

Filip Kulon

Instytut Badań Edukacyjnych

Zespół badawczy EWD

Ocena systemu egzaminów zewnętrznych przez nauczycieli szkół gimnazjalnych i podstawowych

Na przełomie 2009 i 2010 roku Zespół Edukacyjnej Wartości Dodanej przy Centralnej Komisji Egzaminacyjnej rozpoczął panelowe „Badania osiągnięć szkolnych uczniów”¹. Opis koncepcji badania oraz jego realizacji można znaleźć w artykule Zofii Lisieckiej².

Jednym z celów badania jest poznanie tego, co o systemie egzaminów zewnętrznych myślą osoby odgrywające w nim jedną z kluczowych ról – nauczyciele. Niniejsze opracowanie zawiera ilościową analizę danych zebranych za pomocą ankiet kierowanych do nauczycieli szkół gimnazjalnych i podstawowych.

Wszystkie obliczenia wykonano w programie PASW Statistics 18 z uwzględnieniem wag probabilistycznych³. Wagi replikacyjne nie zostały niestety zastosowane, gdyż użyte oprogramowanie nie oferuje takiej funkcji. W związku z tym należy mieć na uwadze to, że opisywane zależności mogą przy użyciu wag replikacyjnych okazać się nieistotne statystycznie. Dokładny opis przyjętego schematu losowania oraz systemu wag można znaleźć w raporcie technicznym z badania.

Opinie o systemie egzaminów zewnętrznych

Jednym z elementów kwestionariuszy było zapytanie wprost nauczycieli o ich opinie dotyczące egzaminów zewnętrznych. Pytanie odnosiło się do dwóch aspektów egzaminów – organizacji egzaminu oraz jakości zadań. W przypadku szkół gimnazjalnych pytano osobno o jakość zadań w części humanistycznej i w części matematyczno-przyrodniczej egzaminu gimnazjalnego, natomiast w przypadku szkół podstawowych nie użyto rozróżnienia na typ zadań. Takie rozwiązanie wiąże się ściśle z podziałem egzaminu gimnazjalnego na osobne części. Nauczyciele mieli ocenić wymienione elementy w perspektywie lat 2002-2009, czyli od początku istnienia systemu egzaminów zewnętrznych. Z racji tego, że egzamin gimnazjalny z języka obcego pojawił się dopiero

¹ Badania panelowe w szkołach podstawowych i gimnazjach realizowane są w ramach projektu Centralnej Komisji Egzaminacyjnej „Badania dotyczące rozwoju metodologii szacowania wskaźników edukacyjnej wartości dodanej (EWD)”, współfinansowanego przez Unię Europejską w ramach Europejskiego Funduszu Społecznego – Program Operacyjny Kapitał Ludzki, Priorytet III Wysoka jakość systemu oświaty, Działanie 3.2 Rozwój systemu egzaminów zewnętrznych.

² Z. Lisiecka *Podłużne badania panelowe w szkołach podstawowych i gimnazjach a rozwój wskaźnika EWD*, [w:] *Teraźniejszość i przyszłość oceniania szkolnego*, red. B. Niemierko, M.K. Szmigel, Kraków, 2010

³ Liczba obserwacji w bazach danych wynosi odpowiednio 2839 dla gimnazjów i 2432 dla szkół podstawowych. Użycie wag skutkuje jednak obniżeniem efektywnej liczby obserwacji do 2245 dla gimnazjów i 2134 dla szkół podstawowych.

w 2009 roku, nie pytano o tę część. Treść pytań oraz rozkład odpowiedzi dla nauczycieli gimnazjów przedstawia tabela 1., a dla nauczycieli szkół podstawowych tabela 2.

Tabela 1. Rozkład odpowiedzi dla pytania A.4. z ankiety dla nauczycieli gimnazjów

Czy w perspektywie lat 2002-2009 egzamin gimnazjalny jest ze względu na:	coraz lepszy	coraz gorszy	mniej więcej taki sam	trudno powiedzieć	brak danych
organizację egzaminu przez system egzaminacyjny	33%	2%	52%	11%	2%
jakość zadań w części humanistycznej	10%	4%	44%	35%	7%
jakość zadań w części matematyczno-przyrodniczej	11%	5%	45%	32%	6%

Tabela 2. Rozkład odpowiedzi dla pytania A.4. z ankiety dla nauczycieli szkół podstawowych

Czy w perspektywie lat 2002-2009 sprawdzian po szkole podstawowej jest ze względu na:	coraz lepszy	coraz gorszy	mniej więcej taki sam	trudno powiedzieć	brak danych
organizację egzaminu przez system egzaminacyjny	30%	3%	48%	18%	0%
jakość zadań egzaminacyjnych	12%	12%	48%	29%	0%

Pobieżna analiza przedstawionych rozkładów odpowiedzi wskazuje, że większość nauczycieli uważa, że system egzaminów zewnętrznych na przestrzeni lat 2002-2009 nie uległ zmianie. Wyraźnie widać jednak większy odsetek odpowiedzi „trudno powiedzieć” w ocenach dotyczących jakości zadań egzaminacyjnych w porównaniu z ocenami organizacji egzaminu (wśród nauczycieli gimnazjów większy jest także odsetek braków odpowiedzi). W przypadku szkół podstawowych różnica ta wynosi około 10 punktów procentowych, a w przypadku gimnazjów ponad 20 punktów procentowych. Dalsza analiza bez wyjaśnienia tych różnic byłaby niewskazana.

