

Warszawa, 3 listopada 2015 r.

Wpływ egzaminów na jakość edukacji

Dlaczego egzaminy są tak ważne? Jakie są wady i zalety systemu egzaminacyjnego? W jaki sposób mogą pomóc w doskonaleniu pracy szkoły? Czy i jak z wyników egzaminacyjnych korzystają dyrektorzy szkół i nauczyciele? – o tym w raporcie o stanie edukacji 2014 Instytutu Badań Edukacyjnych „Egzaminy zewnętrzne w polityce i praktyce edukacyjnej”

Rola testów w edukacji jest jedną z najczęściej dyskutowanych kwestii. Dotyczy to zarówno Polski, jak i praktycznie wszystkich innych krajów, nawet tak „wolnego od testów” kraju jak Finlandia. Większość zarzutów dotyczących testów odnosi się bardziej do tego co robi się z wynikami testów, niż do samych testów i skutków, jakie wywołują, w zamierzony lub niezamierzony sposób, w oświacie. Chodzi tu nie tylko o osławione „uczenie pod testy”, ale też wpływ testów na postrzeganie celów edukacyjnych.

Wpływ egzaminów zależy przede wszystkim od funkcji, jakie wypełniają w danym kraju. Zamierzone i niezamierzone skutki testów są większe w tych systemach edukacyjnych, w których testy odgrywają większą rolę, ponieważ dyrektorzy i nauczyciele dążą do poprawy wyniku testów ze względu na pozytywne lub negatywne konsekwencje. W teorii polscy dyrektorzy i nauczyciele mogą wykorzystać wynik do poprawy jakości swojej pracy, nie obawiając się innych konsekwencji, ale nie będąc też specjalnie do tego zmotywowanymi. W praktyce, różnego rodzaju rankingi i zestawienia szkół powodują, że średnie wyniki egzaminu są traktowane jako ważna informacja o szkole. Różne jednak może być postrzeganie znaczenia wyniku egzaminacyjnego przez dyrektora jedynej szkoły w niedużej gminie, a inne dyrektora szkoły, który konkuruje o uczniów z innymi szkołami – jak to ma miejsce w wielu dużych miastach w Polsce.

Kluczowe pytanie dotyczy tego, czy wykorzystywanie testów i egzaminów wpływa na umiejętności uczniów. Badania pokazują, że uczniowie z krajów, w których istnieją egzaminy zewnętrzne uzyskują lepsze wyniki niż uczniowie w krajach, w których nie ma egzaminów. Czy oznacza to, że wprowadzenie egzaminów poprawia umiejętności uczniów? Niekoniecznie. Egzaminy mogą pomagać w poprawie umiejętności uczniów, ale tylko wtedy, gdy nacisk na egzaminy jest skojarzony z innymi elementami zmian w systemie edukacji. Egzaminy sprawdzają się zwłaszcza w tych krajach, w których oświata jest zdecentralizowana, a szkoły i nauczyciele mają dużą autonomię w doborze treści i sposobów nauczania.

Badania edukacyjne dają dobry obraz negatywnych i pozytywnych skutków egzaminów na poziomie uczniów, szkół i nauczycieli.

Zawężanie treści nauczania

Polega na zwiększaniu czasu, który przeznaczają na nauczanie (nauczyciele) i uczenie się (uczniowie) i większy wysiłek wkładany przez nauczycieli w nauczanie tych przedmiotów, które są uwzględniane w egzaminach. Efekt może prowadzić do zwiększenia efektywności nauczania tych przedmiotów szkolnych, które uznaje się za najważniejsze. Negatywną stroną jest ograniczanie zakresu nauczanych treści do tych, które będą uwzględnione w egzaminach. Wiele wskazuje na to, że zawężanie treści nauczania występuje także w Polsce czego wyrazistym przykładem może być mała



uwaga przykładana do kształtowania umiejętności mówienia na lekcjach języka angielskiego (w egzaminach sprawdzana dopiero na maturze).

