferencji Diagnostyki Edukacyjnej. Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.
- Jasińska, A. i Modzelewski, M. (2014). Testy osiągnięć szkolnych TOS3: przykład narzędzia skonstruowanego z wykorzystaniem modelu Rascha. *Edukacja* 2(127), 85–107.
- Jensen, A. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review*, 39(1), 1–123.
- Jensen, A. R. (1991). Spearman's ρ and the problem of educational equality. *Oxford Review of Education*, 17(2), 169–187.
- Jeynes, W. H. (2007). The relationship between parental involvement and urban secondary school student academic achievement: a meta-analysis. *Urban Education*, 42(1), 82–110.
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527–535.
- Karwowski, M. (red.). (2013). *Ścieżki rozwoju edukacyjnego młodzieży – szkoły ponadgimnazjalne. Trafność wskaźników edukacyjnej wartości dodanej dla szkół maturalnych*. Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii PAN.
- Karwowski, M. i Dziedziewicz, D. (2012). *Test umiejętności na starcie szkolnym. Podręcznik*. Warszawa: Instytut Badań Edukacyjnych.
- Katulska, S. (2010). Edukacyjny wyścig szczurów. Pobrano z <http://szkola.wp.pl/kat,1021419,title,Edukacyjny-wyscig-szczurow,wid,12396648,page,2,wiadomosc.html>
- Klitgaard, R. E. i Hall, G. R. (1975). Are there unusually effective schools? *Journal of Human Resources*, 10(1), 90–106.
- Konarski, R. (2009). *Modele równań strukturalnych: teoria i praktyka*. Warszawa: Wydawnictwo Naukowe PWN.
- Konarzewski, K. (2012). *TIMSS i PIRLS 2011 Osiągnięcia szkolne polskich trzecioklasistów w perspektywie międzynarodowej*. Warszawa: Centralna Komisja Egzaminacyjna. Pobrano z http://www.cke.edu.pl/images/stories/00000000000000000000_pirls_2011/timss_pirls_2011_raport_koncowy.pdf
- Koniewski, M. (2013). Szacowanie efektu nauczyciela na osiągnięcia edukacyjne uczniów z wykorzystaniem hierarchicznego modelowania liniowego. *Edukacja*, 123(3), 37–59.
- Leak, J., Duncan, G. J., Li, W., Magnuson, K., Schindler, H. i Yoshikawa, H. (2010). Is timing everything? How early childhood education program impacts vary by starting age, program duration and time since the end of the program. Referat wygłoszony na Association for Policy Analysis and Management meetings, 4–6 listopada, Boston, MA.

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

- Lee, J. i Fish, R. M. (2010). International and interstate gaps in value-added math achievement: multilevel instrumental variable analysis of age effect and grade effect. *American Journal of Education*, 117(1), 109–137.
- Lenkeit, J. (2013). Effectiveness measures for cross-sectional studies: a comparison of value-added models and contextualised attainment models. *School Effectiveness and School Improvement*, 24(1), 1–25. doi:10.1080/09243453.2012.680892
- Levine, P. B. (2011). How does parental unemployment affect children's educational performance. W: G. J. Duncan i R. J. Murnane (red.), *Whither opportunity? Rising inequality, schools, and children's life chances* (s. 315–335). New York, NY: Chicago: Russell Sage Foundation; Spencer Foundation.
- Lindberg, S. M., Hyde, J. S., Petersen, J. L. i Linn, M. C. (2010). New trends in gender and mathematics performance: a meta-analysis. *Psychological Bulletin*, 136(6), 1123–1135. doi:10.1037/a0021276
- Lisiecka, Z. (2010). *Badania uwarunkowań osiągnięć szkolnych uczniów szkół podstawowych. Sprawozdanie z realizacji I etapu podłużnych badań panelowych w szkołach podstawowych w okresie od 1. 10. 2009 do 10. 06. 2010*. Nieopublikowany raport techniczny. Warszawa: Centralna Komisja Egzaminacyjna. Pobrano z http://2013.ewd.edu.pl//badania-szkoly-podstawowe/SP_raport_techniczny.pdf
- Lynn, R. i Mikk, J. (2009). Sex differences in reading achievement. *Trames. Journal of the Humanities and Social Sciences*, 13(1), 3–13. doi:10.3176/tr.2009.1.01
- Marks, H. M. i Louis, K. S. (1997). Does teacher empowerment affect the classroom? The implications of teacher empowerment for instructional practice and student academic performance. *Educational Evaluation and Policy Analysis*, 19(3), 245–275.
- Martin, A. J. (2009). Age appropriateness and motivation, engagement, and performance in high school: effects of age within cohort, grade retention, and delayed school entry. *Journal of Educational Psychology*, 101(1), 101–114. doi:10.1037/a0013100
- Maslowski, R. (2001). *School culture and school performance: an explorative study into the organizational culture of secondary schools and their effects*. Twente: Twente University Press.
- Mattingly, D. J., Prislín, R., McKenzie, T. L., Rodriguez, J. L. i Kayzar, B. (2002). Evaluating evaluations: the case of parent involvement programs. *Review of Educational Research*, 72(4), 549–576.
- May, D. C. i Kundert, D. K. (1995). Does delayed school entry reduce later grade retentions and use of special education services? *Remedial and Special Education*, 16(5), 288–94.
- Mayer, S. E. i Knutson, D. (1999). Does the timing of school affect how much children learn? W: S. E. Mayer i P. E. Peterson (red.), *Earning and learning: how schools matter* (s. 79–102). Washington, D.C.: Brookings Institution Press.
- McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., Hamilton, L. S. i RAND Education. (2003). *Evaluating value-added models for teacher accountability* (No. 0-8330-3542-8). Santa Monica, CA: RAND Corporation.
- Midgley, C., Feldlaufer, H. i Eccles, J. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, 81(2), 247–258.

- Miller, R. J. i Rowan, B. (2006). Effects of organic management on student achievement. *American Educational Research Journal*, 43(2), 219–253.
- Ministerstwo Edukacji Narodowej (1991). Ustawa z dnia 7 września 1991 r. o systemie oświaty (Dz.U. Nr 56, poz. 458). Pobrano z:
<http://isap.sejm.gov.pl/DetailsServlet?id=WDU19910950425>
- Morrison, F. J. F. J., Griffith, E. M. i Alberts, D. M. (1997). Nature-nurture in the classroom: entrance age, school readiness, and learning in children. *Developmental Psychology*, 33(2), 254–262.
- Murnane, R. J. i Phillips, B. R. (1981). What do effective teachers of inner-city children have in common. *Social Science Research*, 10(1), 83–100.
- Muthén, L. K. i Muthén, B. O. (1998–2012). *Mplus user's guide. Seventh edition*. Los Angeles, CA: Muthén & Muthén.
- Noble, C. i Bradford, W. (2000). *Getting it right for boys and girls*. London: Routledge. Pobrano z
<http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=62877&site=ehost-live>
- Nye, B., Konstantopoulos, S. i Hedges, L. V. (2004). How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26(3), 237–257.
- Organisation for Economic Co-operation and Development (2008). *Measuring improvements in learning outcomes: best practices to assess the value-added of schools*. Paris: OECD.
- Organisation for Economic Co-operation and Development (2009). *PISA data analysis manual* (wyd. 2). Paris: OECD.
- Pinker, S. (2005). *Tabula rasa: spory o naturę ludzką?*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Plomin, R., DeFries, J. C., McClearn, G. E. i McGuffin, P. (2001). *Genetyka zachowania*. Warszawa: Wydawnictwo Naukowe PWN.
- Podstawa programowa z komentarzami. Tom 6. Edukacja matematyczna i techniczna w szkole podstawowej, gimnazjum i liceum* (bdw.). Pobrano z
http://www.oke.krakow.pl/inf/filedata/files/men_tom_6.pdf
- Pokropek, A. (2009). Metody statystyczne wykorzystywane w szacowaniu trzyletnich wskaźników egzaminacyjnych. W: B. Niemierko i M. K. Szmigel (red.), *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Materiały z XV Krajowej Konferencji Diagnostyki Edukacyjnej. Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.
- Pokropek, A. i Żółtak, T. (2012). *Trzyletni wskaźnik egzaminacyjny. Dokumentacja techniczna*, Warszawa: Centralna Komisja Egzaminacyjna. Pobrano z
<http://ewd.edu.pl/downloads/Dokumentacja%20techniczna%20v2.0.pdf>
- Pregler, A. (red.). (2013). *Ogólnopolskie badanie umiejętności trzecioklasistów. Raport z badania OBUT 2013*. Warszawa: Instytut Badań Edukacyjnych.
- Radwańska, J. (2008). Szkolny wyścig szczurów. *Psychologia w Szkole*, 4, 19–29.
- Raudenbush, S. W. (2004). *Schooling, statistics, and poverty: can we measure school improvement?* Princeton, NJ: Educational Testing Service.
- Raudenbush, S. W. i Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. Newbury Park, CA: Sage Publications.

