

**Doc. PhDr. Josef Malach, CSc.**

**RNDr. Martin Malčík, Ph.D.**

**RNDr. Radek Krpec, Ph.D.**

Pedagogická fakulta Ostravské univerzity v Ostravě

## **Edukacyjna wartość dodana a odpowiedzialność nauczycieli i szkół**

(tłumaczenie nadesłane przez autorów)

### **Abstrakt**

Modele wartości dodanej dotyczą przyrostu wiedzy ucznia pomiędzy poszczególnymi etapami edukacji względem wcześniej określonych celów edukacji szkolnej. Parametr ten oraz modele wykorzystywane do jej wyliczenia w różnym stopniu separują od innych wpływów czy też czynników, które również mają wpływ na uczenie się i rozwój uczniów. Władze odpowiedzialne za prowadzenie i nadzorowanie szkół w ramach procesów zarządzania jakością używają względnie obiektywne dane o przyroście wiedzy konkretnych uczniów, edukacyjnej wartości dodanej poszczególnych nauczycieli oraz szkoły jako całości. Należy wyznaczyć konkretne wskaźniki niezawodności parametru wartości dodanej w przedziałach czasowych, które umożliwiłyby wyeliminowanie nieplanowanych oraz incydentalnych wpływów na edukację szkolną i w ten sposób zapobiegły mylnym interpretacjom uzyskanych danych i umożliwiły podjęcie odpowiednich kroków. Autorzy przyczynku, opartego na danych pochodzących z pomiaru edukacyjnej wartości dodanej, dokumentują problem zmienności uzyskanych danych spowodowany zastosowaniem różnych zmiennych kontekstualnych.

### **Słowa kluczowe:**

Edukacyjna wartość dodana, odpowiedzialność nauczycieli i szkół

### **Edukacyjna wartość dodana jako narzędzie ewaluacji pedagogicznej**

Nauki pedagogiczne używają pojęcia *edukacyjna wartość dodana* stosunkowo od niedawna. Po raz pierwszy zastosowane ono zostało na początku 1992 roku w stanie Tennessee, dzisiaj używane jest w 21 stanach USA (2011)<sup>1</sup>. Działają tu zwłaszcza trzy modele: najbardziej rozpowszechniony jest EVAAS (the Educational Value-Added Assessment System), którego autorem jest W. Sanders, kolejnym używanym jest system DVASS (the Dallas Value-Added Assessment System) oraz system REACH (Rate of Expected Academic

---

<sup>1</sup> Value Added Assessment. Center for Greater Philadelphia, University of Philadelphia, USA, 2011. [http://www.cgp.upenn.edu/ope\\_value.html](http://www.cgp.upenn.edu/ope_value.html)

Change) <sup>2</sup>. W Europie w 2002 roku rozpoczęto mierzenie wartości dodanej w Wielkiej Brytanii w postaci określenia netto różnic wyników uczniów pomiędzy poszczególnymi kluczowymi poziomami (Bartmańska, 2006). Później wiele państw, zwłaszcza Norwegia, Francja, Szwecja, Holandia i Polska, transformowało pierwotne modele w specyficzne warianty narodowe. Na wyższych uczelniach mierzenie wartości dodanej w bezpośrednim związku z odpowiedzialnością za efekty nauczania rozpoczyna się w 2004 roku, kiedy to koncepcję tę przedstawił Hersh na konferencji AAHEA (*American Association for Higher Education & Accreditation*) (2004, s. 7).

Praca OECD z 2008 roku <sup>3</sup> definiuje modele wartości dodanej w ten sposób, że porównywane są zawsze dwa okresy, w których mierzone są wyniki nauczania. Modele wartości dodanej mierzą wkład szkoły w postępy ucznia w porównaniu z wcześniej określonymi szkolnymi celami edukacyjnymi. Wkład to wartość nieuwzględniająca innych czynników, które również mają pozytywny wpływ na postępy ucznia. Dane podejście jasno wyznacza potrzebę mierzenia postępów ucznia, tzn. zmianę w czasie, która odzwierciedla działanie szkoły na jednostkę uczącą się. Ważne jest również oddzielenie wkładu szkoły od innych wpływów czy czynników, które również mają wpływ na uczenie się, co jest problemem kompleksowym (Ryśka, 2009).

Zwolennicy aplikowania modeli wartości dodanej podkreślają jej zalety czy mocne strony zwykle w ten sposób, że porównują je z powszechnie używanym ogólnym testowaniem wiedzy i umiejętności uczniów. Braun (2005) konstatuje, że stwierdzenie występowania dodatniego EWD może stworzyć podstawy do obrony oceniania nauczycieli w odróżnieniu od metody polegającej na absolutnym poziomie wyników uczniów lub procencie osiągnięcia stałych standardów. Pozytywnie są oceniane próby uwzględnienia serii tzw. kontekstualnych zmiennych, a więc indywidualnego poziomu rozwoju ucznia, wsparcia udzielanego przez rodziców, motywacji, przyzwyczajęń, dynamiki interpersonalnej, studiowanego kierunku, które były przedmiotem naszych badań

Niemierko (2010) uważa wynik pomiarów edukacyjnej wartości dodanej za informację, która może wzmocnić strategię podnoszenia efektywności edukacji przy pomocy procedury przez autora nazwanej jako *ocenianie orientujące*. Celem tego typu oceniania, które usytuowane jest pomiędzy ocenianiem kształtującym a sumującym, jest ocenianie wyników nauki z punktu widzenia celu, do którego jednostka lub grupa zmierza różnymi drogami i różnym tempem. Chodzi o ocenianie etapowe, które mierzone jest modułowo, przy pomocy części programów lub semestrów. Może być wykorzystane do przyjmowania decyzji o programach nauczania lub metodach realizacji na poziomie klasy, szkoły lub regionu.

