

prof. zw. dr hab. Bolesław Niemierko

Czy egzaminy zakłócają pracę nauczyciela?

Referat oparty na badaniach terenowych Wydziału Badań i Analiz OKE w Gdańsku

Referat podejmuje problem sprzeczności między egzaminowaniem zewnętrznym osiągnięć uczniów a ocenianiem społeczno-wychowawczym w szkołach oraz, co za tym idzie, niechęci wielu nauczycieli w kraju i za granicą do edukacyjnej diagnozy unormowanej opartej na pomiarze dydaktycznym.

Sposobem przezwyciężenia tej sprzeczności może być ukierunkowanie diagnozy nie na stan, lecz na rozwój osiągnięć ucznia. Autor proponuje stosowanie równoważnika klasy jako miary edukacyjnej wartości dodanej bezwzględnej. Pokazuje interpretację równoważnika na przykładzie dwunastu gimnazjów przebadanych przez Wydział Badań i Analiz Okręgowej Komisji Egzaminacyjnej w Gdańsku. Rozważa możliwość zharmonizowania skutecznego kierowania uczeniem się z zastosowaniami pomiaru dydaktycznego w systemie edukacji.

I. Uczenie się i rozwój

Uczenie się jest opanowywaniem – emocjonalno-motywacyjnym i poznawczym – nowych czynności. Jest zarówno biologicznym mechanizmem przystosowawczym jednostki oraz społecznym mechanizmem przystosowawczym grupy, jak i celowo zorganizowanym działaniem pedagogicznym wyprzedzającym aktualne potrzeby młodych pokoleń. W każdej z tych odmian może przynieść rozwój jednostki i społeczeństwa.

Jako **rozwój**, przypomnę, traktować będziemy (1) ciąg zmian długotrwałych i (2) względnie nieodwracalnych, (3) uporządkowanych w sekwencję rosnącą, a (4) spowodowanych wewnętrznymi siłami jednostki lub grupy społecznej (Brzezińska, 2000, rozdz. 2; Sztompka, 2002, s. 440). Rozwój w tym ujęciu jest zatem zawsze (1) powolny, (2) trwały, zdobywany na lata, (3) pozytywny i (4) w wysokim stopniu samodzielny, niemożliwy do zapewnienia z zewnątrz. Ma wiele postaci – dotyczy zarówno ucznia poddawanego oddziaływaniu szkoły, jak i np. kraju, który wstąpił do Unii Europejskiej. Jeśli wiedzy nie można uczniowi przekazać „jak paczki nadanej na pocztę”, to tym bardziej nie potrafimy przekazać mu rozwoju.

II. Uczenie się a egzamin

Egzamin jest sprawdzaniem i ocenianiem osiągnięć ucznia wyodrębnionym w procesie kształcenia i właśnie to wyodrębnienie jest powodem do jego krytyki. W wyobrażeniu uczniów, ich rodziców, nauczycieli i społeczeństwa może jawić się on jako „obce ciało”, drażniące uczestników i degradujące proces współdziałania. Dotyczy to zwłaszcza egzaminów zewnętrznych, przygotowywanych, prowadzonych i interpretowanych przez osoby nieznające osobiście ani uczniów, ani nauczycieli, ani historii ich współpracy: porozumień, zobowiązań, wzajemnych ustępstw, dorobku.

Istotny dla diagnostyki zorientowanej pomiarowo jest ubytek trafności egzaminów zewnętrznych i wszelkich testów standaryzowanych w stosunku do obserwacji i innych metod diagnostycznych stosowanych w sposób ciągły przez nauczycieli. Chodzi tu o **trafność programową**, rozumianą jako zgodność treści pomiaru z **programem zrealizowanym** (*implemented curriculum*), który – jak to boleśnie pokazały doświadczenia ostatniego roku – często nie pokrywa się z **programem zamierzonym** (*intended curriculum*) zadekretowanym w podstawie programowej.

Przypomnę tu piękne dzieło dr Arlety Poręby-Konopczyńskiej (2001), która adaptowała model Petera Knighta (2000) do polskiej terminologii i metaforyki. Przedstawia to Tab. 1.

Tabela 1. Rzetelność i trafność egzaminów szkolnych (według A. Poręby-Konopczyńskiej)

Właściwość egzaminu	Wysoka rzetelność	Umiarkowana rzetelność	Niska rzetelność
Wysoka trafność	Perły profesjonalizmu	Nauczyciel – światły profesjonalista	Nauczyciel obdarzony intuicją pedagogiczną
Umiarkowana trafność	Egzaminy zewnętrzne	Nauczyciel – dobry rzemieślnik	Codzienna praktyka szkolna
Niska trafność	P o l a m i n o w e e g z a m i n o w a n i a !		

Trafność egzaminu zewnętrznego, którego przebieg jest regulowany spoza nauczycielskiego systemu kształcenia, jest ograniczona jego nieuniknionym wyobcowaniem z tego systemu. Zdarza mu się nawet ześlizgnąć na „pole minowe”... Typowe egzaminy wewnętrzne, samodzielnie regulowane przez nauczyciela, odbywają się natomiast przy niskiej lub umiarkowanej rzetelności, ale umiarkowanej lub wysokiej trafności sprawdzania osiągnięć uczniów.

III. Zarzuty wobec edukacyjnej diagnostyki unormowanej

Pod względem metodologicznym możemy rozróżnić **edukacyjną diagnostykę nieformalną**, uprawianą na co dzień przez nauczycieli-wychowawców, oraz **edukacyjną diagnostykę unormowaną**, będącą rozpoznawaniem wyników uczenia się dokonywanym z zastosowaniem standaryzowanych narzędzi pomiaru (Niemierko, w druku). Te narzędzia mogą być dostępne na rynku wydawniczym, w ośrodkach metodycznych i badawczych, a także są stosowane w działalności komisji egzaminacyjnych.

Edukacyjna diagnostyka unormowana w postaci obowiązkowego **badania osiągnięć uczniów** (*high-stakes testing*) jest szeroko stosowana w Stanach Zjednoczonych i w innych krajach anglosaskich. W tych krajach notuje się opór nauczycieli wobec testów standaryzowanych (Goslin, 1967; Plake i in., 1993). Jego przyczyną jest **efekt zwrotny**, polegający na wpływie testu na treść i formę kształcenia poprzedzającego egzamin. Uczniowie, ich rodzice i opinia publiczna przeceniają znaczenie pomiaru dydaktycznego. Domagają się od szkoły „nauczania według testu”, co zwięża treść kształcenia i rujnuje jego metody. Wielu nauczycieli unika standaryzowanej diagnozy osiągnięć uczniów i stara się ograniczyć jej znaczenie edukacyjne. Lista zarzutów wobec tego rodzaju egzaminowania zewnętrznego ukształtowała się następująco (Smith i Rottenerg, 1991; Niemierko, 2007, s. 359):

1. redukcja czasu przydzielonego na „zwykłe” nauczanie,
2. zaniedbywanie materiału pozaegzaminacyjnego przez nauczycieli,
3. dostosowywanie metod kształcenia do zadań egzaminacyjnych,
4. ograniczanie swobody wyboru przedmiotów i uczenia się,
5. ujemny wpływ etyczny na nauczycieli,
6. postrzeganie egzaminu przez uczniów, zwłaszcza młodszych, jako surowej kary,
7. budzenie w uczniach niechęci do przedmiotów szkolnych,
8. obniżanie samooceny uczniów,
9. możliwa dyskryminacja środowiskowa i etniczna uczniów,
10. przecenianie wartości zadań wyboru wielokrotnego.

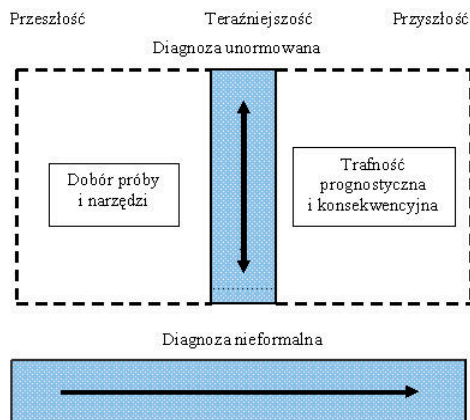
Jak wykazują badania (Plake i in., 1993), niespełna 70% badanych nauczycieli amerykańskich zaliczyło przedmiot „pomiar dydaktyczny” na uczelni lub w toku studiów podyplomowych, ale ich umiejętności pomiarowe niewiele przewyższały umiejętności pozostałych nauczycieli. Ogółem 86% wyraziło przekonanie o dużym znaczeniu nieformalnych testów nauczycielskich dla podnoszenia jakości kształcenia, ale tylko 34% powiedziało to samo o testach standaryzowanych, a kursy pomiaru dydaktycznego nie miały istotnego wpływu na te opinie.

