



You have downloaded a document from
RE-BUŚ
repository of the University of Silesia in Katowice

Title: Wykorzystanie wyników opracowań statystycznych w pracy nauczyciela na przykładzie egzaminów gimnazjalnych i próbnych matur

Author: Halina Widła

Citation style: Widła Halina. (2003). Wykorzystanie wyników opracowań statystycznych w pracy nauczyciela na przykładzie egzaminów gimnazjalnych i próbnych matur. "Neofilolog" (Nr 23, (2003) s. 88-94).



Uznanie autorstwa - Bez utworów zależnych Polska - Ta licencja zezwala na rozpowszechnianie, przedstawianie i wykonywanie utworu zarówno w celach komercyjnych i niekomercyjnych, pod warunkiem zachowania go w oryginalnej postaci (nie tworzenia utworów zależnych).



UNIwersYTET ŚLĄSKI
W KATOWICACH



Biblioteka
Uniwersytetu Śląskiego



Ministerstwo Nauki
i Szkolnictwa Wyższego

Halina Widła

Uniwersytet Śląski

Wykorzystanie wyników opracowań statystycznych w pracy nauczyciela na przykładzie egzaminów gimnazjalnych i próbnych matur.

Nowe problemy, nowe wyzwania

Upowszechnienie podstaw wiedzy na temat pomiaru wśród nauczycieli stało się palącym problemem w gimnazjach i liceach. Od czasu, kiedy wiedzę i kompetencje uczniów mierzymy testami zewnętrznymi, nauczyciele skarżą się na problemy z dogłębną i rzetelną interpretacją wyników. W roku 2001 byli oni zatrzwożając bezradni, czemu trudno się dziwić, zważywszy, że OKE¹⁹ opublikowały wyniki uwzględniając główne mierniki tendencji centralnej i dyspersji. W roku 2002 prezentacje zyskały na klarowności, upowszechniono wśród nauczycieli podstawową wiedzę o pomiarze, tak by nauczyli się na podstawie prawidłowej interpretacji określonych wskaźników korygować metody i techniki nauczania. Do pełnego sukcesu jeszcze daleko. Wielu z nas, filologów, nauczycieli języków obcych, niejednokrotnie z ogromnym i cenionym doświadczeniem zawodowym, nie odczuło dotąd potrzeby poszerzenia swej wiedzy o choćby elementy tak hermetycznej dyscypliny. Moim celem nie jest specjalistyczny wykład, lecz zwrócenie uwagi na dwa istotne aspekty tytułowego zagadnienia: praktyczny i badawczy, oba adresowane teoretycznie do dwóch różnych gremiów, w rzeczywistości jednak, kiedy wielu nauczycieli podejmuje studia doktoranckie, a wielu pracowników naukowych uczy w szkołach, podwójnie do tego samego odbiorcy.

Potocznie statystyka kojarzy się z wykresami kołowym lub słupkowym, mającymi na celu ilustrację danego zjawiska, rozkład procentowy stanu, preferencji, wieku, płci, poglądów lub ich ewolucji w czasie. Z takimi prezentacjami graficznymi spotykamy się w codziennej prasie. Tamże również nabieramy przekonania, że najważniejszym miernikiem statystycznym jest tak zwana średnia arytmetyczna, skandalicznie nadużywana, nagminnie źle interpretowana, no i wyśmiana w dowcipach. Wiadomo przecież, że uśrednić można wszystko, ale z wiedzy takiej niewiele wynika. Żarty o lewej nodze trzymanej w wodzie lodowatej a prawej we wrzątku, czyli w sumie w letniej, lub o mężach bijających żony średnio półtora raza na dzień, wypowiedane nawet przez osoby z cenzusem naukowym w celu ukazania bezsensowności statystycznych analiz, dowodzą zenującej ignorancji. Ten stereotyp powieli się czasem świadomie, zaciemniając prawdę i manipulując danymi, a robi się to przykładając zły miernik do danych. Średnia dochodów społeczeństwa nie informuje o poziomie życia. Takie zabiegi przywodzą mi na myśl chwyt z erystyki

¹⁹ OKE - Okręgowe Komisje Egzaminacyjne

schopenhauerowskiej, i nie ma w tym przesady, bo w obu wypadkach cynicznie wykorzystuje się niewiedzę i brak warsztatu.

