



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

IBE  *entuzjaści
edukacji*

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



Analizy IBE/3/2015

Paweł Penszko

Piotr Zielonka

ANALIZA WPŁYWU PROGRAMU “CYFROWA SZKOŁA” NA WYNIKI SPRAWDZIANU SZÓSTOKLASISTY



Autorzy:
Paweł Penszko
Piotr Zielonka

Konsultacje merytoryczne:
dr Michał Sitek
dr Paweł Strawiński

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

© Copyright by: *Instytut Badań Edukacyjnych, Warszawa 2015*

Publikacja opracowana w ramach projektu systemowego: Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego współfinansowanego przez Unię Europejską ze środków Europejskiego Funduszu Społecznego, realizowanego przez Instytut Badań Edukacyjnych

Egzemplarz bezpłatny

Streszczenie

Rządowy program „Cyfrowa szkoła” został zrealizowany w roku szkolnym 2012/2013 w 399 szkołach podstawowych. Dzięki losowemu doborowi placówek do programu spośród ponad 3,5 tys. wnioskujących możliwe było skonstruowanie grupy kontrolnej, czego dokonano przez odpowiednie przeważenie grupy szkół nieskutecznie wnioskujących, i wiarygodne oszacowanie wpływu udziału szkół w programie na wyniki osiągnięte przez ich uczniów na sprawdzianie szóstoklasisty. W analizach wykorzystano wyniki sprawdzianów z trzech lat: 2012 (przed udziałem w programie), 2013 (ok. 3-5 miesięcy po dostarczeniu sprzętu do szkół) i 2014 (ok. 15-17 miesięcy po dostarczeniu sprzętu do szkół). Badano nie tylko wpływ programu na średni wynik uczniów, ale również na kształt rozkładu wyników w poszczególnych częściach sprawdzianu. Zastosowano testy t i symulacyjne testy Monte Carlo do rozstrzygnięcia, czy obserwowane różnice między grupą zasadniczą a grupą kontrolną można uznać za skutek realizacji „Cyfrowej szkoły”, czy też mogą one być przypadkowe, powstałe w trakcie losowego doboru szkół. Wyniki analiz skłaniają do wniosku, że program przyniósł efekt w postaci polepszenia rezultatów niektórych słabszych uczniów w części „rozumowanie” sprawdzianu z 2013 roku (część ta zawiera zadania o charakterze matematycznym i logicznym). Wpływ ten był jednak niewielki (przełożył się na podwyższenie średniego wyniku o zaledwie 3,5% odchylenia standardowego) ograniczony tylko do jednej części sprawdzianu i krótkotrwały – w wynikach sprawdzianu z 2014 roku nie był już widoczny. Nie zaobserwowano innych przejawów oddziaływania programu na wyniki sprawdzianu szóstoklasisty. Kontrfaktyczne ewaluacje podobnych zagranicznych programów sugerują, że interwencje publiczne tego typu nie przynoszą wyraźnej poprawy wyników egzaminacyjnych w ciągu kilkunastu miesięcy. Zaprezentowane w niniejszym opracowaniu analizy dotyczą jedynie wpływu programu „Cyfrowa szkoła” na kompetencje i umiejętności mierzone przez sprawdzian szóstoklasisty i pozostawiają otwartą kwestię, czy program podwyższył inne kompetencje uczniów (na przykład kompetencje cyfrowe lub umiejętność pracy w grupie).

Summary

The 'Digital School' was a Polish government pilot study involving 399 primary schools, conducted in academic year 2012/2013. It might be regarded as a 'one laptop per child' initiative. More than 3,500 schools applied to join the project and the schools which finally participated in the project were selected on a random basis. The population of pupils from unsuccessfully applying schools was weighted, so as to construct a valid control group and estimate the impact of the 'Digital School' project on final exam results. The results of three exams were included: the first in April 2012 (from before the project began), the second in April 2013 (three to five months after the ICT equipment was delivered), and the final in April 2014 (more than one year after the delivery of ICT equipment). The impact of the intervention on exam score distribution was examined, beyond the usual comparison of means. T-tests and Monte Carlo permutation tests were applied to determine whether the conclusion that the project intervention affected the score distribution could be legitimately drawn or the differences could be purely chance, that is to say, stemming from the random school selection. The results suggest that some low-scoring pupils obtained better scores in the 2013 'reasoning' section of the exam owing to the intervention (this section contains mathematical and logic tasks). However, the effect was weak (mean exam score is estimated to have increased by only 3.5% of standard deviation), limited to one exam section and short term – it was not visible anymore in 2014. No other effects were observed. Counterfactual impact evaluations of similar initiatives from other countries suggest that such public interventions do not significantly affect exam results within 15-18 months. The analyses presented in this paper only regard competencies and skills measured by the final exam and do not answer the question whether other competencies or skills (for example digital competencies or group work skills) were affected by the 'Digital School' project.

Spis Treści

1. Wprowadzenie	6
2. Rządowy program “Cyfrowa szkoła”	8
3. Koncepcja i metodologia analizy	10
3.1. Cel i zakres analizy	10
3.2. Wykorzystane dane.....	11
3.3. Konstrukcja grupy kontrolnej.....	12
3.4. Testy istotności statystycznej.....	17
4. Wyniki analiz.....	19
5. Wnioski	25
6. Literatura cytowana.....	26

1. Wprowadzenie

W ostatniej dekadzie w wielu państwach świata rozpoczęto realizację programów znanych jako „programy 1:1” albo „programy OLPC (*one laptop per child*)”. Ich podstawowym elementem jest przekazanie mobilnego komputera nauczycielowi oraz uczniom (przy zachowaniu zasady, że na jednego ucznia przypada jeden komputer) w celu ich wykorzystania do nauki szkolnej. Przy czym nie chodzi tu o naukę informatyki lub obsługi komputera, lecz o użytkowanie nowoczesnych technologii jako pomocy dydaktycznych w nauczaniu wszystkich przedmiotów szkolnych. W uzasadnieniach celowości finansowania tego rodzaju programów ze środków publicznych pojawiają się trzy typowe argumenty, odnoszące się do trzech oczekiwanych efektów projektu: nabycia przez uczniów umiejętności przydatnych na rynku pracy, przeciwdziałania wykluczeniu cyfrowemu oraz poprawie jakości kształcenia (Valiente, 2010; Severin & Capota, 2011). Wskazuje się też na potencjał tkwiący w nowych technologiach, np. w zakresie indywidualizacji i personalizacji nauczania czy aktywizacji uczniów (Balanskat, Blamire, & Kefala, 2006; Condie & Munro, 2007; The NMC Horizon Report Europe: 2014 Schools Edition, 2014), a także na rozdźwięk między szkołą, bazującą na tradycyjnych pomocach dydaktycznych, a życiem codziennym uczniów, w którym urządzenia elektroniczne i komputerowe są wszechobecne (Keane, Keane, & Blicblau, 2013).

Kierując się wymienionymi względami, w latach 2006–2012 w krajach Ameryki Łacińskiej i rejonu Karaibów uruchomiono co najmniej 18 programów 1:1, w ramach których rozdzielono około 2 milionów laptopów (Severin & Capota, 2011). Światowym liderem w tej dziedzinie okazał się Urugwaj, w którym podjęto decyzję o zapewnieniu komputerów mobilnych wszystkim uczniom publicznych szkół podstawowych i gimnazjów (de Melo, Machado, & Miranda, 2014). Programy 1:1 realizowane były również w Ameryce Północnej, Afryce, Azji i Australii (White, 2008; Severin & Capota, 2011). W Europie w 2012 roku realizowano ponad 19 interwencji publicznych tego typu (Bocconi, Kampylis, & Punie, 2013), w tym polski program „Cyfrowa szkoła”, który jest przedmiotem niniejszego opracowania.

Idea programów 1:1 znalazła więc szerokie poparcie wśród politycznych decydentów. Badanie naukowe nie świadczy jednak tak jednoznacznie na jej korzyść. Samo zastosowanie technologii informacyjno-komunikacyjnych (TIK) nie przynosi automatycznie poprawy osiągnięć edukacyjnych uczniów. Co prawda opublikowana w 2011 roku metaanaliza podsumowująca wyniki ponad 1000 badań przyniosła ustalenie, że średni efekt wykorzystania TIK w tej sferze jest pozytywny, a jego wartość przekracza 0,3 (Tamim, Bernard, Borkohovski, Abrami, & Schmid, 2011), ale w niektórych wcześniejszych metaanalizach nie stwierdzono, by efekt TIK był istotnie różny od zera (Koufogiannakis & Wiebe, 2006). Ponadto wyniki analiz przeprowadzanych na danych z międzynarodowego testu kompetencji PISA sugerują, że intensywne wykorzystanie TIK w szkole jest negatywnie skorelowane z wynikami uczniów lub że zależność między tymi dwoma czynnikami ma kształt odwróconej litery U (Fuchs & Wößmann, 2005; Gil Flores, 2012; Biagi & Loi, 2013; Federowicz, 2014). Negatywny wpływ na oceny szkolne zaobserwowano w przypadku realizowanego w Rumunii programu dotacji na zakup komputerów domowych w ubogich gospodarstwach domowych – zakupione urządzenia były zresztą, wbrew intencjom twórców programu, stosunkowo rzadko używane do celów edukacyjnych (Malamud & Pop-Eleches, 2011). W innych, eksperymentalnych badaniach dotyczących zapewniania uczniom komputerów do domowego użytku nie odnotowano żadnego oddziaływania na osiągnięcia edukacyjne lub oceny uczniów przez nauczycieli (Beuermann, Cristia, Cruz-Aguayo, Cueto, & Malamud, 2012; Fairlie & Robinson, 2013). Jeśli chodzi o programy 1:1 w ścisłym tego słowa znaczeniu, to w opublikowanej już w 2006 roku syntezie wyników badań

na ten temat William Penuel zauważył, że niektórzy z uczonych donoszą o pozytywnym oddziaływaniu tego rodzaju interwencji na kompetencje cyfrowe i umiejętności tworzenia tekstów pisemnych, ale wnioski dotyczące tego drugiego obszaru wydają się słabiej ugruntowane ze względu na brak badań mierzących te umiejętności nie tylko po, ale również przed udziałem w projekcie (Penuel, 2006). W ostatnich latach pojawiły się dwa opracowania wolne od zauważonych przez Penuela niedostatków dzięki zastosowaniu rygorystycznej metodologii kontrfaktycznej, eksperymentalnej lub quasi-eksperymentalnej. W żadnym z nich nie stwierdzono wpływu programu 1:1 na wyniki wypełnianych przez uczniów testów z matematyki i języka nauczania (Cristia, Ibararán, Cueto, Santiago, & Severín, 2012; de Melo, Machado, & Miranda, 2014).