Z zarzutami wobec diagnostyki unormowanej walczą autorzy zagraniczni (Cizek, 2001, patrz: Niemierko, 2006, s. 360) i polscy (Groenwald, 2001), ale i oni muszą przyznać, że jest w nich wiele słuszności. *W reformowanej oświacie* – napisała Maria Groenwald (s. 343) – *eksponuje się technologiczne podejście*

do procesu edukacji, zaś pomija rolę, jaką w tym procesie odgrywają nauczyciele. Nie docenia się wartości ich pracy, a to obniża autorytet, społeczny prestiż i godność pedagogów. Wywiązuje się bezwzględna rywalizacja o wyniki egzaminu, która przed uczniami otwiera drzwi do edukacji na wyższym szczeblu, nauczycieli zaś porządkuje na skali skuteczności nauczania (tamże, s. 345). To wymuszone współzawodnictwo nie sprzyja postępowi pedagogicznemu, wręcz przeciwnie, paraliżuje inicjatywę, twórczość i oddolną aktywność nauczycieli-nowatorów, stwarza w nich silne poczucie beznadziejności i niemocy wobec struktur oświatowych (s. 348). Jeśli nawet uznamy, że te szkody społeczne są przerysowane, to liczyć się musimy z takim właśnie spostrzeganiem egzaminu doniesłego przez wielu nauczycieli nie tylko w pierwszych latach wdrażania systemu.

IV. Potrzeba miar rozwoju

Znaczna część uprzedzeń nauczycieli do egzaminacyjnej ingerencji z zewnątrz wynika ze sprzeczności między diagnozą unormowaną, dokonywaną poprzecznie, w ujęciu przekrojowym, a uprawnianą przez nich na co dzień diagnozą nieformalną, podłużną, wiążącą przeszłość, teraźniejszość i przyszłość uczenia się jednostki i grupy. Zasadniczą różnicę między diagnozą unormowaną i diagnozą nieformalną w edukacji stanowi więc zakres analiz. Można to przedstawić jak na Rys. 1.



Rysunek 1. Porównanie zakresu czasowego diagnozy unormowanej i diagnozy nieformalnej w edukacji

Diagnoza unormowana obejmuje aktualne właściwości badanych obiektów w ujęciu przekrojowym dla odpowiedniej populacji. Opisuje rzeczywistość, posługując się obrazem całego terytorium występowania zjawiska, przedstawiając zarejestrowane fakty na tle właściwości populacji. Ilustruje to pokolorowana kolumna „Teraźniejszość”, a w niej dwukierunkowa strzałka symbolizująca tendencję do poszerzania informacji.

W diagnostyce unormowanej abstrahujemy od wcześniejszej znajomości obiektów, a więc najlepiej jej dokonywać siłami zewnętrznymi. Z przeszłości czerpiemy jedynie wiedzę umożliwiającą właściwy dobór próby obiektów i sytuacji (narzędzi), a przyszłość, stanowiąca problematykę **trafności konsekwencyjnej** diagnozy, jest terenem ostrożnych przewidywań rozwoju zjawisk oraz skutków podejmowanych decyzji.

Diagnoza nieformalna nie ogranicza się do terażniejszości. Obejmuje także przeszłość małej grupy obiektów, o której diagnosta ma wcześniejszą wiedzę umożliwiającą mu przyczynową interpretację poznawanych zjawisk oraz przyszłość, przewidywaną na podstawie znajomości ich kontekstu. Opisuje przeszłość, terażniejszość i przyszłość w powiązaniu, co wzmacnia wyciągane wnioski, lecz wywołuje wrażenie ulegania uprzedzeniom. Wąski zakres takiej diagnozy, wspieranie się przeszłością w jej dokonywaniu i wybieganie w przyszłość ilustruje pokolorowany prostokąt z poziomą strzałką pokazującą kierunek wnioskowania. Przyszłość jest w niej polem osobistej odpowiedzialności pedagoga, co silnie łączy diagnozę nieformalną od działania edukacyjnego wykonawcy, wywołując niekiedy efekt **samospełniającego się proroctwa**.

Nauczyciel może poszerzyć pole oceniania, posługując się unormowanymi testami standaryzowanymi, ale stanowi to dla niego duże obciążenie. Bardziej ekonomiczne wydaje się zbliżenie dwu rodzajów diagnozy edukacyjnej przez nadanie charakteru podłużnego egzaminom zewnętrznym.

V. Wartość dodana bezwzględna

Do niedawna, niemal do końca XX wieku, uważano podłużne porównywanie osiągnięć ucznia za bezpłodne ze względu na odmienną treść kształcenia w kolejnych klasach. Można było, co najwyżej, porównywać pozycję, a więc ustalać **wartość dodaną względną** tych osiągnięć: w centylach, staninach lub innych jednostkach skal różnicowych (Niemierko, 2005b). To podejście znalazło już zastosowanie w analizach wyników egzaminów zewnętrznych w Polsce (Szmigiel i Rappe, 2005a, 2005b, 2006).

Gdy jednak silniej niż materiał przedmiotowy zaczęły być akcentowane ogólnorozwojowe cele kształcenia, nabrała znaczenia **kumulatywność osiągnięć uczniów** rozumiana jako możliwość traktowania osiągnięć łącznych ucznia w wybranym okresie jako sumy osiągnięć etapowych, uzyskiwanych w kolejnych odcinkach czasu. Kumulatywność dotyczy głównie umiejętności, takich jak czytanie, pisanie, liczenie, rozumowanie, matematyzacja zagadnień, interpretacja zjawisk przyrodniczych, interpretacja zjawisk społecznych. Można przyjąć, że te umiejętności są rozwijane w sposób **ciągły**, przez cały okres kształcenia ogólnego.

Osiągnięcia kumulatywne mogą być mierzone jedną wersją testu lub jego wersjami równoległymi w kolejnych klasach lub nawet w szkołach kolejnego szczebla i mogą być przedstawiane we wspólnej skali. Powstaje więc możliwość szacowania **wartości dodanej bezwzględnej**, jako różnicy poziomu umiejętności, ustalonej na podstawie wyników pomiaru tej samej umiejętności w wybranych fazach procesu uczenia się.

Najdokładniejsze oszacowanie wartości dodanej bezwzględnej uzyskujemy przy zastosowaniu tego samego testu we wszystkich klasach (kolejnych latach uczenia się danej umiejętności). Takie podejście jest jednak stosowane rzadko, a to ze względu na duże różnice osiągnięć między kolejnymi klasami powodujące, iż zadania zbyt trudne dla jednych i zbyt łatwe dla drugich traciłyby moc różnicującą. Stosuje się więc zastępczo różne modele **skalowania wertykalnego** (Kolen, 2006, s. 171 – 180), polegające na wiązaniu wyników w kolejnych klasach za pomocą **testek kotwiczących**, czyli części wspólnych dwu lub więcej testów pozwalających na bezpośrednie porównanie wyników dwu grup uczniów o różnym poziomie osiągnięć.

Serie testów standaryzowane przez wielkie przedsiębiorstwa pomiarowe, takie jak Educational Testing Service w Princeton, zawierają testy powiązane testkami kotwiczącymi „na zakładkę” (np. klasa I i II, II i III itd.) od przedszkola do zakończenia szkoły średniej (*K – 12 range*), ale w Polsce takich serii jeszcze nie mamy. Pionierskie próby skalowania wertykalnego testów przeprowadzili Janusz Mulawa i Dariusz Sroka w Instytucie Badań Kompetencji w Wałbrzychu (2003). Zastosowali w klasie VI szkoły podstawowej oraz w klasach I i II gimnazjum test „Moja przedsiębiorczość w Unii Europejskiej”, „Test predyspozycji językowej” i test ogólnoprzyrodniczy. Odnotowali kilkuprocentowe zaledwie różnice średnich osiągnięć uczniów kolejnych klas.