Podstawowe pojęcia opisujące osiągnięcia uczniów

W takich oto warunkach CKE²⁰ i OKE publikują dane statystyczne dotyczące wyników testów gimnazjalnych i próbnych matur. Będą je interpretować wszyscy: uczniowie i ich rodzice w celu wyszukania szkoły na możliwie wysokim poziomie, kuratorzy, by mieć ogłęd sytuacji w regionie i dowiedzieć się, na podstawie obiektywnych danych, kto i jak pracuje. Wreszcie dyrektorzy szkół będą mieć możliwość, by ocenić pracę swoich podwładnych, w konfrontacji z konkurencją. W dobie niżu demograficznego tak właśnie będzie się działo bez względu na ton oficjalnych deklaracji o pomocniczej, informacyjnej roli testów. Czy rzeczywiście, nie radząc sobie z interpretacją danych statystycznych, będą w stanie wyrobić sobie obiektywny pogląd na istniejącą rzeczywistość? A jak my, nauczyciele, będziemy mogli wybronić się od ewentualnych zarzutów złej pracy? Uważam, że choćby pobeżne zapoznanie się z kilku ważnymi pojęciami umożliwi polemikę merytoryczną zamiast demagogicznej. W tym celu, opierając się na materiałach dostarczonych przez komisje egzaminacyjne, przygotowałam kilka charakterystycznych przykładów.

Do podstawowych wskaźników opisujących osiągnięcia uczniów należą:

- ☞ **łatwość zestawu zadań** będąca stosunkiem liczby punktów uzyskanych za rozwiązanie zadań przez wszystkich uczniów do maksymalnej liczby punktów możliwych do uzyskania za zadania;
- ☞ **średnia arytmetyczna** - suma wszystkich uzyskanych wyników podzielona przez ich liczbę. W wypadku języków obcych przydatna zwłaszcza w podsumowaniach, bo dająca możliwość uśrednienia wszystkich kompetencji. Trudno oceniać ogólną znajomość języka w oderwaniu od pewnych sprawności.

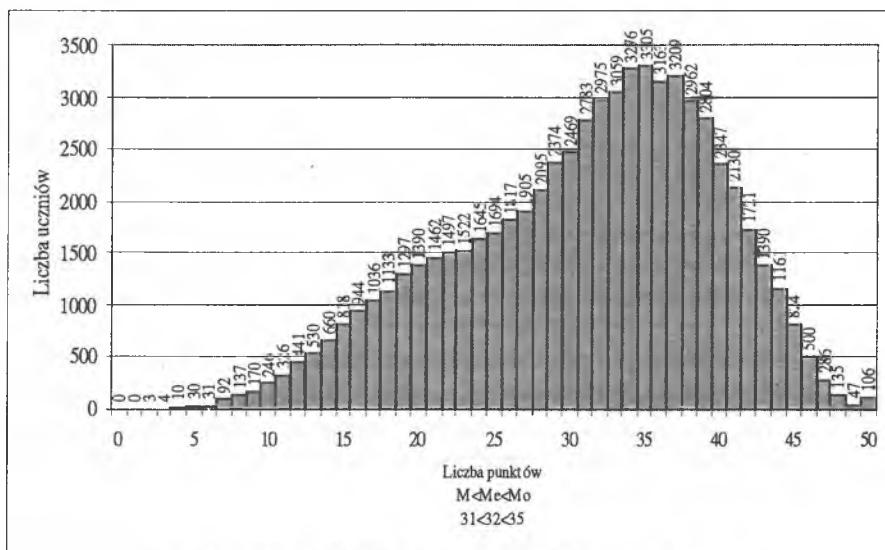
W wypadku oceny poszczególnych umiejętności warto jednak sięgnąć po takie wskaźniki jak:

- ☞ **mediana** - wynik środkowy spośród wyników uczniów uporządkowanych rosnąco lub malejąco. Ważne, by sprawdzić, ile osób sytuuje się po obu stronach tego wskaźnika i odpowiednio dopasować ćwiczenia korygujące;
- ☞ **modalna** - wynik, który w badanej grupie powtarza się najczęściej. To on, a nie średnia arytmetyczna pomaga określić, w jakim stopniu dana sprawność została opanowana przez grupę;
- ☞ **wyniki : najwyższy i najniższy;**
- ☞ **rozstęp wyników** - różnica między najwyższym a najniższym wynikiem osiągniętym przez uczniów;
- ☞ **odchylenie standardowe** - miara rozproszenia wyników uczniów w odniesieniu do wyniku średniego. Ten ostatni wskaźnik jest bardzo istotny, a niezmiernie rzadko brany pod uwagę. Inaczej pracuje się w wypadku, gdy dana umiejętność przyswojona została w sposób zróżnicowany, a inaczej, gdy w sposób podobny, bez względu na poziom.

