

**Marek Zapieraczyński**

Okręgowa Komisja Egzaminacyjna w Łodzi

## **Perspektywy oceniania odpowiedzi do zadań otwartych na egzaminach zewnętrznych - sprawdzianie i egzaminie gimnazjalnym**

### **1. Jak są oceniane odpowiedzi do zadań otwartych ze sprawdzianu i egzaminu gimnazjalnego?**

Uczniowie klas szóstych szkoły podstawowej podczas sprawdzianu rozwiązują 5 do 6 zadań otwartych. Na tę liczbę składają się dwa lub trzy zadania krótkiej odpowiedzi (KO) i jedno rozszerzonej odpowiedzi (RO) z matematyki oraz jedno zadanie RO z języka polskiego (przeważnie jest to opis lub opowiadanie). Niekiedy zdarzały się też zadania polonistyczne KO (np. ogłoszenie). Z całego testu uczeń może uzyskać 40 punktów, 20 za rozwiązania zadań zamkniętych - wielokrotnego wyboru i 20 za odpowiedzi do wspomnianych zadań otwartych.

W części humanistycznej egzaminu gimnazjalnego uczniowie rozwiązują kilka zadań KO (7-8) oraz dwa lub trzy zadania RO, w tym zadanie dotyczące formy użytkowej (np. ogłoszenie) i zadanie wymagające sformułowania dłuższej wypowiedzi, najczęściej rozprawki. Na 50 możliwych do uzyskania punktów aż 30, czyli 60%, mogą uczniowie uzyskać za rozwiązanie zadań otwartych. Sama rozprawka to 15-16 punktów.

W części matematyczno-przyrodniczej egzaminu gimnazjalnego uczniowie rozwiązują kilka zadań KO i kilka zadań RO (razem 9-10 zadań) o treściach: matematycznych, fizycznych, chemicznych, geograficznych oraz biologicznych. Z tego egzaminu uczeń może uzyskać po 25 punktów za zadania zamknięte i otwarte.

Zadania zamknięte nie stanowią problemu w zakresie ich oceny, ponieważ uczeń, wskazując odpowiedź, właściwie sam sobie przyznaje punkt lub się go pozbawia. Problem pojawia się przy punktowaniu odpowiedzi do zadań otwartych, które oceniają przygotowani do tego egzaminatorzy. Uczestniczą oni co roku w dodatkowym szkoleniu przed przystąpieniem do oceniania prac. W trakcie pracy muszą kierować się odpowiednimi kluczami punktowania, które mają za zadanie obiektywizację oceniania.

### **2. Efektywność oceniania odpowiedzi do zadań otwartych zgodnie z przyjętą do tej pory metodą**

Egzaminatorzy oceniający odpowiedzi uczniowskie do zadań otwartych ze sprawdzianu w świetle obowiązujących przepisów i zgodnie z przyjętymi zwyczajami mogą niezależnie od wykształcenia sprawdzać wszystkie zadania (i matematyczne, i polonistyczne). Nauczyciel np. wychowania

fizycznego sprawdza zatem odpowiedzi zarówno polonistyczne, jak i matematyczne. Podobnie inni egzaminatorzy. Jak łatwo wywnioskować, obowiązująca formuła sprawdzania nie sprzyja fachowości i obiektywności oceniania prac. Aby właściwie wypunktować poprawność językową i interpunkcyjną w pracy polonistycznej, trzeba być praktykiem, czyli czynnym nauczycielem języka polskiego. Oczywiście nie można wykluczyć wyjątków - zdarza się, że wzorowo oceniają poprawność językową nauczyciele innych przedmiotów, ale są to wyjątki.