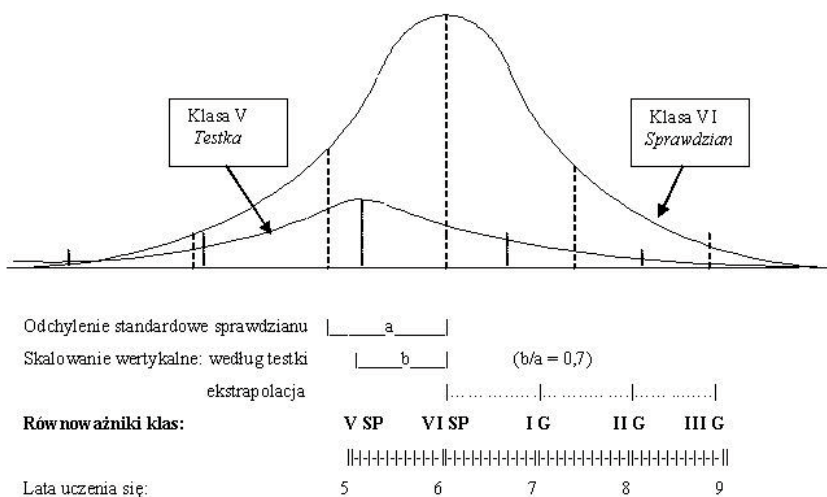
VI. Tymczasowa konstrukcja skali równoważników klasy

Skąd wziąć miarę wartości dodanej osiągnięć mierzonych egzaminem zewnętrznym w Polsce, uzyskanej w toku jednego roku nauki? Potrzebne nam dane przyniosło zastosowanie w charakterze testki kotwiczącej arkuszy sprawdzianu z lat 2003, 2004 i 2005 do prób reprezentatywnych uczniów klasy V szkoły podstawowej o liczebności około 1800 (Niemierko, 2004; Niemierko, 2005a). Ten pomiar okazał się dostatecznie rzetelny ($rtt > 0,85$) na to, by uczynić go podstawą oszacowania **efektu standardowego jednej klasy szkolnej**. Ten efekt miał wartość od 0,67 w roku 2003 do 0,85 w roku 2005, co znaczy, że postęp roczny między klasą V a VI wyniósł średnio około 0,75 odchylenia standardowego wyników sprawdzianu w klasie VI. Oszacowanie 0,7 można przyjąć, przez ekstrapolację, jako miarę wartości dodanej bezwzględnej także i w następnych klasach (od I do III w gimnazjum), z zastrzeżeniem, że stanowi to wielkie uproszczenie, gdyż kolejne przyrosty wyników zapewne maleją (łatwość testu nie może przekroczyć wartości 1), a nadto zależą od przedmiotu kształcenia i od jakości testu.

Oszacowanie efektu standardowego jednej klasy szkolnej pozwala nam na przedstawienie osiągnięć uczniów w skali **równoważników klasy** (*RK*), czyli przez podanie liczby lat i miesięcy uczenia się określonych umiejętności, w których przeciętny uczeń uzyskuje dany poziom osiągnięć. Rok szkolny dzielimy umownie na dziesięć miesięcy (od września do czerwca), zakładając – dla uproszczenia obliczeń – równy przyrost osiągnięć ucznia każdego miesiąca i brak wpływu wakacji na poziom osiągnięć.

Na przykład uczeń Janek Kowalski, który miał wynik sprawdzianu po szkole podstawowej dokładnie 7 staninów, a więc wyższy o jedno odchylenie standardowe od średniej, wyprzedza średnią krajową o $1 : 0,7 = 1,4$ równoważnika klasy, a ponieważ szkoła podstawowa jest sześcioklasowa, równoważnik klasy jego wyniku wynosi $6 + 1,4 = 7,4$. Ten równoważnik znaczy, że Janek uzyskał wynik na poziomie siedmiu lat i czterech miesięcy uczenia się, a więc na poziomie średniego wyniku krajowego w II klasie gimnazjum z końcem grudnia.

Całość wykonanej operacji budowania skali równoważników klasy i ustalania równoważnika klasy osiągnięć Janka w szkole podstawowej przedstawia Rys. 2.



Rysunek 2. Równoważniki klas szkolnych I – III gimnazjum (G) zbudowane na podstawie ekstrapolacji skalowania wertykalnego testki kotwiczącej wyniki klasy V i VI szkoły podstawowej (SP)

Na Rys. 2. widzimy dwa rozkłady wyników testowania: górą – rozkład wyników krajowego sprawdzianu po VI klasie szkoły podstawowej o odchyleniu standardowym oznaczonym jako a , dołem – rozkład wyników testki kotwiczącej klasę V z klasą VI szkoły podstawowej, którego średnia w klasie V jest niższa od średniej w klasie VI o $0,7$ odchylenia standardowego wyników klasy VI. Wielkość $0,7$, oznaczona na rysunku jako b , stanowi efekt standardowy jednej klasy szkolnej. Posługując się tą wielkością, oszacowano równoważniki klas I – III w gimnazjum jako ekstrapolację efektu standardowego u progu gimnazjum na trzy kolejne lata uczenia się mierzonych umiejętności.

Przyjęte założenia i operacje wyrażają następujące wzory:

1.

$$RK_{SP} = \frac{z_S}{0,7} + 6$$

gdzie RK_{SP} to równoważnik klasy ucznia kończącego szkołę podstawową, a z_S to wartość standardowa jego wyniku sprawdzianu,

2.

$$EWD_G = \frac{z_{EG} - r_{EG,S} \cdot z_S}{0,7} + 3$$

gdzie EWD_G to wartość dodana (bezwzględna) w gimnazjum, z_{EG} to wartość standardowa wyniku egzaminu gimnazjalnego, $r_{EG,S}$ – współczynnik korelacji wyników egzaminu gimnazjalnego i sprawdzianu, z_S to wartość standardowa wyniku sprawdzianu,

3.

$$RK_G = RK_{SP} + EWD_G$$

gdzie nowym symbolem jest RK_G , równoważnik klasy absolwenta gimnazjum.

Dysponując skalą równoważników klasy szkolnej dla gimnazjum, możemy oszacować postęp Janka w uczeniu się podstawowych umiejętności w trzech klasach gimnazjum. Gdyby jego wynik egzaminu gimnazjalnego wyniósł 5,0 staninów, co odpowiada średniej 9,0 lat uczenia się, to wartość dodana bezwzględna jego wyniku wyniosłaby tylko $9,0 - 7,4 = 1,6$ klasy szkolnej.

Jakiego równoważnika klasy moglibyśmy oczekiwać od Janka, gdyby utrzymał swoje wcześniejsze tempo rozwoju? Ponieważ wynik oczekiwany Janka (uwzględniający regresję ku średniej) wynosi 6,2 stanina, to $9,0 + 1,4 (6,2 - 5,0) = 10,7$ lat uczenia się. Jego postęp, niewiele większy niż półtora roku (z 7,4 do 9,0) jest o ponad połowę mniejszy od oczekiwanego (z 7,4 do 10,7), wynoszącego ponad trzy lata ($10,7 - 7,4 = 3,3$ roku). Porównując te wartości, należałoby uznać, że w gimnazjum Janek bardzo niewiele dodał do umiejętności, jakie zdobył w szkole podstawowej. Ustalenie przyczyn tego spowolnienia rozwoju powinno być przedmiotem dalszych procedur diagnostycznych.

VII. Niska rzetelność równoważników klasy

Przedstawianie wyników egzaminu w skali równoważników klasy, a więc w latach i miesiącach uczenia się, jest wygodne dla nieprofesjonalnego użytkownika, ale naraża go na **złudzenie** nierealnej precyzji pomiaru.

Przy współczynniku rzetelności sprawdzianu po szkole podstawowej wynoszącym około 0,80 błąd standardowy pomiaru przekracza 6 miesięcy, a więc tylko różnice 1,3 roku i większe są statystycznie istotne. Wynik Janka Kowalskiego (1,4 roku powyżej średniej) niewiele przekracza tę wartość. Wraz z nim tylko około 1/3 egzaminowanych uzyskuje wyniki istotnie różne (wyższe lub niższe) od średniej.

Bardziej precyzyjny jest, czterokrotnie dłuższy, egzamin gimnazjalny, którego współczynnik rzetelności przekracza zwykle 0,90, a więc błąd standardowy kurczy się do 4,5 miesiąca, a od średniej odróżnia się około połowy egzaminowanych. Także i ta wielkość jednak powinna nas powstrzymać od pochopnego wnioskowania o różnicach.

W obydwu egzaminach paromiesięczne (jednostaninowe, nieprzekraczające 5 surowych punktów sprawdzianu i każdej z części egzaminu gimnazjalnego) różnice osiągnięć pojedynczych uczniów nie powinny być brane pod uwagę przy podejmowaniu ważnych decyzji pedagogicznych. A nawet wtedy, gdy mierzone umiejętności ucznia wykazują istotne różnice w stosunku do średniej, wynoszące jedną lub więcej klas szkolnych, wnioskowanie, że przesunięcie go do innej klasy, wyższej lub niższej, byłoby dla niego pożyteczne, jest ze względów wychowawczych (zmiana grupy rówieśników) i dydaktycznych (brak ciągłości materiału programowego) nieuzasadnione.

