

Henryk SZALENIEC  
Maria Krystyna SZMIGEL  
OKE Kraków

## MODEROWANIE OCENIANIA PRAC UCZNIOWSKICH

Jednym z najtrudniejszych zadań podczas oceniania prac z egzaminu jest zapewnienie porównywalności przy jednoczesnej trosce, aby bogactwo uczniowskich rozwiązań zarówno typowych, jak i kreatywnych znalazło wyraz w wyniku oceny. Nie wszystkie zadania pisemnych egzaminów dają uczniowi sposobność ujawnienia umiejętności unikatowego i twórczego rozwiązania, ale jeżeli takie perełki się trafią, to egzaminator powinien je rozpoznać i ocenić właściwie, mimo że pozornie np. nie mieszczą się w schemacie oceniania. Badania kompetencji z języka polskiego i matematyki uczniów klas ósmych w województwie małopolskim, wykorzystujące arkusze egzaminacyjne zbudowane z zadań otwartych, były eksperymentem, który umożliwił testowanie różnych dróg dochodzenia do oceniania w porównywalny sposób przez duże zespoły egzaminatorów.

## RODZAJE MODEROWANIA OCENIANIA PRAC UCZNIOWSKICH

W 2000 r., kiedy to prawie 740 zewnętrznych egzaminatorów (z matematyki i języka polskiego) oceniało około 53 000 prac uczniów, można było zaobserwować, jak istotne znaczenie w tej dziedzinie ma doświadczenie i informacja zwrotna o jakości oceniania dostarczana przez lidera zespołu oceniającego. Działania prowadzące do zapewnienia porównywalności oceniania na egzaminach zewnętrznych nazywane są *m o d e r o w a n i e m* (Gipps 1994).

Procesy obecne w moderowaniu zmieniają się od twardych do miękkich. Twarde moderowanie, zorientowane statystycznie, polega na stosowaniu poprawek do ocen niektórych ocenających i normowaniu rozkładu wyników tak, aby na przykład zapewnić porównywalność wyników pomiędzy szkołami. Miękkie podejścia to dyskusje na temat oceniania w zespole egzaminatorów i moderowanie grup ocenających. Z logicznego punktu widzenia to drugie podejście wydaje się korzystniejsze dla procesu oceniania i zapewnienia rzetelności. Jakikolwiek podejście przyjmiemy, to cel jest jeden – zapewnienie zgodności i jakości oceniania i egzaminowania.

W praktyce egzaminów w wielu krajach stosowane są różne typy moderowania, wśród których możemy wyróżnić takie, jak:

- moderowanie statystyczne,
- skalowanie wyników,
- moderowanie przez wizytację,
- panel moderacyjny,
- moderowanie grupowe,
- moderowanie konsensusu.

Moderowanie statystyczne wymaga zastosowania procedur statystycznych do standaryzacji wyników egzaminów z wykorzystaniem wyników innego pomiaru. Procedura ta nie wymaga zaangażowania do tego celu nauczycieli. Stosowany jest test porównawczy i na jego podstawie są normowane wyniki średnie poszczególnych szkół. Tak więc wewnątrz szkoły hierarchia wyników poszczególnych uczniów nie ulega zmianie, ale wynik szkoły może być podniesiony lub obniżony stosownie do wyniku testu bazowego. Taki system jest stosowany w Australii dla egzaminu, który jest egzaminem wstępnym do szkoły średniej.

Inna metoda to skalowanie. W Australii metoda ta jest stosowana do zapewnienia porównywalności wyników pomiędzy egzaminami z różnych przedmiotów. Niektóre przedmioty, jak matematyka i przedmioty przyrodnicze, są wybierane przez bardziej zdolnych uczniów, dlatego też trudno jest uzyskać ocenę wysoką z tego przedmiotu, jeżeli egzamin opiera się na pomiarze różnicującym. Proces skalowania umożliwia kompensację oceny dla uczniów wybierających na egzaminie taki trudny przedmiot, z którego nielato jest uzyskać wysoką ocenę. Ocena wyrażona stopniem czy też nawet procentem punktów niewiele nam mówi o umiejętnościach zdającego, gdy nie znamy „surowości” egzaminu. Dając więcej ocen zdającemu tworzymy dokładniejszy obraz jego umiejętności i potencjału intelektualnego. Przy zastosowaniu systemu skalowania uczniowie wybierający trudniejszy przedmiot dostają premię punktową zależną od przedmiotu. Często podnosi się fakt, że w wyniku takiej procedury zmienia się ranga wyników poszczególnych zdających i z tego systemu częściej korzystają chłopcy, gdyż oni mają tendencję do wybierania trudniejszych przedmiotów egzaminacyjnych.