²⁰ CKE - Centralna Komisja Egzaminacyjna

Ważną informacją jest ocena wskaźnika łatwości. OKE proponują porównać łatwość zestawu zadań dla uczniów danej szkoły z łatwością w całym województwie; obliczyć, jaki procent uczniów uzyskał wyniki wyższe niż wynik średni w województwie.

Utworzenie kartoteki do zestawu standardowego z wartościami wskaźnika łatwości jest stosunkowo prostą sprawą. Wystarczy zestawić interesujące nas czynności pod kątem stopnia ich opanowania. Przykładowo, w kartotece czynności przewidzianych do opanowania na teście dla gimnazjalistów w roku 2002, wskaźnik łatwości odczytania znaczenia wyrazów wyniósł 0,91, a wyszukiwania informacji - zaledwie 0,36. Okręgowe komisje sporządzają wykresy podsumowujące egzaminy pod względem stopnia trudności, na przykład:



Wykres nr 1. Źródło: Analiza wykonana przez OKE w Jaworznie, 2002

Łatwo zauważyć, że im bardziej wykres przesunięty jest w prawo, tym łatwiejszy był test.

Ideą egzaminów zewnętrznych jest porównanie osiągnięć między uczniami, szkołami, regionami. W tym celu zastosowano skalę znormalizowaną zwaną staninową. Otrzymuje się ją dzieląc uporządkowane rosnąco surowe rezultaty na dziewięć ponumerowanych przedziałów zawierających kolejno 4 – 7 – 12 – 17 – 20 – 17 – 12 – 7 – 4 % wyników. Takie staniny stopniuje się od najniższego, poprzez bardzo niski, niski, niżej średni, średni, wyżej średni, wysoki i bardzo wysoki, aż do najwyższego; na przykład uczeń, plasujący się w siódmym staninie reprezentuje wysoki poziom, a ten, który uzyskał 3 staniny, ma wynik niżej średni.

Interpretacja wyników nie powinna również sprawiać trudności. Tabela 1, pochodząca z opracowań CKE, przedstawiająca wyniki na znormalizowanej skali staninowej ukazuje, że na przykład wynik ucznia znajdujący się w szóstym staninie mieści się w przedziale 34 - 37 punktów, podobnie jak wyniki punktowe 12 952 innych uczniów:

Nazwa stanina – opis wyniku	Najniższy	Bardzo niski	Niski	Niżej średni	Średni	Wyżej średni	Wysoki	Bardzo wysoki	Najwyższy
Staniny	1	2	3	4	5	6	7	8	9
Przedziały punktowe	0 - 13	14 - 18	19 - 23	24 - 29	30 - 33	34 - 37	38 - 40	41 - 43	44 - 50
Liczba uczniów	2 020	4 591	7 168	11 530	11 286	12 953	8 113	5 241	3 059

Tabela 1. Wyniki uczniów na znormalizowanej skali staninowej

Źródło: Analiza wykonana przez CKE w Warszawie, 2002

Na tej podstawie nauczyciel sam łatwo wyliczy, jaki procent jego uczniów uzyskał wyniki niskie, średnie lub wysokie, a jaki poniżej lub powyżej tych wartości. Interesującym dopełnieniem jest propozycja OKE podziału na kategorie standardów oraz ich łatwości dla wyników z poszczególnych przedziałów staninowych.

Nazwa stanina – opis wyniku	Najniższy	Bardzo niski	Niski	Niżej średni	Średni	Wyżej średni	Wysoki	Bardzo wysoki	Najwyższy
Staniny	1	2	3	4	5	6	7	8	9
Czytanie i odbiór tekstów kultury	0,36	0,50	0,59	0,63	0,69	0,75	0,80	0,84	0,89
Tworzenie własnego tekstu	0,07	0,14	0,25	0,43	0,56	0,66	0,75	0,83	0,92

Tabela 2. Kategorie standardów oraz ich łatwości dla wyników z poszczególnych przedziałów staninowych.