Często w badaniach ankietowych odpowiedzi typu „trudno powiedzieć” nie niosą ze sobą ważnych informacji, a stanowią jedynie sposób na uniknięcie braków danych (w przypadku kiedy respondent nie chce bądź nie potrafi udzielić odpowiedzi na pytanie). Stosunkowo częsty wybór odpowiedzi „trudno powiedzieć” w omawianym pytaniu wiąże się jednak z logicznym postępowaniem badanych, a nie z chęcią ucieczki od odpowiedzi. Należy zdać sobie sprawę z tego, że nie wszyscy nauczyciele w jednakowym stopniu uczestniczą w procesie egzaminów zewnętrznych. Stopień zaangażowania w funkcjonowanie systemu egzaminacyjnego jest w dużej mierze zależny od nauczanego przedmiotu. Część przedmiotów nauczanych w szkole nie jest reprezentowana na sprawdzianie po szkole podstawowej czy egzaminie gimnazjalnym, a co za tym idzie nauczyciele tych przedmiotów mogą nie czuć się na tyle kompetentni, aby odpowiadać na pytania dotyczące systemu egzaminów zewnętrznych.

O ile na temat samej organizacji egzaminu może się wypowiedzieć większa grupa nauczycieli, o tyle na temat jakości zadań mogą (a właściwie powinni) wypowiadać się ci, którzy są w stanie ową jakość ocenić.

Powyższe rozumowanie znajduje swoje odzwierciedlenie w analizowanych danych. Dzięki informacjom o nauczanych przedmiocie możliwe było pogrupowanie badanych według rodzaju nauczanego przedmiotu. Nauczyciele gimnazjów zostali podzieleni na grupy przedmiotów odpowiadające częściom egzaminu gimnazjalnego: 1) matematyczno-przyrodnicze, 2) humanistyczne i 3) języki obce. Do grupy przedmiotów matematyczno-przyrodniczych zaliczono: biologię, chemię, fizykę i astronomię, geografę, informatykę i matematykę. Do grupy przedmiotów humanistycznych: historię, język polski i wiedzę o społeczeństwie. W kwestionariuszu oprócz wymienionych wyżej przedmiotów znalazły się także: religia, etyka, sztuka, technika, wychowanie fizyczne oraz kategoria „inne”. Badani mogli zaznaczyć wiele nauczanych przedmiotów. Przyjęto, że nauczanie jakiegokolwiek przedmiotu z danej kategorii pozwala na przydzielenie takiego nauczyciela do danej grupy, jednakże pod warunkiem, że nie naucza przedmiotu z dwóch pozostałych grup. Przykładowo, nauczyciel chemii i historii nie mógł być jednoznacznie przydzielony do którejś z grup, ale nauczyciel chemii i wychowania fizycznego już tak. Taka procedura spowodowała, że do jednej z trzech grup nauczanych przedmiotów nie można było przypisać 10% nauczycieli, przy czym 5% z powodu tego, że nie zaznaczyli oni żadnego przedmiotu w kwestionariuszu, 2% dlatego, że nie uczą oni żadnego z przedmiotów z trzech grup, a 3% z powodu nauczania przedmiotów z różnych grup. Zdecydowano o wyłączeniu z dalszej analizy nauczycieli nieprzypisanych do którejś z kategorii ze względu na małą liczebność ewentualnej czwartej grupy oraz na jej niejednorodność.

Podobną procedurę zastosowano dla nauczycieli ze szkół podstawowych, jednak w tym przypadku wyodrębniono następujące grupy: 1) edukacja wczesnoszkolna, 2) język polski, 3) matematyka, przyroda lub informatyka. Nie było nauczycieli, którzy nauczaliby tylko języków obcych, więc ta kategoria pozostała pusta. Oprócz tych przedmiotów w kwestionariuszu znalazły się także: religia, etyka, sztuka, wychowanie fizyczne oraz kategoria „inne”. Z dalszej analizy wykluczono 6% nauczycieli szkół podstawowych, przy czym prawie wszyscy z nich uczą przedmiotów z różnych grup (tylko jedna osoba nie uczy żadnego przedmiotu z wydzielonych grup).

W tabeli 3. prezentowany jest rozkład odpowiedzi nauczycieli gimnazjów na omawiane pytanie, z uwzględnieniem ich podziału na grupy nauczanych przedmiotów. W tabeli podano również liczebność poszczególnych grup. Wartości „ogółem” różnią się od tych z tabeli 1. z powodu wykluczenia z analizy części badanych. Współczynnik χ^2 wskazuje, że obserwowane różnice są istotne statystycznie przy poziomie istotności $p < 0,05$.

Tabela 3. Rozkład odpowiedzi dla pytania A.4. z ankiety dla nauczycieli gimnazjów w podziale na grupy nauczanych przedmiotów

Czy w perspektywie lat 2002-2009 egzamin gimnazjalny jest ze względu na:		grupa przedmiotów			
		matematyczno- przyrodnicze	humanistyczne	języki obce	ogółem
organizację egzaminu przez system egzaminacyjny	coraz lepszy	37%	32%	29%	34%
	coraz gorszy	2%	3%	2%	2%
	mniej więcej taki sam	53%	57%	50%	54%
	trudno powiedzieć	7%	9%	19%	11%
	liczebność	887	597	497	1981
jakość zadań w części humanistycznej	coraz lepszy	9%	15%	7%	10%
	coraz gorszy	3%	8%	3%	4%
	mniej więcej taki sam	38%	66%	40%	47%
	trudno powiedzieć	51%	11%	51%	38%
	liczebność	790	601	489	1880
jakość zadań w części matematyczno- -przyrodniczej	coraz lepszy	20%	5%	6%	12%
	coraz gorszy	7%	4%	4%	5%
	mniej więcej taki sam	65%	34%	33%	48%
	trudno powiedzieć	8%	58%	57%	34%
	liczebność	891	525	483	1899

Po podzieleniu nauczycieli gimnazjów na grupy nauczanych przedmiotów odpowiadające częściom egzaminu gimnazjalnego wyraźnie widać różnice w ocenach jakości zadań pomiędzy poszczególnymi grupami. W przypadku oceny organizacji egzaminu różnice te nie są aż tak duże, lecz również występują.