VIII. Zależność edukacyjnej wartości dodanej od szkoły

Ponieważ oszacowania statystyczne zyskują na dokładności w miarę zwiększania liczebności zbioru, którego dotyczą, równoważniki klas szkolnych są użytecznymi wskaźnikami osiągnięć grup uczniów wewnątrz szkół i między szkołami. Pionierskie analizy tego rodzaju podjęły Barbara Przychodzeń i Teresa Kutajczyk w Okręgowej Komisji Egzaminacyjnej w Gdańsku. Zestawiły dane uzyskane w jedenastu gimnazjach wiejskich (Kutajczyk i Przychodzeń, 2008) oraz jednym wysoko selekcyjnym gimnazjum z dużego miasta (oznaczonym jako G12). Najważniejsze wyniki tych analiz są zebrane w Tab. 2.

Tabela 2. Wartość dodana w gimnazjum według wyników sprawdzianu po szkole podstawowej w skali równoważników klasy (lat i miesięcy uczenia się)

Szkoła	Liczba uczniów klasy III	Średnia u progu (RKSP)	Edukacyjna wartość dodana według osiągnięć u progu (EWD _i)								Łącznie (RKG)	Korelacja RKSP, EWD
			<3,5	3,5-4,4	4,5-5,4	5,5-6,4	6,5-7,4	7,5-8,4	≥ 8,5	Ogółem		
G03	36	5,1	3,1	2,4	2,2	2,2				2,4	7,5	-0,26
G01	78	5,3	3,2	2,7	2,8	3,0	3,5	3,1		3,0	8,3	-0,04
G05	42	5,4	3,5	3,4	3,0	2,7	3,1	3,6		3,1	8,5	-0,11
G10	29	5,4		2,7	2,6	2,7	2,9			2,7	8,1	0,28
G08	54	5,6	3,0	3,2	3,1	3,9	3,7	3,8		3,5	9,1	0,37
G02	38	5,8		3,0	2,8	2,7	3,0	2,9		2,9	8,7	0,05
G07	21	5,9			2,6	3,0	3,7			3,0	8,9	0,37
G04	87	5,9	3,0	2,1	2,5	2,5	2,6	3,5	3,0	2,6	8,5	0,29
G06	27	6,0		3,2	3,1	3,4	2,9	3,1		3,1	9,1	-0,13
G09	63	6,5		3,1	3,1	2,6	3,2	3,2	3,3	3,0	9,5	0,09
G11	12	6,9			3,1		3,7	3,6		3,5	10,4	0,35
G12	48	7,8					3,6	3,3	3,0	3,3	11,1	-0,40
Ogółem	535	5,9	3,2	2,9	2,8	2,9	3,3	3,3	3,1	3,0	8,9	0,15

Liczba uczniów w grupie	35	69	101	138	87	91	14	535	x	x
Korelacja RKSP, EWD _i /RKG	-,39	,24	,56	,13	,34	,00	-,21	,53	,96	x
Korelacja EWD _i , EWD Ogółem	,15	,80	,87	,88	,78	.48	,08	x	x	x
Korelacja EWD _i , RKG	-,09	,66	,77	,56	,53	,14	-,13	,75	x	x

Tłustym drukiem wyróżniono współczynniki korelacji istotne statystycznie.

W górnej części Tab. 2. gimnazja są uporządkowane według średniego wyniku sprawdzianu po szkole podstawowej (RKSP), od 5 lat i 1 miesiąca (w G03) do 6 lat i 9 miesięcy (w G11) w gimnazjach wiejskich i do 7 lat i 8 miesięcy w elitarnym gimnazjum miejskim (G12). W kolumnach środkowej części zamieszczono tu średnie edukacyjnej wartości dodanej (EWD_i) w grupach uczniów dobranych w przedziały według wyniku sprawdzianu po szkole podstawowej. Wartości grupowe składają się na edukacyjną wartość dodaną ogółem, wynoszącą od 2 lata i 4 miesięcy (w szkole G03) do 3 lat i 5 miesięcy (w szkołach G08 i G11).

Jako średni równoważnik klasy po ukończeniu gimnazjum (RKG) przedstawiono sumę średniego równoważnika klasy po szkole podstawowej i edukacyjnej wartości dodanej w gimnazjum ogółem [RKSP + EWD_{Ogółem}]. Zgodnie z prawem **efektu wachlarzowego** wyników uczenia się, rozrzut średnich osiągnięć po gimnazjum wzrósł w stosunku do szkoły podstawowej i średnie wynoszą teraz od 7 lat i 5 miesięcy (w G03) do 11 lat i 1 miesiąca (w G12).

Przegląd wierszy górnej części Tab. 2. dowodzi, że nie w każdym gimnazjum występuje efekt wachlarzowy uczenia się. Organizując zajęcia wyrównawcze i inne formy pomocy słabszym uczniom, szkoła jest w stanie zapewnić im ponadprzeciętny postęp osiągnięć i złagodzić, a nawet odwrócić ten efekt. Tylko dwa gimnazja (G08 i G04) wykazały istotny statystycznie **dodatni efekt wachlarzowy** mierzony współczynnikiem korelacji równoważnika klasy po szkole podstawowej (RKSP) i wartości dodanej w gimnazjum (EWDG). W elitarnym gimnazjum G12 zarejestrowano istotny statystycznie **ujemny efekt wachlarzowy** , co można przypisać **efektowi pułapu** , czyli brakowi zadań dla najlepszych uczniów, a zapewne także **nadinterpretacji treści zadań** , polegającej na skłonności uczniów do podawania rozwiązań nieprzewidywanych w kluczach i schematach punktowania. Ogółem jednak, wobec 535 uczniów, dodatni efekt wachlarzowy okazał się istotny statystycznie, choć słaby, gdyż współczynnik korelacji równoważnika klasy po szkole podstawowej i wartości dodanej w gimnazjum wyniósł 0,15, a więc był niski.

Trzy gimnazja (G03, G04 i G10; wartość dodana od 2,4 do 2,7) pogłębiły lukę osiągnięć uczniów w stosunku do średniej krajowej, sześć innych (G02, G01, G07, G09, G05 i G06; wartość dodana od 2,9 do 3,1) dotrzymało kroku tej średniej, a trzy najlepsze (G12, G08 i G11; wartość dodana od 3,3 do 3,5) przyspieszyły rozwój uczniów. Na szczególne uznanie zasługuje gimnazjum G08, które przyjęło wielu słabych uczniów i potrafiło wszystkim zapewnić należyty postęp, a kandydaci przeciętni (w skali kraju) i ponadprzeciętni uzyskali w nim bardzo duży wzrost umiejętności.

Dolna (oddzielona) część Tab. 2. dotyczy korelacji średnich wartości dodanych grup uczniów utworzonych według równoważnika klasy „u progu” szkoły (EWDi) z trzema kryteriami rozwoju uczniów: wynikami w szkole podstawowej (RKSP), średnią wartości dodanej (EWD_{Ogółem}) i wynikami po gimnazjum (RKG). Na podstawie tej korelacji spróbujemy odpowiedzieć na pytanie o **zależność postępów ucznia** (edukacyjnej wartości dodanej) **od poziomu pracy szkoły** (od jej wyników „u progu” i wyników egzaminu gimnazjalnego).

Przeglądając kolumnami obie części tabeli, spostrzegamy następujące prawidłowości:

1. Uczniowie o najniższych wynikach w szkole podstawowej ($RKSP < 3,5$; niewiele ponad 5% ogółu) uzyskują w gimnazjum nieco wyższy od przeciętnej postęp (EWDi = 3,2) niezależnie od poziomu pracy szkoły (współczynniki korelacji są nieistotne statystycznie).
2. Uczniowie o niskich i przeciętnych wynikach w szkole podstawowej ($3,5 \leq RKSP < 6,5$; około 60% ogółu) uzyskują w gimnazjum postęp nieco niższy od przeciętnej (EWDi = 2,9) skorelowany z wartością dodaną w szkole i z równoważnikiem klasy po gimnazjum (pięć współczynników korelacji w granicach 0,6 – 0,9).
3. Uczniowie o wysokich wynikach w szkole podstawowej ($6,5 \leq RKSP < 7,5$; około 15% ogółu) uzyskują w gimnazjum wyższy od przeciętnej postęp (EWDi = 3,3) skorelowany z wartością dodaną w szkole i z równoważnikiem klasy po gimnazjum (współczynniki korelacji wynoszą od 0,5 do 0,8).
4. Uczniowie o najwyższych wynikach w szkole podstawowej ($RKSP \geq 7,5$; około 20% ogółu) uzyskują w gimnazjum nieco wyższy od przeciętnej postęp (EWDi = 3,2) niezależnie od poziomu pracy szkoły (współczynniki korelacji z wynikami egzaminów są bliskie zeru).
5. Średnia edukacyjna wartość dodana w gimnazjum nieco niżej koreluje ze średnim równoważnikiem klasy przyjętych kandydatów ($r = 0,53$) niż ze średnim równoważnikiem klasy absolwentów ($r = 0,75$).
6. Korelacja średnich równoważników klasy przyjętych kandydatów ze średnimi równoważnikami klasy po gimnazjum (oszacowanymi z wykorzystaniem regresji liniowej) jest bardzo wysoka ($r = 0,96$).