---

<sup>2</sup> Braun, H.I. *Using Student Progress to Evaluate Teachers: A primer on Value-added models*. Educational Testing Service. Princeton, 2005.

<sup>3</sup> OECD: *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*. Paris, 2008. ISBN 978-92-64-05022-8.

## Wykorzystanie wyników pomiarów wartości dodanej

Powszechnie deklarowane są następujące cele polityczne związane z wykorzystaniem wyników pomiaru wartości dodanej: poprawa poziomu szkół, odpowiedzialność szkół za osiągnięte wyniki i wybór szkoły (OECD 2008). Pierwszy cel dotyczy systemów na poziomie ogólnopństwowym. Na tym poziomie można ocenić efektywność szkół przy wzajemnym porównaniu, ocenić funkcjonowanie specyficznych inicjatyw, wspierających lub innowacyjnych programów lub zmian systemowych (np. dotyczących curriculum). Drugi cel jest potrzebny do wypełniania funkcji szkoły – ponoszenie odpowiedzialności za wyniki pracy edukacyjnej. Wyników analizy danych można użyć do głębszych refleksji nad efektywnością szkoły i do utrzymywania szeregu powiązań pomiędzy efektami edukacji a finansowaniem szkół i ich zarządzaniem. Trzeci cel daje rodzicom możliwość wykorzystania informacji przy decyzji dotyczącej wyboru szkoły, która najlepiej odpowiadałaby potrzebom ich dzieci.

## Publikowanie wyników pomiarów wartości dodanej

Zasadniczą problematyką w systemie pomiaru wartości dodanej jest opisanie wyników pomiarów i przesyłanie uzyskanych wyników do wszystkich zainteresowanych osób i ich ewentualne rozpowszechnianie.

Przykładem przedkładania wyników mierzenia, z wielu przyczyn specyficznym dla USA, jest w stanie Kalifornia kategoryzowana lista wyników pochodzących ze systemu STAR<sup>4</sup> (Standardized Testing and Reporting). Ma wiele kryteriów kategoryzujących i umożliwia dokonywanie kolejnych analiz danych.

Wyniki są na przykład opracowane z uwzględnieniem płci uczniów, uczniów o specjalnych wymaganiach edukacyjnych, dzieci pochodzących z rodzin o gorszych warunkach ekonomicznych, dla wszystkich ras i grup etnicznych w stanie oraz dla wszystkich grup językowych.

W Polsce pierwszą informację o wynikach pomiarów wartości dodanej opublikowano w 2005 roku<sup>5</sup>. Użyto regresyjnego modelu szacowania wartości dodanej, a wyniki były prezentowane w dziewięciostopniowej (tzw. staninowej) skali. Wyniki przeznaczone były dla uczniów, ew. ich rodziców, nauczycieli i dyrektorów szkół.

**Dla uczniów** i ich nauczycieli dane są najbardziej zrozumiałe, jeżeli podane są w postaci tabeli i grafów. Tabela zawiera wyniki ucznia uzyskane podczas testów wstępnych oraz przewidywany wynik drugiego testowania. Różnica między wartością przewidywaną i zmierzoną, która może być dodatnia lub ujemna, jest edukacyjną wartością dodaną konkretnego ucznia. Wartość tę poglądowo może pokazać graf. Osiągnięcia uczniów opisywane są często słownie przy pomocy pięciostopniowej skali, która używa określeń przyrostu jako mały, niżej średni, średni, wyżej średni, duży. Odpowiednie są w tym

<sup>4</sup> Standardised Testing and Reporting (STAR) Program. Dostępne na: <http://star.cde.ca.gov/star2010/aboutSTAR.asp>

<sup>5</sup> M.K. Szmigel, A. Rappe, *Komunikowanie wartości dodanej osiągnięć edukacyjnych uczniom, nauczycielom i dyrektorom szkół* [w:] *Egzamin*, Biuletyn Badawczy 8/2006, Wydział Badań i Ewaluacji CKE, Warszawa, 2006. Dostępne na: [http://www.cke.edu.pl/imagesstories/badania/biul\\_8.pdf](http://www.cke.edu.pl/imagesstories/badania/biul_8.pdf)

przypadku wcześniej przemyślane komentarze do poszczególnych punktów pięciostopniowej skali, które w sobie łączą ocenę sumaryczną z elementami o charakterze formatywnym<sup>6</sup>.

**Dla dyrektorów szkół** podawane są wyniki testowania uczniów po etapach – rocznikach szkolnych, obliczenie przewidywanego wyniku pomiędzy wybranymi rocznikami, później obliczona jest różnica między wartością przewidywaną i zmierzoną i słownie dodany stopień wartości dodanej. Również szkoła otrzymuje szerszy komentarz słowny proponujący zmiany w procesie nauczania – uczenia się w monitorowanych przedmiotach.

**Dla założycieli szkół lub organów prowadzących** przygotowywane są tabele sumaryczne dotyczące wartości dodanej szkół przez nich założonych lub monitorowanych, również w pięciopunktowej skali.

Sposoby publikowania, pokazywania wyników i ich interpretowanie mają bardzo silny wymiar etyczny i mogą przy niewłaściwym użyciu wywołać destrukcję stosunków pomiędzy zainteresowanymi podmiotami edukacji szkolnej, prowadzić do demotywowania uczniów i nauczycieli lub do zmian personalnych.

Przykładem tego twierdzenia mogą być choćby publikowane przez prasę informacje o stwierdzonej wartości dodanej, ew. „efektywności“ poszczególnych nauczycieli i szkół, które wywołały burzliwą dyskusję, gdy stanowiska rodziców i fachowców były przeciwstawne<sup>7</sup>. Niektóre głosy wzywały do natychmiastowego zwolnienia nauczycieli o niższym stopniu zmierzonej osobistej wartości dodanej bez względu na niektóre inne faktory, które mogły mieć wpływ na otrzymany wynik, np. okres praktyki, grupowe charakterystyki ich uczniów i jakość zewnętrznych warunków edukacyjnych.