Niektóre z Okręgowych Komisji Egzaminacyjnych wprowadziły podział na egzaminatorów oceniających wyłącznie odpowiedzi matematyczne i na egzaminatorów oceniających tylko odpowiedzi polonistyczne. Miało to na celu ułatwienie pracy egzaminatorom i przede wszystkim podniesienie jakości ocen przez nich dokonywanych. Na podstawie testów diagnostycznych przeprowadzonych przez Centralną Komisję Egzaminacyjną, np. w 2008 roku, można wnioskować, że przyniosło to zamierzony skutek - komisje egzaminacyjne w Gdańsku i Łodzi (wprowadziły podział egzaminatorów na „polonistów” i „matematyków”) zanotowały znacznie mniej rozbieżności w ocenach egzaminatorów niż pozostałe komisje. W następnych latach kolejne komisje (np. warszawska i krakowska) zdecydowały się na zastosowanie podobnego zabiegu jak w Gdańsku i Łodzi. Niestety, nie możemy powiedzieć, że w 100% problem oceniania sprawdzianu został w ten sposób rozwiązany. Stanowisko takie potwierdzają dodatkowe weryfikacje ocenionych przez egzaminatorów odpowiedzi uczniowskich dokonywane przez pracowników OKE. Błędy w ocenie pewnej ilości prac uczniowskich ujawniane zostają, niestety, również w trakcie weryfikacji prac przed wglądami dokonywanymi przez uczniów i ich rodziców. Dotyczy to wszystkich poziomów egzaminowania. Co prawda, trudno oczekiwać, że tam, gdzie mamy do czynienia z czynnikiem ludzkim, błędów uda się całkowicie uniknąć, ale na pewno trzeba dążyć do minimalizowania tego zjawiska.

Aby zobrazować dobitniej poruszony problem, zaprezentuję niewłaściwą ocenę stosunkowo łatwych do oceny zdań matematycznych ze sprawdzianu w kwietniu 2010 roku.

Przykład 1.

21. Maksymalną długość nart dla skoczka oblicza się, mnożąc wzrost zawodnika przez 1,46. Oblicz maksymalną długość nart dla zawodnika o wzroście 1,5 m. Wynik wyraż w centymetrach.

Zapisz wszystkie obliczenia.

0  $1,5 \cdot 1,46 = 2,19 \text{ m}$


0  
Odpowiedź: Maksymalna..... długość nart..... dla..... zawodnika o.....  
wzroście 1,5 m wynosi 2,19 m.

Uczeń rozwiązał zadanie całkowicie poprawnie, ale zamiast 2 punktów dostał 0. Egzaminatorowi najwyraźniej zabrakło miana „m” w zapisie działania przy liczbie 1,5. Niezapisanie tego miana nie miało najmniejszego wpływu na rozwiązanie zadania. Miano wystąpiło w treści zadania i uczeń był świadom wykonywanych przez siebie czynności.

Przykład 2.

22. Podczas meczu koszykówki Paweł trafił do kosza 5 razy, Leszek miał 2 razy więcej trafień niż Paweł, a Zbyszek o 3 mniej niż Paweł i Leszek razem. Ile razy trafił do kosza Leszek, a ile Zbyszek?

0		$5 \cdot 2 = 10$	$10 + 5 = 15$	<del><math>15 - 3 = 12</math></del>																
0			$15 - 3 = 12$																	

Odpowiedź: Leszek trafił do kosza 10 razy, a Zbyszek 12 razy.

To zadanie także zostało rozwiązane całkowicie poprawnie, a sposób wypunktowania odpowiedzi jest niewłaściwy.

Jeszcze większe problemy występują podczas oceny zadania polonistycznego, czyli np. opowiadania. Uczeń może uzyskać za odpowiedź aż 10 punktów i dlatego konsekwencje związane z błędnymi decyzjami egzaminatora w tym przypadku są dla niego o wiele poważniejsze. Jeżeli egzaminator zbyt pochopnie dojdzie do wniosku, że wypowiedź ucznia jest nie na temat, uczeń może stracić nawet 7 punktów. Takie sytuacje, na szczęście nieliczne, zostały ujawnione podczas weryfikacji prac przed wglądami do prac. Mniej dotkliwe dla uczniów są błędy związane z oceną poprawności językowej, ortograficznej i interpunkcyjnej. Tu uczeń może stracić, a niekiedy zyskać, od jednego do trzech punktów. Oczywiście wypowiedzi polonistyczne implikują podczas oceniania znacznie więcej błędów egzaminatora niż odpowiedzi do zadań matematycznych, co wykazało wewnętrzne badanie przeprowadzone w OKE Łódź w 2007 roku. Okazało się wówczas, że oceniając prace ze sprawdzianu, egzaminatorzy popełniają cztery razy więcej błędów przy zadaniach polonistycznych niż przy ocenie odpowiedzi matematycznych.