Prawidłowości 1 – 4 potwierdzają, niestety, zasadność gorliwego poszukiwania przez uczniów (ich rodziców) gimnazjów o najwyższych pozycjach w rankingu wyników egzaminacyjnych („tabeli ligowej”) w stosunku do uczniów o **niskich, przeciętnych i wysokich** wynikach w szkole podstawowej (około $\frac{3}{4}$ ogółu), ale uczniowie o **najniższych i najwyższych** wynikach (około $\frac{1}{4}$ ogółu) uzyskują duży postęp (w zakresie programowym egzaminu gimnazjalnego) bez względu na tak określanej poziom pracy szkoły. Strategia wyboru najlepszej szkoły oparta na rankingu wyników końcowych może być jednak wzbogacona o kryterium edukacyjnej wartości dodanej (charakterystyczne przypadki gimnazjów G08 i G04).

Prawidłowości 5. i 6. potwierdzają znaczenie doboru kandydatów o możliwie wysokich wynikach sprawdzianu po szkole podstawowej dla pozycji gimnazjum w rankingu wyników egzaminu końcowego. Także i tutaj jednak **wiele zależy od szkoły** (np. w gimnazjum G05 uczniowie o podobnych wynikach sprawdzianu u progu uzyskują wartość dodaną średnio o cztery miesiące uczenia się większą niż w gimnazjum G10).

Ze względu na niewielką liczebność próby szkół (12 gimnazjów), jej ograniczoną reprezentatywność (prawie wyłącznie szkoły wiejskie) i uproszczoną procedurę budowania norm równoważnikowych (Rys. 2.), wszystkie wnioski oparte na referowanych tu badaniach wymagają jeszcze potwierdzenia. By go dokonać, nie musimy operować równoważnikami klas szkolnych, wystarczy przedstawić wyniki punktowe dwu egzaminów w dowolnej skali standardowej. Wartości dodana osiągnięć uczniów w latach i miesiącach uczenia się lepiej jednak przemawia do wyobraźni użytkownika wyników diagnozy niż odchylenie standardowe wyników egzaminu.

IX. Edukacyjna wartość dodana w strefach osiągnięć uczniów

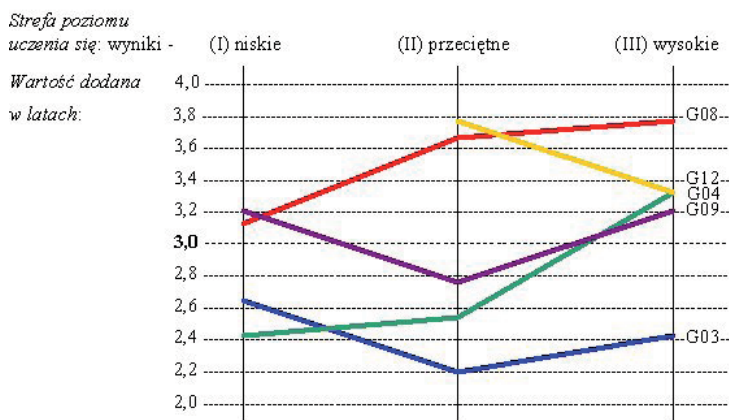
Skoro stwarzanie warunków rozwoju osiągnięć ucznia jest naczelnym zadaniem placówki edukacyjnej, **ocena jakości pracy szkoły** zyska na trafności, gdy dokonamy jej nie na podstawie jednorazowego, końcowego pomiaru osiągnięć uczniów, lecz na podstawie edukacyjnej wartości dodanej. Takie podejście jednak także sprzyja szkołom, które pracują w bardziej korzystnym środowisku i przyjmują lepiej przygotowanych kandydatów, gdyż efekt wachlarzowy powoduje, że rozwijają się oni, statystycznie biorąc, szybciej. Trzeba zatem znaleźć sposób na to, by ocenę pracy szkół uniezależnić od tego efektu.

Rozwiązaniem problemu może być ustalanie wartości dodanej osiągnięć uczniów w podziale uczniów na **strefy poziomu uczenia się**, czyli na grupy uzyskujące istotnie różne wyniki egzaminacyjne. Ze względu na niską rzetelność wartości dodanej Teresa Kutajczyk i Barbara Przychodzeń (2008) wyznaczyły tylko trzy takie strefy: (I) **wyników niskich**, w których równoważnik klasy wyników sprawdzianu po szkole podstawowej był niższy o ponad rok od normy ($RKSP < 5,0$), (II) **wyników przeciętnych**, w których ten równoważnik nie przekraczał \pm jednego roku ($5,0 \leq RKSP \leq 7,0$) i (III) **wyników wysokich**, w których równoważnik był wyższy o ponad rok od normy ($RKSP > 7,0$). W pięciu gimnazjach wybranych z Tab. 2. dało to rozkłady procentowe uczniów przedstawione w Tab. 3.

Tabela 3. Uczniowie według strefy poziomu uczenia się w gimnazjum i równoważnika klasy szkolnej po jego ukończeniu

Gimnazjum	Wyniki sprawdzianu po szkole podstawowej				Równoważnik klasy absolwentów
	(I) niskie	(II) przeciętne	(III) wysokie	razem	
G03	42%	53%	5%	100%	7,5
G08	31%	52%	17%	100%	9,1
G04	21%	63%	16%	100%	8,5
G09	16%	49%	35%	100%	9,5
G12	0%	12%	88%	100%	11,1
Ogółem	25%	49	26%	100%	8,8

Jak widzimy w Tab. 3., wyniki egzaminu gimnazjalnego, przedstawione jako średnie równoważniki klasy absolwentów poszczególnych szkół, dość dokładnie odpowiadają rozkładom procentowym uczniów według poziomów uczenia się. Prawidłowość narusza tylko Gimnazjum Nr 8, w którym jest więcej uczniów o niskich wynikach sprawdzianu niż uczniów o wysokich wynikach sprawdzianu, a średni równoważnik klasy absolwentów przewyższa normę 9,0 (dziewięć lat uczenia się ogółem). Czy jednak wartość dodana w strefach poziomu uczenia się układa się równie harmonijnie? Odpowiedź na to pytanie przedstawia Rys. 3.



Rysunek 3. Wartość dodana w gimnazjum według strefy poziomu uczenia się

Na podstawie Rys. 3. możemy poczynić następujące spostrzeżenia o wartości dodanej osiągnięć uczniów (*EWD*) w pięciu gimnazjach:

1. W dwu szkołach, G03 i G04, edukacyjna wartość dodana ogółu absolwentów jest niższa niż w najlepszych pod tym względem szkołach - G12 i G08. Ta różnica jest istotna statystycznie, a więc wybór gimnazjum **ma wpływ** na rozwój osiągnięć ucznia.

2. W **strefie I** (wyniki niskie) rozrzut wartości *EWD* jest stosunkowo mały, ale tylko uczniowie dwu szkół – G09 i 08 nie zwiększają swojej luki rozwojowej w stosunku do przeciętnej ($EWD > 3,0$).

3. W najliczniejszej **strefie II** (wyniki przeciętne) rozrzut wartości *EWD* zgodnie z prawem efektu wachlarzowego rośnie, a szkoły G12 i G08 zdecydowanie górują nad resztą. Tam więc należałoby kierować średnio przygotowanego (i średnio uzdolnionego) ucznia w nadziei na przyśpieszenie rozwoju jego osiągnięć. Zauważmy, że szkoły G09 i G04, przyjmujące lepiej przygotowanych kandydatów (Tab. 3., kolumna *RK_{SP}*), pozostają daleko w tyle za G08.

4. W **strefie III** (wyniki wysokie) dużą przewagę *EWD* zyskała szkoła G08. Zaskakujący jest zaledwie przeciętny wskaźnik (3,3) elitarnego G12. Wytłumaczyć to można tylko tym, że w dużej grupie (42 uczniów), jaka w tej szkole stanowi strefę „wysokiego poziomu uczenia się”, mogło znaleźć się sporo uczniów, którzy uzyskali wysoki wynik sprawdzianu niesamodzielnie lub po dostaniu się do prestiżowej szkoły stracili zainteresowanie uczeniem się, zwłaszcza przedmiotów humanistycznych (ich średnia *EWD* w tej grupie przedmiotów egzaminacyjnych wyniosła tylko 2,85).