Krytyczne refleksje dotyczące publikowania wyników wartości dodanej nauczycieli przedstawiono w siedmiu punktach:

1. Punkty oznaczające wartość dodaną nie mogą być użyte jako pierwotny środek do podejmowania wielce ryzykownych decyzji, jak na przykład stała praca lub zwolnienie.
2. Przyjęcie status quo jest niedopuszczalne.
3. Twierdzenia: „z dwojga złego należy wybierać mniejsze zło“ nie powinno się używać.
4. Nie przestawać uwzględniać wartości danych pochodzących z pomiaru wartości dodanej dla oceny formatywnej pracy edukacyjnej.
5. Bardzo ważna jest wiarygodna i zasadna obserwacja przez szkolonych oceniających.
6. Musimy pójść dalej niż tylko koncentrować się na najlepszych i najgorszych nauczycielach.
7. Efektywność nauczyciela jest kontekstualna.

---

<sup>6</sup> Ibidem, s. 53.

<sup>7</sup> *Adding Value to the Value-Added Debate* Tuesday, *Los Angeles Times*, August 31, 2010. Dostępne 23.9.2010 na <http://eduooptimists.blogspot.com/2010/08/adding-value-to-value-added-debate.html>.

Pokazują one złożoność interpretacji publikowanych informacji i ich wymiary etyczne. Każde państwo powinno w narodowym historyczno-kulturowym kontekście rozważyć wszystkie okoliczności związane z publikowaniem wyników pomiarów wartości dodanej i podjąć odpowiedzialną decyzję. Musi wziąć pod uwagę również pytanie, czy opublikowanie wyników szkół nie prowadzi w stopniu niepożądanym do konkurencji pomiędzy szkołami i nie stanie się tym samym przeszkodą w korzystaniu z przykładów dobrej praktyki, która ma mocny potencjał podnoszenia ogólnej jakości edukacji. Wiele krajów publikuje wyniki testowania uczniów, z drugiej strony wiele najlepszych państw według badań TIMSS i PISA, np. Singapur i Finlandia, tego nie robią. Niedawno zmienił swą politykę publikowania wyników Hong-Kong.<sup>8</sup>

W Republice Czeskiej obecnie trwa ostra dyskusja między Ministerstwem Szkolnictwa i wszystkimi wojewodami, którzy są założycielami większości szkół średnich, na temat publikowania wyników państwowej części egzaminu maturalnego, która po raz pierwszy przebiegła w RC w 2011 roku. Ministerstwo nie ma zamiaru udostępnić danych i opublikowało za zgodą szkół tylko wyniki dziesięciu najlepszych liceów w Republice (TOP 10 Gy), dziesięciu pozostałych szkół (TOP 10 pozostałe szkoły) oraz 3 najlepszych szkół w 14 województwach (TOP 3 w województwach). Województwa grożą państwu rozprawą sądową. Matury nie zdało 19,5 proc. studentów, a więc około 17 000 osób. Mogą one teraz żądać powtórnego sprawdzenia wyników, na razie tak uczyniło około 200 osób. Egzamin poprawkowy będą we wrześniu. Matury państwowe przygotowano 14 lat, kosztowały ogółem 680 milionów koron<sup>9</sup>. Opublikowane aktualne fakty świadczą o tym, że dzisiaj relatywnie obiektywne zmierzenie edukacyjnej wartości dodanej nie stanowi tak wielkiego problemu, ale problemem jest raczej nieprzygotowanie nauczycieli, managementu szkół, założycieli i społeczeństwa do publikowania uzyskanych wyników.

### **Edukacyjna wartość dodana i odpowiedzialność szkoły**

Nasze badania koncentrują się na tym celu pomiaru wartości dodanej, który prezentuje odpowiedzialność szkół za wyniki swej pracy edukacyjnej. Systemy odpowiedzialności szkół umożliwiają dokładne określenie, które obiekty są odpowiedzialne za jakość udzielanych usług i wyników edukacji, jak również za konsekwencje swego podejścia, i to nie tylko prawnie w stosunku do założycieli, sponsorów, rady szkolnej, wizytatorów, płatników podatków, ale zwłaszcza wobec rodziców, uczestników kształcenia i innych zaangażowanych osób, by bilansowo wykazać odpowiednie wykorzystanie środków na dające się skontrolować wyniki docelowe (Pol, 2007).

Podstawowe cele procesu umacniania odpowiedzialności szkół są trafnie sformułowane w tzw. jednolitym modelu reformy w USA<sup>10</sup>, który próbuje zapewnić dzieciom wszystkich narodowości okazję do uzyskania najwyższego poziomu wykształcenia. System odpowiedzialności musi być sprawiedliwy

<sup>8</sup> McKinsey&Company Klesající výsledky českého základního a středního školství: Fakta a řešení, 2010

<sup>9</sup> Lidovky. Zpravodajský server Lidových novin. [http://www.lidovky.cz/kraje-zadaji-zverejneni-vysledku-maturit-vyhrozuji-soudem-pqq-/ln\\_domov.asp?c=A110624\\_155823\\_ln\\_domov\\_ape](http://www.lidovky.cz/kraje-zadaji-zverejneni-vysledku-maturit-vyhrozuji-soudem-pqq-/ln_domov.asp?c=A110624_155823_ln_domov_ape)

<sup>10</sup> Comprehensive Reform Model. [http://www.cgp.upenn.edu/ope\\_value.html](http://www.cgp.upenn.edu/ope_value.html)

wobec edukatorów i płatników podatków. Edukacyjna wartość dodana i jednolite ramy odpowiedzialności mogą być przydatne w krajach i regionach jako model do a) utworzenia systemu oceniania i zastępowania poszczególnych nauczycieli i osób na stanowiskach kierowniczych, które będzie do przyjęcia przez polityków i do zaakceptowania, o ile chodzi o stronę ekonomiczną, b) udzielania bardzo potrzebnego profesjonalnego rozwoju i wsparcia dla edukatorów, c) poprawy dyscypliny nauczycieli pracujących w szkołach o niskim budżecie d) poprawy pracy personalnej w szkole<sup>11</sup>.