Z podobnymi problemami w ocenianiu spotykamy się podczas egzaminu gimnazjalnego, którego znaczenie w kontekście wyniku branego pod uwagę przy rekrutacji do szkoły ponadgimnazjalnej jest nieporównywalnie większe niż sprawdzianu. Szczególnie dużo kontrowersji wzbudza ocenianie rozprawki przez nauczycieli historii i plastyki na egzaminie gimnazjalnym w części humanistycznej. Z kolei w części matematyczno-przyrodniczej wątpliwości może wzbudzać ocenianie zadań matematycznych czy fizycznych przez geografów i biologów. Oczywiście nie chodzi tu o deprecjonowanie kogokolwiek, ale niestety nie wszyscy egzaminatorzy chcą i potrafią realnie ocenić swoje możliwości oceniania zadań, z którymi w rzeczywistości szkolnej nie mają właściwie prawie nic wspólnego.

### 3. Skutki wynikające ze stosowania aktualnej formuły oceniania odpowiedzi do zadań otwartych

Wyniki otrzymywane przez uczniów podlegają niezamierzonemu przez system wpływowi zwanemu efektem egzaminatora. Efekt ten można zaobserwować, analizując wyniki prac uczniów, które zostały ocenione przez różne zespoły egzaminatorów. Wystarczy porównać liczbę punktów uzyskanych za zadania zamknięte w obu grupach uczniów, a następnie odnieść je do liczby punktów, które przydzielono uczniom za odpowiedzi do zadań otwartych. Jeżeli uczniowie z grupy I otrzymali średnio np. 13,2 pkt z zadań zamkniętych, a uczniowie grupy II analogicznie za te zadania 12,5 pkt, to jest mało prawdopodobne, aby uczniowie z grupy II mieli wynik wyższy za zadania otwarte<sup>1</sup>. Jednak, takie przypadki się zdarzają i to niekiedy na tyle drastyczne, że np. grupa II może otrzymać średnio nawet 1,5 pkt więcej od grupy I. W efekcie, sumując wyniki za zadania zamknięte i otwarte, otrzymujemy komunikat, że uczniowie z grupy II uzyskali średni wynik o 0,8 pkt lepszy od uczniów z grupy I. W rzeczywistości najprawdopodobniej tak nie jest, a uczniowie z grupy pierwszej są lepsi od uczniów z grupy II także w odpowiedziach na zadania otwarte. Można nawet postawić hipotezę, że ta różnica wynosi około 1 punktu, szacując na podstawie wyniku z zadań zamkniętych. Konkludując, przy zachowaniu porównywalności oceniania zadań otwartych grupa I powinna osiągnąć średni wynik około 1,5 - 1,7 wyższy za cały test od grupy II. Jednak „dzięki” efektowi egzaminatora otrzymaliśmy informację, że to grupa II lepiej napisała test, jej średni wynik, jak wspominałem, był wyższy o 0,8 pkt.

Istnieje obawa, że komunikowane w całej Polsce wyniki, które budzą przytoczone wyżej wątpliwości, przyczyniają się do wprowadzania w błąd zarówno opinii publicznej, jak i decydentów oświatowych.

Zaobserwowane przez lata reakcje egzaminatorów mogą poważnie wpływać na wynik pojedynczego ucznia. W świetle obserwacji poczynionych w sesjach 2009-2010 w OKE Łódź na podstawie testów diagnostycznych, jakie zastosowano wobec egzaminatorów sprawdzianu, ocena odpowiedzi ucznia do zadań otwartych różniła się nawet o 4 punkty.

Nie można zapominać, że wyniki uzyskane przez uczniów odzwierciedlają efekty pracy nauczycieli, a zatem informują o poziomie nauczania w danej szkole. Suma wyników placówek obrazuje poziom nauczania w gminie, powiecie, województwie, okręgu. W ten sposób kształtowany jest pogląd opinii publicznej, władzy (również centralnej) o osiągnięciach egzaminowanych populacji w poszczególnych regionach. I tak dowiadujemy się na przykład, że trzynastolatkowie z województwa podkarpackiego osiągają wyższe wyniki ze sprawdzianu w szóstej klasie niż ich rówieśnicy z województwa pomorskiego. Wzbudza to uznanie dla tych pierwszych tym bardziej, że w roku 2010 tylko uczniowie z mazowieckiego i małopolskiego byli jeszcze lepsi.