Uczenie pod test

Z zawężaniem treści nauczania wiąże się inne zjawisko – nauczanie pod konkretny test, co polega np. na ćwiczeniu rozwiązywania zadań o podobnej treści i formacie co te występujące w egzaminie. Z tego rodzaju zjawiskiem mamy też do czynienia w Polsce. W naszym kraju arkusze egzaminacyjne wraz ze schematami oceniania, a często przykładami rozwiązań są dostępne powszechnie na stronie internetowej CKE. W rezultacie stanowią one dostępny materiał nie tylko dla uczniów, ale także dla nauczycieli. Takie repozytorium zadań wykorzystywane jest przez niektórych nauczycieli jako łatwy sposób „treningu” uczniów przed egzaminem. Podporządkowywanie nauczania konkretnemu sposobowi mierzenia umiejętności może powodować, że uczeń świetnie radzi sobie z zadaniami określonego rodzaju, np. z matematyki, ale uzyskałby dużo słabszy wynik w inaczej skonstruowanym teście mierzącym ten sam zakres umiejętności.

Koncentrowanie się na wybranych kategoriach uczniów

Skrajnym przykładem negatywnego wpływu na praktykę szkolną są próby oszukiwania systemu. Dyrektorzy i nauczyciele mogą próbować wpływać na to, którzy uczniowie przystępują do egzaminów. Szkoły mogą też unikać przyjmowania uczniów mających słabe wyniki lub mających inne cechy, które są skorelowane z wynikami (np. uczniów z gorszym pochodzeniem społecznym). Sporadycznie przypadki takie odnotowywane są w Polsce w przypadku egzaminu maturalnego.

Różnicowanie się szkół

Publikowanie i korzystanie z wyników testów może także wpływać na różnicowanie się szkół. Dobrzy i jednocześnie zamożni uczniowie idą do dobrych klas i szkół (a to, które klasy i szkoły są dobre, można m.in. poznać po wynikach egzaminów, które są upubliczniane), a słabi i mniej zamożni – odwrotnie. W Polsce analizy prowadzone przez prof. Romana Dolatę pokazały, że wprowadzenie systemu egzaminów zewnętrznych uruchomiło w dużych miastach proces selekcji i autoselekcji na progu gimnazjum. W konsekwencji wielkomiejskie gimnazja coraz bardziej różnią się od siebie pod względem wyników egzaminacyjnych. Procesy te można wiązać z lepszą dostępnością danych egzaminacyjnych, choć związek ten nie jest oczywisty. Zwiększanie dostępności informacji o jakości szkół poprawia sytuację mniej zamożnych, którzy zwykle mają słabszy dostęp do tego rodzaju informacji. Ale jednocześnie bardziej zamożni rodzice są bardziej skłonni korzystać z tego rodzaju informacji i są lepiej przygotowani do jej interpretacji.

Pozytywna rola egzaminów

Ale trzeba też pamiętać o pozytywnej roli egzaminów. Egzaminy także mogą wspierać pozytywne zmiany w procesie uczenia się, jeżeli:

- zarówno treść zadań zastosowanych w egzaminie, jak i ich format pozwalają sprawdzać szerokie spektrum umiejętności wymagających kompetencji rozwijanych w szkole;
- wyniki egzaminu wykorzystywane są tylko do tych celów, do których dany egzamin został zaplanowany;
- komunikowanie wyników jest zoptymalizowane w taki sposób, aby szkoła mogła je wykorzystać do doskonalenia nauczania;
- szkoły i nauczyciele mają możliwość uzyskania wsparcia, jak interpretować i wykorzystywać wyniki egzaminów – w tym zakresie jest szerokie pole do działania dla placówek doskonalenia nauczycieli;



- wyniki egzaminów są przedmiotem analiz efektów szkoły, ale ważne decyzje są podejmowane w kontekście szerokiej gamy różnorodnej działalności szkoły i osiągnięć jej uczniów.