Plan odpowiedzialności szerokiej inicjatywy „Efektywna edukacja publiczna“ w stanie Pensylwania jest symetryczny – uwzględnia menedżerów szkolnych, a także nauczycieli. System, zgodnie z daną inicjatywą, ocenia menedżerów szkolnych według tego, jak efektywnie przekazują wysoki poziom wykształcenia wszystkim studentom, jak wykorzystują dane o edukacji studentów do podejmowania decyzji i jak budują kulturę szkolną o wysokim standardzie i nieprzerwany profesjonalny rozwój<sup>12</sup>.

Odpowiedzialność szkoły w zakresie spełniania jej funkcji edukacyjnej w zasadzie jest odpowiedzialnością za udzielanie jakościowych usług przejawiającą się wysokim poziomem osiągniętego wykształcenia. Złożone zjawisko jakości szkoły o wielu zmiennych wymaga pewnego uproszczenia, by mogło być badane, a następnie poprawiane. Autorzy niniejszego referatu pracują z modelem ewaluacji (jakości) szkoły CIPP D. Stufflebeama (2003), który należy do stałych i respektowanych modeli.

Wyniki procesu stwierdzania edukacyjnej wartości dodanej mają pewien potencjał, by służyć do ewaluacji wszystkich czterech składników modelu. Dla szkoły są jednak najważniejsze informacje dotyczące procesu kształcenia i jego wyników. Jeżeli weźmiemy pod uwagę ogólnie przyjmowane twierdzenie, że zasadniczy wpływ na wyniki kształcenia uczniów ma praca nauczyciela (Barber, Mourshed 2007), dane dotyczące edukacyjnej wartości dodanej konkretnej klasy lub szkoły pozostają w związku przyczynowym z dydaktyczną pracą nauczycieli poszczególnych przedmiotów i uczniowie kształceni przez wysoko efektywnych nauczycieli osiągają trzykrotnie lepsze wyniki niż uczniowie nieefektywnych nauczycieli. Nie można ignorować faktu znanego z badań PISA i TIMSS, że na wyniki i inne efekty kształcenia ma wpływ szerszy społeczny, ekonomiczny i kulturowy kontekst złożony z wielu czynników (Vesely 2011).

Na wykorzystaniu modeli pomiaru wartości dodanej przy ocenie nauczycieli skoncentrował się Braun (2005). Wymienia dwa główne rodzaje wykorzystania: po pierwsze, identyfikacja nauczycieli, którzy z największym prawdopodobieństwem wymagać będą doskonalenia zawodowego i którzy będą pytani lub obserwowani, by została określona konkretna sfera najbardziej potrzebnego wsparcia; po drugie, identyfikacja szkół, których wyniki są słabsze i potrzebny będzie audyt do stwierdzenia, czy potrzebują specyficznego rodzaju

---

<sup>11</sup> Value-Added Assessment. [http://www.cgp.upenn.edu/ope\\_value.html](http://www.cgp.upenn.edu/ope_value.html)

<sup>12</sup> *Teacher and Principal Evaluation Information 2009-2010 School Year Reports* Pennsylvania Department of Education. <http://www.education.state.pa.us/portal/server.pt/community/certifications/7199/TPE/964624>



pomocy. Zarazem zwraca uwagę na to, że wyniki pomiarów wartości dodanej nie powinny być jedyną postawą do przyjmowania późniejszych decyzji dotyczących nauczycieli (na przykład płace, awans lub sankcje). Poleca, by pomiary wartości dodanej były częścią składową całej grupy narzędzi ewaluacyjnych. Dlatego Nowy System Odpowiedzialności<sup>13</sup> pracuje z dwoma zasadniczymi podstawami do ewaluacji nauczycieli: używa wyników zmierzonej wartości dodanej i standaryzowanej oceny pracy przez organizację nadzorującą. Sama metodologia stwierdzania wartości dodanej zawiera niektóre gwarancje bezpieczeństwa w celu sprawiedliwego osądzania nauczyciela. Bardzo ważne jest używanie danych pochodzących z kilku lat, by nauczyciel nie był karany na przykład z powodu kryzysu z przyczyn osobistych lub przejściowych problemów. Poza tym zmierzone dane są łączone z obserwacją pracy nauczycieli w klasach, która również redukuje możliwość niesprawiedliwej oceny<sup>14 i 15</sup>.

Problem stosunku wartości dodanej i odpowiedzialności szkół, która tworzy odpowiedzialność wszystkich nauczycieli i menedżera szkoły, można podsumować w ten sposób, że wartość dodana jest tylko jednym z materiałów przy ocenianiu odpowiedzialności, a nawet może mieć pewne znaczenie w systemach oceny nauczycieli klasyfikowanych według stopnia wykształcenia.<sup>16</sup>

Warunkiem rzeczywistego efektu wyników uzyskanej wartości dodanej na odpowiedzialność nauczycieli za swą pracę jest akceptowanie tego sposobu oceny przez samych nauczycieli. Mons (2009) podaje, że nauczyciele mają różny punkt widzenia dotyczący standaryzowanego oceniania. Akceptują zasady ogólne, ale krytykują mechanizmy dotyczące interesów, które wywierają zbyt duży wpływ na podejście do kształcenia. Prace z kilku krajów pokazują, że nauczyciele wspierają zasady standardów dotyczących osiągnięć w nauczaniu, a większość menedżerów, zwłaszcza dyrektorów szkół i założycieli szkół standaryzowane testowanie uważa za właściwe.