Przegląd literatury skłania do wniosku, że w technologiach informacyjno-komunikacyjnych tkwi potencjał, który może zostać wykorzystany do poprawy jakości nauczania. Powszechnie podkreśla się, że sama obecność technologii informacyjno-komunikacyjnych nie prowadzi do zmiany technik nauczania i lepszych osiągnięć edukacyjnych. Technologie te są jedynie narzędziem, które może przynieść pozytywne efekty tylko pod warunkiem odpowiedniego podejścia i umiejętnego wykorzystania go przez nauczycieli (Valiente, 2010; Hennessy & London, 2013). Realizowane programy 1:1 często nie wywierały wpływu na umiejętności uczniów mierzone wystandaryzowanymi testami lub wynikami egzaminów zewnętrznych. Ponieważ wpływ ten może zależeć od kształtu interwencji publicznej oraz od środowiska, w jakim jest wdrażana, każdy program 1:1 wymaga jednak odrębnej ewaluacji oddziaływania. Tego rodzaju analiza dotycząca polskiego programu „Cyfrowa szkoła” stanowi właśnie temat niniejszego opracowania.

2. Rządowy program „Cyfrowa szkoła”

Rządowy program rozwijania kompetencji uczniów i nauczycieli w zakresie stosowania technologii informacyjno-komunikacyjnych – „Cyfrowa szkoła” został wprowadzony rozporządzeniem Rady Ministrów z dnia 3 kwietnia 2012 roku. Zakładał on interwencję publiczną w czterech obszarach:

- 1) Obszar „e-nauczyciel” (współfinansowany z Europejskiego Funduszu Społecznego w ramach Programu Operacyjnego Kapitał Ludzki) obejmował rozwój umiejętności nauczycieli w zakresie wykorzystania TIK w nauczaniu, komunikacji z uczniami i rodzicami oraz prowadzeniu dokumentacji szkolnej. Środkiem do nabywania tych umiejętności były powołane międzyszkolne sieci współpracy nauczycieli, grupujące nauczycieli z kilku relatywnie bliskich geograficznie szkół. Członkowie sieci brali udział w szkoleniu e-learningowym oraz mieli okazję do wymiany wiedzy i doświadczeń na czterech osobistych spotkaniach sieci. W ramach sieci prowadzone były również lekcje otwarte, udostępniane scenariusze lekcji i wykonywane wraz z uczniami projekty edukacyjne. Kluczową rolę w transmisji kompetencji odgrywał szkolny e-koordynator, który brał udział w szkoleniach, a następnie miał przekazywać nabytą wiedzę i umiejętności innym nauczycielom w swojej szkole.
- 2) Obszar „e-zasoby edukacyjne” (współfinansowany z Europejskiego Funduszu Społecznego w ramach Programu Operacyjnego Kapitał Ludzki) obejmował tworzenie elektronicznych materiałów edukacyjnych zamieszczanych na portalu Scholaris, audycji oświatowych publikowanych przez Telewizję Polską oraz ogólnie dostępnych e-podręczników.
- 3) Obszar „e-szkoła” (finansowany z budżetu krajowego przy wkładzie własnym organów prowadzących szkoły) obejmował dotacje celowe dla szkół na zakup sprzętu informacyjno-komunikacyjnego: przenośnych komputerów dla nauczycieli, tablic interaktywnych, projektorów, wizualizerów, urządzeń i oprogramowania do przeprowadzania wideokonferencji, systemów do zbierania i analizowania odpowiedzi, ruterów, elementów sieci bezprzewodowej, drukarek i innych urządzeń peryferyjnych.
- 4) Obszar „e-uczeń” (finansowany z budżetu krajowego przy wkładzie własnym organów prowadzących szkoły) obejmował dotacje celowe dla szkół na zakup komputerów mobilnych dla uczniów.

Szczegółowa lista nabywanych urządzeń była określana przez uczestniczącą w programie szkołę, jednak musiała ona w jego ramach stworzyć co najmniej jeden zestaw składający się z przenośnego komputera dla nauczyciela i przenośnych komputerów dla uczniów (w liczbie nie mniejszej niż liczba uczniów), szafki oraz rutera zapewniającego bezpieczne połączenie z Internetem. „Cyfrową szkołę” można zatem uznać za program 1:1, w którym zakup sprzętu został połączony z doskonaleniem zawodowym nauczycieli oraz opracowywaniem elektronicznych zasobów edukacyjnych.

Program był skierowany do uczniów klas IV-VI publicznych szkół podstawowych¹, a więc jego realizacja przebiegała na II etapie edukacyjnym. W rozporządzeniu przewidywano jednak wykorzystanie zakupionego sprzętu w miarę możliwości również przez uczniów klas I-III. Program był

¹ Teoretycznie udział w programie mogły również brać udział ogólnokształcące szkoły muzyczne, jednak w praktyce żadna taka szkoła nie została zakwalifikowana do programu.

realizowany w dwóch wariantach. Wariant I zakładał, że zakupione przenośne komputery dla uczniów będą wykorzystywane wyłącznie w szkole na zajęciach lekcyjnych i w czasie wolnym od zajęć dydaktycznych. Wariant II zakładał, że przenośne komputery będą dodatkowo udostępniane uczniom klasy IV do korzystania w domu. W praktyce na realizację wariantu II przystało tylko 20% szkół biorących udział w programie, a wśród nich wiele obwarowało możliwość wypożyczenia sprzętu poza szkołę kłopotliwymi formalnościami i zastrzeżeniami. W rezultacie do faktycznego korzystania przez uczniów z zakupionych komputerów przenośnych w domu doszło w przypadku niecałych 15% uczestniczących placówek.

Program „Cyfrowa szkoła” był w założeniu pilotażem przed analogiczną interwencją publiczną przeprowadzaną na znacznie szerszą skalę. Dlatego jego celem było przetestowanie przyjętego modelu interwencji, sprawdzenie możliwości absorpcji środków przez szkoły i identyfikacja ewentualnych barier uczestnictwa. Liczono również na to, że realizacja programu pozwoli na wypracowanie dobrych praktyk, które będą mogły zostać upowszechnione. Same działania prowadzone w ramach interwencji miały skutkować podniesieniem umiejętności nauczycieli i uczniów w zakresie stosowania technologii informacyjno-komunikacyjnych w edukacji oraz rzeczywistym wdrożeniem tych technologii na lekcjach poszczególnych przedmiotów szkolnych. Miało to wspierać zmianę modelu uczenia w kierunku rozwijania kreatywności, umiejętności kooperacji oraz krytycznego myślenia, w tym wyszukiwania, oceny i twórczego wykorzystywania dostępnych źródeł wiedzy. Oczekiwano, że technologie informacyjno-komunikacyjne pomogą w upowszechnieniu indywidualizacji kształcenia i położeniu większego nacisku na pracę grupową, a ponadto będą sprzyjały zwiększeniu motywacji i zaangażowania uczniów. Zakładano, że wszystkie te zmiany przełożą się na podwyższenie jakości i skuteczności edukacji. Końcowym, pośrednim efektem miało być w związku z tym podniesienie kompetencji podstawowych, społecznych i twórczych uczniów.

Program został zrealizowany w roku szkolnym 2012/2013. Organy prowadzące składały wnioski o udział szkół w programie w kwietniu 2012 roku. Na maj zaplanowano wybór szkół uczestniczących w programie. Spośród ponad 3,5 tys. zgłoszonych szkół do programu zakwalifikowano 402 (z których 399 faktycznie wzięło udział). O udziale w programie decydowało losowanie, przy czym pula środków finansowych dostępnych dla szkół z danego województwa, o danej wielkości i realizujących dany wariant programu była wcześniej ustalana według specjalnego algorytmu. Na wydatkowanie kwoty dotacji szkoły miały czas do końca roku kalendarzowego 2012. W praktyce więc zakupiony sprzęt najczęściej trafiał do szkół w grudniu 2012 roku lub w styczniu 2013 roku. Nieco opóźniona była realizacja działań w obszarze „e-nauczyciel” - pierwsze osobiste spotkania w ramach sieci współpracy odbywały się z reguły dopiero w kwietniu 2013 roku. W czerwcu szkoły składały sprawozdania z realizacji programu.

Oprócz wymienionych dotychczas komponentów, w programie zaplanowano również tzw. komponent badawczy oraz ewaluację. Komponent badawczy, realizowany przez Ministerstwo Administracji i Cyfryzacji, miał na celu przetestowanie innowacyjnych modeli wykorzystywania TIK w edukacji. Wzięły w nim udział 33 szkoły podstawowe, nie biorące udziału w zasadniczej części programu „Cyfrowa szkoła”, w tym 9 szkół w charakterze próby kontrolnej. Z kolei ewaluacja służyła zbadaniu sposobów i efektywności wykorzystania TIK w 399 uczestniczących szkołach. Została ona powierzona Instytutowi Badań Edukacyjnych. Raport z ewaluacji ex-post programu został opublikowany w 2013 roku na stronie internetowej instytutu (Instytut Badań Edukacyjnych, 2013). Znajduje się w nim między innymi pierwsza wersja analizy wpływu udziału placówki w programie „Cyfrowa szkoła” na wyniki osiągnięte przez jej uczniów na sprawdzianie szóstoklasisty. Niniejsze opracowanie stanowi rozszerzenie tej analizy i uzupełnienie jej dzięki wykorzystaniu wyników sprawdzianu szóstoklasisty z 2014 roku.