Inny sposób zwiększenia rzetelności egzaminu poprzez podniesienie porównywalności oceniania to moderowanie przez wizytację (inspekcję). Wiąże się ono z ponownym sprawdzeniem pewnej liczby prac uczniowskich, uprzednio ocenionych przez nauczycieli danej placówki. Oceny dokonują pracownicy instytucji odpowiedzialnej za egzaminy. Sprawdzeniu podlegają nie tylko prace uczniowskie, ale i procedura przeprowadzania egzaminu. Moderowanie przez wizytację stosowane jest w Polsce podczas egzaminów maturalnych. W programie „Nowa Matura” w Regionalnej Komisji Egzaminów Szkolnych w Krakowie eksperymentalnie rozszerzono ten model moderowania również na ocenę prac uczniowskich (Bajorski 1999). Moderowanie przez wizytację dotyczy małej liczby szkół. Ma ono wiele zalet dla profesjonalnego rozwoju nauczycieli w dziedzinie egzaminowania. Doświadczenia uzyskane w objętych moderowaniem przez wizytację szkołach mogą być upowszechniane we wszystkich placówkach poprzez spotkania organizowane przez kuratorium oświaty oraz warsztaty prowadzone przez placówki doskonalenia nauczycieli. Moderowanie przez wizytowanie jest rozpowszechnione w wielu krajach na egzaminach zewnętrznych na poziomie kolegów.

Innym rodzajem moderowania jest panel moderacyjny, który polega na spotkaniu ekspertów do spraw oceniania z nauczycielami. Oceniający dochodzą do konsensusu w zakresie przydzielania odpowiedniej liczby punktów na skali za określoną część pracy pisemnej. Jest to dochodzenie do konsensusu w interpretacji i posługiwaniu się uzgodnionym schematem oceniania. Celem zajęć panelowych jest raczej upewnienie się, że oceny będą przydzielane tak, jak wspólnie uzgodniono, niż wpływanie na sam proces oceniania. Obecnie panel moderacyjny stosowany jest do podniesienia porównywalności oceniania

prac maturalnych („stara” matura) w szkołach leżących na terenie działania OKE w Krakowie. Uczestnictwo w panelu jest dobrowolne, a jego szeroki zasięg uzyskiwany jest przez kaskadę. I tak np. 10 maja 2000 r. pracownicy OKE w Krakowie prowadzili panel moderacyjny dla nauczycieli specjalistów przedmiotowych rekrutujących się z poszczególnych delegatur kuratoriów oświaty w obrębie działania OKE w Krakowie. Tak przygotowani nauczyciele w swoich delegaturach prowadzą panele moderacyjne dla nauczycieli.

Ciekawą i skuteczną modyfikacją panelu moderacyjnego jest moderowanie grupowe. Podczas moderowania grupowego nauczyciele oceniający oraz specjaliści w dziedzinie oceniania dyskutują na temat wybranej pracy uczniowskiej, która podlega ocenie. Celem takich zajęć jest dojście do takiego samego rozumienia kryteriów oceniania. Brany jest pod uwagę zarówno sam proces, jak i wyniki oceniania pracy uczniowskiej przez wszystkich członków grupy. Spotkania przedstawicieli wszystkich szkół mają zapewnić zgodność oceniania w całym systemie edukacyjnym na danym poziomie szkolnym. Takie rozwiązanie jest bardziej kosztowne niż spotkania wewnątrz szkół, ale ma również ogromne znaczenie w doskonaleniu nauczycielskich umiejętności oceniania. Ten typ moderowania został zastosowany we wrześniu i październiku 1998 r. podczas próbnych badań kompetencji z języka polskiego uczniów klas ósmych szkół zlokalizowanych w miejscowościach, które od stycznia 1999 r. weszły w skład województwa małopolskiego.