### **Niezawodność edukacyjnej wartości dodanej**

Wprowadzenie pomiaru wartości dodanej do procesu sterowania jakością i wzmocnienie odpowiedzialności stwarza również potrzebę sprawdzania przewidywanych efektów poprzez odpowiednie analizy i badania. Są już do dyspozycji narzędzia do ewaluacji narzędzi ewaluacyjnych tzn. metaewaluacji. Newton (2007) oprócz kryterium dokładności technicznej (*technical accuracy*), której wskaźnikami są przede wszystkim wiarygodność i zasadność, wprowadza jeszcze kolejnych pięć potrzebnych kryteriów. Do przeforsowania użycia stwierdzonej wartości dodanej i jej utrzymania się w procesie sterowania jakością i odpowiedzialności szkół, wszystkie kryteria wzajemnie się uzupełniają. Są poruszane również sprawy zasadności użycia narzędzi oceny, zwłaszcza testów (Stobart 2008).

<sup>13</sup> A New System of Accountability . Center for Greater Philadelphia, University of Philadelphia, USA, 2011. [http://www.cgp.upenn.edu/ope\\_value.html](http://www.cgp.upenn.edu/ope_value.html)

<sup>14</sup> Value – Added Assessment. [http://www.cgp.upenn.edu/ope\\_value.html](http://www.cgp.upenn.edu/ope_value.html)

<sup>15</sup> Standards for Educational and Psychological Testing, 1999.

<sup>16</sup> OPE Pennsylvania Legislation. [http://www.cgp.upenn.edu/pdf/OPE\\_PA\\_Legislation.pdf](http://www.cgp.upenn.edu/pdf/OPE_PA_Legislation.pdf)

Ogólnie biorąc, wiarygodność (reliability) to poziom, na którym wyniki studentów przy powtarzanej procedurze oceny są spójne (Nitko 2004). To znaczy, że wiarygodność to poziom, do którego wyniki oceny studentów są zgodne, jeżeli a) studenci rozwiązują to samo zadanie, b) jeden lub dwaj nauczyciele oceniają to samo zadanie, c) studenci rozwiązują dwa lub więcej równoważnych zadań w tej samej lub różnej sytuacji. Nitko zarazem konstatuje, że wiarygodność jest niezbędna dla zasadności. Wiarygodność oceny zapewnia zasadność (validity) decyzji. Zakładamy, że zasadność użytego narzędzia, a więc testów do powtórnych pomiarów, po pewnym czasie będzie dostateczna i pilotażowo sprawdzona.

Podane pojmowanie wiarygodności pomiaru wyników jest odpowiednim punktem wyjścia do jego przeniesienia do warunków wiarygodności zyskanej edukacyjnej wartości dodanej w stosunku do odpowiedzialności szkół.

## **Cel badań**

Głównym celem badań opisywanym w niniejszym artykule jest określenie, jaki wpływ ma użycie socjoekonomicznych czynników na wartość dodaną uczniów, szkół i czy włączenie danych czynników do obliczania wartości dodanej znajdzie odzwierciedlenie w wynikach uczniów i całej szkoły.

## **Badana grupa i metoda gromadzenia danych**

Dane używane w niniejszym artykule zostały uzyskane w ramach projektu Testowanie uczniów 1. roczników szkół kończących się egzaminem maturalnym oraz Testowanie uczniów 3. roczników szkół kończących się egzaminem maturalnym z celem stwierdzenia relatywnego przyrostu wiedzy we wszystkich szkołach średnich w województwie morawskośląskim w okresie 2007 - 2010. Gromadzenie danych w 1. roczniku objęło grupę

8000 studentów i odbywało się na początku studiów, parę tygodni po rozpoczęciu przez studentów nauki w szkole średniej. Gromadzenie danych w 3. roczniku odbywało się wiosną konkretnego roku, zwykle na przełomie maja i czerwca i objęło grupę 6000 studentów. Do obliczania edukacyjnej wartości dodanej używano pary danych uzyskanych w obu okresach pomiarowych dla około 6000 studentów.

W ramach testowania w szkołach przeprowadzono również ankietę. Ankieta miała na celu badanie podstawowych czynników socjoekonomicznych i zadowolenia uczniów. Studenci byli testowani elektronicznie, za pomocą metody relatywnego przyrostu osiągnięć uczniów (Malach, Malčík 2010).

Pomiar edukacyjnej wartości dodanej, który realizowało Metodické i Ewaluačné centrum Uniwersytetu Ostrawskiego w Ostrawie (Metodické a evaluační centrum Ostravské univerzity v Ostravě) jest pierwszym swego typu pomiarem przeprowadzonym w Republice Czeskiej, gdzie dotychczas nie wprowadzono nawet standaryzowanego narodowego testowania uczniów ani pomiaru wartości dodanej<sup>17</sup>. Problematyka testowania narodowego, odbycie się państwowej części egzaminu maturalnego (po raz pierwszy zrealizowane

---

<sup>17</sup> National Testing of Pupils in Europe. Objectives, Organisation and Use Results. EACEA, Brussels, 2009.



w czerwcu 2011) jest bardziej przedmiotem debat politycznych niż fachowym dyskursem. Autorzy aktykułu więc z zainteresowaniem śledzą poczynania polskich kolegów, którzy intensywnie i na wysokim poziomie zajmują się problematyką edukacyjnej wartości dodanej (Stożek 2010, Dolata 2006, Dolata 2007, Dolata, Pokropek 2007).

### Regresywne modele względnego przyrostu wiedzy

Metoda względnego przyrostu jest metodą, która przy obliczaniu wartości dodanej korzysta z regresji linearnej pomiędzy wartościami wejściowymi i wyjściowymi – zmierzonej wiedzy uczniów według poniższego wzoru:

$$y_{ij} = a_0 + a_1x_{ij} + b_1f_{ij1} + \dots + b_nf_{ijn}, \quad (1)$$

gdzie:

$i$  – oznaczenie ucznia w ramach  $j$ -tej szkoły,  $y$  – wyjście,  $a_0$  – konstanta,  $a_1$  – człon regresyjny  $x$  – wejście

$b_1$ – $b_n$  – regresyjne człony socjoekonomicznych i pozostałych czynników

$f_1$ – $f_n$  – socjoekonomiczne i pozostałe czynniki

RSS – Residual Sum of Square – suma kwadratów rezydualów (Krpec, Burda, 2011).

