

**EWA STOŻEK**  
OKE Łódź

## O CZYM MÓWI EFEKT STANDARDOWY?

W sprawozdaniu Centralnej Komisji Egzaminacyjnej „Sprawdzian 2004”<sup>1</sup> zastosowano wskaźnik zwany efektem standardowym, dla porównania wyników uzyskanych przez uczniów w gminach i miastach oraz dla porównania wyników chłopców i dziewcząt. W niniejszym komunikacie przybliżono pojęcie efektu standardowego. Zostały podane sposoby jego obliczania oraz interpretacji. Przytoczono przykłady z badań edukacyjnych ilustrujące zastosowanie efektu standardowego.

### 1. Co to jest efekt standardowy?

Efekt standardowy (ang. *effect size*) wykorzystuje się do mierzenia wpływu pewnego czynnika (działania) na wynik ogólny grupy, np. wpływu liczebności klasy na wyniki uczniów w czytaniu albo wpływu nowego leku na zmniejszenie dolegliwości pacjentów. Dwie porównywane grupy przyjęło się nazywać grupą kontrolną (bazową) i grupą eksperymentalną.

Efekt standardowy wyrażany jest wzorem<sup>2</sup>:

$$d = \frac{\overline{x_1} - \overline{x_2}}{s},$$

gdzie:

$\overline{x_1}$  – średnia grupy pierwszej (eksperymentalnej, którą będziemy porównywać),

$\overline{x_2}$  – średnia grupy drugiej (kontrolnej, bazowej, do której będziemy porównywać),

---

<sup>1</sup> Brożek A., Grabowska D., Jędrasik H., Walczak J., *Sprawozdanie CKE „Sprawdzian 2004”*, Warszawa 2004.

<sup>2</sup> Cohen J., *Statistical power analysis for the behavioral sciences*, 2<sup>nd</sup>ed., Hillsdale, NJ 1988, s. 44.

$s$  – odchylenie standardowe;  
w ogólnym przypadku:

$$s = s_{pooled} = \sqrt{\frac{(N_1 - 1)s_1^2 + (N_2 - 1)s_2^2}{N_1 + N_2 - 2}},$$

gdzie:

$N_1, N_2$  – liczebności pierwszej i drugiej grupy odpowiednio,

$s_1, s_2$  – odchylenie standardowe dla pierwszej i drugiej grupy odpowiednio.

Zauważmy, że w przypadku niewielkiej różnicy w liczebności grup oraz podobnych odchyleń standardowych, można przyjąć  $s = s_2$ . Analogicznie, gdy grupy różnią się znacząco odchyleń standardowymi, ale są podobne co do liczebności,

można stosować  $s = \sqrt{\frac{s_1^2 + s_2^2}{2}}$ .

Efekt standardowy jest wielkością niemianowaną; różnica średnich jest wyrażona w odchyleniach standardowych.

Cohen<sup>3</sup> zaproponował następującą skalę siły efektu (wzmocnienia):

$d = 0,2$	$d = 0,5$	$d = 0,8$
mały efekt (wpływ)	średni efekt (wpływ)	duży efekt (wpływ)

## 2. Przykłady efektów standardowych w badaniach edukacyjnych

Wskaźnik efektu standardowego jest szeroko stosowany w badaniach społecznych, w tym także edukacyjnych. Dla lepszego zilustrowania, jakie wartości może on przyjmować, przytaczam za pracą Cohena<sup>4</sup> przykłady efektów standardowych w badaniach edukacyjnych. Uzupełniłam ponadto tabelę wybranymi wskaźnikami ze sprawozdania CKE „Sprawdzian 2004”.