3. Koncepcja i metodologia analizy

3.1. Cel i zakres analizy

Celem analizy przedstawionej w niniejszym opracowaniu było ustalenie, czy i jak udział w programie „Cyfrowa szkoła” wpływa na wyniki osiągnięte przez jej uczniów w sprawdziane szóstoklasisty. Jak wspomniano w poprzednim rozdziale, końcowym efektem interwencji miało być między innymi podniesienie kompetencji podstawowych. Zdefiniowany w programie wskaźnik doprecyzowywał, że pomiarowi mają podlegać kompetencje w zakresie pisania, czytania i liczenia. Można zakładać, że podniesienie takich kompetencji powinno znaleźć odzwierciedlenie w wynikach sprawdzianu szóstoklasisty. Analiza wpływu uczestnictwa w „Cyfrowej szkole” na te wyniki jest więc sposobem ustalenia, czy program przyniósł oczekiwane efekty.

Trzeba jednocześnie podkreślić, że analiza ta dotyczy jedynie części tych efektów. Całkowicie poza jej zasięgiem leżą kwestie kompetencji nauczycieli i stosowanych przez nich technik dydaktycznych. Sprawdzian szóstoklasisty z założenia nie mierzy również kompetencji cyfrowych, informacyjnych i społecznych uczniów, a element kompetencji twórczych jest w zasadzie zawężony do umiejętności tworzenia krótkich wypowiedzi pisemnych zgodnie z podanym w arkuszu poleceniem. Jeśli chodzi o kompetencje podstawowe, są one uwzględnione w analizie tylko w takim zakresie i stopniu, w jakim są mierzone przez sprawdzian szóstoklasisty. W związku z tym nie należy traktować wyników niniejszych analiz jako rozstrzygających o sukcesie lub niepowodzeniu programu. Pozwalają one zobaczyć jedynie fragment całego obrazu efektów interwencji – ważny i dobrze poddający się mierzeniu, ale niewielki. Rezultaty badań dotyczących innych aspektów oddziaływania „Cyfrowej szkoły” znajdują się w raporcie z ewaluacji ex-post (Instytut Badań Edukacyjnych, 2013) oraz w późniejszych opracowaniach.

Warto zauważyć, że analiza dotyczy wpływu udziału szkoły w programie, a nie wpływu objęcia programem konkretnego ucznia. Rozróżnienie to ma pewne znaczenie, ponieważ w 7 spośród wszystkich 399 szkół biorących udział w programie nie wykorzystywano sprzętu zakupionego w ramach „Cyfrowej szkoły” w klasie VI. Dodatkowo w przypadku 26 nie udało się potwierdzić, czy sprzęt jest używany przez uczniów klasy VI. Spośród pozostałych 366 szkół w 353 sprzęt był wykorzystywany we wszystkich oddziałach klasy VI, a w 13 tylko w niektórych oddziałach. Ze względu na trudność w pełni dokładnego ustalenia, które szkoły stosowały sprzęt w klasie VI (26 braków danych), zdecydowano, że pomiar będzie dotyczył wpływu na wynik sprawdzianu udziału w programie „Cyfrowa szkoła”, a nie wykorzystania sprzętu w zajęciach z uczniami. Oznacza to, że w pewnym sensie analiza dotyczyła bardziej wpływu zamiaru udzielenia wsparcia przez włączenie do programu (*intention to treat*) niż samego wsparcia w postaci udostępnienia sprzętu uczniom (*treatment*). Nie wykluczano zatem z analiz uczniów szkół, którzy nie mieli okazji wykorzystać sprzętu zakupionego w ramach „Cyfrowej szkoły” na zajęciach. Może to „rozmywać”, czyli prowadzić do zaniżenia szacowanego efektu interwencji, ale w niewielkim stopniu, ponieważ przeszło 90% szkół uczestniczących w programie wykorzystywało zakupiony sprzęt w klasie VI. Trzeba też zauważyć, że nawet jeśli uczniowie klasy VI nie używali na zajęciach sprzętu nabytego w ramach „Cyfrowej szkoły”, to udział szkoły w programie mógł na nich w inny sposób wywierać wpływ (np. w związku z lepszą dostępnością technologii informacyjno-komunikacyjnych w świetlicy lub większą aktywnością i kreatywnością nauczycieli).

Do oszacowania wpływu programu „Cyfrowa szkoła” na mierzone kompetencje uczniów wykorzystano wyniki sprawdzianu szóstoklasisty z trzech lat: 2012, 2013 i 2014. Wyniki z roku 2012, czyli przed

wdrożeniem programu, posłużyły do sprawdzenia, czy grupa kontrolna jest dobrze dopasowana do grupy szkół uczestniczących w programie (przed programem nie powinny występować różnice w wynikach sprawdzianu między tymi dwiema grupami). Wyniki z roku 2013 pozwoliły ustalić, czy wystąpiły jakieś efekty po kilku miesiącach realizacji programu (około 3 do 5 miesięcy po dostawie sprzętu), a wyniki z roku 2014 – po kilkunastu miesiącach od jego wdrożenia. Dzięki takiemu schematowi badania możliwa była identyfikacja oddziaływania krótko- i średniofalowego, jak również efektów odroczonego.

Większość tego rodzaju analiz naukowych ogranicza się do zbadania wpływu danej interwencji publicznej na średni wynik uczniów w teście lub egzaminie. Ponieważ jednak jednym z uzasadnień realizacji programów 1:1 jest przeciwdziałanie wykluczeniu cyfrowemu, a ponadto wcześniejsze badania sugerują, że technologie informacyjno-komunikacyjne mogą mieć większy wpływ na umiejętności słabszych i średnich uczniów niż na umiejętności uczniów lepszych (Cheung & Slavin, 2012), postanowiono przyrzeć się kwestii oddziaływania programu „Cyfrowa szkoła” na cały rozkład wyników sprawdzianu szóstoklasisty. Pozwoliło to sprawdzić, czy wpływ programu jest silniejszy albo występuje tylko w przypadku uczniów osiągających gorsze wyniki na sprawdzianie.

3.2. Wykorzystane dane

W charakterze źródła danych zdecydowano się wykorzystać wyniki sprawdzianu szóstoklasisty. Sprawdzian ten jest obowiązkowy i przeprowadzany w klasach szóstych na początku kwietnia. Każdego roku możliwa do uzyskania przez ucznia liczba punktów mieści się w przedziale od 0 do 40. Przy ocenie wypełnionego arkusza testu wyróżnia się pięć części: czytanie, pisanie, korzystanie z informacji, rozumowanie i wykorzystanie wiedzy w praktyce. Liczba punktów do uzyskania w każdej z części jest również taka sama w poszczególnych latach.

Aby uzyskać dokładniejszy obraz wpływu „Cyfrowej szkoły”, analizy przeprowadzono odrębnie dla każdej z części sprawdzianu. Ponadto zaprezentowano wyniki analiz dla całego sprawdzianu traktowanego łącznie. Szczegółowe informacje na temat parametrów rozkładu analizowanych zmiennych zawiera Tabela 1.

W przedstawionych analizach wykorzystano surowy wynik sprawdzianu, czyli liczbę punktów, jaką dany uczeń zdobył zgodnie z arkuszem oceny. Przeprowadzono również analizy na wynikach wyskalowanych za pomocą metody IRT (dwuparametrycznego modelu Rascha), jednak nie zostały one zaprezentowane w niniejszym opracowaniu, ponieważ nie doprowadziły do sformułowania żadnych nowych wniosków, a ich przedstawienie wymagałoby dodatkowych wyjaśnień.

Należy zastrzec, że w ramach niniejszego opracowania nie mieści się analiza rzetelności i trafności sprawdzianu szóstoklasisty jako narzędzia pomiaru kompetencji. Dlatego przyjęto sprawdzian w takiej postaci, w jakiej został on zrealizowany, i wykorzystano wszystkie pytania, które w założeniach twórców arkusza egzaminacyjnego mierzyły daną kompetencję. Nie poruszano kwestii, jakie dokładnie kompetencje mierzy ten sprawdzian i na ile rzetelnie. Warto zauważyć, że im mniej rzetelne jest narzędzie zastosowane przy badaniu kompetencji, tym trudniej zaobserwować istotny wpływ programu na kompetencje, nawet jeśli wpływ ten występuje.

Jak wspomniano wyżej, wykorzystano wyniki sprawdzianów szóstoklasisty z trzech lat: 2012, 2013 i 2014. Dane te połączono z informacją o udziale szkoły w programie i preferowanym przez nią wariantcie programu, a także z innymi danymi charakteryzującymi szkołę i jej otoczenie, pozyskanymi z Systemu Informacji Oświatowej i ogólnie dostępnych zasobów Głównego Urzędu Statystycznego.

Tabela 1. Parametry rozkładu wyników sprawdzianu w całej populacji zdających w bazie danych

zmienna	parametr rozkładu					
	średnia	odch. std.	minimum	mediana	maksimum	
2012	czytanie	6,245	2,075	0	6	10
	pisanie	5,883	2,139	0	6	10
	korzystanie z informacji	2,795	1,094	0	3	4
	rozumowanie	4,111	1,962	0	4	8
	wykorzystanie wiedzy	3,719	2,662	0	4	8
	cały sprawdzian	22,364	7,448	0	22	39
2013	czytanie	7,303	2,239	0	8	10
	pisanie	6,336	2,335	0	7	10
	korzystanie z informacji	2,477	1,151	0	3	4
	rozumowanie	4,165	2,286	0	4	8
	wykorzystanie wiedzy	3,749	2,780	0	3	8
	cały sprawdzian	24,028	8,381	0	24	40
2014	czytanie	7,761	2,096	0	8	10
	pisanie	5,809	2,665	0	6	10
	korzystanie z informacji	2,731	1,056	0	3	4
	rozumowanie	4,549	2,385	0	5	8
	wykorzystanie wiedzy	4,977	2,155	0	5	8
	cały sprawdzian	25,826	8,006	0	27	40

3.3. Konstrukcja grupy kontrolnej

O udziale szkół w programie „Cyfrowa szkoła” decydowało losowanie. Ta okoliczność jest bardzo korzystna z punktu widzenia oceny efektów programu, ponieważ stwarza sytuację zbliżoną do eksperymentalnej. Grupa szkół, które składały wniosek o udział w programie, lecz nie zostały wylosowane („nieskutecznie aplikujące”), stanowi najlepsze źródło do skonstruowania grupy kontrolnej, która dostarcza informacji o tzw. stanie kontrfaktycznym, czyli o tym, jakie byłyby wyniki pomiaru efektów w grupie zasadniczej składającej się ze szkół uczestniczących w programie, gdyby nie realizowano programu. Przy doborze losowym jedyna systematyczna różnica między podmiotami wylosowanymi i niewylosowanymi do programu polega bowiem na udziale w programie, a reszta różnic ma charakter losowy, czyli dobrze poddający się kontroli za pomocą metod statystycznych. Dzięki temu po przyjęciu założeń podejścia kontrfaktycznego różnica między grupą zasadniczą a kontrolną jest podstawą do nieobciążonego oszacowania wpływu udziału szkoły w programie na wyniki uczniów (Angrist & Pischke, 2008; Murnane & Willet, 2011; Gertler, Martinez, Premand, Rawlings, & Vermeersch, 2011).