Nauczyciele należący do grupy języków obcych dwa razy częściej niż pozostałe grupy wybierali odpowiedź „trudno powiedzieć” przy ocenie organizacji egzaminu. Najprawdopodobniej ma to związek z tym, że mieli oni za sobą dopiero pierwszy egzamin gimnazjalny z języka obcego i przygotowywali się do drugiej edycji (badania przeprowadzono w pierwszym kwartale 2010 roku). Profile procentowe poszczególnych grup są jednak podobne – większość nauczycieli (połowa lub więcej) niezależnie od nauczanego przedmiotu oceniła, że organizacja egzaminu przez system egzaminacyjny na przestrzeni lat 2002-2009 jest mniej więcej taka sama. Około jedna trzecia uczących twierdzi, że organizacja egzaminu jest coraz lepsza, a zaledwie 2-3%, że jest coraz gorsza.

W przypadku oceny jakości zadań w części humanistycznej egzaminu gimnazjalnego odpowiedzi nauczycieli z grupy przedmiotów humanistycznych bardzo wyraźnie odróżniają się od dwóch pozostałych grup. Różnice między grupą matematyczno-przyrodniczą i języków obcych prawie nie istnieją.

Aż połowa nauczycieli z grup, którzy nie przygotowują uczniów do tej części egzaminu, wybrała odpowiedź „trudno powiedzieć”. Biorąc pod uwagę pozostałe odpowiedzi, widać, że nadal większość badanych uważa, że egzamin jest mniej więcej taki sam, nieco mniej, że jest lepszy, a jeszcze mniej, że jest gorszy. Nauczycieli dostrzegających poprawę jest dwa do trzech razy więcej niż dostrzegających pogorszenie.

Przypatrując się rozkładowi odpowiedzi dotyczących jakości zadań w części matematyczno-przyrodniczej egzaminu, można zauważyć, że w tym przypadku to grupa nauczycieli przedmiotów matematyczno-przyrodniczych częściej wybierała odpowiedzi inne niż „trudno powiedzieć”, co potwierdza słuszność takiego analizowania danych. Znowu najczęściej wybierana była odpowiedź, że jakość zadań jest mniej więcej taka sama, rzadziej, że jest lepsza, a najrzadziej, że gorsza – nie ma tutaj niespodzianek. Nauczyciele przedmiotów matematyczno-przyrodniczych oceniają jednak częściej, że jakość zadań jest coraz lepsza, niż miało to miejsce w przypadku części humanistycznej i nauczycieli przedmiotów humanistycznych – ocen, że jakość zadań jest coraz lepsza, jest trzykrotnie więcej niż ocen, że jest coraz gorsza. Dwie pozostałe grupy w podobnym stopniu oceniają zadania jako coraz lepsze i coraz gorsze, co również odróżnia tę część egzaminu od części humanistycznej.

Rozkład odpowiedzi nauczycieli szkół podstawowych w podziale na grupy nauczanych przedmiotów (tabela 4.) prowadzi do podobnych wniosków jak w przypadku nauczycieli gimnazjów. Nauczyciele edukacji wczesnoszkolnej rzadziej formułują opinię niż nauczyciele języka polskiego czy matematyki, przyrody lub informatyki. Ci ostatni też jednak dość często unikają wyraźnych opinii w odniesieniu do jakości zadań na sprawdzianie. Ponownie współczynnik χ^2 wskazuje, że obserwowane różnice są istotne statystycznie przy poziomie istotności $p < 0,05$.

Tabela 4. Rozkład odpowiedzi dla pytania A.4. z ankiety dla nauczycieli szkół podstawowych w podziale na nauczane przedmioty

Czy w perspektywie lat 2002-2009 sprawdzian po szkole podstawowej jest ze względu na:		nauczany przedmiot			
		edukacja wczesnoszkolna	język polski	matematyka, przyroda lub informatyka	ogółem
organizację egzaminu przez system egzaminacyjny	coraz lepszy	30%	25%	35%	30%
	coraz gorszy	3%	3%	4%	3%
	mniej więcej taki sam	42%	60%	51%	48%
	trudno powiedzieć	26%	13%	11%	18%
	liczebność	993	398	616	2007
jakość zadań egzaminacyjnych	coraz lepszy	12%	10%	13%	12%
	coraz gorszy	10%	13%	12%	11%
	mniej więcej taki sam	37%	60%	59%	48%
	trudno powiedzieć	41%	18%	16%	29%
	liczebność	992	398	616	2006

Analizując oceny organizacji egzaminów, warto zwrócić uwagę na fakt, że nauczyciele matematyki, przyrody i informatyki częściej niż nauczyciele języka polskiego twierdzą, że owa organizacja jest coraz lepsza. Nauczyciele języka polskiego częściej natomiast uważają, że jakość organizacji sprawdzianu przez system egzaminacyjny jest mniej więcej taka sama.