Analiza wyników w tabeli pokazuje, że rzadko w badaniach edukacyjnych mamy do czynienia ze wzmocnieniem średnim i dużym (wg skali Cohena). Obserwowane wielkości efektu standardowego większe niż 0,2 są traktowane jako statystycznie znaczące wzmocnienie. Stosowanie testów próbnych jest czynnikiem znacząco podnoszącym wynik testu egzaminacyjnego (efekt standardowy +0,32). Stosowanie w nauczaniu klas otwartych nie podnosi osiągnięć uczniów (efekt standardowy -0,06), ale nieznacznie podwyższa poziom akceptacji szkoły przez ucznia (efekt standardowy +0,17). Wyniki dziewcząt na sprawdzianie w

<sup>3</sup> Cohen J., op. cit., s. 25.

<sup>4</sup> Coe R., *What is an „Effect Size”. A guide for users*, 2000 – [www.cemcentre.org/ebeuk/research/effectsize/ESguide.htm](http://www.cemcentre.org/ebeuk/research/effectsize/ESguide.htm)

zakresie pisania są zdecydowanie wyższe niż wyniki chłopców (efekt standardowy +0,50).

Tab. 1. Wybrane przykłady efektów standardowych w badaniach edukacyjnych

Czynnik (wpływ czego)	Co mierzono? (wpływ na co)	<i>d</i>
zmniejszenie klasy z 23 do 15 uczniów <sup>5</sup>	wyniki uczniów w czytaniu	+0,30
	wyniki uczniów z matematyki	+0,32
testy próbne <sup>6</sup>	wynik testu egzaminacyjnego	+0,32
korepetycje rówieśnicze <sup>7</sup>	wyniki uczniów-korepetytorów	+0,33
	wyniki uczniów biorących korepetycje	+0,40
„otwarta klasa” <sup>8</sup>	osiągnięcia uczniów	-0,06
	stosunek ucznia do szkoły	+0,17
płeć <sup>9</sup> grupa bazowa chłopcy, grupa eksperym. dziewczęta	wynik ogólny sprawdzianu	+0,29
	pisanie	+0,50
	rozumowanie	+0,10
lokalizacja szkoły <sup>10</sup> grupa bazowa – gminy, grupa eksperyment. – miasta	wynik ogólny sprawdzianu	+0,28
	czytanie	+0,17
	pisanie	+0,24

### 3. Efekt standardowy specyficznych trudności w uczeniu się

Policzmy efekt standardowy dla dwóch grup: pierwszą (eksperymentalną) będą stanowić uczniowie ze specyficznymi trudnościami w uczeniu się (dysleksja rozwojowa), dla której  $N_1 = 4353$ , drugą zaś uczniowie bez tych trudności ( $N_2 = 43272$ ). Wyniki tych grup na sprawdzianie 2005 r. (dane OKE Łódź) prezentuje tab. 2.

Uczniowie ze specyficznymi trudnościami w uczeniu się uzyskują słabsze wyniki niż uczniowie bez dysleksji rozwojowej w zakresie czytania, korzystania z informacji, rozumowania i wykorzystania wiedzy w praktyce. Efekt standardowy dla tych umiejętności jest rzędu -0,2, co kwalifikuje go jako mało znaczący. Znaczący efekt standardowy obserwujemy w przypadku umiejętności pisania – uczniowie ze specyficznymi trudnościami w uczeniu się uzyskują lepsze wyniki

<sup>5</sup> Finn J.D., Achilles C.M., *Answers and questions about class size. A statewide experiment*, „American Educational Research Journal” 1990, nr 27, s. 557–577.

<sup>6</sup> Kulik J.A., Kulik C.C., Bangert R.L., *Effects of practice on aptitude and achievement test scores*, „American Education Research Journal”, nr 21, s. 435–447.

<sup>7</sup> Cohen P.A., Kulik J.A., Kulik C.C., *Educational outcomes of tutoring: a meta-analysis of findings*, „American Educational Research Journal” 1982, nr 19, s. 237–242.

<sup>8</sup> Giaconia R.M., Hedges L.V., *Identifying features of effective open education*, „Review of Educational Research” 1982, nr 52, s. 570–602.