Bardziej zaawansowanym organizacyjnie rozwiązaniem jest moderowanie konsensusu, które obejmuje zarówno panel moderacyjny, jak i działania, które mają doprowadzić do ustaleń w procedurze oceniania i posługiwaniu się kryteriami. W pracach uczestniczą zarówno eksperci, jak i nauczyciele oceniający. Takie moderowanie zbliżone jest do moderowania grupowego. Wprowadzano go stopniowo przez trzy lata w badaniach kompetencji uczniów klas ósmych w województwie małopolskim podczas koordynacji oceniania prac uczniowskich przez kilkuset nauczycieli powołanych w Rejonowych Komisjach Oceniających. Okręgowa Komisja Egzaminacyjna w Krakowie stosowała moderowanie konsensusu podczas oceniania 8000 prac uczniów biorących udział w badaniach pilotażowych (nowych egzaminów) w lutym i marcu 2000 r.

## PANEL MODERACYJNY NA PRZYKŁADZIE JEDNEGO ZADANIA OTWARTEGO Z MATEMATYKI

Aby bliżej zapoznać się z ideą moderowania konsensusu, prześledzimy pracę zespołów oceniających wyniki uczniów z badań kompetencji z matematyki podczas dwóch wydarzeń badawczych, które miały miejsce w 1988 i w 2000 r. Za każdym razem ograniczymy się tylko do jednego zadania. Do analizy wybraliśmy zadania sprawdzające podobne czynności — umiejętność zastosowania reguły kolejności działań podczas mnożenia i dodawania liczb wymiernych różnych znaków. Aby można było doskonalić tę metodę poprzez analizę kolejnych etapów postępowania, praca zespołów była rejestrowana amatorską kamerą video.

Prezentowany materiał dokumentacyjny pochodzi z panelu moderacyjnego, który odbył się 19 maja 1998 r. w siedzibie Kuratorium Oświaty w Krakowie. Pracami zespołu kierowali przewodniczący zespołu przedmiotowego matematyków, który zatwierdzał na-

rządza do badania oraz autorka zespołu przygotowującego testy do badań. W panelu brało udział 10 przewodniczących Rejonowych Komisji Oceniających, którzy później kierowali ocenianiem prac w swoich zespołach.

Panele prowadzono także w kolejnych latach podczas badań kompetencji uczniów klas ósmych z matematyki i języka polskiego w województwie krakowskim, a następnie małopolskim, rozszerzając liczbę osób w nim uczestniczących do 30. Przewodniczący każdej komisji był odpowiedzialny za zachowanie porównywalności oceniania 2000 prac uczniów.

W wyniku doświadczeń przeszliśmy od prostego panelu moderacyjnego, poprzez moderowanie grupowe (w 1999 r.), do moderowania konsensusu w 2000 r.

Pierwszy panel rozpoczął się 20 minut po tym, jak złamano pieczęcie na kopertach z testami, a uczniowie pogrążyli się w rozwiązywaniu zadań. Zgromadzeni w jednej sali Kuratorium Oświaty w Krakowie przewodniczący zespołów oceniających weszli w rolę uczniów i dostali do rozwiązania te same zadania co ósmoklasiści. Po rozwiązaniu wszystkich zadań zapoznali się ze schematem oceniania zaproponowanym przez autorów narzędzi do badań, schematem oceniania w kontekście sprawdzanych czynności zadaniami testu. Upewnili się co do sposobu i kolejności rejestracji oceny oraz poprawności wykonania złożonej czynności badanej zadaniem I.

Materiał zarejestrowany na taśmie video wyraźnie oddaje, jak zróżnicowane są wewnętrzne modele odniesienia, które są wpisane w struktury poznawcze każdego z nauczycieli. Zostały one ukształtowane podczas wieloletniej praktyki szkolnej. Opierając się na tych subiektywnych wewnętrznych modelach odniesienia poszczególni nauczyciele chcą kreować model oceniania pracy egzaminacyjnej.

Dopiero w toku dyskusji nad schematem oceniania w kontekście zapisanych własnych rozwiązań zadania przewodniczący zespołów oceniających uświadomili sobie, że te modele są różne. Dyskusja poparta przykładami zakotwiczonymi w praktyce własnej klasy ujawniła też genezę tych naturalnych i nieuniknionych różnic. Uczestnicy panelu uświadomili sobie, że po wyjściu z sali przejdą do zespołów z uzgodnionym schematem i modelem oceniania prac, bogatsi o doświadczenia nie tylko własne, ale i koleżanek oraz kolegów, przekonani o potrzebie konsekwentnego trzymania się schematu oceniania w trakcie oceniania prac uczniowskich.