Ocena jakości zadań na sprawdzianie wypada nieco inaczej niż analogiczna ocena dla egzaminu gimnazjalnego. Nauczyciele edukacji wczesnoszkolnej częściej wybierali odpowiedź „trudno powiedzieć” niż pozostałe grupy, co nie powinno dziwić. Ważne jest jednak to, że odsetek odpowiedzi „coraz lepiej” i „coraz gorzej” jest zbliżony we wszystkich grupach, co nie miało miejsca w przypadku pytania o egzamin gimnazjalny. Pomimo braku podziału na typ zadań można przypuszczać, że nauczyciele odnoszą się do zadań sprawdzających umiejętności, których uczą. Oznaczałoby to większy krytycyzm nauczycieli szkół podstawowych w stosunku do jakości zadań na sprawdzianie niż nauczycieli gimnazjów w stosunku do zadań na egzaminie gimnazjalnym.

Z analizy ocen egzaminów przez nauczycieli wyłania się raczej obraz pozytywny. Ważnym elementem jest to, że opinie formułują osoby stykające się z danym zagadnieniem – czy jest to organizacja egzaminu, czy też jakość zadań. Choć pytanie było sformułowane dość ogólnie i nie zawierało warunku nauczania przedmiotu reprezentowanego na egzaminie, to spora grupa nauczycieli postanowiła nie wypowiadać się w kwestiach, z którymi nie mają prawdopodobnie bezpośredniej styczności (mowa tu głównie o jakości zadań). Dzięki temu można przypuszczać, że przedstawione opinie są przemyślane i poparte doświadczeniem. Z drugiej strony niewielkie zainteresowanie zagadnieniami związanymi z systemem egzaminacyjnym nauczycieli, którzy nie przygotowują uczniów bezpośrednio do egzaminów, może nieco smucić. Pozostaje mieć nadzieję, że takie podejście nie przekłada się na brak poczucia odpowiedzialności za wyniki egzaminacyjne uczniów, które zależą od pracy wszystkich nauczycieli w szkole.

Niepokojąca jest ocena jakości zadań na sprawdzianie i zbliżony procent oceniających ją coraz lepiej, jak i coraz gorzej. Jedna dziesiąta nauczycieli twierdząca, że jakość zadań jest coraz gorsza, może nie stanowić bardzo dużej grupy, ale w porównaniu z analogicznym pytaniem dla egzaminu gimnazjalnego, gdzie dwa lub trzy razy więcej nauczycieli uważa, że jakość zadań jest coraz lepsza niż coraz gorsza, powinna stanowić sygnał, że istnieje możliwość poprawy.

Postawy wobec systemu egzaminów zewnętrznych

W kwestionariuszach dla nauczycieli biorących udział w badaniu znalazły się baterie dwudziestu czterech stwierdzeń dotyczących egzaminów. Każde ze stwierdzeń badani mogli ocenić na pięciostopniowej skali odpowiedzi: 1) zdecydowanie się zgadzam, 2) raczej się zgadzam, 3) raczej się nie zgadzam, 4) zdecydowanie się nie zgadzam i 5) trudno powiedzieć. Nie jest to może typowy układ, w którym na środku powinna znaleźć się kategoria neutralna, ale dzięki takiemu rozwiązaniu niejako wymuszano odpowiedzi inne niż neutralne. Oczywiście można postawić zarzut, że odpowiedź „trudno powiedzieć” różni się w swoim wydźwięku np. od odpowiedzi „ani się zgadzam, ani się nie

zgadzam”, przyjmując, że odpowiedź „trudno powiedzieć” wyraża brak opinii w danej kwestii, a odpowiedź neutralna umieszczona na środku skali pytań posiadanie ambiwalentnego stosunku do ocenianego elementu. Rozważania dotyczące konstrukcji skali odpowiedzi zostaną jednak tutaj pominięte, gdyż mogłyby z powodzeniem stanowić materiał na osobne opracowanie.

Dla celów przeprowadzonej analizy zmieniono kolejność odpowiedzi na pytania, umieszczając odpowiedź „trudno powiedzieć” na środku skali i przesuwając w prawo odpowiedzi negatywne. Otrzymało w ten sposób następującą skalę odpowiedzi: 1) zdecydowanie się zgadzam, 2) raczej się zgadzam, 3) trudno powiedzieć, 4) raczej się nie zgadzam i 5) zdecydowanie się nie zgadzam.

Celem umieszczenia stwierdzeń w ankietach było między innymi utworzenie z nich skal postaw wobec systemu egzaminów zewnętrznych. Oczywiście można je także analizować osobno, nie budując skal czy indeksów. Taką analizę danych z opisywanych badań i danych zebranych wśród egzaminatorów OKE w Krakowie przeprowadziły Anna Rappe i Joanna Peter⁴. Poniżej zostaną zaprezentowane wyniki analizy omawianych stwierdzeń za pomocą metody głównych składowych.

W badaniu pilotażowym w kwestionariuszach umieszczono więcej stwierdzeń, lecz po analizie zebranych danych zdecydowano o ograniczeniu ich liczby do dwunastu pozytywnych i dwunastu negatywnych. W założeniu miały one tworzyć dwie skale – zalet egzaminów (stwierdzenia pozytywne) i zagrożeń egzaminacyjnych (stwierdzenia negatywne). Listy stwierdzeń (wraz z ładunkami na wyodrębnionych w przeprowadzonej na danych z badania zasadniczego analizie składowych głównych) znajdują się w aneksie. Stwierdzenia z kwestionariuszy dla nauczycieli gimnazjów znajdują się w tabeli 5., a dla nauczycieli szkół podstawowych w tabeli 6. Są to stwierdzenia analogiczne – dostosowano je tylko do egzaminu, którym kończy się dana szkoła. Tabele zawierają tylko dwadzieścia trzy stwierdzenia. Brak w nich pozycji „Wynik egzaminu/sprawdzianu jest w dużej mierze przypadkowy”, gdyż została ona wyłączona z analizy, o czym dalej.