4. Trafność proponowanego kontekstowego modelu oceny efektywności nauczania

- Raudenbush, S. W. i Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20(4), 307–335.
- Raudenbush, S. W., Rowan, B. i Kang, S. J. (1991). A multilevel, multivariate model for studying school climate with estimation via the EM algorithm and application to U.S. high-school data. *Journal of Educational and Behavioral Statistics*, 16(4), 295–330.
- Rosenholtz, S. J. (1986). Organizational conditions of teacher learning. *Teaching and Teacher Education*, 2(2), 91–104.
- Rowan, B., Chiang, F. S. i Miller, R. J. (1997). Using research on employees' performance to study the effects of teachers on students' achievement. *Sociology of Education*, 70(4), 256–284.
- Scheerens, J. i Bosker, R. (1997). *The foundations of educational effectiveness*. New York, NY: Pergamon.
- Schein, E. H. (2010). *Organizational culture and leadership*, fourth edition. San Francisco, CA: Jossey-Bass.
- Schütz, G., Ursprung, H. i Woessmann, L. (2005). *Education policy and equality of opportunity*. Munich: CESifo Group. Pobrano z http://ideas.repec.org/p/ces/ceswps/_1518.html
- Scientific Software International (2012). Pobrano z <http://www.ssicentral.com/hlm/example6-2.html>
- Singh, K., Bickley, P. G., Trivette, P., Keith, T. Z., Keith, P. B. i Anderson, E. (1995). The effects of four components of parental involvement on eighth-grade student achievement: structural analysis of NELS-88 data. *School Psychology Review*, 24(2), 299–317.
- Smith, J. (2009). Can regression discontinuity help answer an age-old question in education? The effect of age on elementary and secondary school achievement. *The B.E. Journal of Economic Analysis & Policy*, 9(1). doi:10.2202/1935-1682.2221
- Smylie, M. A., Lazarus, V. i Brownlee-Conyers, J. (1996). Instructional outcomes of school-based participative decision making. *Educational Evaluation and Policy Analysis*, 18(3), 181–198.
- Steinmayr, R. i Spinath, B. (2008). Sex differences in school achievement: what are the roles of personality and achievement motivation? *European Journal of Personality*, 22(3), 185–209. doi:10.1002/per.676
- Sternberg, R. J., Grigorenko, E. i Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly*, 47(1), 1–41. doi:10.1353/mpq.2001.0005
- Stevens, A. H. i Schaller, J. (2009). *Short-run effects of parental job loss on children's academic achievement*. National Bureau of Economic Research Working Paper Series, No. 15480. Pobrano z <http://www.nber.org/papers/w15480>
- Stoet, G. i Geary, D. C. (2013). Sex differences in mathematics and reading achievement are inversely related: within- and across- nation assessment of 10 years of PISA data. *PLoS ONE*, 8(3), e57988. doi:10.1371/journal.pone.0057988
- Świdurska, A. (2009). Kształtowanie się przemysłu wysokiej techniki w Polsce. *Prace Komisji Geografii Przemysłu*, 13, 56–67.
- Szumski, G. i Karwowski, M. (2012). Funkcjonowanie psychospołeczne uczniów z lekką niepełnosprawnością intelektualną na tle uczniów sprawnych intelektualnie – znaczenie formy kształcenia i płci. *Ruch Pedagogiczny*, (3), 33–55.

- Teddlie, Ch., Stringfield, S. i Reynolds, D. (2000). Context issues within school effectiveness research. W: Ch. Teddlie i D. Reynolds (red.), *The international handbook of school effectiveness research* (s. 160–185). London–New York: Falmer Press.
- Teo, A., Carlson, E., Mathieu, P., Egeland, B. i Sroufe, L. A. (1996). A prospective longitudinal study of psychosocial predictors of achievement. *Journal of School Psychology, 34* (3), 285–306. doi:10.1016/0022-4405(96)00016-7
- Timmermans, A. C., Doolaard, S. i Wolf, I. de (2011). Conceptual and empirical differences among various value-added models for accountability. *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice, 22*(4), 393–413. doi:10.1080/09243453.2011.590704
- Tołwińska, B. (2011). Kierowanie szkołą: rola dyrektora – partycypacja nauczycieli. W: S. Kwiatkowski, S. i J. Michalak (red.), *Przywódstwo edukacyjne w szkole i jej otoczeniu*. Warszawa: Difin.
- Treiman, D. J. (1977). *Occupational prestige in comparative perspective*. New York, NY: Academic Press.
- Tschannen-Moran, M. i Woolfolk Hoy, A. W. (2001). Teacher efficacy: capturing an elusive construct. *Teaching and Teacher Education, 17*(7), 783–805. doi:10.1016/S0742-051X(01)00036-1
- Tsereteli, M., Martskvishvili, K. i Aptarashvili, I. (2011). The impact of public high school management system on academic achievement. *Problems of Education in the 21st Century, 32*, 121–128.
- Zawistowska, A. (2012). *Horyzontalne nierówności edukacyjne we współczesnej Polsce*. Warszawa: Wydawnictwo Naukowe Scholar.
- Zhu, C., Devos, G. i Li, Y. (2011). Teacher perceptions of school culture and their organizational commitment and well-being in a Chinese school. *Asia Pacific Educational Review, 12*(2), 319–328.

Raport dla szkoły

Z BADANIA PODŁUŻNEGO W SZKOŁACH PODSTAWOWYCH

Efektywność nauczania na I etapie edukacyjnym



Dane szkoły

Nr identyfikacyjny:

Nazwa szkoły:

Miejscowość:

Adres:

Instytut Badań Edukacyjnych instytut badawczy

ul. Górczewska 8, 01-180 Warszawa | tel.: +48 22 241 71 00 | ibe@ibe.edu.pl | www.ibe.edu.pl

NIP 525-000-86-95 | Regon 000178235 | KRS 0000113990 Sąd Rejonowy dla m.st. Warszawy w Warszawie

Wstęp

Od kilku lat prowadzone są w Polsce, współfinansowane z Europejskiego Funduszu Społecznego, badania mające na celu wypracowanie dobrych wskaźników efektywności nauczania oraz monitorowanie ewaluacyjnej funkcji egzaminów zewnętrznych. Stanowią one kluczową część pracy badawczej Pracowni Edukacyjnej Wartości Dodanej. Realizowane są między innymi w 180 losowo wybranych szkołach podstawowych z całej Polski.

Państwa Placówka jest jedną ze szkół, które wzięły udział w badaniu. Przystąpili do niego uczniowie, którzy w roku szkolnym 2009/10 rozpoczęli naukę w klasach pierwszych. Wtedy odbył się pierwszy etap badania, w trakcie którego przeprowadzono między innymi badanie ankietowe rodziców, mające na celu poznanie cech środowiska rodzinnego, z którego pochodzą uczniowie. Losy tych uczniów były śledzone w następnych latach, kiedy realizowano kolejne etapy badania. Na początku roku szkolnego 2012/13, kiedy badani uczniowie uczęszczali do czwartych klas, przeprowadzono pomiar osiągnięć szkolnych podsumowujący etap edukacji wczesnoszkolnej. Badanie będzie kontynuowane do momentu zakończenia przez uczniów nauki w szkole podstawowej. Podczas badania zebrano dane od ponad 5 tysięcy uczniów i ich rodziców.

Niniejsza broszura jest kolejną przekazaną Państwu, prezentującą wybrane wyniki Państwa Szkoły. Tym razem podsumowujemy niektóre zagadnienia dotyczące edukacji wczesnoszkolnej. Chcemy się z Państwem podzielić informacją na temat wskaźników efektywności nauczania na pierwszym etapie edukacyjnym.

Podkreślamy, iż wyniki konkretnych szkół i klas nie są prezentowane w żadnych opracowaniach poza niniejszym, kierowanym wyłącznie do Państwa rąk.

Zapraszamy do lektury!

Podsumowanie badania w szkole

W raporcie zaprezentowane zostaną wyniki zbiorcze dla szkoły. W przypadku szkół jedno- i dwuoddziałowych (w momencie rozpoczęcia badania na etapie edukacji wczesnoszkolnej) wskaźniki zostały wyliczone na podstawie danych ze wszystkich oddziałów na danym etapie nauczania. Jednak w przypadku szkół z większą niż 2 liczbą oddziałów, do badania wybrano losowo 2 oddziały. Poniżej prezentujemy, w których oddziałach przeprowadzono badanie w Państwa Szkole.

Badaniem objęto następujące oddziały: A.