W prezentowanej tu analizie oszacowano rzeczywisty wpływ udziału w programie na wyniki uczniów ze szkół uczestniczących (tzw. *treatment effect on the treated*). Ponieważ grupa szkół

uczestniczących w programie była losowo dobrana spośród szkół aplikujących, możliwe byłoby również oszacowanie oczekiwanego wpływu na ogół szkół zgłaszających się do programu, gdyby wszystkie te szkoły zostały do niego włączone. Natomiast nie jest uprawnione wnioskowanie na tej podstawie o potencjalnym wpływie programu na ogół szkół podstawowych w Polsce, ponieważ szkoły aplikujące nie są losową próbą ogółu szkół w Polsce. Takie wnioskowanie byłoby jednak przydatne tylko wówczas, gdyby podobna interwencja miała objąć obowiązkowo wszystkie szkoły podstawowe w kraju. Tak długo, jak przyszłe interwencje tego typu będą miały charakter dobrowolny, a warunkiem udziału będzie przesłanie zgłoszenia przez organ prowadzący lub samą placówkę edukacyjną, istotne będzie oczekiwane oddziaływanie na szkoły, które chcą uczestniczyć w programie tego typu. Przeprowadzone analizy dotyczą właśnie tak określonego oddziaływania.

Kilka czynników komplikowało schemat badania. Poniżej omówiono kolejno te czynniki wraz z opisem kroków podjętych w reakcji na nie.

1. W programie nie zastosowano doboru losowego prostego, lecz dobór losowy warstwowy, w ramach warstw wyodrębnionych ze względu na województwo, preferowany wariant programu i kategorię wielkości szkoły (kategorie: „mała” licząca do 100 uczniów w klasach I-VI, „średnia” 101-300 uczniów, „duża” powyżej 300 uczniów). W rezultacie, jak pokazuje Tabela 2, struktura grupy szkół wylosowanych do programu pod względem tych trzech zmiennych różniła się od struktury szkół aplikujących, lecz niewylosowanych.

W celu porównania średnich i rozkładów między grupą zasadniczą a kontrolną przeważono grupę kontrolną w taki sposób, aby suma wag uczniów z grupy kontrolnej dla każdej kombinacji województwa, wariantu i wielkości szkoły była taka sama, jak suma wag uczniów szkół uczestniczących w programie o tej samej kombinacji. Czyli aby spełniony był warunek:

$$\forall a, b, c \sum_j w_{0abcj} = \sum_i w_{1abci}$$

gdzie: w_{0abcj} = waga j -tego ucznia ze szkoły nieskutecznie aplikującej z województwa a , preferującej wariant b i o kategorii wielkości c ;

w_{1abci} = waga i -tego ucznia ze szkoły biorącej udział w programie z województwa a , preferującej wariant b i o kategorii wielkości c .

Wartości wag dla każdej kombinacji wartości zmiennych warstwujących obliczono, dzieląc pierwotną sumę wag dla uczniów szkół uczestniczących w programie przez liczebność dla uczniów szkół nieskutecznie aplikujących.

W związku z powyższym do grupy kontrolnej nie włączono 81 (czyli ok. 0,08%) uczniów pochodzących ze szkół aplikujących nieskutecznie o takiej kombinacji wartości trzech zmiennych warstwujących, która nie występowała wśród uczestników. Z drugiej strony z porównania wyłączono 256 (czyli ok. 2,0%) uczniów szkół uczestniczących w programie o takiej kombinacji wartości trzech zmiennych warstwujących, która nie występowała w grupie szkół nieskutecznie aplikujących.

Dzięki temu zabiegowi łączny rozkład trzech zmiennych warstwujących w przeważonej grupie kontrolnej został zrównany z łącznym rozkładem tych zmiennych w grupie zasadniczej (tj. grupie uwzględnionych w analizach uczniów szkół biorących udział w programie). Można zaś założyć, że różnice w rozkładach wszystkich pozostałych zmiennych mają charakter losowy (z zastrzeżeniem punktu 2).

Tabela 2. Porównanie rozkładów zmiennych warstwujących między szkołami zakwalifikowanymi do programu a szkołami nieskutecznie wnioskującymi

szkoły:		nieskutecznie wnioskujące (N=3123)	zakwalifikowane (N=402)
województwo	dolnośląskie	7,1%	5,0%
	kujawsko-pomorskie	5,3%	6,2%
	lubelskie	5,6%	7,2%
	lubuskie	2,0%	2,5%
	łódzkie	5,1%	5,7%
	małopolskie	10,9%	10,4%
	mazowieckie	17,8%	12,4%
	opolskie	1,5%	3,0%
	podkarpackie	6,3%	7,0%
	podlaskie	2,4%	3,2%
	pomorskie	7,0%	6,0%
	śląskie	10,4%	10,2%
	świętokrzyskie	3,3%	4,0%
	warmińsko-mazurskie	2,7%	4,0%
	wielkopolskie	9,9%	9,2%
	zachodniopomorskie	2,7%	4,0%
	wariant	wariant I	93,7%
wariant II		5,8%	20,9%
brak danych		0,5%	0,0%
wielkość szkoły	mała	30,7%	33,1%
	średnia	41,4%	43,0%
	duża	27,7%	23,9%
	brak danych	0,3%	0,0%

2. Losowanie szkół do programu przeprowadzono do momentu, w którym wyczerpała się pula środków finansowych przydzielonych do danej warstwy. W niektórych województwach losowanie ograniczano do wniosków, opiewających na kwotę nie przekraczającą pozostałej jeszcze do rozdzielania puli. Może to powodować, że szkoły, które wnioskowały o mniejszą kwotę, miały większe prawdopodobieństwo udziału w programie. W badaniu nie dysponowano informacją o kwocie środków, o którą wnioskowały szkoły aplikujące nieskutecznie, dlatego nie było możliwe sprawdzenie, czy występuje taki efekt, ani uwzględnienie go w przeprowadzanych analizach. Można jednak przypuszczać, że jego skala była ograniczona przez to, że nie występował we wszystkich województwach, oraz że występował jedynie w sytuacji, gdy większa część puli była już rozdzielona, a więc dotyczył tylko szkół dołosowanych jako ostatnie.

3. W województwie lubuskim nie przeprowadzono losowania dla wariantu II, ponieważ liczba wnioskujących o udział w nim szkół była na tyle mała, że wszystkie zostały zakwalifikowane. W związku z tym wyłączono z ważenia i analiz wariant II w województwie lubuskim.
4. Wśród szkół nieskutecznie aplikujących do programu znalazło się 15 szkół, które nie spełniały kryteriów udziału w nim. Szkoły te oraz ich uczniów wyłączono z analiz.
5. Trzy szkoły zakwalifikowane do programu ostatecznie nie wzięły w nim udziału, a pięć aplikujących nieskutecznie uczestniczyło w komponencie badawczym „Cyfrowej szkoły”. Ponieważ zjawiska te dotyczą znikomej liczby szkół (mniej niż 1%), zdecydowano się nie wprowadzać zmiennej instrumentalnej, lecz wykluczono z ważenia i analiz powyższe szkoły.
6. W przypadku 15 szkół aplikujących nieskutecznie otrzymana baza nie zawierała informacji o preferowanym wariantcie. Szkoły te wyłączono z ważenia i analiz.
7. W niektórych ze szkół aplikujących w danym roku szkolnym nie było żadnych uczniów klas szóstych. Oczywiście nie było zatem również wyników sprawdzianu szóstoklasisty. Szkoły takie wyłączono z ważenia i analiz.
8. Nie dysponowano wynikami sprawdzianu dla 89% aplikujących szkół specjalnych oraz 8% aplikujących szkół ogólnodostępnych. W związku z tym wyłączono z ważenia i analiz szkoły specjalne, przeprowadzając je jedynie na szkołach ogólnodostępnych.

Dodatkowo brak danych o wynikach sprawdzianu wśród szkół ogólnodostępnych nie był rozłożony losowo, lecz zależał od województwa, liczby mieszkańców gminy i wielkości szkoły. Powodowało to, że grupa szkół biorących udział w programie, w przypadku których dysponowano wynikami sprawdzianu, nie była reprezentatywna dla ogółu uczestniczących placówek. Aby zapewnić możliwość lepiej uzasadnionej ekstrapolacji wyników w celu formułowaniu wniosków o całościowym efekcie programu w grupie wszystkich realizujących go ogólnodostępnych szkół, grupę zasadniczą przeważono pod względem trzech wymienionych wyżej cech – województwa, liczby mieszkańców gminy (w 5 kategoriach) i wielkości szkoły (w 3 kategoriach) – w taki sposób, aby jej struktura odpowiadała strukturze całej badanej populacji placówek uczestniczących w „Cyfrowej szkole”. Zastosowano w tym celu metodę ważenia wieńcowego.

Powyższy zabieg pociągnął za sobą konieczność dodatkowego przeważenia grupy kontrolnej ze względu na liczbę mieszkańców gminy, tak aby zachować jej porównywalność z grupą zasadniczą. Struktura ze względu na województwo i wielkość szkoły była zrównywana ważeniem w ramach opisywanej wcześniej procedury konstrukcji grupy kontrolnej. Aby przeważać próbę jednocześnie przez rozkład łączny województwa, wariantu i wielkości szkoły oraz rozkład kategorii liczby mieszkańców gminy, zastosowano metodę ważenia wieńcowego.

Po wykonaniu powyższych kroków dla każdego rocznika sześcioklasistów uzyskano grupę zasadniczą i porównywalną grupę kontrolną. Liczebności tych grup zostały przedstawione poniżej (Tabela 3). Zaprezentowano również porównanie rozkładu obu grup pod względem kluczowych zmiennych (Tabela 4).