Analizę głównych składowych rozpoczęto od sprawdzenia rozwiązania z określeniem liczby składowych na podstawie kryterium Kaisera wartości własnej składowej większej od 1. Przy takim podejściu dla nauczycieli gimnazjów wyodrębniono pięć głównych składowych, a dla nauczycieli szkół podstawowych 4. Analiza wykresów osypiska sugerowała rozwiązanie z 2-4 składowych dla gimnazjów i 3 dla szkół podstawowych. Niestety tak wyodrębnione główne składowe (niezależnie od rotacji) nie posiadają łatwej interpretacji i trudno odnaleźć powody, dla których stwierdzenia ładują poszczególne składowe.

Podjęto więc decyzję o narzuceniu rozwiązania z dwoma głównymi składowymi, aby umożliwić interpretację tak otrzymanych wyników. W ten sposób zostało utrzymane założenie przyjęte podczas doboru stwierdzeń do kwestionariusza o utworzeniu z nich dwóch skal. Ostatecznie z analizy usunięto stwierdzenie „Wynik egzaminu/sprawdzianu jest w dużej mierze przypadkowy”, gdyż łądowało ono obydwie składowe na zbliżonym poziomie (choć z jedną z nich związek był dodatni, a z drugą ujemny) i użyto rotacji Varimax.

⁴ A. Rappe, J. Peter, *Opinie nauczycieli i egzaminatorów o systemie egzaminów zewnętrznych*, [w:] *Teraźniejszość i przyszłość oceniania szkolnego*, red. B. Niemierko, M.K. Szmigel, Kraków, 2010

Otrzymane w ten sposób rozwiązanie wyjaśnia 31% wariancji dla nauczycieli gimnazjów i 32% dla nauczycieli szkół podstawowych. Nie jest to wielkość duża, ale zadowalająca, zwłaszcza, jeśli weźmie się pod uwagę łatwość interpretacji. Rozwiązania wstępne, o których była mowa wyżej, wyjaśniały około 10% wariancji więcej i nie warto dla takiej wartości rezygnować z klarowności. Zarówno dla danych z gimnazjów, jak i ze szkół podstawowych współczynnik KMO dopasowania modelu do danych wynosi 0,88, co jest godne pochwały. Do przeprowadzenia analizy użyto 2054 obserwacji w bazie z gimnazjów i 2128 obserwacji w bazie dla szkół podstawowych.

Na podstawie opisanego rozwiązania można utworzyć dwie skale: zalet egzaminów i zagrożeń egzaminacyjnych. Pierwotnie skala zalet miała być tworzona przez stwierdzenia pozytywne, a skala zagrożeń przez negatywne. Przeprowadzona analiza skutkuje jednak nieco innym doбором stwierdzeń. Jedno z negatywnych stwierdzeń zostało usunięte z analizy, natomiast inne („Ocena pracy szkoły i nauczycieli na podstawie wyników egzaminu/sprawdzianu jest niesprawiedliwa.”) ma niski ładunek na skali zagrożeń, a wyższy, choć ujemny, na skali zalet. W związku z tym skala zagrożeń składa się z dziesięciu negatywnych stwierdzeń, a skala zalet z dwunastu pozytywnych i jednego negatywnego (związanego ze skalą ujemnie).

Dla tak dobranych zmiennych przeprowadzono analizę rzetelności. Skala zalet składająca się z trzynastu elementów (odwrócono odpowiedzi dla negatywnego stwierdzenia) dla danych z gimnazjów posiada α Cronbacha równe 0,821, a dla danych ze szkół podstawowych α równe 0,811. Usunięcie jakiegokolwiek z pozycji nie spowodowałoby wzrostu rzetelności tak skonstruowanej skali. W przypadku skali zagrożeń składającej się z dziesięciu pozycji rzetelność jest nieco niższa. Dla danych z gimnazjów α wynosi 0,703, a dla danych ze szkół podstawowych α jest równe 0,752. Podobnie usunięcie jakiegokolwiek pozycji nie powoduje wzrostu rzetelności.

Dalszym krokiem w analizie było obliczenie ocen (wartości) na obydwu składowych dla wszystkich badanych⁵. Dzięki temu możliwe stało się porównanie pozycji różnych grup nauczycieli na opisanych skalach. Do porównań wybrano podział ze względu na lokalizację szkoły, nauczany przedmiot (a właściwe przynależność do grup opisanych wyżej) oraz fakt bycia egzaminatorem lub nie. Ze względu na wielkość miejscowości, w której znajduje się szkoła wyodrębniono cztery grupy: 1) wieś, 2) miasto do 20 tys. mieszkańców, 3) miasto od 20 do 100 tys. mieszkańców i 4) miasto powyżej 100 tys. mieszkańców. Dla każdej ze zmiennych przeprowadzono jednoczynnikową analizę wariancji z ocenami na skalach jako zmiennymi zależnymi.

W przypadku faktu bycia egzaminatorem surowe dane umożliwiały podział na egzaminatorów sprawdzianu, poszczególnych części egzaminu gimnazjalnego oraz innego egzaminu. Ze względu na niewielkie liczebności grup postanowiono jednak wziąć pod uwagę sam fakt bycia egzaminatorem, niezależnie od rodzaju egzaminu.

⁵ Użyto w tym celu metody regresji. Oceny są standaryzowane, zatem dla wszystkich badanych średnia wynosi 0, a odchylenie standardowe 1.