<sup>1</sup> Zapieraczyński M., Sapanowski S., Zbidi A., *Zależność wyników uzyskiwanych przez uczniów z zadań otwartych od liczby punktów uzyskanych za zadania zamknięte, [w:] Uczenie się i egzamin w oczach uczniów pod red. Niemierki B. i Szmigiel M.K., XIII Krajowa Konferencja Diagnostyki Edukacyjnej, Łomża, 5 - 7 października 2007, Łomża 2008, s. 344 - 352.*

Należy podkreślić, że władze oświatowe z południowo-wschodniej Polski podejmowały w latach 2003-2010 różnorakie działania w celu poprawy wyników sprawdzianu i egzaminów gimnazjalnych<sup>2</sup>. Znamy jednak makroprawidłowości dotyczące wyników uzyskiwanych przez uczniów w całej Polsce. Czego one dotyczą? Ograniczmy się do dwóch. Po pierwsze, uczniowie z dużych i średnich miast uzyskują wyniki znacząco wyższe od uczniów z małych miast i wsi. Relacje te dla województw pomorskiego i podkarpackiego kształtują się następująco:

#### **Odsetek szóstoklasistów w szkołach w miejscowościach różnej wielkości**

Województwo	Szóstoklasiści z dużego i średniego miasta	Szóstoklasiści z małego miasta i ze wsi
Pomorskie	46,1%	53,9%
Podkarpackie	26,3%	73,7%

**Źródło:** *Osiągnięcia uczniów kończących szkołę podstawową w 2009 roku. Sprawozdanie ze sprawdzianu 2009. Centralna Komisja Egzaminacyjna w Warszawie.*

#### **Wyniki sprawdzianu w zależności od lokalizacji szkoły**

	Liczba uczniów	Średni wynik
Wieś	161 024	23,5
Małe miasto	60 241	24
Średnie miasto	73 975	25,2
Duże miasto	88 388	26,4

**Źródło:** *Wstępne informacje o wynikach sprawdzianu 2010, Centralna Komisja Egzaminacyjna, Warszawa, 28 maja 2010 r.*

Jak łatwo zauważyć, makroprawidłowości ogólnopolskie nie przekładają się na średni wynik uczniów z omawianych dwóch województw. Budzi to zainteresowanie, dlaczego tak się stało. Czy jest to anomalia, a może jednak prawidłowość? A może należy to dokładniej przeanalizować, może przeprowadzić dodatkowe badanie lub badania? Być może sprawa nie jest warta zachodu. Ale tu pojawia się następna wątpliwość: zasobność danej grupy obywateli przekłada się również na osiągnięcia egzaminacyjne ich dzieci – im większe dochody, tym wyższe wyniki<sup>3</sup>. Dochody mieszkańców podkarpackiego są niższe od tych, które osiągają mieszkańcy pomorskiego. Ilustruje to mapka zamieszczona poniżej. Region pomorski jest o wiele bardziej zasobny od podkarpackiego, który okazuje się najbiedniejszym w całej Polsce. Rozwija się także wolniej od pomorskiego<sup>4</sup>.

<sup>2</sup> Pismo podpisane przez panią A. Pieniążek Dyrektora Wydziału Ewaluacji KO w Rzeszowie „Działania Podkarpackiego Kuratora Oświaty podjęte w ramach nadzoru pedagogicznego w celu poprawy wyników sprawdzianów i egzaminów gimnazjalnych – od 2003 roku. Rzeszów, 2010-09-03”.

<sup>3</sup> Jakubowski M. *Wpływ czynników ekonomicznych na wyniki egzaminów zewnętrznych*, [w:] Biuletyn Badawczy CKE „Egzamin”, 11/2007.