Tabela 3. Liczebność grupy zasadniczej i kontrolnej

	2012		2013		2014	
	szkół	uczniów	szkół	uczniów	szkół	uczniów
Grupa zasadnicza	368	13.492				
Grupa kontrolna	2.858	111.084	2.746	104.406	2.832	107.223

Tabela 4. Struktura grupy zasadniczej i kontrolnej (na poziomie uczniów)

	2012		2013		2014		
	zasadnicza (N=13492)	kontrolna (N=111084)	zasadnicza (N=12731)	kontrolna (N=104406)	zasadnicza (N=12985)	kontrolna (N=107223)	
województwo	dolnośląskie	5,7%	5,7%	6,2%	6,2%	5,5%	5,5%
	kujawsko-pomorskie	4,4%	4,4%	4,6%	4,6%	4,7%	4,7%
	lubelskie	7,0%	7,0%	6,4%	6,4%	6,5%	6,5%
	lubuskie	2,5%	2,5%	2,6%	2,6%	2,6%	2,6%
	łódzkie	5,8%	5,8%	5,6%	5,6%	6,1%	6,1%
	małopolskie	9,7%	9,7%	9,9%	9,9%	10,1%	10,1%
	mazowieckie	14,6%	14,6%	14,7%	14,7%	14,1%	14,1%
	opolskie	3,0%	3,0%	2,5%	2,5%	2,5%	2,5%
	podkarpackie	6,4%	6,4%	6,7%	6,7%	7,2%	7,2%
	podlaskie	3,9%	3,9%	3,2%	3,2%	3,1%	3,1%
	pomorskie	6,3%	6,3%	6,5%	6,5%	6,1%	6,1%
	śląskie	10,0%	10,0%	10,3%	10,4%	10,2%	10,2%
	świętokrzyskie	3,4%	3,4%	3,7%	3,7%	3,7%	3,7%
	warmińsko-mazurskie	4,9%	4,9%	4,5%	4,5%	4,3%	4,3%
wielkopolskie	7,5%	7,5%	8,3%	8,2%	8,3%	8,3%	
zachodniopomorskie	4,9%	4,9%	4,3%	4,3%	5,0%	5,0%	
w.	wariant I	83,1%	83,1%	83,6%	83,6%	82,9%	82,9%
	wariant II	16,9%	16,9%	16,4%	16,4%	17,1%	17,1%
wielk. szkoły	mała	10,9%	10,9%	10,2%	10,2%	10,6%	10,6%
	średnia	37,4%	37,4%	37,6%	37,6%	37,5%	37,5%
	duża	51,8%	51,8%	52,2%	52,1%	51,9%	51,9%
liczba mieszk. gminy	do 5 tys.	4,7%	4,7%	4,6%	4,5%	4,3%	4,3%
	od 5 tys. do 20 tys.	39,6%	39,6%	40,5%	40,2%	41,1%	41,1%
	od 20 tys. do 50 tys.	24,7%	24,7%	24,3%	24,5%	23,9%	23,9%
	od 50 tys. do 100 tys.	6,7%	6,7%	6,9%	7,3%	7,1%	7,1%
	powyżej 100 tys.	24,4%	24,3%	23,7%	23,6%	23,6%	23,6%

3.4. Testy istotności statystycznej

W prowadzonych analizach posłużono się testami istotności statystycznej do rozstrzygnięcia, czy zaobserwowane różnice między grupą zasadniczą a kontrolną świadczą o rzeczywistym wpływie programu „Cyfrowa szkoła” na wyniki sprawdzianu, czy też mogą być zupełnie przypadkowe, czyli stanowiące jedynie konsekwencję określonego rezultatu losowania szkół do programu. Zastosowano dwa rodzaje testów statystycznych. Pierwszym były popularne testy asymptotyczne – testy t dla prób niezależnych w odniesieniu do średnich oraz test Kołmogorowa-Smirnowa w celu określenia istotności różnic między całymi rozkładami wyników. Drugim rodzajem były symulacyjne (permutacyjne) testy Monte Carlo.

Pierwszym testem zastosowanym przy porównaniu średniego wyniku uczniów między grupą zasadniczą a grupą kontrolną był test t dla prób niezależnych. Użyta wersja testu uwzględniała zespołowy charakter obu grup – czyli fakt, że dobór grupy zasadniczej odbywał się na poziomie szkół, a nie na poziomie uczniów². Przy porównaniu rozkładów wyników w pierwszym kroku zastosowano natomiast test Kołmogorowa-Smirnowa, w którym statystyką testową jest maksymalna odległość między dwoma rozkładami skumulowanymi. Test ten ma jednak poważne wady. Nie uwzględnia tego, że porównywane grupy są ważne, a dobór uczniów do grup zespołowy. Zespołowy charakter doboru (dobór odbywał się na poziomie szkół, a nie pojedynczych uczniów) prawdopodobnie zwiększa błąd standardowy estymatorów, w związku z czym test Kołmogorowa-Smirnowa prawdopodobnie niedoszacowuje wartość p. Oznacza to, że jego zastosowanie może prowadzić do wniosku, że różnice między porównywanymi grupami są istotne statystycznie, podczas gdy w rzeczywistości różnice te nie są istotne. Z tego względu zdecydowano się nie uwzględniać w dalszej części niniejszego opracowania wyników tego testu, a zamiast tego odwołać do bardziej wiarygodnego sposobu rozstrzygnięcia o istotności różnic, jakim są testy Monte Carlo³.

Testy permutacyjne Monte Carlo polegały na przeprośdzeniu wielokrotnie (dokładnie 10 000 razy) symulacji doboru szkół do udziału w programie⁴. W każdej takiej symulacji losowano więc grupę „pseudo-zasadniczą” spośród wszystkich szkół, które złożyły uprawnione wnioski o udział w programie. Starano się przy tym odwzorować zasady doboru szkół faktycznie zastosowane w programie „Cyfrowa szkoła”. Ponieważ jednak nie dysponowano informacją o kwocie wsparcia, o jaką wnioskowała dana szkoła, liczbę szkół wylosowaną w każdej warstwie (czyli dla każdej kombinacji województwa, preferowanego wariantu programu i wielkości szkoły) ustalono na taką wartość, jaka faktycznie została wylosowana. W związku z tym w każdej symulacji do grupy „pseudo-zasadniczej” włączono 399 szkół. Następnie zastosowano wszystkie procedury konstrukcji grupy kontrolnej opisane w rozdziale 3.3, w tym wyłączenie części szkół z analizy i przeważenie pozostałych. Uzyskano w ten sposób grupę „pseudo-zasadniczą” i grupę „pseudo-kontrolną”. Dla każdej części sprawdzianu i dla sprawdzianu ogółem obliczono różnicę w średnim wyniku sprawdzianu szóstoklasisty oraz maksymalną odległość między skumulowanymi rozkładami wyniku sprawdzianu w grupie zasadniczej i w grupie kontrolnej. Po przeprowadzeniu 10 000 symulacji dysponowano dziesięcioma tysiącami takich zestawów różnic. Warto zauważyć, że różnice te wynikały z losowo-warstwowego schematu doboru próby, natomiast nie były systematycznym efektem udziału w „Cyfrowej szkole”, ponieważ w obrębie danej warstwy szkoły faktycznie uczestniczące miały takie same prawdopodobieństwo zakwalifikowania się do grupy pseudo-zasadniczej jak szkoły

² Do przeprowadzenia testu wykorzystano komendę *regress* w pakiecie statystycznym Stata.

³ Pomysł zastosowania takich testów wyszedł od dr. Pawła Strawińskiego, któremu niniejszym składamy podziękowania.

⁴ Testy Monte Carlo zaprogramowano w środowisku statystycznym R.

w rzeczywistości wnioskujące nieskutecznie. Dlatego uzyskane rozkłady różnic można uznać za oszacowane rozkłady statystyki testowej w sytuacji prawdziwości hipotezy o braku wpływu programu na wyniki sprawdzianu szóstoklasisty. Dzięki temu można było zastosować typową procedurę statystycznej weryfikacji hipotez z wyznaczeniem obszaru krytycznego. W przypadku średniej był to obszar dwustronny: faktycznie obserwowana różnica średniego wyniku sprawdzianu szóstoklasisty okazywała się istotna statystycznie na poziomie istotności $p=0,05$, jeśli była większa lub mniejsza niż ponad 97,5% symulowanych różnic. W przypadku odległości między skumulowanymi rozkładami obszar krytyczny był jednostronny - różnica była uznawana za istotną, jeśli faktycznie obserwowana odległość była większa od ponad 95% symulowanych odległości.

W analizach nie wprowadzono korekty ze względu na liczbę przeprowadzonych testów. Przeprowadzanie testów dla wielu lat i wielu części sprawdzianu nie wywołuje takiej konieczności, jeśli ustalanie istotności różnic dla każdej części i roku potraktować jako weryfikację odrębnej hipotezy. Pewną komplikacją jest natomiast sprawdzanie istotności dla całego sprawdzianu łącznie, ponieważ test dla całego sprawdzianu nie jest niezależny od testów dla poszczególnych części sprawdzianu. Formułując wnioski z przeprowadzonych analiz oparto się na wynikach testów dla poszczególnych części sprawdzianu. W niniejszym opracowaniu ze względu na chęć ukazania pełniejszego obrazu kwestii będących przedmiotem badania zaprezentowano jednak również wyniki dla całego sprawdzianu.

4. Wyniki analiz

Po skonstruowaniu grupy zasadniczej i grupy kontrolnej porównanie surowych wyników sprawdzianu między nimi może być źródłem do wnioskowania o efektów programu. Poniżej zamieszczono szczegółowe porównanie średnich (Tabela 5. Średni wynik sprawdzianu szóstoklasisty w grupie zasadniczej i kontrolnej. Tabela 5) i rozkładów częstości (Wykres 1, Wykres 2, Wykres 3, Tabela 6). Wnioski są nieco zróżnicowane w zależności od części sprawdzianu.