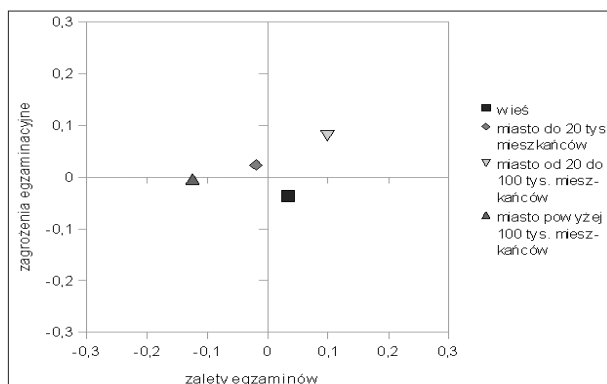
Wśród nauczycieli gimnazjów bycie lub niebycie egzaminatorem nie powoduje zróżnicowania na skali zagrożeń, natomiast różnice na skali zalet są istotne statystycznie przy poziomie istotności $p < 0,05$. Egzaminatorzy lokują się na tej skali poniżej średniej, a nie-egzaminatorzy powyżej. Różnica ta jednak jest znikoma – współczynnik η wynosi zaledwie 0,07. Liczebność grup wynosiła odpowiednio 1302 dla nie-egzaminatorów i 752 dla egzaminatorów.

Nauczyciele szkół podstawowych nie wykazują z kolei zróżnicowania na skali zalet, a na skali zagrożeń egzaminatorzy lokują się powyżej średniej, a nie-egzaminatorzy poniżej średniej. Różnice są istotne statystycznie przy poziomie istotności $p < 0,05$, jednak również są one minimalne – η wynosi 0,08.

Można pokusić się o stwierdzenie, że egzaminatorzy widzą mniej zalet oraz więcej wad egzaminów zewnętrznych niż nie-egzaminatorzy (choć zależy to od rodzaju egzaminu). Nie powinno dziwić to, że większe zaangażowanie w działanie systemu egzaminacyjnego, czyli bycie egzaminatorem, powoduje inne jego postrzeganie. Niepokojące jest jednak to, że osoby, które lepiej poznały funkcjonowanie tego systemu, mają o nim gorsze zdanie niż pozostali.

Przy podziale nauczycieli gimnazjów ze względu na wielkość miejscowości, w której mieści się szkoła, na skali zagrożeń nie występuje zróżnicowanie (podobnie jak było to w przypadku faktu bycia egzaminatorem). Na skali zalet występują różnice istotne statystycznie przy poziomie istotności $p < 0,05$. W związku z tym, że mamy tutaj do czynienia z podziałem na cztery kategorie przydatne jest przeprowadzenie testów post hoc mających na celu wskazanie, pomiędzy którymi grupami owo zróżnicowanie występuje. Test Schaffé'a wskazuje, że istotne statystycznie różnice występują tylko pomiędzy grupą nauczycieli ze szkół w miastach od 20 do 100 tys. mieszkańców i grupą ze szkół z miast powyżej 100 tys. mieszkańców.

Wykres średnich ocen w poszczególnych grupach przedstawia rysunek 1. Widać na nim wyraźnie, że nauczyciele z największych miast znajdują się najniżej na skali zalet egzaminów, a nauczyciele ze średnich miast najwyżej. Wielkość tych różnic znow jest jednak minimalna – η wynosi 0,07.

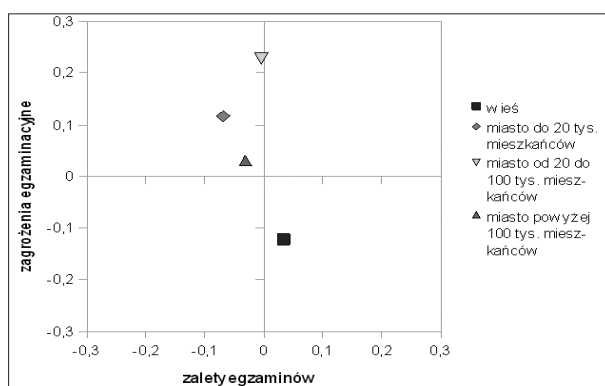


Rys. 1. Średnie oceny na skalach dla nauczycieli gimnazjów w podziale na grupy według wielkości miejscowości

Podział nauczycieli szkół podstawowych ze względu na wielkość miejscowości, w której znajduje się szkoła, nie powoduje różnic na skali zalet (tak jak w przypadku podziału na egzaminatorów i nie-egzaminatorów). Na skali zagrożeń występują jednak istotne statystycznie różnice (przy $p < 0,05$). Test Scheffe'a wskazuje, że różnice te, podobnie jak w przypadku nauczycieli gimnazjów, występują pomiędzy nauczycielami z dużych i średnich miasto. Ponadto w tym przypadku nauczyciele ze wsi różnią się od pozostałych grup.

Na rysunku 2. widać, że nauczyciele ze szkół wiejskich na skali zagrożeń znajdują się poniżej średniej, podczas gdy pozostali nauczyciele lokują się powyżej średniej. Nauczyciele z dużych miast wskazują najwięcej zagrożeń. W tym wypadku różnice te są nieco większe niż wcześniej, choć nadal pozostają na bardzo niskim poziomie – η wynosi 0,13.