<sup>4</sup> Podkarpackie to region biedy. Z danych Głównego Urzędu Statystycznego wylania się obraz typowego mieszkańca tego regionu - chorego, biednego, bardziej balującego niż uczącego się. Główny Urząd Statystyczny, porównując wszystkie województwa w kraju, przedstawił najnowsze dane. Podkarpackie jest najbiedniejsze w Polsce. PKB (Produkt Krajowy Brutto) na jednego mieszkańca woj. podkarpackiego wynosi 20 829 złotych, czyli 67,5 proc. średniej krajowej. Ciągłe też mieszkańcy Podkarpacia zarabiają mniej niż wynosi średnia w innych regionach kraju. Dwa lata temu miesięczny dochód na jedną osobę wynosił około 660 zł. To o 250 zł mniej niż średnia w Polsce. Może dlatego bardzo dużo osób korzysta ze świadczeń pomocy społecznej. Wyliczono, że na 10 tysięcy osób pomoc społeczna zajmuje się aż 834 osobami. To o 96 osób więcej niż średnia w kraju. [www.interia.pl]

#### 4. Jak poprawić efektywność oceniania zadań otwartych?

Podstawową przyczyną zakłóceń w punktacji tego typu zadań są oczywiście indywidualne skłonności egzaminatorów, którzy mimo konieczności stosowania kluczy odpowiedzi wykazują tendencje do surowości bądź łagodności w ich stosowaniu. Do przyczyn leżących po stronie indywidualnych skłonności ocenających dochodzą systemowe czynniki zakłócające obiektywność oceniania. Są one o wiele groźniejsze, ponieważ ich działanie jest bardziej wymierne. Wspomniałem o indywidualnych skłonnościach egzaminatorów do surowości bądź łagodności; niestety zachodzi obawa, że zjawisko to może występować również wśród koordynatorów egzaminów w poszczególnych okręgach oraz województwach. Czy dotyczy to województw pomorskiego i podkarpackiego? Nie przesądzam sprawy. Dopuszczam możliwość pomyłki i dlatego chętnie poddam moje przypuszczenia próbie weryfikacyjnej. Gdyby okazało się, że niestety mam rację, to wyniki egzaminów zewnętrznych w poszczególnych regionach Polski mogą być nieadekwatne do rzeczywistych możliwości populacji, które opisują.

W jaki sposób sprawdzić stan faktyczny? Wydaje się, że dobrą okazję stwarza podjęta przez dr. H. Szaleńca próba wprowadzenia do systemu egzaminowania tzw. e-oceniania.<sup>5</sup> Zakłada ono skanowanie wszystkich prac uczniowskich i sprawdzanie przez egzaminatorów poszczególnych zadań, oczywiście tych, z którymi najlepiej sobie radzą; sprzyja to podniesieniu jakości pracy egzaminatora, ponieważ jego praca ma być także stale monitorowana. W przypadku błędnego oceniania egzaminator może zostać odsunięty od sprawdzania, a zadania, które ocenił, można poddać dodatkowej weryfikacji. Równie istotną zaletą tego sposobu oceniania prac jest możliwość ich nieograniczonego dystrybuowania do różnych egzaminatorów i przy wykorzystaniu maksymalnych możliwości systemu elektronicznego doprowadzenie do sytuacji, w której egzaminator sprawdzi kilkaset rozwiązań uczniowskich, każde z nich będzie z innej szkoły i innego powiatu (gminy w kraju). Te dwa czynniki, czyli podniesienie jakości oceniania i zniwelowanie efektu subiektywizmu regionalnego, mogą, nawet powinny, w bardzo radykalny sposób wpłynąć jakościowo na wyniki indywidualne oraz okręgowe (wojewódzkie, powiatowe) i bardzo poważnie je zobiektywizować, a co za tym idzie, uwiarygodnić.

Jedynym czynnikiem zakłócającym obiektywną ocenę potencjału poszczególnych uczniów oraz większych populacji byłby już tylko czynnik ściągania (podpowiadania), czyli łamanie procedur przeprowadzania egzaminów. To, że zjawisko to występuje, nie można mieć wątpliwości. Świadczą o tym unieważnione egzaminy oraz sytuacje, których nie udało się jednoznacznie zinterpretować i tylko dlatego nie zostały one napiętnowane. Pytaniem pozostaje skala tego zjawiska. Oczywiście trudno to określić ze względu na powszechność egzaminów oraz brak możliwości monitorowania tego procesu. Niezależnie jednak od ilości ściągających zmiana procedur przeprowadzania egzaminów (zespoły nadzorujące powinny składać się z osób, które nie są bezpośrednio zainteresowane wynikami uczniów) na pewno przyczyniłaby się do zmarginalizowania wspomnianego problemu. Po zastosowaniu tych zabiegów pozyskalibyśmy