<sup>9</sup> Brożek A., Grabowska D., Jędrasik H., Walczak J, op. cit., s. 55.

<sup>10</sup> Ibidem, s. 54.

z efektem standardowym +0,31. Efekt ten nie wydaje się być naturalny i jest związany z dostosowaniem kryteriów oceniania wypracowania dla uczniów dyslektycznych. Wynik ogólny charakteryzuje się praktycznie zerowym efektem standardowym, co dało podstawę do wnioskowania, że sprawdzian dla uczniów z dysleksją *dostosowano w sposób właściwy do ich dysfunkcji*<sup>11</sup>. Uważam, że wniosek ten jest zbyt daleko idący: właściwe dostosowanie powinno oznaczać porównywalność wyników we wszystkich obszarach wymagań egzaminacyjnych, a nie tylko w wyniku ogólnym.

Tab. 2. Efekt standardowy specyficznych trudności w uczeniu się

Umiejętność	$\bar{x}_1$	$\bar{x}_2$	$s_1$	$s_2$	$s_{pooled}$	$d$
Czytanie	7,93	8,29	1,65	1,64	1,64	-0,22
Pisanie	7,07	6,33	2,17	2,45	2,43	+0,31
Rozumowanie	5,52	5,86	2,05	2,04	2,04	-0,17
Korzystanie z informacji	2,8	3,00	1,19	1,18	1,18	-0,17
Wykorzystywanie wiedzy	5,6	5,91	1,87	1,86	1,86	-0,17
Wynik ogólny	28,91	29,39	6,84	7,34	7,30	-0,07

#### 4. Efekt standardowy a wynik standardowy

Efekt standardowy i wynik standardowy są bardzo zbliżonymi wskaźnikami. Wynik standardowy wyrażany jest wzorem<sup>12</sup>:

$$z = \frac{x - \bar{x}}{s},$$

gdzie:

$x$  – pojedynczy wynik,

$\bar{x}$  – średni wynik grupy,

$s$  – odchylenie standardowe grupy.

Wynik standardowy pozwala na usytuowanie pojedynczego wyniku (np. ucznia) na skali niezależnej od wielkości średniej i odchylenia standardowego grupy (np. klasy), bowiem wyniki standardowe mają średnią 0 i odchylenie standardowe 1. Wynik standardowy jest wielkością niemianowaną. Przedstawienie wyniku ucznia na skali wyników standardowych daje nam informację o tym, w jakiej odległości od średniej (od 0) mierzonej w odchyleniach standardowych znajduje się jego wynik. Podobną informację można uzyskać w innej znormalizowanej skali, jaką jest powszechnie stosowana skala staninowa (ang. *standard nine*).

<sup>11</sup> Brożek A., Grabowska D., Jędrasik H., Walczak J, op. cit., s. 56.

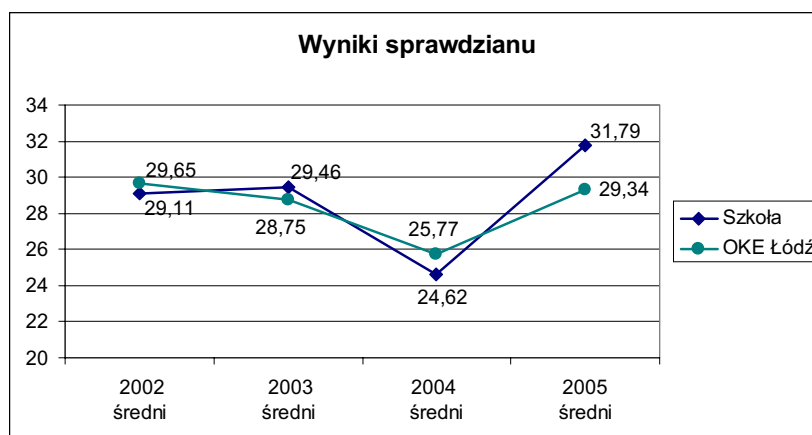
<sup>12</sup> Ferguson G.A., Takane Y., *Analiza statystyczna w psychologii i pedagogice*, PWN, Warszawa 2003, s. 92–94.