Przyjrzyjmy się temu prostemu zadaniu:

#### ZADANIE 1.

$$2\frac{1}{4} + \frac{1}{2} \cdot (-1\frac{1}{2}) =$$

Schemat oceniania obejmował odpowiedź odniesienia oraz punktację. Indywidualną odpowiedź odniesienia tworzył każdy nauczyciel, rozwiązując to zadanie na początku panelu.

Odpowiedź odniesienia		Punktacja 0-2		
$2\frac{1}{4} + \frac{1}{2} \cdot (-1\frac{1}{2}) = \frac{9}{4} + \frac{1}{2} \cdot (-\frac{3}{4}) = \frac{6}{4} = 1\frac{2}{4} = 1\frac{1}{2}$		1A - 1 punkt 1B - 1 punkt 0 w innych niż opisane niżej przypadkach*		
I	1A Zastosować regułę kolejności wykonania działań oraz pomnożyć liczby wymierne różnych znaków	77,2%	75,8%	0 pkt 19,0 1 pkt 10,2 2 pkt 70,0
	1B Dodać liczby wymierne różnych znaków	75,5%		

\* Uczeń uzyskuje tylko 1 punkt, gdy:

- a) poprawnie pomnożył dane ułamki, a następnie błędnie dodał otrzymane liczby lub  
b) pomylił się w mnożeniu, ale poprawnie obliczył otrzymaną sumę.

W obu ww. przypadkach otrzymana w wyniku wartość liczbową wyrażenia była oczywiście błędna. Zapis na karcie wynikowej 1, 0.

W wyniku oceny prac uczniowskich stwierdzono, że najczęściej uczniowie popełniali błędy w mnożeniu liczb wymiernych o różnych znakach oraz (lub) błędy rachunkowe w dodawaniu liczb wymiernych o różnych znakach. Raport z badań uwzględnia wszystkie błędne sposoby rozwiązania zadania.

Zastosowane dwa lata później, w roku szkolnym 1999/2000, zadanie było następujące:

$$\text{Oblicz: } 14 + 56 : 7$$

#### Schemat oceniania

Odpowiedź odniesienia		Punktacja 0-2		
$14 + 56 : 7 = 14 + 8 = 22$		1A - 1 punkt 1B - 1 punkt 0 w innych niż opisane niżej przypadkach*		
I	1A Zastosować regułę kolejności wykonania działań	94%	94%	
	1B Wykonać dzielenia i dodawania	96%		

Na zakończenie zwróćmy jeszcze uwagę na kilka charakterystycznych sformułowań, które padały podczas dyskusji nad ocenianiem pierwszego zadania w trakcie panelu moderacyjnego.

- Nie interesuje mnie, jak on to rozwiązał, przecież uzyskał dobry wynik.
- Jeżeli uczeń uzyskał w zadaniu dobry wynik, to co to znaczy? Jaką informację o swoich umiejętnościach przekazał uczeń?
- Czy można karać ucznia, że większość obliczeń wykonuje w pamięci?
- Przecież to jest praca pisemna, uczeń ma się wypisać.
- Ja bym dała 1 punkt, jeżeli uczeń uzyskał poprawny wynik.
- Nie, jeżeli tylko wynik poprawny, a brak zapisu, to uczeń powinien otrzymać 0 punktów.

- Ja bym też dała 0 punktów, tak oceniamy na szkolnej klasówce.
- Mam poważne wątpliwości, czy w tym wypadku powinno być tylko 0.
- Ja bym dała 1 punkt, może poszczególne etapy obliczeń były w brudnopisie, a tutaj uczeń w pośpiechu przeniósł tylko wynik.
- Poczekajmy, aż ocenimy inne prace. Może większość bardzo dobrych uczniów tak rozwiązuje.
- Jeżeli w tekście zadania było napisane, że uczeń ma obliczyć, to powinien obliczyć, a nie zgadywać. Gdyby miał odgadnąć wynik, to w zadaniu byłoby — zgadnij.
- Może on obliczył, tylko nie zapisał?
- A co ze słowną odpowiedzią?
- Przecież uczeń na podstawie zapisu zadania nie wie, kiedy jest konieczna słowna odpowiedź, a kiedy nie.
- Ale przecież tak uczymy.
- A może przyjmijmy inne rozwiązanie, dajmy 1 punkt za metodę rozwiązania i 1 za poprawny wynik.
- Czy jednak zastosowana metoda i wynik powinny być równo traktowane?
- Wróćmy do tego, co miało sprawdzać to zadanie i czy uczeń uwidoczniał w rozwiązaniu opanowanie tej umiejętności.
- Są kanony w matematyce, od których odstępować nie można, i jest nim reguła kolejności działań. Jeżeli jej nie stosuje, to wszelkie dalsze przyznawanie punktów za to zadanie nie powinno być możliwe.
- Ale to jest pierwsze zadanie na rozgrzewkę, a jak okaże się trudne, jeżeli będziemy w ten sposób punktować?
- A jeżeli stosuje regułę, ale pomylił się w mnożeniu?”