Omówmy najpierw wyniki porównania dla części „czytanie”, „pisanie” i „korzystanie z informacji”. Średni wynik tych części sprawdzianu szóstoklasisty 2012 jest bardzo zbliżony w grupie zasadniczej i kontrolnej – różnica, na korzyść grupy zasadniczej, nie przekracza 0,04 surowego punktu, czyli około 2% odchylenia standardowego (Tabela 5). Oba zastosowane testy (asymptotyczny test t i symulacyjny test Monte Carlo) potwierdzają, że różnica ta jest nieistotna statystycznie. Wykres 1 i Wykres 2 pokazują, że nie tylko średnia, ale i cały rozkład jest zbliżony w obu grupach (linie dla grupy zasadniczej i kontrolnej niemal idealnie się pokrywają). Ponieważ porównanie dotyczy okresu sprzed realizacji programu prowadzi to do wniosku, że grupa kontrolna jest dopasowana do grupy zasadniczej i stanowi dobre źródło do oszacowania stanu konfaktycznego. Również dla lat 2013-2014 rozkłady wyników sprawdzianu są w podobnym stopniu zbliżone między grupą zasadniczą a grupą kontrolną. Oznacza to, że udział w programie „Cyfrowa szkoła” nie wpłynął na wyniki osiągnięte przez uczniów w trzech wymienionych częściach sprawdzianu szóstoklasisty.

W przypadku części „rozumowanie” wyniki sprawdzianu szóstoklasisty 2012 są zbliżone między obiema porównywanymi grupami uczniów. Również w tym przypadku grupa kontrolna okazuje się zatem dobrze dopasowana do grupy zasadniczej. Jednak w roku 2013, a zatem kilka miesięcy po wdrożeniu programu „Cyfrowa szkoła”, różnica pod względem średniego wyniku sprawdzianu między grupą zasadniczą a grupą kontrolną jest nieco większa – wynosi 0,08 surowego punktu sprawdzianu (około 3,5% odchylenia standardowego). Oba zastosowane testy statystyczne wskazują na to, że różnica ta nie jest istotna na poziomie istotności 0,05, ale jest bliska istotności (i okazałaby się istotna, gdyby przyjąć bardzo liberalny poziom istotności 0,1). Bardziej precyzyjny obraz różnic między grupą zasadniczą a kontrolną daje porównanie rozkładów wyników sprawdzianu w części „rozumowanie”. Jak pokazuje Wykres 2, w odróżnieniu od poprzednich części linie rozkładów w grupie zasadniczej i w grupie kontrolnej nie pokrywają się. Mianowicie w grupie zasadniczej mniej jest uczniów, którzy zdobyli 0, 1 lub 2 punkty, a więcej uczniów, którzy zdobyli 3, 4, 5 lub 6 punktów. Sugeruje to, że dzięki programowi niektórzy uczniowie, którzy w razie braku realizacji programu osiągnęliby stosunkowo niski wynik, przesunęli się do grupy uczniów o wynikach ze środka skali. Tabela 6 zawiera informację, że zaobserwowana różnica w skumulowanych rozkładach znajduje się na progu istotności ($p=0,05$). W związku z tym trudno jednoznacznie rozstrzygnąć, czy zaobserwowana różnica jest efektem programu, czy też wynika z czynników o charakterze losowym. Ta pierwsza interpretacja cechuje się jednak dużym prawdopodobieństwem. Jeżeli jest słuszna, to warto dodatkowo zauważyć, że zaobserwowane przypuszczalnie oddziaływanie programu dotyczyło jedynie niektórych uczniów osiągających na sprawdzianie, w części „rozumowanie” wynik poniżej przeciętnego (dla lepszych wyników widoczne na wykresie linie zaczynają się zbliżać), przez co jest w mniejszym stopniu widoczne w średniej wartości wyniku.

Tabela 5. Średni wynik sprawdzianu szóstoklasisty w grupie zasadniczej i kontrolnej.

			czytanie	pisanie	korzy- stanie z infor- macji	rozumo- wanie	wykorzy- -stanie wiedzy	cały spra- wdzian
2012	A. Grupa zasadnicza	średnia	6,32	5,92	2,83	4,13	3,79	22,99
	B. Grupa kontrolna	średnia	6,27	5,91	2,81	4,11	3,73	22,83
	Różnica (A-B)	różnica	+0,04	+0,01	+0,02	+0,02	+0,07	+0,16
	Istotność (p=)	test t	0,31	0,84	0,26	0,49	0,17	0,32
		Monte Carlo	0,27	0,99	0,23	0,56	0,23	0,34
2013	A. Grupa zasadnicza	średnia	7,34	6,38	2,48	4,22	3,77	24,19
	B. Grupa kontrolna	średnia	7,32	6,34	2,47	4,14	3,70	23,97
	Różnica (A-B)	różnica	+0,02	+0,04	+0,01	+0,08	+0,07	+0,22
	Istotność (p=)	test t	0,54	0,53	0,65	0,08	0,30	0,26
		Monte Carlo	0,70	0,83	0,67	0,08	0,43	0,37
2014	A. Grupa zasadnicza	średnia	7,79	5,83	2,74	4,59	5,01	25,96
	B. Grupa kontrolna	średnia	7,76	5,85	2,73	4,54	4,96	25,84
	Różnica (A-B)	różnica	+0,03	-0,02	+0,01	+0,05	+0,05	+0,12
	Istotność (p=)	test t	0,42	0,72	0,43	0,31	0,30	0,51
		Monte Carlo	0,58	0,61	0,62	0,36	0,37	0,66

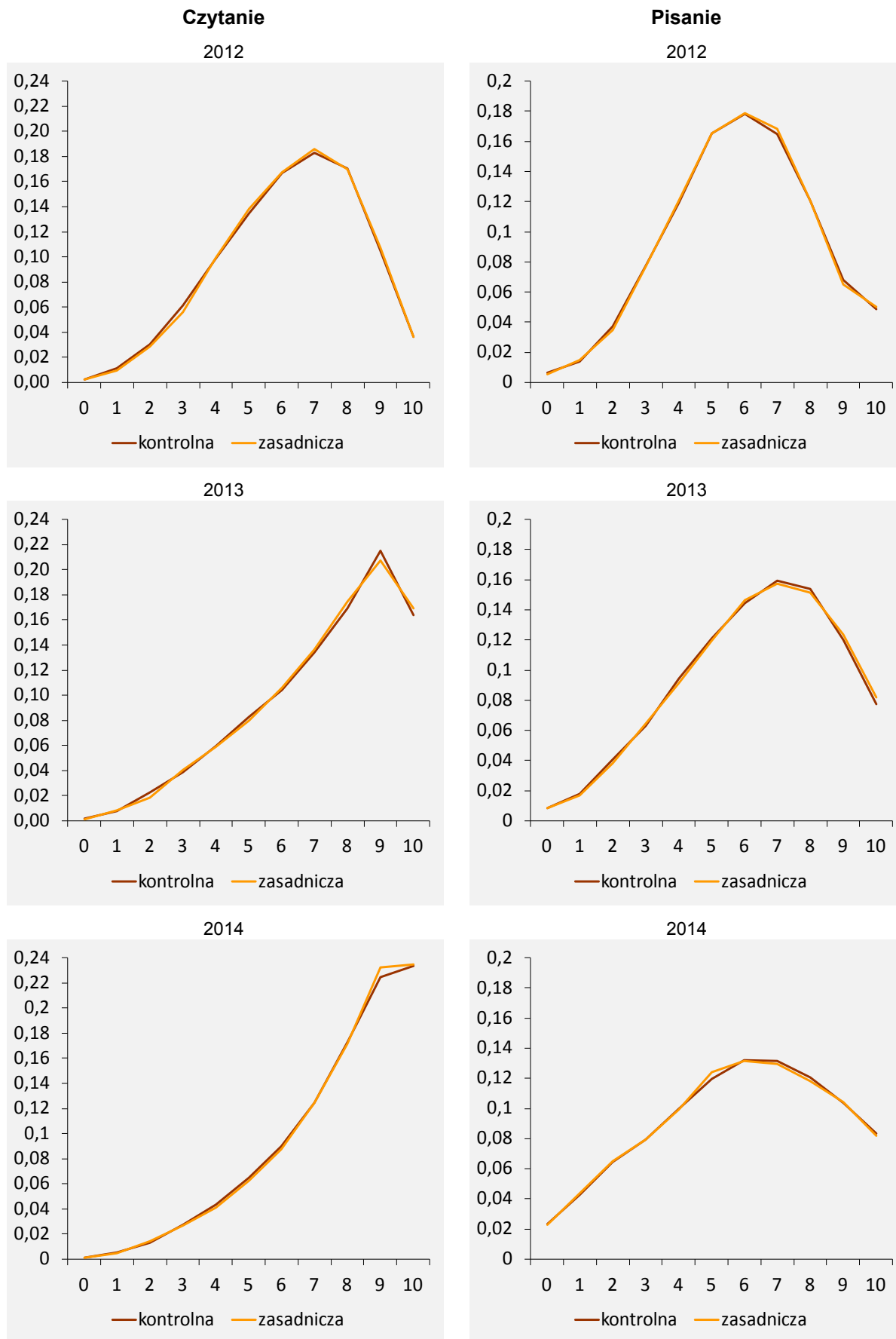
Średnie i różnice podano w zaokrągleniu, dlatego sumowanie podanych wartości może dać niedokładny wynik.

Tabela 6. Istotność różnic między skumulowanymi rozkładami surowych wyników sprawdzianu w grupie zasadniczej i grupie kontrolnej.