Źródło: Analiza wykonana przez CKE w Warszawie, 2002

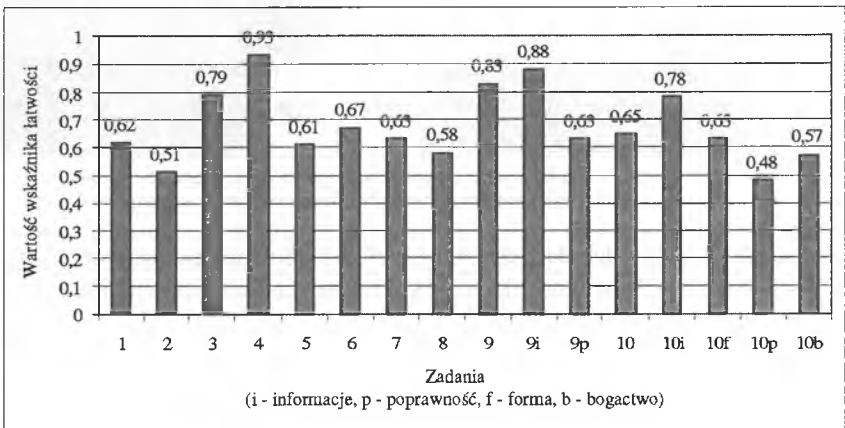
Inną ciekawą formą globalnej charakterystyki wyników jest próba ich porównania na obu poziomach z rozbiem na arkusze testujące poszczególne umiejętności:

Liczba zdających	Liczba punktów możliwych do zdobycia	Łatwość zestawu	Rozstęp	Średnia arytmetyczna	Mediana	Modalna	Odchylenie standardowe
<i>Ogółem na poziomie podstawowym</i>							
27	100	0,70	38 - 96	69,8	70	-	16,85
Arkusz I – Rozumienie ze słuchu							
27	30	0,64	8 - 28	19,2	18	18	5,33
Arkusz II – Rozumienie tekstu czytanego							
27	40	0,73	18 - 38	29,3	28	36	6,23
Arkusz III – Wypowiedź pisemna							
27	30	0,71	5 - 30	21,3	23	-	7,23
<i>Ogółem na poziomie rozszerzonym</i>							
5	100	0,63	41 - 91	62,8	61	-	18,2
Arkusz IV – Rozumienie ze słuchu							
5	30	0,55	10 - 26	16,4	12	-	7,13
Arkusz V – Rozumienie tekstu czytanego							
5	34	0,65	16 - 29	22,2	21	-	4,97
Arkusz VI – Wypowiedź pisemna							
5	36	0,67	14 - 36	24,2	24	-	9,86

Tabela 3. Charakterystyka wyników osiągniętych przez zdających
Źródło: Analiza wykonana przez OKE w Jaworznie, 2002

Komisje zgodnie potwierdzają, że matura z języka francuskiego na poziomie podstawowym okazała się łatwa. Najlepiej radzono sobie z zadaniami mierzącymi rozumienie tekstu czytanego. Rozszerzona matura z języka francuskiego okazała się umiarkowanie trudna. Najstabilniej wypadło rozumienie ze słuchu na obu poziomach.

Łatwość zadań z poziomu podstawowego ilustrowana jest przy pomocy wykresów bądź przedstawiona w tabeli. Poniższy wykres ukazuje, że większość zadań z poziomu podstawowego była dla zdających umiarkowanie trudna: trzy zadania były łatwe i jedno - bardzo łatwe.



Wykres nr 2. Łatwość zadań z poziomu podstawowego.
Źródło: Analiza wykonana przez OKE w Jaworznie, 2002

Poniższa tabela jest chyba pełniejszą formą prezentacji mimo braku wizualizacji omawianego zagadnienia. Przedstawia dla odmiany zestawienie wskaźników łatwości dla poziomu rozszerzonego.

Arkusz IV – Rozumienie ze słuchu			
Zadanie 11	Zadanie 12	Zadanie 13	
0,44	0,64	0,56	
Arkusz V – Rozumienie tekstu czytanego			
Zadanie 14	Zadanie 15	Zadanie 16	
1	0,40	0,68	
Arkusz VI – Wypowiedź pisemna			
Zadanie 17			
0,67			
Zadanie 17 (treść)	Zadanie 17 (forma)	Zadanie 17 (poprawność)	Zadanie 17 (bogactwo)
0,69	0,67	0,62	0,71

Tabela 4. Łatwość zadań z poziomu rozszerzonego
 Źródło: Analiza wykonana przez OKE w Jaworznie, 2002

Nowe możliwości badawcze

Powyższe ilustracje są tylko drobnym fragmentem obszernych opracowań, jakie wykonują centralna i okręgowe komisje egzaminacyjne. Po raz pierwszy pojawia się szansa nie tylko zdiagnozowania poziomu zdających według jednolitych kryteriów, ale obiektywnego porównania osiągnięć z uwzględnieniem wielu kryteriów. Pomiar dydaktyczny okazał się więc palącą koniecznością w dobie reformy szkolnictwa, próbnych matur i egzaminów gimnazjalnych.