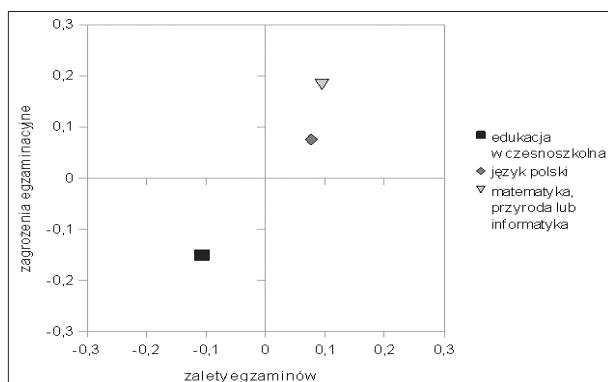
Zaobserwowane zależności są interesujące, ale brak łatwego sposobu ich interpretacji. Nie jest na przykład tak, że wraz ze wzrostem wielkości miejscowości pozycja na obydwu skalach staje się wyższa czy też, że na skali zalet egzaminów staje się niższa, a na skali zagrożeń egzaminacyjnych wyższa. Taką liniową zależność „psują” nauczyciele ze szkół wiejskich. Należałoby zastanowić się nad specyfiką szkół z poszczególnych kategorii miejscowości i wtedy dopiero próbować wyjaśnić zaobserwowane różnice.



Rys. 2. Średnie oceny na skalach dla nauczycieli szkół podstawowych w podziale na grupy według wielkości miejscowości

Nauczyciele gimnazjum podzieleni według grup nauczanych przedmiotów opisanych w pierwszej części tego opracowania nie wykazują zróżnicowania na żadnej ze skal. Analogiczny podział nauczycieli szkół podstawowych uwidacznia istotne statystycznie różnice na obydwu skalach (przy $p < 0,05$). Test Schaffe'a wskazuje, że różnice występują pomiędzy nauczycielami edukacji wczesnoszkolnej i pozostałymi grupami.

Na rysunku 3. można zobaczyć, że nauczyciele edukacji wczesnoszkolnej znajdują się poniżej średniej na obydwu skalach, podczas gdy pozostałe grupy nauczycieli lokują się powyżej średniej na obydwu skalach. Różnice te są znikome, dla skali zalet η wynosi 0,1, a dla skali zagrożeń 0,15.



Rys. 3. Średnie oceny na skalach dla nauczycieli szkół podstawowych w podziale na grupy według nauczanego przedmiotu

Podsumowanie

Ogólny obraz systemu egzaminów zewnętrznych jest raczej dobry – więcej nauczycieli wskazuje, że na przestrzeni lat ulega on poprawie niż pogorszeniu. Jedynie w przypadku jakości zadań na sprawdzianie taka sama liczba nauczycieli ocenia, że są one coraz lepsze i coraz gorsze.

Wyraźnie ocena systemu egzaminacyjnego zależy od stopnia zaangażowania w jego funkcjonowanie. Nauczyciele nieprzygotowujący bezpośrednio uczniów do egzaminów rzadziej wyrażają o nich opinię. Na przykładzie analizowanych skal widać, że również wielkość miejscowości, w której znajduje się szkoła, jest powiązana z postrzeganiem systemu egzaminacyjnego.

Z przeprowadzonych analiz nie wynika, aby nauczyciele byli wyraźnie podzieleni, jeśli chodzi o ocenę systemu egzaminów zewnętrznych. Istnieją oczywiście pewne różnice, ale nie są one bardzo duże. To, iż większość nauczycieli uważa, że system egzaminacyjny początku jego istnienia jest mniej więcej taki sam, świadczyć może o jego stabilności. Nie oznacza to oczywiście, że jest on pozbawiony wad.

Zasadne wydaje się przeprowadzenie pogłębionych analiz jakościowych związanych z oceną zagadnień związanych z egzaminami zewnętrznymi. Dałoby to szansę na zidentyfikowanie czynników mających istotny wpływ na wyrażane o nich opinie. Wstępem do takich pogłębionych badań jakościowych jest na pewno opracowanie Marka Piotrowskiego⁶, w którym wykorzystano dane z tego samego badania zespołu EWD.

⁶ M. Piotrowski, *Ewaluacyjna funkcja egzaminów zewnętrznych w opiniach nauczycieli szkół podstawowych i gimnazjów*, [w:] *Teraźniejszość i przyszłość oceniania szkolnego*, red. B. Niemierko, M.K. Szmigel, Kraków, 2010

Aneks

Tabela 5. Ładunki składowych głównych dla poszczególnych stwierdzeń z ankiety dla nauczycieli gimnazjów (dwie składowe, rotacja Varimax, stwierdzenia posortowane według wartości ładunków, pominięto ładunki mniejsze od |0,3|)