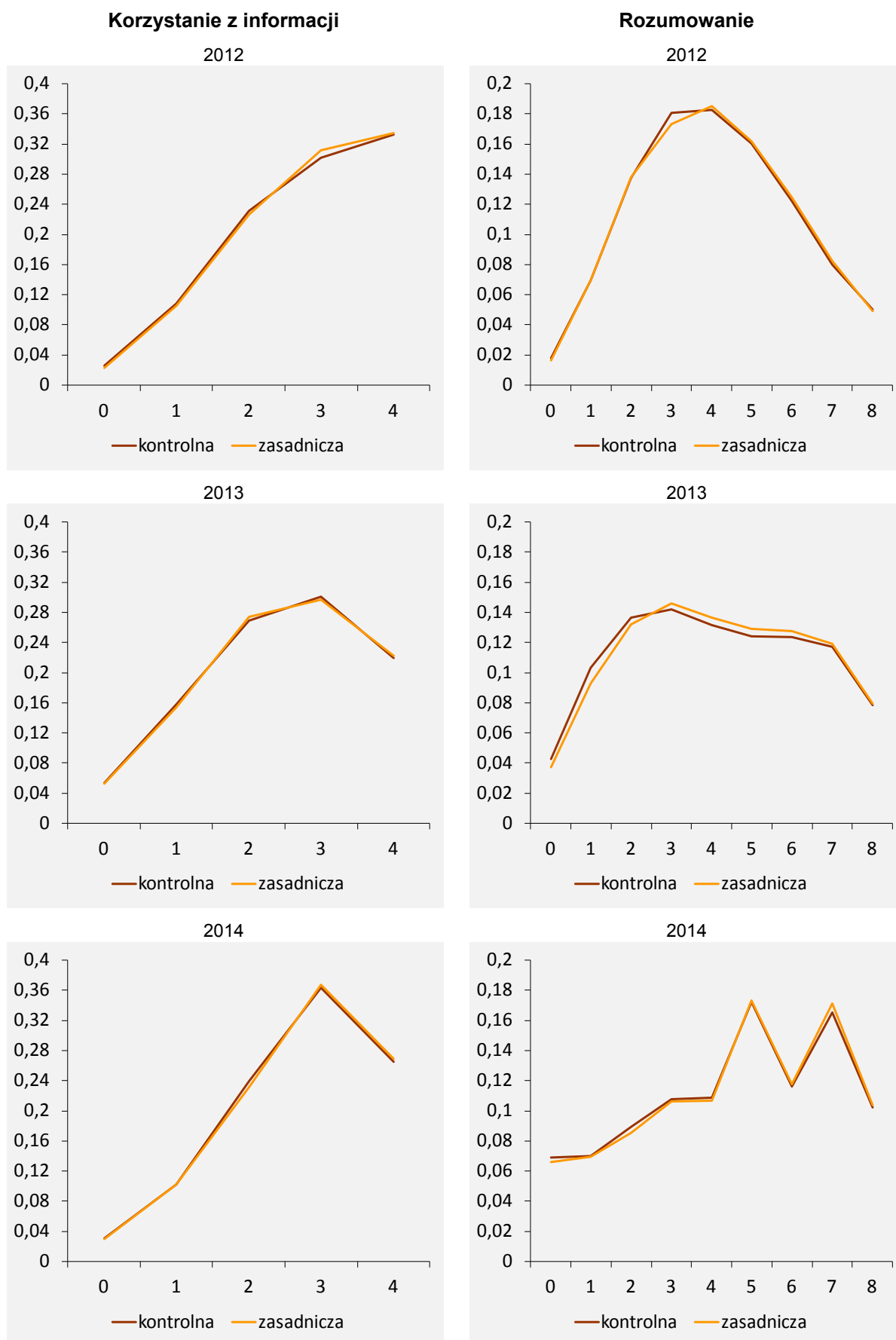
	czytanie	pisanie	korzystanie z informacji	rozumowanie	wykorzystanie wiedzy	cały sprawdzian
2012	0,34	0,96	0,18	0,51	0,17	0,33
2013	0,84	0,99	0,83	0,05	0,56	0,38
2014	0,62	0,87	0,39	0,61	0,33	0,56

Zastosowano testy symulacyjne (permutacyjne) Monte Carlo.

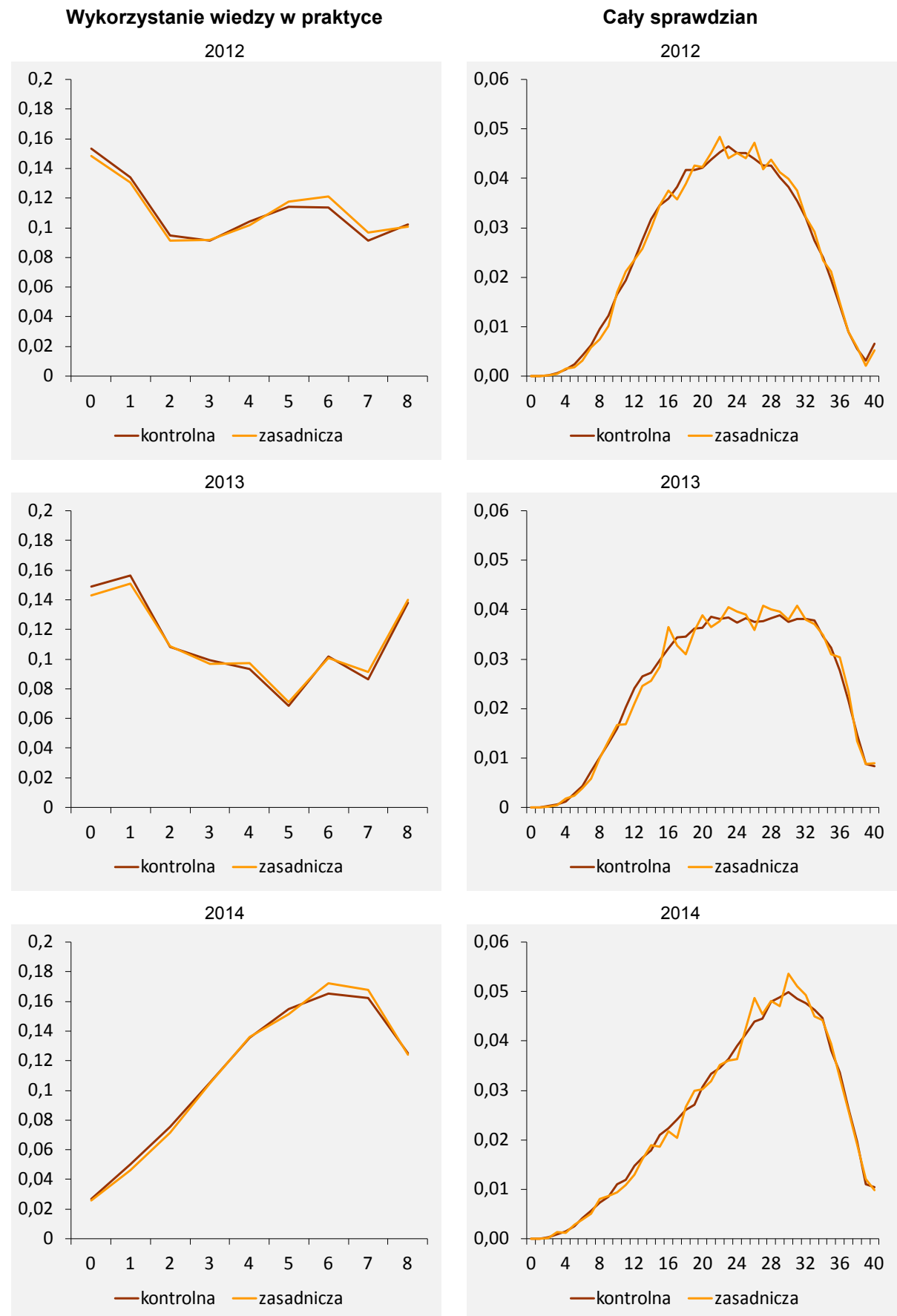
Wykres 1. Rozkład wyników sprawdzianu szóstoklasisty w grupie zasadniczej i kontrolnej



Wykres 2. Rozkład wyników sprawdzianu szóstoklasisty w grupie zasadniczej i kontrolnej



Wykres 3. Rozkład wyników sprawdzianu szóstoklasisty w grupie zasadniczej i kontrolnej



W przypadku sprawdzianu szóstoklasisty z 2014 roku różnica w średnich osiągniętych w części „rozumowanie” między grupą zasadniczą a kontrolną jest już mniejsza – wynosi 0,05 surowego punktu (ok. 2% odchylenia standardowego). Patrząc na Wykres 2 ma się wrażenie, że rozkłady częstości są w obu grupach praktycznie takie same. A zatem jeśli w roku 2013 rzeczywiście uwidocznił się efekt programu, to był to efekt krótkotrwały, ponieważ rok później nie został już zaobserwowany.

Ze specyficzną sytuacją mamy do czynienia w przypadku części „Wykorzystanie wiedzy w praktyce”. Już w 2012 roku, a więc przed udziałem w programie, różnica w średnim wyniku tej części między grupą zasadniczą a kontrolną była nieco większa niż w przypadku pozostałych części i wynosiła 0,07 surowego punktu (około 2,6% odchylenia standardowego). Wykres 3 pokazuje, że w grupie zasadniczej więcej było uczniów, którzy zdobyli w tej części 6 lub 7 punktów, a więc osiągnęli wysoki, choć nie maksymalny rezultat. Sprawia to wrażenie, że grupa kontrolna jest gorzej dopasowana do grupy zasadniczej niż w przypadku poprzednich części. Jednak wszystkie zastosowane testy statystyczne zgodnie orzekają o nieistotności tych różnic, a więc mogły one powstać w sposób czysto losowy i nie świadczą o niedoskonałości metody doboru grupy kontrolnej.

Jeżeli spojrzeć na wykres rozkładu wyników z części „wykorzystanie wiedzy w praktyce” w roku 2013, to może się wydawać, że różnice między grupą zasadniczą a kontrolną są mniejsze niż w roku 2012. Różnica między wartościami średnimi, którą podaje Tabela 5, okazuje się jednak taka sama – 0,07 surowego punktu (2,5% odchylenia standardowego). Z kolei na wykresie dla roku 2014 zdaje się powtarzać wzór z roku 2012, ponieważ w grupie zasadniczej więcej jest uczniów, którzy zdobyli 6 lub 7 punktów. Różnica między średnimi jest minimalnie mniejsza – wynosi 0,05 surowego punktu (2,3% odchylenia standardowego). Wszystkie zastosowane testy statystyczne wskazują, że różnice te są dalekie od istotności. Konkluzja jest więc taka, że na przestrzeni lat 2012-2014 nie zaobserwowano żadnych zmian w wynikach części „wykorzystanie wiedzy”, które mogłyby świadczyć o wpływie udziału szkoły w programie „Cyfrowa szkoła”.