Uwaga! Jeśli na danym poziomie nauczania było w szkole więcej klas niż dwie wyżej wymienione, uzyskane wyniki należy interpretować jako łączną informację o nauczaniu w wymienionych klasach.

Uwaga! Jest to oznaczenie oddziału dla uczniów rozpoczynających naukę w szkole podstawowej w roku szkolnym 2009/10.

Dlaczego potrzebujemy wskaźników efektywności nauczania?

Efektywność nauczania w danej szkole często jest utożsamiana z jej wynikami w testach osiągnięć szkolnych. Wyniki testów zależą jednak nie tylko od pracy szkoły, ale także od cech uczniów takich jak ich zdolności poznawcze, czy cech środowiska rodzinnego takich jak status społeczno-ekonomiczny czy aspiracje rodziców względem edukacji ich dzieci. Dlatego wysoki lub niski wynik w testach osiągnięć może być nie tylko efektem pracy szkoły, ale także konsekwencją potencjału uczniowskiego.

Gdyby wszystkie szkoły w Polsce pracowały z uczniami o takich samych cechach mających znaczenie dla ich osiągnięć oraz w takim samym otoczeniu społecznym, to wyniki testów osiągnięć faktycznie mogłyby być dobrą miarą efektywności nauczania. Każda szkoła miałaby bowiem taką samą szansę na osiągnięcie wysokich wyników. Jednak tak nie jest. Szkoły w Polsce różnią się pod względem tego, z jakimi uczniami pracują. Przeprowadzone badanie pozwoliło oszacować skalę tego zróżnicowania. Szkoły podstawowe najbardziej różnią się między sobą ze względu na poziom wykształcenia rodziców uczniów, status społeczny rodziny oraz poziom inteligencji uczniów. Mniejsze, choć nadal znaczące, zróżnicowanie międzyszkolne zaobserwowano dla zasobności materialnej rodziny ucznia, prestiżu wykonywanego przez rodziców zawodu oraz aspiracji rodziców względem wykształcenia ich dzieci. Badania pokazały, że aby sprawiedliwie ocenić wkład szkoły w osiągnięte wyniki, należy więc uwzględnić to, że szkoły pracują z uczniami o różnych cechach.

Prezentowane Państwu wskaźniki efektywności nauczania uwzględniają zróżnicowanie szkół ze względu na cechy uczniów oraz środowiska, w jakim pracuje szkoła. W ten sposób pokazują wyniki szkoły, na tle innych, pracujących z podobnymi uczniami. O szkołach, które uzyskały wyniki wyższe niż inne szkoły pracujące w podobnym środowisku, będziemy mówić, że uczą z ponadprzeciętną efektywnością. Są to bowiem szkoły, które wypracowały z uczniami więcej niż inne placówki, do których uczęszczali podobni uczniowie. Natomiast szkoły, które uzyskały wyniki niższe, niż inne pracujące

podobnymi uczniami, będziemy nazywać szkołami pracującymi z niższą niż przeciętna efektywnością. Uzyskały one bowiem słabsze rezultaty, niż inne szkoły pracujące z podobnymi uczniami.

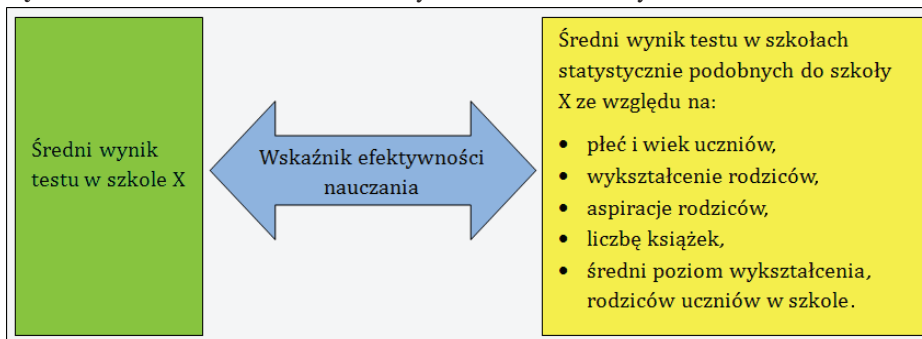
Wskaźnik efektywności nauczania jest więc miarą względną, pokazującą czy dana szkoła osiąga wyższe, czy niższe wyniki, niż inne placówki pracujące z podobnymi uczniami. Nie mówi ona jednak, jaki jest poziom wiedzy uczniów. O tym ostatnim informują wyniki testów osiągnięć. Dopiero obie te miary: wynik testu osiągnięć i wskaźnik efektywności nauczania, analizowane łącznie, dają pełniejszy obraz pracy szkoły. Dlatego w niniejszym raporcie prezentujemy Państwu zarówno informację o wynikach w testach osiągnięć uzyskanych przez Państwa uczniów, jak i o wyznaczonych na ich podstawie wskaźnikach efektywności nauczania.

Co to jest kontekstowy wskaźnik efektywności nauczania?

Prezentowane Państwu wskaźniki efektywności nauczania zostały wyliczone na podstawie danych z badania przeprowadzonego na reprezentatywnej próbie 180 szkół podstawowych w Polsce. Do ich wyznaczenia wykorzystano informacje o środowisku rodzinnym, z którego pochodzą uczniowie. Wykorzystano następujące wskaźniki statusu społeczno-ekonomicznego: wykształcenie rodziców, liczba posiadanych książek oraz aspiracje rodziców. Przeprowadzone analizy pokazały, że wskaźniki te w wystarczającym stopniu pozwalają opisać status rodziny ucznia. Uwzględniono też informację o płci i wieku ucznia (wraz z informacją, czy jest on uczniem z głównego rocznika, czy jest starszy lub młodszy). Środowisko, w którym pracuje szkoła, opisano także za pomocą średniego wykształcenia rodziców uczniów w szkole. Wszystkie te cechy są istotnie związane z osiągnięciami szkolnymi uczniów. Prezentowane wskaźniki nazywamy kontekstowymi, bo pozwalają na porównanie szkół podobnych do siebie ze względu na wiele ważnych cech opisujących kontekst pracy szkoły.

Do wyliczenia wskaźników wykorzystano modele statystyczne, na podstawie których obliczamy, jakie średnio wyniki uzyskały szkoły pracujące z uczniami o podobnych cechach. Szkoły te nazywamy szkołami statystycznie podobnymi do danej szkoły.

Rysunek 1. Idea szacowania kontekstowych wskaźników efektywności nauczania



Wyniki szkół statystycznie podobnych traktujemy jako punkt odniesienia dla wyników danej szkoły. Następnie patrzymy, czy w tej szkole uczniowie uzyskali wyniki wyższe, czy niższe niż w szkołach pracujących z podobnymi uczniami.

Jeśli uzyskano wyniki wyższe, wskaźnik efektywności nauczania będzie przyjmował **wartość dodatnią** i będzie świadczył o **ponadprzeciętnej efektywności**. Jeśli uzyskano wyniki niższe, będzie przyjmował **wartość ujemną** i będzie świadczył o **poniżej przeciętnej efektywności** nauczania. Jeśli uzyskano wyniki porównywalne ze statystycznie podobnymi szkołami, wskaźnik będzie przyjmował wartość w okolicach zera i będzie świadczył o tym, że szkoła naucza z **przeciętną efektywnością**, czyli taką, jak większość szkół w kraju.

Wskaźnik efektywności nauczania jest więc liczony w taki sposób, że szansa na jego wysoką wartość nie zależy od tego, z uczniami o jakich cechach (uwzględnionych w modelu) pracowała szkoła. Szkoła, która świetnie pracowała z uczniami z niekorzystnego środowiska społecznego, osiągnie dodatnią wartość wskaźnika, mimo że jej średnie wyniki w teście osiągnięć będą prawdopodobnie niższe niż szkoły, która słabo pracowała z uczniami pochodzącymi z „dobrych domów”.

Wskaźniki te zostały policzone dla trzech obszarów wyników nauczania odpowiadających wykorzystanym testom osiągnięć: umiejętności czytania, świadomości językowej i umiejętności matematycznych.

Dlaczego wyniki komunikujemy z przedziałem ufności?