5. Kolorowe krzywe na wykresie przedstawiające **profile strefowe** szkół przybierają różne kształty. Silne, istotne statystycznie przyrosty osiągnięć uczniów obserwujemy między strefami I a II i III w G08 i między strefami I i II a III w G04. Te przyrosty wskazują na efekt wachlarzowy w tych szkołach: lepiej przygotowani uczniowie rozwijają się w nich szybciej. Inne „zygzaki” nie nadają się do uogólnień, ale mogą zainteresować kadre pedagogiczną placówek i wpłynąć na jej systemy pedagogiczne.

Jaka informacja o wynikach kształcenia jest bardziej potrzebna szkołom: o końcowych osiągnięciach uczniów czy o wartości dodanej (w ewentualnym układzie strefowym)? W serii badań tego zagadnienia Maria Krystyna Szmigel i Anna Rappe (2005a, 2005b, 2006) stwierdziły, że dyrektorzy szkół (gimnazjów) na ogół przychylnie traktują edukacyjną wartość dodaną jako nowy system wskaźników jakości pracy szkoły. Dostrzegają jego wartość dla szkół o dużej liczbie uczniów z rodzin dysfunkcyjnych, wyrażają jednak obawy o nadmiar statystyki edukacyjnej, o naciski rodziców, by tworzyć szkoły i klasy elitarne oraz o błędne interpretacje wartości dodanej ze strony nadzoru pedagogicznego. Autorki tak kończą jeden z raportów (2005a, s. 135):

Mądrze wykorzystana informacja o wartości dodanej może stać się elementem kształtowania polityki oświatowej, w tym służyć przeciwdziałaniu niekontrolowanej selekcji uczniów tak na szczeblu gimnazjum, jak szkolnym czy klasowym. [...] Prawdopodobnie informacja taka pozwoli na wyprowadzenie z cienia niektórych szkół i regionów, które słabo wypadają w rankingach tworzonych na podstawie średnich wyników egzaminu zewnętrznego.

X. Kierowanie szkołą i jego skutki

Kierowanie instytucją jest sprawowaniem **władzy**, to jest wywieraniem wpływu na innych pracowników przewyższającego ich wpływ na kierującego (Wojciszke, 2002, s. 403). Dyrektor szkoły może żądać od nauczycieli wykonania określonych zadań, gdyż ponosi za ich pracę odpowiedzialność wobec społeczeństwa i wobec swoich zwierzchników. Ta odpowiedzialność obejmuje przede wszystkim osiągnięcia emocjonalne i poznawcze uczniów wyznaczone obowiązującym programem kształcenia.

Dyrektor stanowi w szkole **władzę prawnocną**, co znaczy, że podejmowanie decyzji dotyczących działania instytucji jest jego prawem i obowiązkiem. Inne typy i źródła władzy przez niego sprawowanej mają mniejsze znaczenie i węższy zakres. Są to (Touhy, 2002, s. 112): **władza z nagradzania** (wysoką oceną, premią finansową), **władza z karania** (niską oceną, sankcjami), **władza z przyzwolenia** (uznanie wśród kadry), **władza z kompetencji** (rozwiązywanie problemów instytucji).

Dyrektor szkoły nie kieruje nią jednoosobowo. *Od pedagogów, podobnie jak od przedstawicieli innych inteligentnych zawodów – pisze Richard Arends w swym podręczniku „Uczymy się nauczać” (1994, s. 387) – oczekuje się nie tylko wypełniania podstawowej funkcji zawodowej (w tym przypadku nauczania uczniów), lecz także uczestniczenia w kierowaniu instytucją.* Tę funkcję spełniają rady pedagogiczne, ale o stylu kierowania placówką decyduje zasadniczo jej dyrektor.

Wyniki amerykańskich i brytyjskich badań nad „efektywnością szkoły” jednoznacznie wskazują na znaczenie kierowania placówką i szczególną rolę jej dyrektora. Czynniki skuteczności kształcenia okazują się bowiem (MacBeath i in., 2003, s. 93):

1. merytoryczne kierownictwo,
2. wspólna wizja i wspólne cele,
3. dobrze zorganizowane środowisko dydaktyczne,
4. koncentracja na uczeniu się i nauczaniu,
5. wysokie oczekiwania, ambitne zadania,
6. wzmocnienia pozytywne jako informacja zwrotna,
7. systematyczne ocenianie osiągnięć uczniów i szkoły,
8. poczucie odpowiedzialności wśród uczniów i ich wysoka samoocena,
9. jasność celów i struktury zajęć edukacyjnych,
10. podnoszenie kwalifikacji przez nauczycieli,
11. pozytykiwanie rodziców do partnerstwa ze szkołą.

XI. Dwa style kierowania

Ponieważ każda instytucja polega na (a) wykonywaniu pewnych zadań przez (b) zespół pracowników, najprostszą klasyfikację stylów kierowania nią stanowi rozróżnienie: (a) nastawienia na zadania i (b) nastawienia na zespół pracowników (Wojciszke, 2002, s. 408 – 410).

Styl zadaniowy (dominacyjny, dyrektywny) kierowania instytucją jest zorientowany na osiąganie celów jej działania i wykazuje podobieństwo do technologicznego systemu kształcenia. **Styl relacyjny** (demokratyczny, integracyjny) jest zorientowany na korzystne stosunki (relacje) w zespole pracowników i wykazuje podobieństwo do humanistycznego systemu kształcenia. Te dwa style mogą być w różnych proporcjach związane, ale połączenie intensywnego wysiłku z miłymi stosunkami w pracy bywa trudne i rzadko spotykane. Bodaj częstsze są skrajności i wypaczenia tych stylów: rygorystyczno-autokratyczne zarządzanie instytucją bez względu na potrzeby, a nawet przeciw potrzebom pracowników i liberalno-pobłażliwe zachowania zwierzchników, będące w istocie zaniechaniem kierowania (Janowski, 2002, s. 82).

Jak stwierdza Bogdan Wojciszke (tamże), poprawny styl zadaniowy kierowania instytucją owocuje **skutecznością zadaniową** w postaci osiągania jej celów, a więc w szkole – rozwoju osiągnięć uczniów na każdym poziomie uczenia się, natomiast styl relacyjny kierowania instytucją owocuje **skutecznością relacyjną** w postaci dobrej atmosfery pracy i miłych stosunków międzyludzkich, a więc w szkole – uspołecznienia nauczycieli i uczniów oraz wspólnego zadowolenia z przebiegu uczenia się.

Badania wykazują, że te dwa rodzaje skuteczności są bardzo słabo, choć zwykle dodatnio powiązane. Pogląd, że „dobra szkoła, to szkoła, którą uczeń lubi” nazywa Krzysztof Konarzewski „jednym z najbardziej uporczywych stereotypów pedagogicznych”. *Jest oczywiste – tłumaczy (2004, s. 146) – że na ogół bardziej lubiana jest szkoła nastawiona na współzycie niż na osiągnięcia. Czy jednak należy stąd wnosić, że ta pierwsza jest także bardziej efektywnym narzędziem przekazywania wiedzy? [...] Klasa nastawiona na osiągnięcia zmniejsza wprawdzie samoistną przyjemność przebywania w szkole, ale za to uruchamia inne siły, które zmuszają młodego człowieka do wytrwałego uczestnictwa w pracy szkolnej, nierozsądnie byłoby więc zakładać, że pod względem dydaktycznym będzie mniej efektywna.*

XII. U źródeł wysokiej wartości dodanej szkoły

Może okazać się, że wskaźniki wartości dodanej bezwzględnej przedstawianej za pomocą równoważników klasy wprowadzają nas w nowe układy zależności, mało widoczne w tradycyjnych badaniach stanu osiągnięć uczniów. Z tego względu interesująca będzie analiza uwarunkowań wysokiej wartości dodanej w wiejskim gimnazjum **G08** (Tab. 2., Rys. 3.). Z opracowania „Czynniki skuteczności kształcenia ogólnego w gimnazjach wiejskich” (Kutajczyk i Przychodzeń, 2008) wybieramy te zmienne niezależne, których wskaźniki dla tego gimnazjum były najwyższe lub najniższe spośród jedenastu badanych gimnazjów wiejskich, albo bliskie tym skrajnym pozycjom. Przedstawia je Tab. 4.