Stwierdzenia z ankiety dla nauczycieli gimnazjum	składowa 1	składowa 2
Wyniki egzaminu pozwalają na ocenę skuteczności nauczania w szkole.	0,70	
Wyniki egzaminów pozwalają organom prowadzącym racjonalnie oceniać szkoły.	0,62	
Zadania na egzaminie dobrze sprawdzają to, co w edukacji jest najważniejsze.	0,59	
Wynik egzaminu jest dobrym kryterium rekrutacji uczniów do szkoły ponadgimnazjalnej.	0,59	
Informacje o wynikach szkoły na egzaminie pozwalają lepiej planować działania dydaktyczne.	0,59	
Egzamin motywuje uczniów do systematycznej nauki.	0,59	
Egzamin zwiększa zaangażowanie nauczycieli w pracę.	0,56	
Egzamin mobilizuje szkoły do ciągłego wysiłku i doskonalenia metod nauczania.	0,55	
Egzamin przygotowuje uczniów do radzenia sobie w trudnych sytuacjach.	0,55	
Jednolity w całym kraju egzamin pozwala porównywać osiągnięcia szkolne uczniów.	0,52	
Egzamin zwiększa społeczne zainteresowanie oświatą.	0,47	
Ocena pracy szkoły i nauczycieli na podstawie wyników egzaminu jest niesprawiedliwa.	-0,44	
Wynik egzaminu zależy przede wszystkim od pracy nauczyciela.	0,44	
Szkoły zamieniają ostatnie miesiące nauki w ciągłe przerabianie zadań testowych i niekończący się egzamin próbny.		0,63
Nauczyciele pomijają ważne treści programowe nieobecne na egzaminie.		0,58
Szkoły pod presją egzaminów coraz mniej chętnie przyjmują słabszych uczniów.		0,57
Egzamin zamiast skłaniać szkoły do lepszej pracy, napędza rynek korepetycji i różnego typu kursów.		0,57
Egzamin osłabia autorytet nauczyciela.		0,51
Egzamin zmniejsza szanse uczniów wywodzących się z biednych rodzin na dostanie się do dobrej szkoły ponadgimnazjalnej.		0,50
Nauczyciele pracujący ze słabszymi uczniami odczuwają frustrację, wynikającą z niskich wyników swoich uczniów na egzaminie.		0,48
Szkoły skupiające najbardziej uzdolnionych uczniów często spoczywają na laurach.		0,47
Egzamin osłabia bieżącą motywację uczniów do nauki, skłaniając do uczenia się na ostatnią chwilę.	-0,34	0,45
Wynik egzaminu w zbyt dużym stopniu decyduje o dalszych losach szkolnych ucznia.		0,31

Tabela 6. Ładunki składowych głównych dla poszczególnych stwierdzeń z ankiety dla nauczycieli szkół podstawowych (dwie składowe, rotacja Varimax, stwierdzenia posortowane według wartości ładunków, pominięto ładunki mniejsze od |0,3|)

Stwierdzenia z ankiety dla nauczycieli szkoły podstawowej	składowa 1	składowa 2
Wyniki sprawdzianu pozwalają na ocenę skuteczności nauczania w szkole.	0,66	
Informacje o wynikach szkoły na egzaminie pozwalają lepiej planować działania dydaktyczne.	0,64	
Sprawdzian motywuje uczniów do systematycznej nauki.	0,61	
Sprawdzian zwiększa zaangażowanie nauczycieli w pracę.	0,61	
Sprawdzian mobilizuje szkoły do ciągłego wysiłku i doskonalenia metod nauczania.	0,59	
Sprawdzian przygotowuje uczniów do radzenia sobie w trudnych sytuacjach.	0,58	
Zadania na sprawdzianie dobrze sprawdzają to, co w edukacji jest najważniejsze.	0,57	
Wynik sprawdzianu jest dobrym kryterium rekrutacji uczniów do szkoły gimnazjalnej.	0,56	
Wyniki sprawdzianów pozwalają organom prowadzącym racjonalnie oceniać szkoły.	0,55	
Jednolity w całym kraju sprawdzian pozwala porównywać osiągnięcia szkolne uczniów.	0,48	
Sprawdzian zwiększa społeczne zainteresowanie oświatą.	0,47	
Ocena pracy szkoły i nauczycieli na podstawie wyników sprawdzianu jest niesprawiedliwa.	-0,37	
Wynik sprawdzianu zależy przede wszystkim od pracy nauczyciela.	0,36	
Nauczyciele pomijają ważne treści programowe nieobecne na sprawdzianie.		0,65
Szkoły pod presją sprawdzianów coraz mniej chętnie przyjmują słabszych uczniów.		0,63
Sprawdzian, zamiast skłaniać szkoły do lepszej pracy, napędza rynek korepetycji i różnego typu kursów.		0,62
Sprawdzian zmniejsza szanse uczniów wywodzących się z biednych rodzin na dostanie się do dobrej szkoły gimnazjalnej.		0,62
Szkoły zamieniają ostatnie miesiące nauki w ciągłe przerabianie zadań testowych i niekończący się próbny sprawdzian.		0,62
Szkoły skupiające najbardziej uzdolnionych uczniów często spoczywają na laurach.		0,53
Sprawdzian osłabia autorytet nauczyciela.		0,47
Sprawdzian osłabia bieżącą motywację uczniów do nauki, skłaniając do uczenia się na ostatnią chwilę.		0,47
Nauczyciele pracujący ze słabszymi uczniami odczuwają frustrację, wynikającą z niskich wyników swoich uczniów na sprawdzianie.		0,43
Wynik sprawdzianu w zbyt dużym stopniu decyduje o dalszych losach szkolnych ucznia.		0,40

Bibliografia:

1. Górniak J., *Analiza czynnikowa i analiza głównych składowych*, „ASK. Społeczeństwo, Badania, Metody” 7/1998, IFiS PAN.
2. Lisiecka Z. *Podłużne badania panelowe w szkołach podstawowych i gimnazjach a rozwój wskaźnika EWD*, [w:] *Teraźniejszość i przyszłość oceniania szkolnego*, red. B. Niemierko, M.K. Szmigel, Kraków, 2010
3. Nawojczyk M. *Przewodnik po statystyce dla socjologów*, SPSS Polska, Kraków 2002.
4. Piotrowski M. *Ewaluacyjna funkcja egzaminów zewnętrznych w opiniach nauczycieli szkół podstawowych i gimnazjów*, [w:] *Teraźniejszość i przyszłość oceniania szkolnego*, red. B. Niemierko, M.K. Szmigel, Kraków, 2010
5. Rappe A., Peter J. *Opinie nauczycieli i egzaminatorów o systemie egzaminów zewnętrznych*, [w:] *Teraźniejszość i przyszłość oceniania szkolnego*, red. B. Niemierko, M.K. Szmigel, Kraków, 2010