Kiedy chcemy opisać jakąś cechę otaczającej nas rzeczywistości (jak np. wynik szkoły w zakresie umiejętności jej uczniów czy efektywność nauczania), skazani jesteśmy na niepewność oszacowania wartości danej miary. Niepewność ta wynika z dwóch źródeł błędów. Pierwsze związane jest z samym pomiarem interesujących nas charakterystyk takich jak poziom umiejętności uczniów. Wyniki uczniów w testach mogą bowiem zależeć nie tylko od poziomu ich umiejętności, ale także od dyspozycji w danym dniu, zgadywania odpowiedzi w zadaniach zamkniętych, ściągania, błędów podczas oceniania testów, przypadkowości w doborze zadań testowych itp. Drugi błąd związany jest z faktem, że pomiarem zawsze obejmuje się tylko próbę uczniów szkoły, na podstawie której wypowiada się o wynikach jej pracy. Mamy tu na myśli zarówno fakt, że pomiarem mogą nie zostać objęci niektórzy uczniowie z danego poziomu nauczania, jak i fakt, że uczniowie z określonego rocznika są tylko przykładowymi uczniami danej szkoły. Być może, gdyby w tym samym czasie uczęszczali do niej inni uczniowie, uzyskane wyniki byłyby odmienne. Powoduje to konieczność wyznaczenia przedziału ufności dla oszacowanych wartości (zarówno średniego wyniku w teście osiągnięć, jak i wskaźnika efektywności nauczania). Przedział ufności, to przedział, który z określonym prawdopodobieństwem zawiera prawdziwą wartość wskaźnika.

Wielkość przedziału ufności zależy między innymi od liczby uczniów, na podstawie wyników których wyznaczamy dany wskaźnik. Im więcej uczniów zostało uwzględnionych w analizach, tym jesteśmy bardziej pewni uzyskanego wyniku, dlatego przedział ufności może być mniejszy. Wielkość przedziału ufności zależy także od przyjętego przez nas stopnia pewności, z jaką chcemy komunikować dany wynik. Im większą chcemy posiadać pewność, że prawdziwa wartość wskaźnika znajduje się w wyznaczonym przedziale, tym będzie on szerszy. W niniejszym raporcie komunikujemy wyniki z 90-procentowymi przedziałami ufności. Wskazywać on będzie przedział, w którym z 90-procentowym prawdopodobieństwem znajduje się prawdziwa wartość danego wskaźnika szkoły.

Oznacza to, że jeśli np. chcemy powiedzieć, że z 90-procentową pewnością efektywność pracy szkoły jest powyżej przeciętnej, cały przedział ufności powinien znajdować się powyżej punktu oznaczającego przeciętną efektywność, czyli powyżej zera. Jeśli natomiast chcemy powiedzieć, że dwie szkoły różnią się efektywnością nauczania, przedziały ufności oszacowanych dla nich wskaźników muszą być rozłączne. Podobnie oczywiście wykorzystujemy przedziały ufności przy wyciąganiu wniosków na temat wyników w testach osiągnięć.

Wyniki Państwa szkoły

W dalszej części raportu przedstawione zostały wyniki Państwa Szkoły dla każdego z trzech obszarów wyników nauczania. Najpierw przedstawiono wyniki szkoły w danym teście, uzupełniając je o dodatkowe informacje, w stosunku do tego, co zostało Państwu przekazane w poprzednim raporcie z badania. Następnie zaprezentowano rezultaty szkoły w zakresie efektywności nauczania. Na zakończenie pokazano łącznie informację o wynikach szkoły w teście osiągnięć i efektywności nauczania w tym zakresie na tle wszystkich szkół biorących udział w badaniu. Informacja o tym, co mierzył każdy test została Państwu przedstawiona w poprzednim raporcie, dlatego nie będziemy jej tutaj powtarzać.

Wyniki z zakresu umiejętności czytania

Wyniki testu osiągnięć

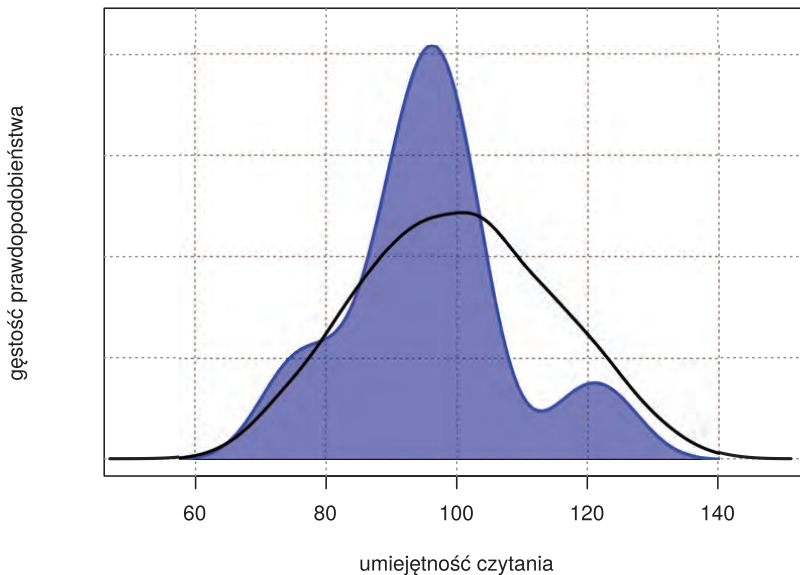
Liczba uczniów, których wyniki zostały wykorzystane przy wyznaczeniu wskaźnika dla szkoły: 17.

Uwaga! W analizach wykorzystano dane dla wszystkich uczniów, którzy rozwiązali dany test. Jeśli jednak liczba ta jest mniejsza niż 10, prezentowane wyniki należy interpretować z dużą ostrożnością, mogą być one bowiem mało wiarygodne.

Na poniższym wykresie zaprezentowano rozkład wyników uczniów Państwa Szkoły na tle wyników w populacji. Czarną linią wyrysowano rozkład wyników w populacji uczniów. Jego średnia wynosi 100 punktów, a odchylenie standardowe 15 punktów. Niebieskim kolorem zaznaczono rozkład wyników w teście umiejętności czytania uczniów Państwa Szkoły. Porównując wzajemne położenie obu tych rozkładów, mogą Państwo ocenić, czy uczniowie Państwa Szkoły uzyskali lepsze czy słabsze wyniki niż średnio uczniowie w kraju. Jeśli rozkład wyników w Państwa Szkole jest przesunięty w prawo w stosunku do rozkładu populacyjnego, oznacza to, że uczniowie uzyskali wyniki powyżej przeciętnych. Jeśli jest przesunięty w lewo – uzyskali wyniki poniżej średniej. Dodatkowo mogą Państwo porównać kształty obu rozkładów. Jeśli rozkład wyników w Państwa Szkole jest szczuplejszy i wyższy niż rozkład populacyjny, oznacza to, że wyniki uczniów w Państwa Szkole są mniej zróżnicowane niż średnio w populacji. Jeśli jest on szerszy i niższy, oznacza to, że uczniowie w Państwa Szkole uzyskali wyniki bardziej zróżnicowane niż

średnio w populacji. Ponadto, jeśli daje się zauważyć znaczące odstępstwa kształtu rozkładu wyników w szkole od rozkładu normalnego (którego przybliżeniem jest dzwonowaty kształt rozkładu populacyjnego), może to świadczyć o istnieniu w szkole grup uczniów bardzo różniących się wynikami.

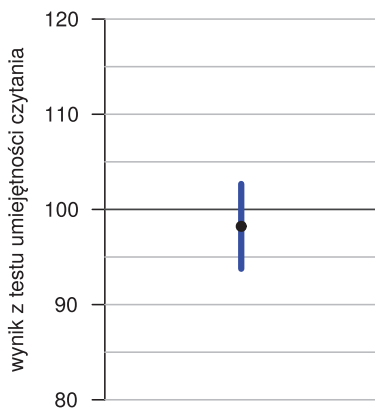
Rysunek 2. Rozkład wyników w teście umiejętności czytania w szkole na tle rozkładu wyników w populacji.



Na kolejnym wykresie przedstawiono średni wynik szkoły w teście umiejętności czytania wraz z przedziałem ufności. Oś pionowa to skala wyników uzyskanych w teście umiejętności czytania. Średni wynik szkoły jest reprezentowany przez czarny punkt. Przedział ufności jest zaznaczony niebieskimi odcinkami. W przedziale tym z 90-procentowym prawdopodobieństwem znajduje się prawdziwy wynik Państwa Szkoły. Wyższy wynik wskazuje wyższy poziom danej umiejętności.