Tabela 4. Cechy charakterystyczne Gimnazjum Nr 8 w opinii uczniów

Pozycja ankiety	Wskaźnik aprobaty	Porównanie wskaźnika między szkołami
Rozumiem, co tłumaczy nauczyciel na lekcjach języka polskiego.*	4,0	najwyższy
Czytam lektury szkolne.	3,9	najwyższy
Interesują mnie zagadnienia z zakresu języka polskiego.	3,8	najwyższy
<i>Lekcje języka polskiego są dla mnie ciekawe.</i>	3,4	najwyższy
Uczenie się języka polskiego sprawia mi przyjemność.	3,6	drugi z najwyższych
Jestem uzdolniony w dziedzinie języka polskiego.	3,4	drugi z najwyższych
Zdobytą wiedzę stosujemy w praktyce.	3,6	drugi z najniższych
Wspólnie planujemy i organizujemy imprezy klasowe.	3,2	drugi z najniższych
W ważnych wydarzeniach klasowych uczestniczą rodzice.	2,9	drugi z najniższych
Lubię wychowawcę.	3,5	najniższy
Relacje pomiędzy nauczycielami a uczniami są dobre.	3,4	najniższy
Nauczyciele interesują się, jak radzę sobie ze stawianymi mi wymaganiami.	3,2	najniższy
W szkole panuje właściwa atmosfera.	3,5	najniższy
Nauczyciele udzielają wskazówek, jak uczyć się skutecznie.	3,1	najniższy

*) Tłustym drukiem zaznaczono pozycje, w których G08 mocno różni się od pozostałych szkół.

Obraz, jaki wyłania się ze wskaźników zebranych w Tab. 4., w równym stopniu ucieszy i zmartwi pedagoga humanistę:

1. Wszystkie pozycje, w których Gimnazjum Nr 8 góruje nad pozostałymi (sześć początkowych w tabeli) dotyczą języka polskiego. Mimo że wartość dodana w części matematyczno-przyrodniczej egzaminu (3,7) przewyższała wartość dodaną w części humanistycznej (3,3), to wytworzenie przez szkołę wiejską, przyjmującą raczej słabych kandydatów ($RKSP = 5,6$), autentycznego zaangażowania w język ojczysty i literaturę jest niewątpliwie sukcesem zatrudnionych w niej polonistów.

2. Pozostałe pozycje (w liczbie ośmiu), w których Gimnazjum Nr 8 pozostaje w tyle za innymi wiejskimi gimnazjami, mogą wprawić w zdumienie entuzjastę postępu pedagogicznego. W oczach uczniów ich szkoła jest mało sympatyczna, mało zainteresowana uczniami i ich rodzicami, nie stwarza miłej atmosfery. Malują typowy portret **szkoły tradycyjnej**, z przewagą działań poznawczych nad emocjonalno-motywacyjnymi i teorii nad praktyką, stawiającej uczniom wysokie wymagania i pozostawiającej ich samym sobie. Warto zaznaczyć jednak, że we wszystkich badanych gimnazjach uczniowie o wyższych wynikach egzaminu byli bardziej krytyczni w ocenie klimatu społecznego swojej klasy i szkoły niż uczniowie o niższych wynikach egzaminu (Kutajczyk i Przychodzeń, 2008, s. 54 i 56, Tab. 27. i 28.). Statystycznie biorąc, słabsi gimnazjaliści bardziej lubią szkołę i słabsze gimnazja bywają bardziej lubiane, co potwierdza poprzednio cytowane zdanie Krzysztofa Konarzewskiego.

XIII. Dwa gimnazja – dwa rodzaje skuteczności?

Dla uwypuklenia różnic między samopoczuciem uczniów w szkołach o różnej edukacyjnej wartości dodanej zestawimy (a) średnie wskaźniki aprobaty (w skali 0 – 5) dziesięciu twierdzeń ankietowych przez uczniów *Gimnazjum Nr 3* (G03), najsłabszego w badaniach edukacyjnej wartości dodanej (EWD) przeprowadzonych przez Przychodzeń i Kutajczyk (2008), oraz przez uczniów *Gimnazjum Nr 8* (G08) – najmocniejszego (patrz Rys. 3.), a także (b) współczynniki korelacji wskaźników *EWD* z ocenami tych właściwości dokonanymi przez uczniów jedenastu szkół wiejskich.

Tabela 5. Ocena szkoły i nauczycieli w gimnazjach: G03 – o najniższej średniej wartości dodanej (2,4) i G08 – o najwyższej średniej wartości dodanej (3,5) oraz współczynniki korelacji pozycji ankiety z edukacyjną wartością dodaną 11 gimnazjów

Pozycja ankiety	G03	G08	Korelacja
Lubię wychowawcę.	4,8	3,5	-0,35
Wspólnie planujemy i przygotowujemy imprezy.	4,4	3,2	0,24
W ważnych wydarzeniach klasowych uczestniczą rodzice.	3,8	2,9	0,22
Chętnie chodzę do szkoły.	4,1	3,4	-0,21
W szkole czuję się bezpiecznie.	4,7	4,2	-0,06
W szkole panuje właściwa atmosfera.	4,4	3,5	-0,11
Relacje pomiędzy nauczycielami a uczniami są dobre.	4,5	3,4	-0,34
Nauczyciela udzielają wskazówek, jak uczyć się skutecznie.	4,6	3,1	-0,36
Nauczyciele interesują się tym, jak radzę sobie z wymaganiami.	4,3	3,2	-0,16
Zdobytą wiedzę stosujemy w praktyce.	4,5	3,6	-0,21
Ogółem	4,4	3,4	-0,13

W postawach uczniów dwu gimnazjów występują systematyczne różnice, a większość współczynników korelacji między opiniami uczniów o szkole a *EWD* jest ujemna. Gdyby oceniać „jakość pracy szkoły” G03 i G08 głównie na podstawie ankiet, jak w metodzie nazwanej „**mierzeniem pracy szkoły**” (Wlazło, 2002), to Gimnazjum Nr 3, wykazujące się **skutecznością relacyjną** kształcenia, zdystansowałoby swojego rywala. Gdyby zaś oceniać tę jakość głównie na podstawie edukacyjnej wartości dodanej, co odpowiada **skuteczności zadaniowej** kształcenia, to Gimnazjum Nr 8 należałoby do najlepszych szkół wiejskich w kraju, a Gimnazjum Nr 3 – do najsłabszych. Dobrze byłoby więc pogodzić te dwa kryteria i zastosować je wspólnie, ale to nie jest łatwe, jak na to wskazuje ujemna korelacja (-0,13) między oceną szkoły i nauczycieli przez uczniów a wynikami egzaminu.

XIV. Jakim szkołom egzaminu nie przeszkadzają?

Brak nam systematycznych badań dotyczących zależności między opiniami nauczycieli i uczniów o egzaminie zewnętrznym od stylu kierowania szkołą a wynikami, jakie szkoła uzyskuje. Można domyślać się, że taka zależność istnieje. Na podstawie wyników badań Przychodzeń i Kutajczyk oraz ogólnej wiedzy o systemach kształcenia można domyślać się następujących prawidłowości:

1. W szkołach o **relacyjnym stylu kierowania** zakłócenie spokoju uczenia się powodowane przez egzaminu zewnętrznego jest większe. Pojawiają się bezwzględne obowiązki, obcy ludzie, niespodziewane sytuacje. Egzaminu są odczuwane jako naruszenie harmonijnego układu społecznego, zagrożenie obniżeniem samooceny, brak zaufania.

2. W szkołach o **zadaniowym stylu kierowania** podejmuje się więcej działań przygotowujących do egzaminu. W badaniach Przychodzeń i Kutajczyk syntetyczny wskaźnik intensywności przygotowań do egzaminu gimnazjalnego okazał się dodatnio skorelowany (około 0,4 dla części humanistycznej i około 0,5 dla części matematyczno-przyrodniczej egzaminu) z edukacyjną wartością dodaną szkoły, a znane już nam Gimnazjum Nr 8 uzyskało wskaźnik najwyższy (Kutajczyk i Przychodzeń, 2008, s. 78).

3. Istnieją systematyczne różnice w **metodach przygotowań** do egzaminu zewnętrznego: w jednych szkołach (m. in. G03) dominuje wspólne rozwiązywanie zadań egzaminacyjnych z poprzednich lat, a w innych (najwyraźniej w G08) na przygotowania składają się przede wszystkim próbne egzaminu, zaś udział w dodatkowych zajęciach pozostawia się decyzji samego ucznia (tamże, s. 81).

Z tych analiz nie wynika, że uczniowie i nauczyciele w szkołach o wyższej wartości dodanej osiągnięć bardziej „lubią” egzaminu zewnętrznego niż w szkołach o niższej wartości dodanej. Uczniowie o wyższych wynikach są zwykle bardziej krytyczni wobec szkoły. Zbigniew Kwieciński nazwał to zjawisko **inwersją symboliczną** i tak je opisał (1990, s. 312): *Młodzież, która doznała niepowodzeń w szkole podstawowej, sobie samej przypisuje przyczynę znalezienia się na ścieżce wyłączenia z kultury i drogi ku wyższemu statusom społeczno-zawodowym. Młodzież ta wyżej ocenia kompetencje i wysiłek nauczycieli niż własną pracę i wykorzystanie zdolności oraz czasu.* Ta młodzież, średnio biorąc, bardziej szanuje szkołę.