Ileż wątpliwości wyrazili uczestnicy panelu, a przecież były to najłatwiejsze zadania w teście, za które większość uczniów uzyskiwała 2 punkty. Lepiej, że takie wątpliwości są uwidoczniane i analizowane przed przystąpieniem do oceniania. Im więcej czasu poświęcimy na dochodzenie do konsensusu, tym większe prawdopodobieństwo, że uda nam się rzetelniej oceniać uczniowskie prace.

## PODSUMOWANIE

Dyskusji podczas panelu moderacyjnego musimy się uczyć. Bardzo ważne jest słuchanie każdej wypowiedzi, ponieważ czasem odosobniony punkt widzenia może być na wagę rzetelności całego egzaminu. Podczas panelu moderacyjnego w badaniach pilotażowych z języka niemieckiego prowadzonych przez OKE w lutym i marcu 2000 r. tylko dzięki stanowczej postawie jednego z oceniających, który miał inne zdanie niż pozostali członkowie zespołu, udało się wyeliminować błąd ukryty w schemacie oceniania. Fakt ten świadczy o tym, że nie zawsze zdanie większości ekspertów bliższe jest poprawnemu rozwiązaniu. Dobrze przeprowadzone moderowanie konsensusu daje głębsze zrozumienie schematu oceniania, ocenę stopni swobody dla egzaminatora oraz pozwala lepiej zaasymilować schemat oceniania do wewnętrznego modelu oceniania każdego z nauczycieli biorącego w nim udział. Dyskusja podczas panelu pozwala także lepiej osądzić własne nauczycielskie

doświadczenie w ocenianiu w kontekście praktyki innych specjalistów. Poznanie poglądów innych praktyków na ten temat znacznie wzmacnia istotnie nasze indywidualne doświadczenie, które jest tak bardzo potrzebne zewnętrznemu egzaminatorowi.

Aby jednak udało się podczas egzaminów próbnych, które odbędą się jesienią 2001 r., zdobyć konieczne doświadczenie w tej dziedzinie, warto dokumentować, analizować i korzystać z wszystkich dotychczasowych doświadczeń, które już miały miejsce w tej dziedzinie w długim procesie dochodzenia do egzaminów zewnętrznych w polskiej oświacie.

## LITERATURA

- Bajorski A., 1999, *Badanie pisemnych prac maturalnych*, [w:] *W kierunku nowych egzaminów szkolnych*. Kraków.
- Gipps C. V., 1984, *Beyond testing. Towards a theory of educational assessment*, London.
- Gipps C. V., Brown M., McCallum B., 1995, *Intuition or evidence, teachers and national assessment of seven-year-olds*, Philadelphia, Buckingham.
- Niemierko B., 1999, *Pomiar wyników kształcenia*, Warszawa.
- Noizet G., Caverni J. P., 1988, *Psychologiczne aspekty oceniania osiągnięć szkolnych*, Warszawa.
- Szalaniec H., Szmigel M. K., 2000, *Informator [OKE Kraków] o wynikach badań kompetencji uczniów klas ósmych 2000*, cz. 1, Kraków.
- Szalaniec H., Szmigel M. K., 2001, *Egzaminy zewnętrzne. Podnoszenie kompetencji nauczycieli w zakresie oceniania zewnętrznego*, Kraków.
- Szalaniec H., Szmigel M. K., 2001, *Raport z badań kompetencji*, maszynopis znajdujący się w Kuratorium oświaty i OKE w Krakowie.
- Szalaniec H., Szmigel M. K., 2001, *Z prac nad porównywalnością wyników oceniania zewnętrznego*, [w:] K. Wenta (red.), *Pomiar edukacyjny jako kompetencje pedagogiczne*, materiały z konferencji naukowej zorganizowanej przez Instytut Pedagogiki US 28–30 maja 2000 r., Szczecin, s. 357–367.