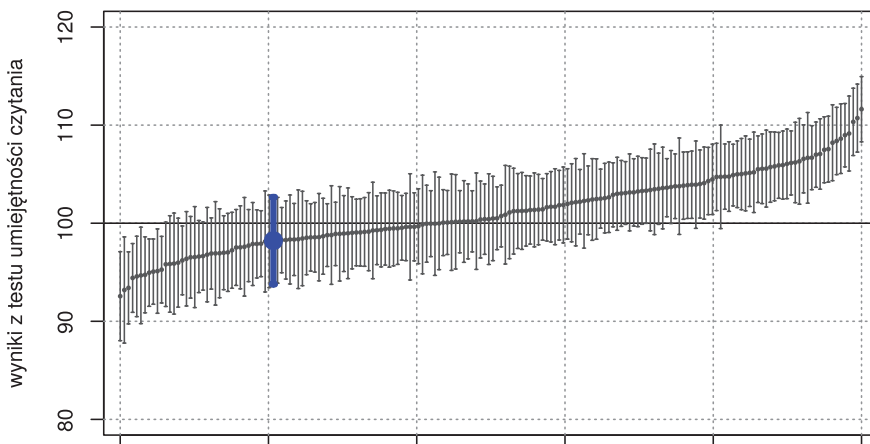
Średni wynik w populacji uczniów wynosi 100 punktów. Dlatego, jeśli cały przedział ufności znajduje się ponad wynikiem 100, uczniowie uzyskali w danym teście wynik wyższy niż średni wynik w populacji. Jeśli cały przedział ufności znajduje się poniżej wyniku 100, uczniowie mieli wynik w danym teście niższy niż średni wynik w populacji uczniów. Jeśli przedział ufności przecina 100 punktów, oznacza to, że uczniowie uzyskali wynik porównywalny ze średnim wynikiem w populacji uczniów.

Rysunek 3. Średni wynik szkoły w teście umiejętności czytania wraz z przedziałem ufności.



Na kolejnym wykresie ten sam wynik przedstawiono na tle wyników wszystkich szkół objętych badaniem. Wynik Państwa Szkoły wraz z przedziałem ufności został pogrubiony i wyróżniony kolorem niebieskim. Wyniki wszystkich szkół zostały uporządkowane rosnąco według wartości średniej. Na wykresie tym widać, że większość szkół osiągnęła porównywalne wyniki (ich przedziały ufności nie są rozłączne), które są bliskie średnim wynikom w populacji uczniów. Przypominamy, że o dwóch szkołach możemy powiedzieć, że różnią się średnimi wynikami, jeśli ich przedziały ufności nie pokrywają się.

Rysunek 4. Średnie wyniki szkoły w teście umiejętności czytania na tle wyników wszystkich badanych szkół.



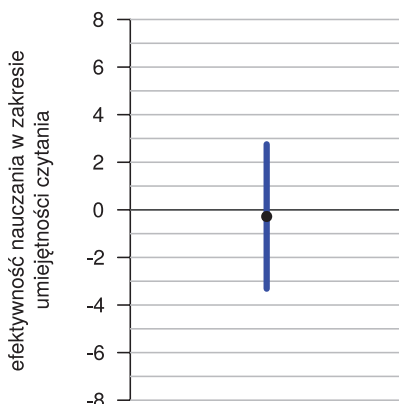
Efektywność nauczania

Liczba uczniów, których wyniki zostały wykorzystane przy wyznaczeniu wskaźnika dla szkoły: 15.

Uwaga! W analizach wykorzystano dane dla wszystkich uczniów, którzy rozwiązali dany test. Jeśli jednak liczba ta jest mniejsza niż 10, prezentowane wyniki należy interpretować z dużą ostrożnością, mogą być one bowiem mało wiarygodne.

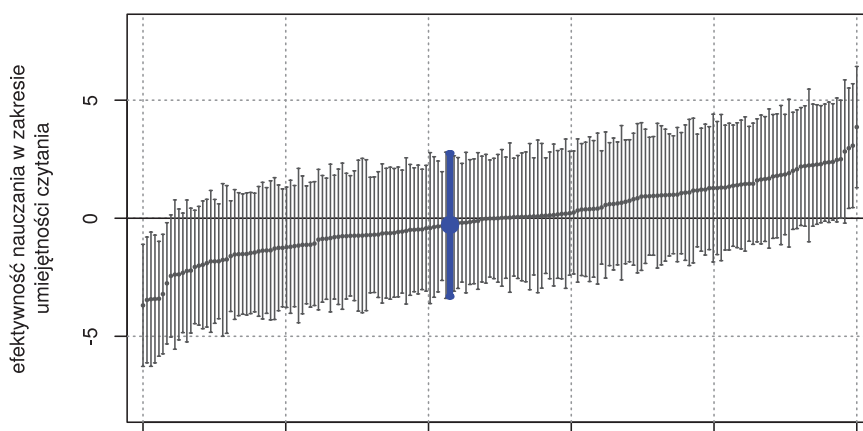
Na poniższym wykresie zaprezentowano wskaźnik efektywności nauczania w zakresie umiejętności czytania (jego średnią oznaczono punktem) wraz z 90-procentowym przedziałem ufności (oznaczonym niebieskimi odcinkami). Oś pionowa to skala efektywności nauczania. Im wyższy na niej wynik, tym wyższa efektywność nauczania. Wynik 0 oznacza średnią efektywność nauczania w populacji. Jeśli cały przedział ufności wskaźnika znajduje się powyżej punktu 0, oznacza to, że Państwa Szkoła naucza z powyżej przeciętną efektywnością. Uzyskuje więc wyniki wyższe niż inne szkoły pracujące z podobnymi uczniami. Jeśli przedział ufności znajduje się poniżej punktu 0, oznacza to, że Państwa Szkoła naucza z poniżej przeciętną efektywnością nauczania. Inne szkoły pracujące z podobnymi uczniami uzyskały średnio lepsze rezultaty. Jeśli przedział ufności przecina wartość 0, oznacza to, że Państwa Szkoła pracuje z przeciętną efektywnością, czyli taką jak większość szkół w kraju.

Rysunek 5. Efektywność nauczania w zakresie umiejętności czytania.



Na kolejnym wykresie efektywność nauczania w Państwa Szkole przedstawiono na tle wyników wszystkich szkół objętych badaniem. Wynik Państwa Szkoły wraz z przedziałem ufności został pogrubiony i wyróżniony kolorem niebieskim. Wyniki wszystkich szkół zostały uporządkowane rosnąco według wartości średniej. Wartość 0 oznacza średnią efektywność nauczania. Wykres ten pokazuje, że większość szkół naucza z porównywalną, przeciętną efektywnością (ich przedziały ufności pokrywają się). Przypominamy, że o dwóch szkołach możemy powiedzieć, że różnią się efektywnością nauczania, jeśli ich przedziały ufności są rozłączne.

Rysunek 6. Efektywność nauczania w zakresie umiejętności czytania na tle wszystkich badanych szkół.



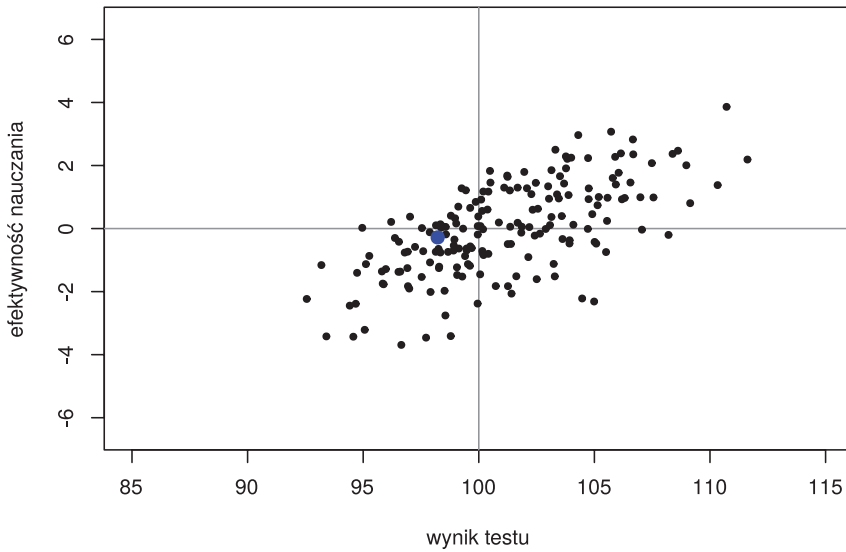
Łączna informacja o wyniku i efektywności nauczania

Na ostatnim wykresie przedstawiono równocześnie informację o średnim wyniku w teście i efektywności nauczania w Państwa Szkole na tle wyników wszystkich szkół biorących udział w badaniu. Dzięki takiemu zestawieniu informacji, mogą Państwo zobaczyć, gdzie ze względu na wartości obu wskaźników łącznie znajduje się Państwa Szkoła w porównaniu do innych szkół.

Na osi poziomej przedstawiono wyniki uzyskane w teście umiejętności czytania. Średni wynik w populacji uczniów to 100 punktów, a odchylenie standardowe wynosi 15 punktów. Na osi pionowej przedstawiono wskaźnik efektywności nauczania. Punkt 0 oznacza średnią efektywność nauczania umiejętności czytania w populacji szkół. Wynik Państwa Szkoły został wyróżniony pogrubieniem i kolorem niebieskim. Na wykresie

przedstawiono wyniki bez przedziałów ufności, by zwiększyć jego czytelność. Trzeba jednak pamiętać, że są one obarczone niepewnością pomiarową, więc nie należy przeceniać małych różnic.