W ramach analizy danych przygotowaliśmy i analizowaliśmy kilka modeli i badaliśmy, jak dane modele, które zawierają nie tylko proste wyniki uczniów, ale również inne czynniki socjometryczne, wpływają na wyniki uczniów. Stopniowo użyliśmy następujących modeli:

- Model podstawowy
- Model pytania
- Model uwzględniający kierunek edukacji
- Model uwzględniający wykształcenie rodziców
- Model uwzględniający odpowiedzi na pytania postawione w ankiecie.

Przez pojęcie *model* należy rozumieć pewien obiekt  $M$ , który w pewnym celu zastępuje inny obiekt  $O$ . Wymagamy przy tym, by osądy wyprowadzone z obiektu  $M$  były ważne z jakąś dopuszczalną bliskością również dla obiektu  $O$ . Za pierwotną uważana jest sytuacja, kiedy obiekt  $O$  jest pewną częścią realnego świata, tzn. ontologicznym obiektem, a  $M$  jest wiedzą o nim. W takim razie  $M$  nazywamy *kognitywnym modelem* części rzeczywistości (Křemen 2007).

### Model podstawowy

W modelu podstawowym uwzględniamy tylko wejściowe i wyjściowe wyniki testów bez uwzględnienia jakichkolwiek innych czynników. To znaczy, że pozostałe czynniki regresywne są zerowe.

$$b_1 - b_n = 0$$

Wyniki obliczeń według modelu podstawowego można zestawić w tabeli 1.

Tabela 1. Współczynniki regresji liniowej w modelu podstawowym

	$a_0$	$a_1$	RSS
język czeski	-1,81	0,81	574746
matematyka	2,6	0,82	1799358
język angielski	25,39	0,47	734116

### Model kierunku

W ramach testowania pierwotnego wzięto pod uwagę tylko jeden faktor, którym był kierunek studiów. Model ten był definiowany jednak inaczej niż wszystkie wyżej podane modele. Faktor kierunku studiów nie został uwzględniony w ramach modelu wielokrotnej regresji liniowej

$$y = a_0 + a_1x + b_1f_1 + \dots + b_n f_n \quad (2)$$

lecz w ramach każdego kierunku były określone parametry prostej regresji liniowej

$$y_i = a_{i0} + a_{i1}x_i$$

gdzie  $i$  identyfikacyjny numer kierunku studiów.

W podobny sposób próbowaliśmy określać w ramach poszczególnych kierunków wielokrotne liniowe modele regresyjne z uwzględnieniem faktorów typu wykształcenie rodziców, odpowiedzi na pytania itp.

$$y = (a_{10} + a_{11}x)f_1 + (a_{20} + a_{21}x)f_2 + (a_{30} + a_{31}x)f_3 + (a_{40} + a_{41}x)f_4 + (a_{50} + a_{51}x)f_5 + (a_{60} + a_{61}x)f_6 + (a_{70} + a_{71}x)f_7 \quad (3)$$

gdzie  $f_i$  jest faktorem kierunku studiów. To znaczy, jeżeli zmierzone wartości  $x$ ,  $y$  kierunku wynoszą 3, w takim razie  $f_1, f_2, f_4, f_5, f_6, f_7$  mają wartość zerową, faktor  $f_3$  wynosi jeden, określane są więc parametry regresji  $a_{30}$  i  $a_{31}$ . To jednak zupełnie to samo, co określanie parametrów regresji dla każdego kierunku z osobna.

### Model wykształcenie rodziców

W poniższym modelu bierzemy pod uwagę wykształcenie rodziców. W skład testu wchodził kwestionariusz, w którym uczniowie odpowiadali na następujące pytania, podawali między innymi wykształcenie rodziców. W przypadku wykształcenia każdego z rodziców wybierali z możliwości: podstawowe, zasadnicze zawodowe, średnie, wyższe. Przy poszukiwaniu modelu regresyjnego jako zasadnicze przejawy się faktory: średnie wykształcenie matki, wyższe wykształcenie matki oraz wyższe wykształcenie ojca. Pozostałe faktory nie miały sygnifikacyjnego wpływu na model końcowy.

f1 – matka średnie, f2 – matka wyższe, f3 – ojciec wyższe

**Tabela 2. Współczynniki regresji liniowej w modelu wykształcenie rodziców**

	$a_0$	$a_1$	$b_1$	$b_2$	$b_3$	RSS
język czeski	-1,2	0,79	1,08	3,22	2,32	560275
matematyka	2,32	0,78	2,6	5,34	3,67	1760989
język angielski	25,01	0,45	1,51	3,11	2,97	718107

### Model pytania

W ramach modelu „pytania“ za faktor regresji uważane są odpowiedzi na niektóre z pytań w ankiecie. Jednak w przypadku różnych przedmiotów jako sygnifikacyjne okazują się inne pytania, więc nie można określić wspólnego modelu dla wszystkich przedmiotów. Przy porównywaniu RSS w przypadku poszczególnych przedmiotów z modelem wykształcenie stwierdzamy, że nie mają one wcale tak wielkiego znaczenia.