<sup>5</sup> Projekt I „PRZYGOTOWANIE SYSTEMU INFORMATYCZNEGO DO E-OCENIANIA” Działanie 3.2 „Rozwój systemu egzaminów zewnętrznych”, Priorytet III „Wysoka jakość systemu oświaty”, Program Operacyjny KAPITAŁ LUDZKI, koordynator H. Szaleńiec.

wreszcie wyniki z egzaminów zewnętrznych, o których można by powiedzieć z dużą pewnością, że odzwierciedlają stan rzeczywisty wiedzy i umiejętności uczniów. W końcu egzaminy po to są przeprowadzane, a nie, aby zapewnić dobre samopoczucie „najzaradniejszym” nauczycielom i szkołom.

## 5. Możliwości eksploracji wyników egzaminów zewnętrznych po zastosowaniu e-oceniania

Dane wpisane do systemu w wyniku sprawdzenia prac uczniowskich dają olbrzymie możliwości ich analizowania i przewidywania w przyszłości określonych skutków podjętych działań. Od dwudziestu lat prowadzone są na szeroką skalę działania w kierunku pozyskiwania wiedzy z danych (*Data Mining*). W Polsce eksploracja danych pozyskiwanych podczas egzaminów zewnętrznych wydaje się być wymarzonym obszarem, który po odpowiednim opracowaniu może dać odpowiedzi na wiele pytań. Uzyskana wiedza może umożliwić wreszcie realną ocenę stanu polskiej oświaty. Być może wyciągnięte w ten sposób wnioski, poparte innymi badaniami, pozwolą obiektywnie rozstrządać np. to, czy dobry wynik, jaki uzyskali na egzaminie gimnazjalnym uczniowie szkoły, jest rzeczywiście efektem realizacji znakomitej oferty edukacyjnej dostarczonej im przez nauczycieli, czy zabiegów rodziców, które sprowadzają się do zapewniania dzieciom dodatkowych zajęć i ponoszenia ogromnych kosztów z tym związanych. Pytania tego typu od wielu lat nurtują zainteresowanych. Krążą także niezwerifikowane opinie, że dobre zdanie o wielu szkołach wynika z pracy osób, które z placówkami tymi nie mają nic wspólnego. Inne z kolei, pracując w przeciętnych środowiskach, uzyskują niezłe wyniki edukacyjne i nie narażają uczniów oraz ich rodziców na dodatkowe zajęcia pozaszkolne i związane z tym koszty. W efekcie często jest tak, że splendor i efekty finansowe uzyskują ci, którym się one nie należą, a rzetelni nauczyciele mogą być najwyżej sfrustrowani. Chyba, że wystarczy im satysfakcja z dobrze wykonanej pracy i wdzięczność młodzieży. Niezależnie od tego instytucje zajmujące się nadzorem pedagogicznym czy wykonujące badania diagnostyczne są zobowiązane z natury rzeczy do właściwego i realistycznego oceniania poczynań dydaktycznych poszczególnych placówek.

### Bibliografia:

1. Larose D.T., *Odkrywanie wiedzy z danych. Wprowadzenie do eksploracji danych*, Warszawa 2006.
2. Larose D.T. *Metody i modele eksploracji danych*, Warszawa 2008.
3. Zapieraczyński M., Sapanowski S., Zbidi A., 2. *Zależność wyników uzyskiwanych przez uczniów z zadań otwartych od liczby punktów uzyskanych za zadania zamknięte*, [w:] *Uczenie się i egzamin w oczach uczniów pod red. Niemierki B. i Szmigiel M.K., XIII Krajowa Konferencja Diagnostyki Edukacyjnej*, Łomża 2007.
4. Zapieraczyński M., *Wiarygodność wyników egzaminu zewnętrznego – sprawdzian w szkole podstawowej*, [w:] *Uczenie się i egzaminy w oczach nauczycieli*, red. B. Niemierko, Opole 2008.
5. *Sprawozdania ze sprawdzianów w szkole podstawowej w latach 2007 - 2009*. Centralna Komisja Egzaminacyjna w Warszawie.