Rysunek 8. Łączna informacja o wyniku w teście osiągnięć i efektywności nauczania w zakresie umiejętności czytania.



Wyniki z zakresu świadomości językowej

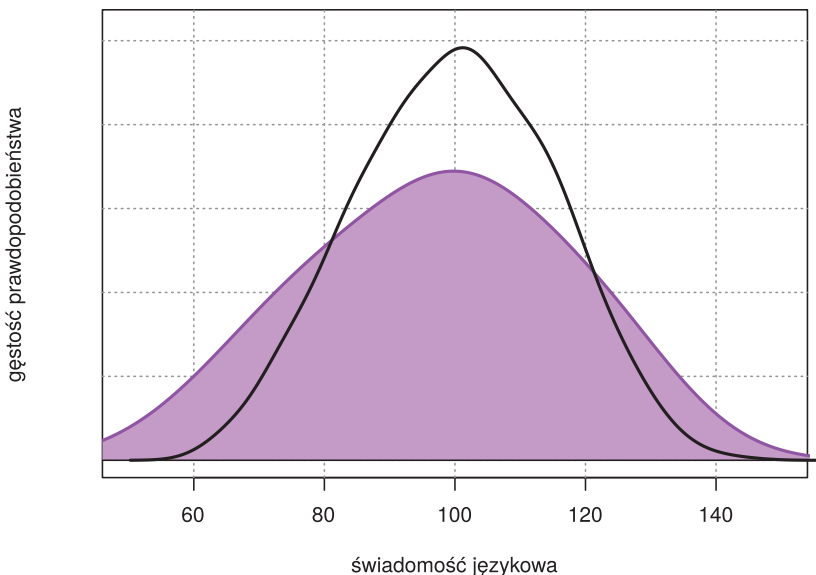
Wyniki testu osiągnięć

Liczba uczniów, których wyniki zostały wykorzystane przy wyznaczeniu wskaźnika dla szkoły: 17.

Uwaga! W analizach wykorzystano dane dla wszystkich uczniów, którzy rozwiązyali dany test. Jeśli jednak liczba ta jest mniejsza niż 10, prezentowane wyniki należy interpretować z dużą ostrożnością, mogą być one bowiem mało wiarygodne.

Na poniższym wykresie zaprezentowano rozkład wyników uczniów Państwa Szkoły na tle wyników w populacji. Czarną linią wyrysowano rozkład wyników w populacji uczniów. Jego średnia wynosi 100 punktów, a odchylenie standardowe 15 punktów. Fioletowym kolorem zaznaczono rozkład wyników uczniów Państwa Szkoły w teście świadomości językowej. Jeśli rozkład wyników w Państwa Szkole jest przesunięty w prawo w stosunku do rozkładu populacyjnego, oznacza to, że uczniowie uzyskali wyniki powyżej przeciętnych. Jeśli jest przesunięty w lewo – uzyskali wyniki poniżej średniej. Jeśli rozkład wyników w Państwa Szkole jest szczuplejszy i wyższy niż rozkład populacyjny, oznacza to, że wyniki uczniów w Państwa Szkole są mniej zróżnicowane niż średnio w populacji. Jeśli jest on szerszy i niższy, oznacza to, że uczniowie w Państwa Szkole uzyskali wyniki bardziej zróżnicowane niż średnio w populacji.

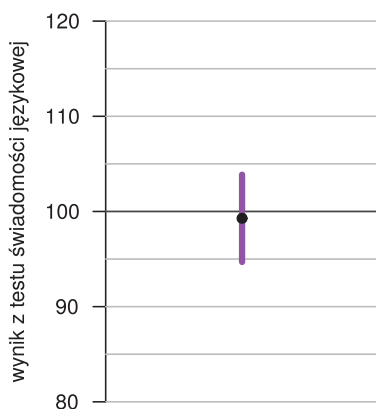
Rysunek 9. Rozkład wyników w teście świadomości językowej w szkole na tle rozkładu wyników w populacji.



Na kolejnym wykresie przedstawiono średni wynik szkoły w teście świadomości językowej wraz z przedziałem ufności. Oś pionowa to skala wyników w teście świadomości językowej. Średni wynik szkoły jest reprezentowany przez czarny punkt. Przedział ufności jest zaznaczony niebieskimi odcinkami. W przedziale tym z 90-procentowym prawdopodobieństwem znajduje się prawdziwy wynik Państwa Szkoły. Wyższy wynik wskazuje na wyższy poziom nauczania danej umiejętności.

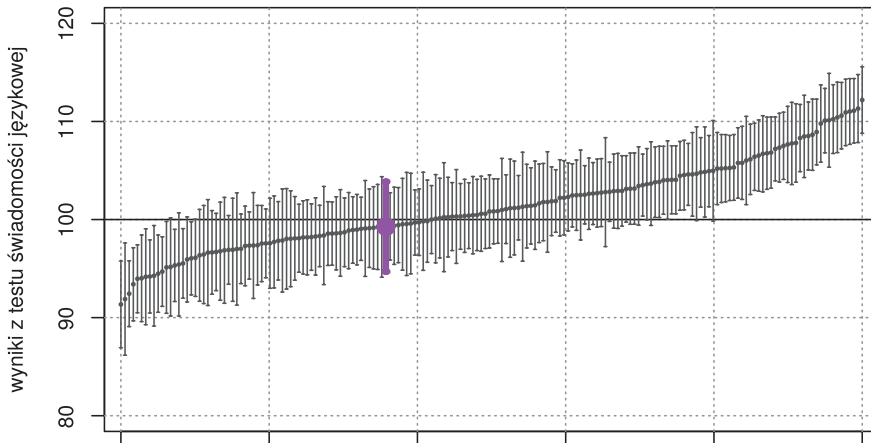
Średni wynik w populacji uczniów wynosi 100 punktów. Dlatego, jeśli cały przedział ufności znajduje się ponad wynikiem 100, uczniowie uzyskali w danym teście wynik wyższy niż średni wynik w populacji. Jeśli cały przedział ufności znajduje się poniżej wyniku 100, uczniowie mieli wynik w danym teście niższy niż średni wynik w populacji uczniów. Jeśli przedział ufności przecina 100 punktów, oznacza to, że uczniowie uzyskali wynik porównywalny ze średnim wynikiem w populacji uczniów.

Rysunek 10. Średni wynik szkoły w teście świadomości językowej wraz z przedziałem ufności.



Na kolejnym wykresie ten sam wynik przedstawiono na tle wyników wszystkich szkół objętych badaniem. Wynik Państwa Szkoły wraz z przedziałem ufności został pogrubiony i wyróżniony kolorem fioletowym. Wyniki wszystkich szkół zostały uporządkowane rosnąco według wartości średniej. Przypominamy, że o dwóch szkołach możemy powiedzieć, że różnią się średnimi wynikami, jeśli ich przedziały ufności nie pokrywają się.

Rysunek 11. Średnie wyniki szkoły w teście świadomości językowej na tle wyników wszystkich badanych szkół.



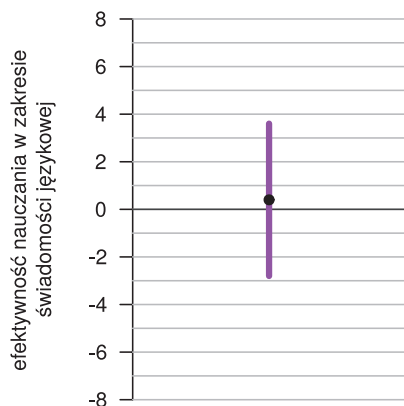
Efektywność nauczania

Liczba uczniów, których wyniki zostały wykorzystane przy wyznaczeniu wskaźnika dla szkoły: 15.

Uwaga! W analizach wykorzystano dane dla wszystkich uczniów, którzy rozwiązali dany test. Jeśli jednak liczba ta jest mniejsza niż 10, prezentowane wyniki należy interpretować z dużą ostrożnością, mogą być one bowiem mało wiarygodne.