Z ankety, która prócz socjoekonomicznych faktorów śledziła również niektóre elementy zadowolenia uczniów w szkole, jako mające znaczenie określono następujące pytania:

język czeski

- Czy przyjęto mnie do szkoły, gdzie chciałem studiować?
- Czy mogę korzystać z korepetycji w ramach działalności szkoły, jeżeli z jakiegokolwiek powodu mam braki w przerabianym materiale?
- Czy boję się odpytywania i prac pisemnych w szkole?

matematyka

- Czy w zasadzie radzę sobie z wykonywaniem zadań podczas lekcji?
- Jeżeli mam problem, czy mogę zwrócić się do nauczyciela/nauczycielki?
- język angielski
- Czy mogę korzystać z korepetycji w ramach działalności szkoły, jeżeli z jakiegokolwiek powodu mam braki w przerabianym materiale?

**Tabela 3. Współczynniki regresji liniowej w modelu pytania**

	$a_0$	$a_1$	$b_1$	$b_2$	$b_3$	RSS
język czeski	-2,94	0,81	1,74	0,77	-0,94	572110
matematyka	-2,75	0,82	3,36	2,4		1794375
język angielski	24,68	0,47	1,09			732925

### Model wykształcenie rodziców + pytania

Jako ostatni model zrobiliśmy kombinację faktorów wykształcenia i pytań. W tabeli 4. są zestawione współczynniki regresji danego modelu.

**Tabela 4. Współczynniki regresji liniowej w modelu pytania i wykształcenie rodziców**

	$a_0$	$a_0$	$a_1$	$b_1$	$b_2$	$b_3$	$b_4$	$b_6$	RSS
język czeski	-2,25	0,78	1,02	3,2	2,26	1,57	0,75	-0,85	558048
matematyka	-0,37	0,78	2,61	5,39	3,67	2,98	2,83		1755709
język angielski	24,31	0,45	1,5	3,1	2,97	1,08			716952

W ramach języka czeskiego analizowano następujące czynniki:  $f_1$  – matka średnie,  $f_2$  – matka wyższe,  $f_3$  – ojciec wyższe, pyt. a), pyt. c), pyt. e) (zob. model pytania).

W ramach matematyki analizowano następujące czynniki:  $f_1$  – matka średnie,  $f_2$  – matka wyższe,  $f_3$  – ojciec wyższe, i pytania pyt. b) pyt. f) (zob. model pytania).

W ramach języka angielskiego analizowano następujące czynniki:  $f_1$  – matka średnie,  $f_2$  – matka wyższe,  $f_3$  – ojciec wyższe, i pytania pyt. c) (zob. model pytania).

### Różnice w kolejności szkół pomiędzy poszczególnymi modelami

Wartość dodana szkoły została określona jako przeciętna arytmetyczna wartości dodanych uczniów szkoły. Do dalszej analizy podzieliliśmy szkoły według kolejności w poszczególnych modelach (np. według kierunku nauczania, według wykształcenia) zawsze równomiernie na 4 poziomy, tj. na cztery części: poziom 1. ...0 – 25%, poziom 2. ...25 – 50%, poziom 3. ...50 – 75%, poziom 4. ...75 – 100%. Jako przykład przytaczamy dwie tabele zmian w kolejności szkół. W tabelach 5. i 6. wartości pokazane są zawsze w ten sposób, że podają, ile procent tych, którzy w modelu pierwotnym byli na poziomie 1. umieszcilo się na poszczególnych poziomach przy zastosowaniu innego modelu (np. kierunku studiów).

Jeśli chodzi o zmiany w matematyce, 33% z poziomu 2. w modelu „kierunek” przesunęło się na poziom 1., podobnie 24% z poziomu 3. poprawiło wyniki na poziomie 2., i odwrotnie 33% szkół z poziomu 1. spadło na poziom 2., podobnie 33% szkół z poziomu 3. spadło na poziom 4., natomiast 14% szkół z poziomu 4. poprawiło swe wyniki na poziomie 2.

Z tabeli 6. można wywnioskować, że wyniki w modelu wykształcenie nie różnią się zbytnio od modelu pierwotnego.

**Tabela 5. Różnice w kolejności szkół pomiędzy modelem podstawowym a modelem kierunek w matematyce**

matematyka		model kierunek			
		poziom 1.	poziom 2.	poziom 3.	poziom 4.
model bez dodatkowych regresorów	poziom 1.	57%	33%	10%	0%
	poziom 2.	33%	29%	19%	19%
	poziom 3.	10%	24%	33%	33%
	poziom 4.	0%	14%	38%	48%

**Tabela 6. Różnice w kolejności szkół pomiędzy modelem podstawowym a modelem wykształcenie w języku angielskim**

język angielski		model wykształcenie			
		poziom 1.	poziom 2.	poziom 3.	poziom 4.
model bez dodatkowych regresorów	poziom 1.	90%	10%	0%	0%
	poziom 2.	10%	81%	10%	0%
	poziom 3.	0%	5%	95%	0%
	poziom 4.	0%	0%	0%	100%

### Współczynniki zmiany absolutnej i całkowitej oraz podział szkół na poziomy

Jeżeli oznaczymy elementy pierwszej linijki symbolami  $a_{11}$ ,  $a_{12}$ ,  $a_{13}$ ,  $a_{14}$  itd. po elementy czwartej linijki  $a_{41}$ ,  $a_{42}$ ,  $a_{43}$ ,  $a_{44}$ , potem możemy wprowadzić współczynnik:

$$k_{az} = \frac{(a_{12} + a_{23} + a_{34} + a_{21} + a_{32} + a_{43})}{6} + 2 \cdot \frac{a_{13} + a_{24} + a_{31} + a_{42}}{4} + 3 \cdot \frac{a_{14} + a_{41}}{2} \quad (4)$$

$k_{az}$  nazwiemy współczynnikiem zmiany absolutnej.

Pierwszy ułamek jest wartością przeciętną z procentu szkół, które przesunęły się o jeden poziom.

Ułamek drugi jest wartością przeciętną z procentu szkół, które przesunęły się o dwa poziomy.

Ułamek trzeci jest wartością przeciętną z procentu szkół, które przesunęły się o trzy poziomy.