Wiążą się z tym dwa zasadnicze problemy: teoretyczny i praktyczny. Jedyną w pełni opracowaną w Polsce metodą pomiaru dydaktycznego jest metoda prof. Bolesława Niemierki. Zachodzi pytanie, czy i na ile sprawdza się ona w wypadku analizy wyników nauczania języków obcych. To pole do popisu dla wszystkich metodyków z zacięciem badawczym. Do zbadania mają oni ogromny obszar, a zainteresowanie wynikami ich pracy – gwarantowane. Biorąc jednak pod uwagę prace komisji, a także ich wysiłek w upowszechnienie wiedzy o pomiarze, należy mieć nadzieję, że mimo trudności reforma będzie sukcesem. Na uwagę zasługują znakomite prezentacje, również w wersji elektronicznej, co w dobie obecnej komputeryzacji szkół nie pozwala na usprawiedliwienie niewiedzy brakiem dostępu do informacji.

Zastanawia też inny problem, wynikający z analizy zarówno próbnych matur z roku 2001, jak i egzaminów z roku 2002. Jak brzmią oficjalne komentarze, młodzież nasza, w świetle ostatnich wyników nowych matur, słaba jest w naukach ścisłych, chemii, biologii oraz historii, a dobra w języku polskim i świetna w znajomości języków obcych. W roku 2001 w niektórych województwach z niektórych przedmiotów próbną nową maturę zdało kilkanaście procent uczniów, a z języków obcych - prawie wszyscy. Można powiedzieć, że autorzy testów ułożyli je pod poziom

przeciętnego ucznia, lub też, że mają lepszą znajomość realiów od pozostałych kolegów. Zachodzi pytanie, czy wszystkie testy zostały wystandaryzowane. Oznacza to, że układa się całe serie ćwiczeń i dla niepoznaki wplata w różne sprawdziany w różnych szkołach i regionach, tak by otrzymać zestawy o średnim wskaźniku łatwości potrzebnym do zdania matury. W wypadku innych przedmiotów tego nie zrobiono, a wyniki okazały się zastraszające. Rodzi się więc kolejne pytanie: jak dobrać kryteria trudności sprawdzianów - do celów jakie sobie stawiamy, czy do faktycznych osiągnięć uczniów, jakie by one nie były? W pierwszym wypadku za dużo osób nie zda matury - i będzie to świadczyło albo o nierealistycznych celach, albo o fatalnej pracy, a w drugim - za dużo ją zda, i będzie to świadczyło, albo o żenująco niskim poziomie wymagań, albo o geniuszu lingwistycznym młodych Polaków i niezwyklej kompetencji dydaktycznej filologów. Jak zwykle nie ma złotego środka. Co więc należy uczynić? Teoretycy wymyślą test na każde zamówienie, ale co na to praktycy?

Bibliografia

- Amyotte Luc, 1996: *Méthodes quantitatives. Application à la recherche en sciences humaines*. Bibliothèque Nationale de Québec, Canada
- Brzeziński Jerzy, 2000: *Badania eksperymentalne w psychologii i pedagogice*. Warszawa, Scholar
- Ferguson George A, Takane Yoshio, 2002: *Analiza statystyczna w psychologii i pedagogice*. Warszawa, PWN
- Niemierko B, 1991: *Między oceną szkolną a dydaktyką*. Warszawa. WSiP
- Niemierko B, 1999: *Pomiar wyników kształcenia*. Warszawa. WSiP
- Zaczyński W, 1995: *Praca badawcza nauczyciela*. Warszawa. WSiP
- Materiały OKE w Jaworznie, 2002: www.oke.jaw.pl
- Materiały CKE w Warszawie, 2002: www.cke.edu.pl

Summary

The issue centres on the so-called didactic measurement, its principles, utility, place and application at the time of the present education reform. The purpose of this paper is to draw attention to two crucial aspects of the problem referred to in the title: the practical and the research ones. Theoretically, they are both addressed to two different bodies, yet in reality, as many teachers take up doctoral studies and many academic workers teach in schools, they are addressed twice to the same recipients.