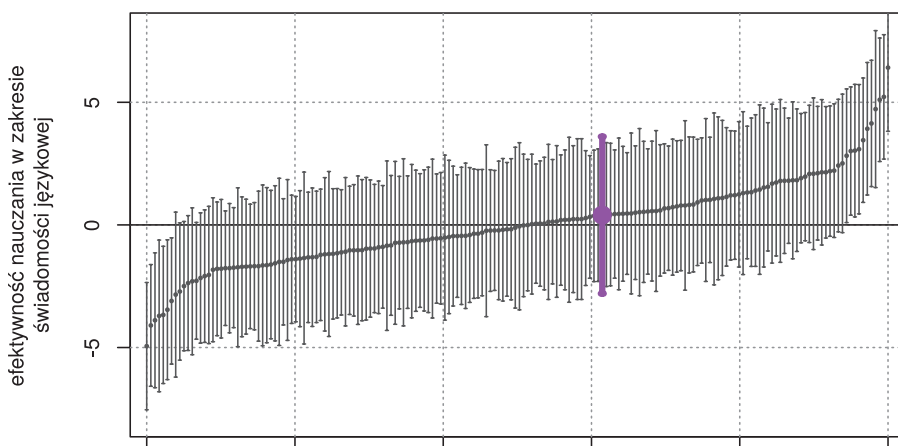
Na poniższym wykresie zaprezentowano wskaźnik efektywności nauczania w zakresie świadomości językowej (jego średnią oznaczono punktem) wraz z 90-procentowym przedziałem ufności (oznaczonym fioletowym odcinkami). Oś pionowa to skala efektywności nauczania. Im wyższy na niej wynik, tym wyższa efektywność nauczania. Wynik 0 oznacza średnią efektywność nauczania w populacji. Jeśli cały przedział ufności wskaźnika znajduje się powyżej punktu 0, oznacza to, że Państwa Szkoła naucza z powyżej przeciętną efektywnością. Jeśli przedział ufności znajduje się poniżej punktu 0, oznacza to, że Państwa Szkoła naucza z poniżej przeciętną efektywnością nauczania. Jeśli przedział ufności przecina wartość 0, oznacza to, że Państwa Szkoła pracuje z przeciętną efektywnością, czyli taką jak większość szkół w kraju.

Rysunek 12. Efektywność nauczania w zakresie świadomości językowej.



Na kolejnym wykresie efektywność nauczania w Państwa Szkole przedstawiono na tle wyników wszystkich szkół objętych badaniem. Wynik Państwa Szkoły wraz z przedziałem ufności został pogrubiony i wyróżniony kolorem fioletowym. Wyniki wszystkich szkół zostały uporządkowane rosnąco według wartości średniej. Wartość 0 oznacza średnią efektywność nauczania. Przypominamy, że o dwóch szkołach możemy powiedzieć, że różnią się efektywnością nauczania, jeśli ich przedziały ufności są rozłączne.

Rysunek 13. Efektywność nauczania w zakresie świadomości językowej na tle wszystkich badanych szkół.

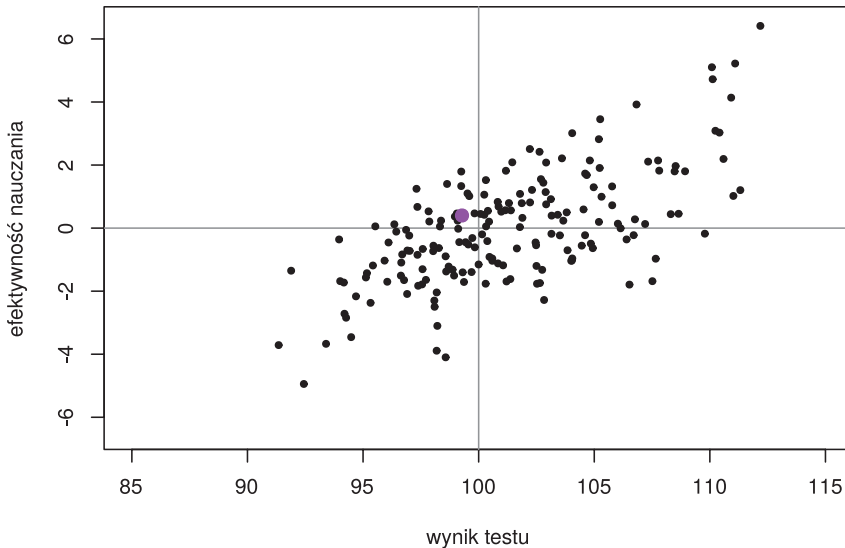


Łączna informacja o wyniku i efektywności nauczania

Na ostatnim wykresie przedstawiono równocześnie informację o średnim wyniku w teście i efektywności nauczania w Państwa Szkole na tle wyników wszystkich szkół biorących udział w badaniu. Dzięki takiemu zestawieniu informacji, mogą Państwo zobaczyć, gdzie ze względu na wartości obu wskaźników łącznie znajduje się Państwa Szkoła w porównaniu do innych szkół.

Na osi poziomej przedstawiono wyniki uzyskane w teście świadomości językowej. Średni wynik w populacji uczniów to 100 punktów, a odchylenie standardowe wynosi 15 punktów. Na osi pionowej przedstawiono wskaźnik efektywności nauczania. Punkt 0 oznacza średnią efektywność nauczania świadomości językowej w populacji szkół. Wynik Państwa Szkoły został wyróżniony pogrubieniem i kolorem fioletowym. Na wykresie przedstawiono wyniki bez przedziałów ufności, by zwiększyć jego czytelność. Trzeba jednak pamiętać, że są one obarczone niepewnością pomiarową, więc nie należy przeceniać małych różnic.

Rysunek 15. Łączna informacja o wyniku w teście osiągnięć i efektywności nauczania w zakresie świadomości językowej.



Wyniki z zakresu umiejętności matematycznych

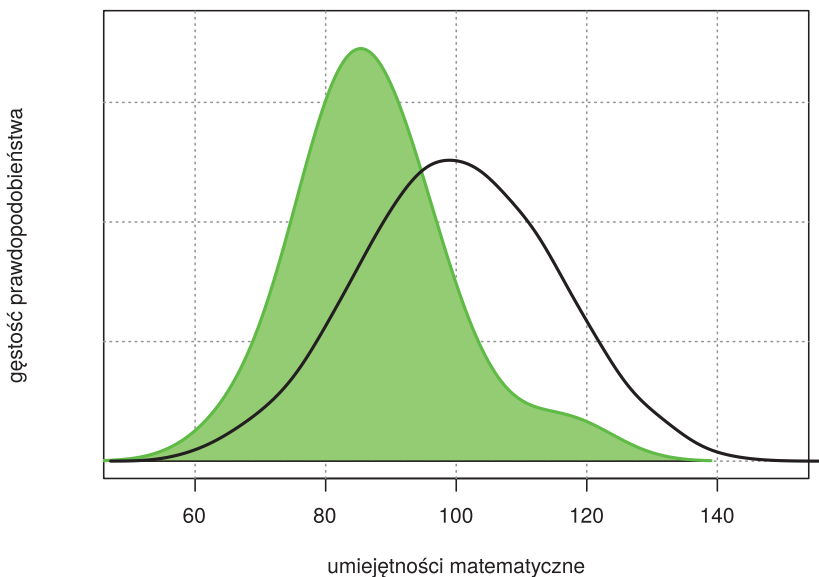
Wyniki testu osiągnięć

Liczba uczniów, których wyniki zostały wykorzystane przy wyznaczeniu wskaźnika dla szkoły: 16.

Uwaga! W analizach wykorzystano dane dla wszystkich uczniów, którzy rozwiązyali dany test. Jeśli jednak liczba ta jest mniejsza niż 10, prezentowane wyniki należy interpretować z dużą ostrożnością, mogą być one bowiem mało wiarygodne.

Na poniższym wykresie zaprezentowano rozkład wyników uczniów Państwa Szkoły na tle wyników w populacji. Czarną linią wyrysowano rozkład wyników w populacji uczniów. Jego średnia wynosi 100 punktów, a odchylenie standardowe 15 punktów. Zielonym kolorem zaznaczono rozkład wyników uczniów Państwa Szkoły w teście umiejętności matematycznych. Jeśli rozkład wyników w Państwa Szkole jest przesunięty w prawo w stosunku do rozkładu populacyjnego, oznacza to, że uczniowie uzyskali wyniki powyżej przeciętnych. Jeśli jest przesunięty w lewo – uzyskali wyniki poniżej średniej. Jeśli rozkład wyników w Państwa Szkole jest szczuplejszy i wyższy niż rozkład populacyjny, oznacza to, że wyniki uczniów w Państwa Szkole są mniej zróżnicowane niż średnio w populacji. Jeśli jest on szerszy i niższy, oznacza to, że uczniowie w Państwa Szkole uzyskali wyniki bardziej zróżnicowane niż średnio w populacji.