Wynik standardowy znajduje zastosowanie wtedy, gdy chcemy porównać pojedyncze wyniki w obrębie grupy (np. uczniów danej klasy) pod względem różnych charakterystyk (np.: wynik ogólny sprawdzianu, wyniki w obrębie różnych obszarów wymagań egzaminacyjnych).

Efekt standardowy natomiast ma zastosowanie przede wszystkim wtedy, gdy chcemy porównać wyniki różnych grup (np. klas A i B) pod względem tej samej charakterystyki (np. umiejętności czytania ze zrozumieniem).

## 5. Efekt standardowy a porównywanie średnich wyników szkół

W 2005 r. po raz czwarty został przeprowadzony sprawdzian po VI klasie szkoły podstawowej. Wzrasta wiedza nauczycieli i dyrektorów szkół o sprawdzianie oraz zainteresowanie analizami wyników. Niestety, najczęściej porównywaną wielkością nadal jest średnia szkoły. Dla względnego (w stosunku do innych szkół) porównania wyników należy wykorzystać jedną ze znormalizowanych skal, np.: skalę staninową (standardowej dziesiątki), skalę standardowej piątki lub skalę wyniku standardowego. Jeszcze dokładniejszą informację o postępie (lub regresie) szkoły można uzyskać, licząc efekt standardowy.



Rys. 1. Średnie wyniki szkoły X ze sprawdzianu w latach 2002–2005

Rozpatrzmy przykład szkoły podstawowej X z terenu działania OKE Łódź. Wykres 1. przedstawia średnie wyniki ze sprawdzianu tej szkoły z lat 2002–2005, odniesione do średniej OKE Łódź i wyrażone w pięciostopniowej skali średnich wyników szkół z terenu działania OKE Łódź. Porównanie w tej skali pozwala jedynie stwierdzić, że wynik szkoły mieści się w przedziale wyników średnich.

Wynik standardowy, czyli różnica średniej szkoły i średniej OKE Łódź odniesiona do odchylenia standardowego średnich wyników szkół, pokazuje, że wyniki szkoły przez pierwsze trzy lata wahały się wokół wartości średniej od  $-0,37$  do  $+0,27$  odchylenia standardowego. W 2005 r. średni względny wynik

szkoły zdecydowanie wzrósł i różnica między wynikiem szkoły a średnim wynikiem w OKE Łódź stanowiła +0,84 odchylenia standardowego.

Do wyliczenia efektu standardowego przyjęto jako grupę bazową wszystkich uczniów przystępujących do sprawdzianu na terenie działania OKE Łódź, a jako grupę eksperymentalną – uczniów szkoły X. Wielkość efektu standardowego (por. tab. 3.) wskazuje, że znaczący postęp (efekt standardowy +0,36) szkoła osiągnęła w 2005 r., natomiast wyniki z poprzednich lat charakteryzowały się małym, statystycznie nieznaczącym efektem.

Tab. 3. Wynik standardowy i efekt standardowy sprawdzianu dla szkoły X w latach 2002–2005

	2002	2003	2004	2005
wynik standardowy	-0,19	+0,27	-0,37	+0,84
efekt standardowy	-0,17	+0,13	-0,15	+0,36

### Pytania na zakończenie

1. Interpretacja efektu standardowego opiera się na założeniu, że rozkłady wyników grup kontrolnej i eksperymentalnej są rozkładami normalnymi. Czy interpretacje co do siły efektu pozostaną takie same, jeśli rozkłady wyników będą różnić się od rozkładu normalnego?
2. Czy skala Cohena jest przydatna dla komunikowania wyników egzaminów zewnętrznych?
3. Czy statystycznie znaczące efekty standardowe mają praktyczne znaczenie również dla polityki edukacyjnej?