Jest prawdopodobne zatem jedynie, że uczniowie G08 reagują na egzaminu inaczej niż uczniowie G03: mniej emocjonalne, a bardziej zadaniowo i, być może, mniej zbiorowo, a bardziej indywidualnie.

XV. Wizja przyszłości

Wraz z postępem pomiarowo-wydawniczym i pomiarowo-technologicznym uzyskiwanie zobiektywizowanej informacji o stanie osiągnięć i oszacowań wartości dodanej w zakresie podstawowych umiejętności będzie dla ucznia coraz łatwiejsze. Nastąpi dalsze rozluźnienie **systemu klasowo-lekcyjnego i szkoły jednolitej**, a wyjątkowo duże zróżnicowania szkół polskich pod względem średniego wyniku kształcenia (Białecki i in., 2003, s. 47 – 48) będzie stopniowo zastępowane zróżnicowaniem wyników wewnątrz szkoły na zasadzie **szkoły rozszerzonej** (*comprehensive school*).

Jednym z modeli przyszłego systemu dydaktycznego jest **kształcenie według wymagań** (*standards-based education*), w którym uczniowie pracują na wybranym przez siebie poziomie (Niemierko, 2007, s. 163 – 164). Stopniowe **usamodzielnianie uczniów** nie tylko w uczeniu się, lecz także w kontrolowaniu własnego rozwoju i w przygotowaniu do pokazania go na zewnątrz, może wyzwolić psychicznie nauczyciela z reżimu egzaminacyjnego. Bo czy mecze sportowe na własnym i cudzym boisku przeszkadzają trenerowi w pracy ze swoimi zawodnikami?

Podsumowanie referatu

1. **Uczenie się**, gdy jest właściwie kierowane, przynosi **rozwój**.
2. **Egzamin**, wewnętrzny i zewnętrzny, sprawdza wynik uczenia się.
3. W wielu krajach notuje się **opór** nauczycieli wobec egzaminów zewnętrznych.
4. Egzamin powinien rejestrować **rozwój** osiągnięć ucznia.
5. Rozwój osiągnięć ucznia stanowi **edukacyjną wartość dodaną**.
6. Rozwój osiągnięć ucznia można przedstawić w skali **równoważników klasy**.
7. Równoważnik klasy osiągnięć pojedynczego ucznia ma duży **błąd pomiaru**.
8. Szkoły mocno **różnią się** co do średniej edukacyjnej wartości dodanej
9. Edukacyjna wartość dodana jest różna w różnych **strefach** poziomu osiągnięć.
10. **Kierowanie szkołą** ma wielkie znaczenie dla osiągnięć jej uczniów.
11. Kierowanie **relacyjne** i kierowanie **zadaniowe** cechuje odmienna skuteczność.
12. Szkoła o **wysokiej** edukacyjnej wartości dodanej może być **wymagająca i surowa**.
13. Szkoła o **niskiej** edukacyjnej wartości dodanej może być **lubiana** przez uczniów.
14. **Reakcja szkoły** na egzamin zewnętrzny może być relacyjna lub zadaniowa.
15. W przyszłości informacja o rozwoju będzie kierowana bezpośrednio **do ucznia**.

Bibliografia:

1. Arends R.I., *Uczymy się nauczać*, WSiP, Warszawa 1994.
2. Bednarek K., Lempa C., *Utracone/utrącone mierzenie*, [w:] B. Niemierko, G. Szyling (red.), *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*, Fundacja Rozwoju Uniwersytetu Gdańskiego, Gdańsk 2005.
3. Białecki I., Blumsztajn A., Cyngot D., *PISA – Program Międzynarodowej Oceny Umiejętności*, ZNP, Warszawa 2003.

4. Brzezińska A., *Spoleczna psychologia rozwoju*, Scholar, Warszawa 2000.
5. Cizek G., More unintended consequences of high-stakes testing. *Educational Measurement: Issues and Practice*, 4 (2001).
6. Goslin D. A., *Teachers and testing*, Russell Sage Foundation, New York 1967.
7. Groenwald M., Czy godność nauczyciela jest zagrożona egzaminowaniem zewnętrznym? (w:) B. Niemierko i W. Małecki (red.), *Dawne i nowe formy egzaminowania*, Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław 2001b.
8. Janowski A., *Pedagogika praktyczna: zarys problematyki, zdrowy rozsądek, wyniki badań*, Fraszka Edukacyjna, Warszawa 2002.
9. Knight P.T., *Summative assessment in higher education: An appraisal in disarray*. Maszynopis.
10. Kolen M. J., Scaling and norming, [w:] Brennan (red.), *Educational measurement. Fourth edition*. Westport: American Council on Education – Praeger.
11. Konarzewski K., *O wychowaniu w szkole*, [w:] K. Kruszewski (red.), *Sztuka nauczania. Czynności nauczyciela*, PWN, Warszawa 2004.
12. Kutajczyk T., Przychodzeń B., *Czynniki skuteczności kształcenia ogólnego w gimnazjach wiejskich*, OKE Gdańsk 2008.
13. Kwieciński Z., *Błędne koło ubóstwa kulturowego*, [w:] Z. Kwieciński, L. Witkowski (red.), *Ku pedagogii pogranicza*, Wyd. UMK, Toruń 1990.
14. MacBeath J., Schratz M., Meuret D., Jakobsen L., *Czy nasza szkoła jest dobra?* WSiP, Warszawa 2003.
15. Mulawa J., Sroka D., *Trafność pomiarowa empirycznej normy osiągnięć*, [w:] B. Niemierko, J. Sygniewicz i W. Walczak (red.), *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*, Wyd. WSHE, Łódź 2003.
16. Niemierko B., *Zrównywanie wyników sprawdzianu 2004 do wyników sprawdzianu 2003. Ekspertyza wykonana na użytek Centralnej Komisji Egzaminacyjnej w Warszawie*. Maszynopis, 2004d.
17. Niemierko B., *Zrównywanie wyników sprawdzianu 2005 do wyników sprawdzianu 2003. Ekspertyza wykonana na użytek Centralnej Komisji Egzaminacyjnej w Warszawie*. Maszynopis, 2005a.
18. Niemierko B., Wartość dodana osiągnięć uczniów, szkół i regionów. Studium wstępne do wykorzystania przez CKE. *Biuletyn Badawczy CKE Nr 3*, 2005b.
19. Niemierko B., *Kształcenie szkolne. Podręcznik skutecznej dydaktyki*, WAiP, Warszawa 2007.
20. Niemierko B., *Diagnostyka edukacyjna. Podręcznik akademicki*, PWN, Warszawa (w druku).
21. Plake B. S., Impara J. C., Fager J. J., Assessment competencies of teachers. A national survey. *Educational Measurement: Issues and Practice* 1993 nr 4.
22. Poręba-Konopczyńska A., *Moje refleksje nad ocenianiem na podstawie układu rzetelności i trafności wyników sprawdzania zaproponowanego przez Petera T. Knighta*, (w:) B. Niemierko i M. K. Szmigel, *Teoria i praktyka oceniania zewnętrznego*, PANDIT, Kraków 2001.
23. Przychodzeń B., Kutajczyk T., *Analizy wyników egzaminu gimnazjalnego w 12 szkołach*. Maszynopis.
24. Szmigel M. K., Rappe A., *Przydatność uproszczonej (staninowej) metody szacowania osiągnięć uczniów, szkół i jednostek administracyjnych oświaty*, [w:] B. Niemierko, G. Szyling (red.), *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*, Fundacja Rozwoju Uniwersytetu Gdańskiego, Gdańsk 2005a.
25. Szmigel M. K., Rappe A., *Komunikowanie wartości dodanej osiągnięć szkolnych uczniom, nauczycielom i dyrektorom szkół*, [w:] B. Niemierko, G. Szyling (red.), *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*, Fundacja Rozwoju Uniwersytetu Gdańskiego, Gdańsk 2005b.
26. Szmigel M. K., Rappe A., *Stołość wskaźnika edukacyjnej wartości dodanej i opinie dyrektorów o jego znaczeniu*, [w:] B. Niemierko, M. K. Szmigel (red.), *O wyższą jakość egzaminów szkolnych*, PTDE, Lublin 2006.
27. Sztompka P., *Socjologia. Analiza społeczeństwa*, Znak, Kraków 2002.
28. Touhy D., *Dusza szkoły. O tym co sprzyja zmianie i rozwojowi*, PWN, Warszawa 2002.
29. Wlazło S., *Jakościowy rozwój szkoły*, MarMar, Wrocław 2002.
30. Wojciszke B., *Człowiek wśród ludzi. Zarys psychologii społecznej*, Scholar, Warszawa 2002.