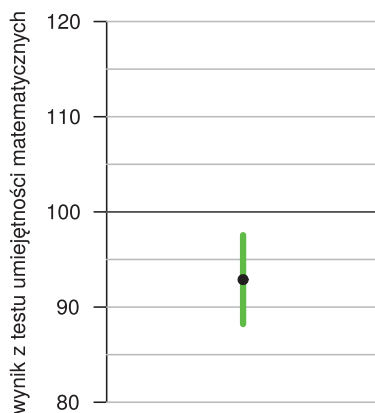
Rysunek 16. Rozkład wyników w teście umiejętności matematycznych w szkole na tle rozkładu wyników w populacji.



Na kolejnym wykresie przedstawiono średni wynik szkoły w teście umiejętności matematycznych wraz z przedziałem ufności. Oś pionowa to skala wyników w teście umiejętności matematycznych. Średni wynik szkoły jest reprezentowany przez czarny punkt. Przedział ufności jest zaznaczony niebieskimi odcinkami. W przedziale tym z 90-procentowym prawdopodobieństwem znajduje się prawdziwy wynik Państwa Szkoły. Wyższy wynik wskazuje na wyższy poziom nauczania danej umiejętności.

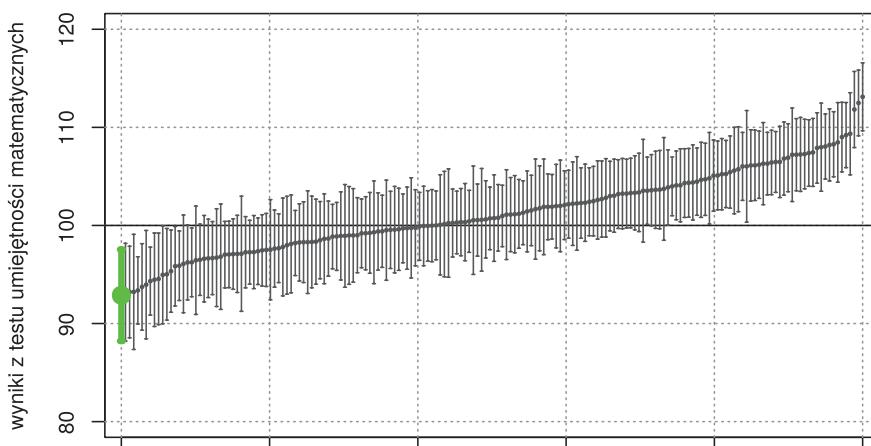
Średni wynik w populacji uczniów wynosi 100 punktów. Dlatego, jeśli cały przedział ufności znajduje się ponad wynikiem 100, uczniowie uzyskali w danym teście wynik wyższy niż średni wynik w populacji. Jeśli cały przedział ufności znajduje się poniżej wyniku 100, uczniowie mieli wynik w danym teście niższy niż średni wynik w populacji uczniów. Jeśli przedział ufności przecina 100 punktów, oznacza to, że uczniowie uzyskali wynik porównywalny ze średnim wynikiem w populacji uczniów.

Rysunek 17. Średni wynik szkoły w teście umiejętności matematycznych wraz z przedziałem ufności.



Na kolejnym wykresie ten sam wynik przedstawiono na tle wyników wszystkich szkół objętych badaniem. Wynik Państwa Szkoły wraz z przedziałem ufności został pogrubiony i wyróżniony kolorem zielonym. Wyniki wszystkich szkół zostały uporządkowane rosnąco według wartości średniej. Przypominamy, że o dwóch szkołach możemy powiedzieć, że różnią się średnimi wynikami, jeśli ich przedziały ufności nie pokrywają się.

Rysunek 18. Średnie wyniki szkoły w teście umiejętności matematycznych na tle wyników wszystkich badanych szkół.



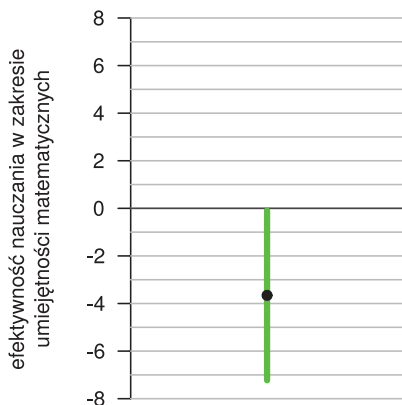
Efektywność nauczania

Liczba uczniów, których wyniki zostały wykorzystane przy wyznaczeniu wskaźnika dla szkoły: 14.

Uwaga! W analizach wykorzystano dane dla wszystkich uczniów, którzy rozwiązali dany test. Jeśli jednak liczba ta jest mniejsza niż 10, prezentowane wyniki należy interpretować z dużą ostrożnością, mogą być one bowiem mało wiarygodne.

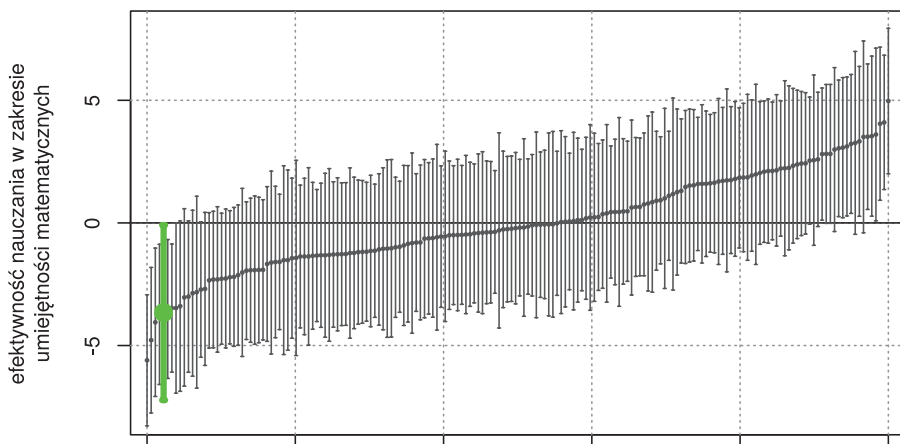
Na poniższym wykresie zaprezentowano wskaźnik efektywności nauczania w zakresie umiejętności matematycznych (jego średnią oznaczono punktem) wraz z 90-procentowym przedziałem ufności (oznaczonym zielonym odcinkami). Oś pionowa to skala efektywności nauczania. Im wyższy na niej wynik, tym wyższa efektywność nauczania. Wynik 0 oznacza średnią efektywność nauczania w populacji. Jeśli cały przedział ufności wskaźnika znajduje się powyżej punktu 0, oznacza to, że Państwa Szkoła naucza z powyżej przeciętną efektywnością. Jeśli przedział ufności znajduje się poniżej punktu 0, oznacza to, że Państwa Szkoła naucza z poniżej przeciętną efektywnością nauczania. Jeśli przedział ufności przecina wartość 0, oznacza to, że Państwa Szkoła pracuje z przeciętną efektywnością, czyli taką jak większość szkół w kraju.

Rysunek 19. Efektywność nauczania w zakresie umiejętności matematycznych.



Na kolejnym wykresie efektywność nauczania w Państwa Szkole przedstawiono na tle wyników wszystkich szkół objętych badaniem. Wynik Państwa Szkoły wraz z przedziałem ufności został pogrubiony i wyróżniony kolorem zielonym. Wyniki wszystkich szkół zostały uporządkowane rosnąco według wartości średniej. Wartość 0 oznacza średnią efektywność nauczania. Przypominamy, że o dwóch szkołach możemy powiedzieć, że różnią się efektywnością nauczania, jeśli ich przedziały ufności są rozłączne.

Rysunek 20. Efektywność nauczania w zakresie umiejętności matematycznych na tle wszystkich badanych szkół.



Łączna informacja o wyniku i efektywności nauczania

Na ostatnim wykresie przedstawiono równocześnie informację o średnim wyniku w teście i efektywności nauczania w Państwa Szkole na tle wyników wszystkich szkół biorących udział w badaniu. Dzięki takiemu zestawieniu informacji, mogą Państwo zobaczyć, gdzie ze względu na wartości obu wskaźników łącznie znajduje się Państwa Szkoła w porównaniu do innych szkół.

Na osi poziomej przedstawiono wyniki uzyskane w teście umiejętności matematycznych. Średni wynik w populacji uczniów to 100 punktów, a odchylenie standardowe wynosi 15 punktów. Na osi pionowej przedstawiono wskaźnik efektywności nauczania. Punkt 0 oznacza średnią efektywność nauczania umiejętności matematycznych w populacji szkół. Wynik Państwa Szkoły został wyróżniony pogrubieniem i kolorem zielonym. Na wykresie przedstawiono wyniki bez przedziałów ufności, by zwiększyć jego czytelność. Trzeba jednak pamiętać, że są one obarczone niepewnością pomiarową, więc nie należy przeceniać małych różnic.

Rysunek 22. Łączna informacja o wyniku w teście osiągnięć i efektywności nauczania w zakresie umiejętności matematycznych.