Znaczenie jest oceniane według przesunięcia, w przypadku pierwszego ułamka wynosi jeden, w przypadku drugiego ułamka dwa, a w przypadku trzeciego ułamka trzy.

Wskaźnik ten mógłby umożliwić w pewien sposób ocenę, do jak wielkich przesunięć w osiągniętych pozycjach na poszczególnych poziomach doszło pomiędzy pierwszym, pierwotnym modelem a modelem drugim. Współczynnik  $k_{az}$  jest dodatni, a więc nie możemy określić, czy doszło do wyraźniejszego przesunięcia w kierunku plus czy minus.

Dlatego spróbujemy wprowadzić jeszcze jeden współczynnik:

$$k_{cz} = \frac{-a_{12} - a_{23} - a_{34} + a_{21} + a_{32} + a_{43}}{6} + 2 \cdot \frac{-a_{13} - a_{24} + a_{31} + a_{42}}{4} + 3 \cdot \frac{-a_{14} + a_{41}}{2} \quad (5)$$

$k_{cz}$  nazwiemy współczynnikiem zmiany całkowitej.

Przy pomocy współczynnika  $k_{cz}$  moglibyśmy określać całkowitą zmianę, do której doszło. Jeżeli wartość współczynnika jest ujemna, oznacza to, że doszło do przesunięcia szkoły w kierunku minus, a wartość dodatnia – oznacza przesunięcie szkoły w kierunku plus.

**Tabela 7. Wartości współczynnika zmiany absolutnej i całkowitej dla poszczególnych przedmiotów i modeli**

współczynnik	kierunek		wykształcenie		kierunek, wykształcenie		wykształcenie pytania		kierunek, pytania, wykształcenie	
	$k_{az}$	$k_{cz}$	$k_{az}$	$k_{cz}$	$k_{az}$	$k_{cz}$	$k_{az}$	$k_{cz}$	$k_{az}$	$k_{cz}$
język czeski	0,15	-0,02	0,02	0	0,15	-0,02	0,02	0	0,15	-0,02
matematyka	0,14	-0,02	0,01	0	0,14	0	0,012	0	0,14	-0,02
język angielski	0,188	-0,03	0,01	-0,02	0,18	-0,02	0,07	0	0,188	-0,02

## Wyniki pracy

Do wyraźnych różnic w kolejności szkół dochodzi pomiędzy modelem podstawowym i modelem kierunek, z czego wnioskujemy, że kierunek studiów w znacznej mierze wpływa na względny przyrost wiedzy studentów. Porównując kolejność szkół w poszczególnych przedmiotach w ramach modelu bez faktora kierunek, widzimy, że różnice nie są takie duże. Szkoły przesuwają się najwyżej o 10 miejsc, zwykle chodzi o jedną, dwie szkoły. Wykształcenie rodziców przejawia się najwyraźniej w zmierzonej wartości dodanej w języku czeskim. Model wykształcenie rodziców + pytania wykazuje podobny wpływ we wszystkich przedmiotach, jednak w różnych przedmiotach znaczenie sygnifikacyjne mają inne pytania. Modele poprzez wzajemną kombinację regresorów w większości przypadków wzmacniają wpływ regresorów na wynik wartości dodanej szkoły.

## Wnioski

Przedłożone wyniki były uzyskane na znaczącej grupie ponad 6000 uczniów. Jednoznacznie zostało potwierdzone, że wliczenie socjoekonomicznych i pozostałych parametrów do obliczania edukacyjnej wartości dodanej metodą względnego przyrostu wiedzy ma sygnifikacyjny wpływ na wyniki. Najwyraźniejszy wpływ ma kierunek studiów uczniów, co jednak jest zrozumiałe, ponieważ w ramach studiowanego kierunku przejawia się wyraźnie socjalny podział uczniów, a zarazem poszczególne kierunki w ramach swego szkolnego programu nauczania przeznaczają na nauczanie monitorowanych przedmiotów różną ilość czasu i poświęcają im mniej lub więcej uwagi.

Bardzo interesujące jest stwierdzenie, że chociaż absolutna odchyłka jest często wyraźna, całkowita odchyłka nie jest tak duża, a więc „lepsze” szkoły przesuwają się na „gorsze” pozycje i na odwrót. Kombinacja obu współczynników określa zawsze wielkość zmiany i to, czy zmiany była na plus czy na minus.

Na końcu można sformułować pewne przesłanki, które wymagają jednak kolejnego empirycznego sprawdzenia, kiedy to wyniki wartości dodanej będą dostatecznie wiarygodne i będą mogły być użyte w procesie wykorzystania odpowiedzialności nauczycieli i szkół.

- Wyniki będą spójne podczas powtarzanych pomiarów (autorzy artykułu będą mieli wkrótce do dyspozycji wyniki drugich pomiarów na podstawie danych z lat od 2008 do 2011).



- Wyniki nie będą wykazywać wielkich różnic przy aplikowaniu różnych modeli uwzględniających tzw. kontekstualne zmienne.
- Wyniki będą w znacznym stopniu zgodne z innymi narzędziami ewaluacji, np. z orzeczeniami superwizora.
- Wyniki powinny współgrać z oceną własną szkoły.
- Wyniki nie będą znacząco spadać, jeżeli szkoła rozpoczęła działania naprawcze i korelacyjne.
- Wyniki uzyskane podczas powtarzanych międzynarodowych badań będą współgrać z wynikami wartości dodanej.

Droga do sprawdzenia przewidywanych korelacji wartości dodanej z kolejnymi podanymi wyjściami ewaluacyjnymi będzie dosyć złożona, jednak obiektywne wyniki badań mogą podkreślić wiarygodność i efektywność wartości dodanej lub zwrócić uwagę na jego ograniczone wykorzystanie w sferze odpowiedzialności nauczycieli i szkół za wyniki ich starań edukacyjnych.