Pozostaje zobaczyć, jaki jest obraz sytuacji przy porównaniu wyników całego sprawdzianu, bez podziału na poszczególne części. W 2012 roku średni wynik sprawdzianu był wyższy o 0,16 surowego punktu (2,1% odchylenia standardowego). W 2013 roku dystans ten wzrósł do 0,22 punktu (2,6% odchylenia standardowego), by w 2014 roku zmaleć do 0,12 punktu (1,5% odchylenia standardowego). Jak pokazuje Wykres 3, kształt rozkładu wyników całego sprawdzianu wydaje się w obu grupach podobny, z tą jedynie różnicą, że w grupie kontrolnej jest on bardziej „wygładzony”, a w grupie zasadniczej bardziej nieregularny, co można tłumaczyć większą podatnością na przypadkowe oscylacje wynikającą z mniejszego rozmiaru tej drugiej. Wszystkie zastosowane testy statystyczne wskazują na nieistotność zaobserwowanych różnic.

5. Wnioski

Dzięki losowemu doborowi szkół do udziału w programie „Cyfrowa szkoła” możliwe było wiarygodne oszacowanie wpływu tego udziału na wyniki osiągnięte przez uczniów na sprawdzianie szóstoklasisty. Wyniki analiz sugerują, że program spowodował podwyższenie osiągnięć niektórych uczniów, którzy w razie jego niezrealizowania osiągnęliby wynik poniżej przeciętnej, w części „rozumowanie” sprawdzianu. Efekt ten jest jednak stosunkowo niewielki (nie przekłada się na istotną statystycznie różnicę średnich, zaś różnica w skumulowanych rozkładach częstości jest na progu istotności), ograniczony do niektórych uczniów i tylko do jednej części sprawdzianu, a ponadto krótkotrwały – w 2014 roku, czyli kilkanaście miesięcy od dostarczenia sprzętu informacyjno-komunikacyjnego do szkół, nie był już widoczny. Nie zaobserwowano innych zmian w wynikach sprawdzianu, które mogłyby być skutkiem „Cyfrowej szkoły”.

Kiedy przystępowano do analiz, wysunięto roboczą hipotezę, że w 2013 roku nie zostaną jeszcze zaobserwowane żadne efekty programu, ponieważ czas, jaki upłynął od jego wdrożenia (kilka miesięcy) jest zbyt krótki, żeby wprowadzone zmiany przełożyły się na wyniki sprawdzianu. Spodziewano się, że taki wpływ będzie widoczny dopiero w 2014 roku. Wyniki analiz pokazały, że w rzeczywistości było odwrotnie – jedyny zaobserwowany przypuszczalny efekt programu wystąpił w 2013 roku i zanikł do momentu przeprowadzenia sprawdzianu w 2014 roku. Z jednej strony sugeruje to, że już okres 3-5 miesięcy realizacji programu 1:1 wystarczy, aby zaobserwować pewne zmiany w osiągnięciach edukacyjnych uczniów. Z drugiej pokazuje, że zmiany te są efemeryczne. Skłania to do wysunięcia hipotezy, że w przypadku „Cyfrowej szkoły” wystąpił „efekt nowości”. Podobnie jak w wielu innych badaniach, również w tym przypadku nauczyciele informowali, że zastosowanie TIK zwiększa atrakcyjność prowadzonych lekcji dla uczniów, a przez to rośnie ich zaangażowanie i motywacja do aktywnego udziału w lekcji (Instytut Badań Edukacyjnych, 2013). Można podejrzewać, że jest to związane między innymi z tym, że zastosowanie TIK jest czymś nowym i nietypowym, dzięki czemu pobudza uwagę uczniów. W miarę czasu może jednak spowszednieć i przestać być nadzwyczajną atrakcją, a stać się elementem codzienności. Wyjaśniałoby to krótkotrwałość zaobserwowanych efektów. Więcej światła na tę kwestię mogą rzucić wyniki szerszej zakrojonych badań sposobu wykorzystania sprzętu informacyjno-komunikacyjnego oraz technik dydaktycznych stosowanych przez nauczycieli ze szkół uczestniczących w programie „Cyfrowa szkoła”. Badania takie są aktualnie prowadzone przez Instytut Badań Edukacyjnych, jednak w momencie opracowania niniejszego materiału nie są jeszcze znane ich rezultaty.

Jak wspomniano w części wprowadzającej, w dwóch niedawno przeprowadzonych w krajach Ameryki Łacińskiej kontrfaktycznych ewaluacjach programów 1:1 stwierdzono, że nie przekładają się one na wyniki uczniów w egzaminach z języka urzędowego i z matematyki (Cristia, Ibararán, Cueto, Santiago, & Severín, 2012; de Melo, Machado, & Miranda, 2014). Również w omówionych w niniejszym opracowaniu analizach nie zaobserwowano żadnego wpływu na wyniki sprawdzianu szóstoklasisty z 2014 roku. Prowadzi to do wniosku, że raczej nie należy się spodziewać, by wdrażane w przyszłości w Polsce programy 1:1, stanowiące kontynuację „Cyfrowej szkoły”, w znaczący sposób oddziaływały na wyniki egzaminów zewnętrznych – przynajmniej w perspektywie kilkunastu miesięcy, bo taka została uwzględniona w trzech wymienionych badaniach. Kwestią wymagającą dalszych badań jest to, czy program ten miał wpływ na kompetencje uczniów niemierzone sprawdzianem szóstoklasisty, takie jak kompetencje cyfrowe czy też umiejętności pracy w grupie.

6. Literatura cytowana

- Angrist, J. D., & Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.
- Balanskat, A., Blamire, R., & Kefala, S. (2006). *The ICT Impact Report: A review of studies of ICT impact on schools in Europe*. European Communities.
- Beuermann, D., Cristia, J., Cruz-Aguayo, Y., Cueto, S., & Malamud, O. (2012). *Home Computers and Child Outcomes: Short-Term Impacts from a Randomized Experiment in Peru*. New York: Inter-American Development Bank.
- Biagi, F., & Loi, M. (2013). Measuring ICT Use and Learning Outcomes: evidence from recent econometric studies. *European Journal of Education*, 48 (1), 28-42.
- Bocconi, S., Kampylis, P., & Punie, Y. (2013). Framing ICT-enabled Innovation for Learning: the case of one-to-one learning initiatives in Europe. *European Journal of Education*, 48 (1), 113-130.
- Cheung, A. C., & Slavin, R. E. (2012). How features of educational technology applications affect student reading outcomes: A meta-analysis. *Educational Research Review* (7), 198-215.
- Condie, R., & Munro, B. (2007). *The impact of ICT in schools - a landscape review*. Coventry: Becta.
- Cristia, J. P., Ibararán, P., Cueto, S., Santiago, A., & Severín, E. (2012). *Technology and Child Development: Evidence from the One Laptop per Child Program*. New York: Inter-American Development Bank.
- de Melo, G., Machado, A., & Miranda, A. (2014). *The Impact of a One Laptop per Child Program on Learning: Evidence from Uruguay*. IZA Discussion Paper No. 8489.
- Fairlie, R. W., & Robinson, J. (2013). *Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren*. IZA Discussion Paper No. 7211.
- Federowicz, M. (Red.). (2014, lipiec). *OECD PISA: Wyniki badania 2013 w Polsce*. Pobrano 02 04, 2015 z lokalizacji http://www.ifspan.waw.pl/pliki/pisa-2012-raport_09_07-best.pdf
- Fuchs, T., & Wößmann, L. (2005). *Computers and Student Learning: Bivariate and Multivariate Evidence on the Availability and Use of Computers at Home and at School*. Ifo Working Paper No. 8.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. (2011). *Impact Evaluation in Practice*. Washington DC: World Bank.
- Gil Flores, J. (2012). Utilización del ordenador y rendimiento académico entre los estudiantes españoles de 15 años. *Revista de Educación* (357), 375-396.

Hennesy, S., & London, L. (2013). *Learning from International Experiences with Interactive Whiteboards: The Role of Professional Development in Integrating the Technology*. OECD Education Working Papers, No. 89, OECD Publishing.

Instytut Badań Edukacyjnych. (2013, 08 03). *Ewaluacja ex-post rządowego programu rozwijania kompetencji uczniów i nauczycieli w zakresie stosowania technologii informacyjno-komunikacyjnych – „Cyfrowa szkoła”*. (P. Penszko, Red.) Pobrano 02 2015, 09 z lokalizacji <http://eduentuzjasci.pl/images/stories/publikacje/ibe-raport-cyfrowa-szkola.pdf>

Keane, T., Keane, W. F., & Blicblau, A. S. (2013). The Use of Educational Technologies to Equip Students with 21st Century Skills. *WCCE 2013. 10th IFIP World Conference on Computers in Education*. 1, strony 74-82. Toruń: Nicolaus Copernicus University Press.

Koufogiannakis, D., & Wiebe, N. (2006). Effective Methods for Teaching Information Literacy Skills to Undergraduate Students: A Systematic Review and Meta-Analysis. *Evidence Based Library and Information Practice* , 1 (3), 3-43.

Malamud, O., & Pop-Eleches, C. (2011). Home Computer Use and the Development of Human Capital. *Quarterly Journal of Economics* (126), strony 987-1027.

Murname, R. J., & Willet, J. B. (2011). *Methods Matter: Improving Causal Inference in Educational and Social Science Research*. New York: Oxford University Press.

Penuel, W. R. (2006). Implementation and Effects Of One-to-One Computing Initiatives: A Research Synthesis. *Journal of Research on Technology in Education* , 38 (3), 329-248.

Severin, E., & Capota, C. (2011). *One-to-One Laptop Programs in Latin America and the Caribbean*. New York: Inter-American Development Bank.

Tamim, R. M., Bernard, R. M., Borkohovski, E., Abrami, P. C., & Schmid, R. F. (2011). What Forty Years of Research Says About the Impact of Technology on Learning: A Second-Order Meta-Analysis and Validation Study. *Review of Educational Research* , 81 (1), 4-28.

(2014). *The NMC Horizon Report Europe: 2014 Schools Edition*. European Commission / The New Media Consortium.

Valiente, O. (2010). *1-1 in Education: Current Practice, International Comparative Research Evidence and Policy Implications*. OECD Education Working Papers, No. 44, OECD Publishing.

White, G. K. (2008, 06). *ICT Trends in Education*. Pobrano 02 03, 2015 z lokalizacji Digital Learning Research: http://research.acer.edu.au/digital_learning/2/