Co nam mówią wyniki egzaminów

Jak jednak oceniać pracę szkoły? Na przykład, licea ogólnokształcące można oceniać na podstawie odsetka uczniów, którzy zdali egzamin maturalny, używając do tego średniego wyniku maturalnego lub wskaźników edukacyjnej wartości dodanej. Generalnie odchodzi się od prostego porównania wyników egzaminacyjnych, które faworyzuje szkoły działające w bogatszych gminach, gdzie uczą się dzieci dobrze wykształconych rodziców. Znaczenie ma także wielkość szkoły – w dużych szkołach wyniki uczniów mają w przybliżeniu rozkład normalny: najczęściej jest uczniów przeciętnych, mniej uczniów bardzo słabych i bardzo dobrych. W szkołach mających niewielu uczniów wyniki będą się bardziej różnić, co powoduje, że małe szkoły łatwiej mogą znaleźć się w grupie szkół osiągających najlepsze i najgorsze wyniki. Pożyteczne wykorzystanie wyników wymaga często sięgania po rezultaty, które są odpowiednio przetworzone, a ich interpretacja wymaga wiedzy o ograniczeniach przyjętej metodologii.

Wyników testu nie powinno się wykorzystywać jako jedynej informacji w decyzjach mających konsekwencje dla osób i instytucji. Największy pożytek daje łączenie informacji pochodzących z testów z innymi informacjami. Wiele badań wskazuje także na to, że cennym źródłem informacji, nadspodziewanie dobrze przewidującym dalsze sukcesy ucznia, są oceny szkolne. Ważne są z punktu widzenia oceny efektywności nauczania, mają także dane dotyczące losów absolwentów.

Porównywalne wyniki egzaminów

Dużą słabością polskich egzaminów jest brak możliwości porównywania wyników w czasie. Egzaminy z kolejnych lat różnią się trudnością i nie możemy przez to powiedzieć, czy wyniki kolejnych roczników są lepsze lub gorsze od poprzednich. Możemy jedynie porównywać między sobą uczniów i szkoły oraz wyliczać wskaźniki edukacyjnej wartości dodanej i w ten sposób wnioskować o poprawie lub pogorszeniu jakości edukacji. Słabość ta utrudnia wykorzystywanie danych egzaminacyjnych przez szkoły, organy prowadzące czy Ministerstwo Edukacji do doskonalenia pracy szkoły czy oceny zmian w jakości edukacji. Problem ten jest szczególnie widoczny w przypadku egzaminu maturalnego, bo w każdym roku faktyczny próg zdawalności jest różny – wyniki egzaminacyjne z różnych lat i sesji egzaminacyjnych są nieporównywalne.

W celu osiągnięcia porównywalności wyników między latami zespół Analiz Osiągnięć Uczniów (ZAOU) działający w Instytucie Badań Edukacyjnych podjął w ramach projektu „Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego” prace nad badaniem, które umożliwi porównanie wyników egzaminacyjnych

Wynik egzaminu zależy od trudności arkusza testowego oraz umiejętności uczniów. Gdyby umiejętności uczniów w krajowej populacji nie zmieniały się, to średni krajowy wynik egzaminacyjny byłby informacją o trudności testu. Gdyby trudność arkuszy testowych w kolejnych latach była podobna, to na podstawie wyników egzaminacyjnych można byłoby wnioskować o tym, jak zmienia się w czasie poziom umiejętności uczniów. Skonstruowanie arkuszy testowych o podobnych parametrach jest jednak zadaniem bardzo trudnym i nie zawsze zakończonym powodzeniem. Porównywalność wyników między latami można osiągnąć kontrolując trudność arkusza (procedury zrównywania wyników) lub zakładając stały poziom umiejętności w populacji (normalizacja i standaryzacja rozkładów, np. do skali staninowej). W pierwszym przypadku uzyskujemy



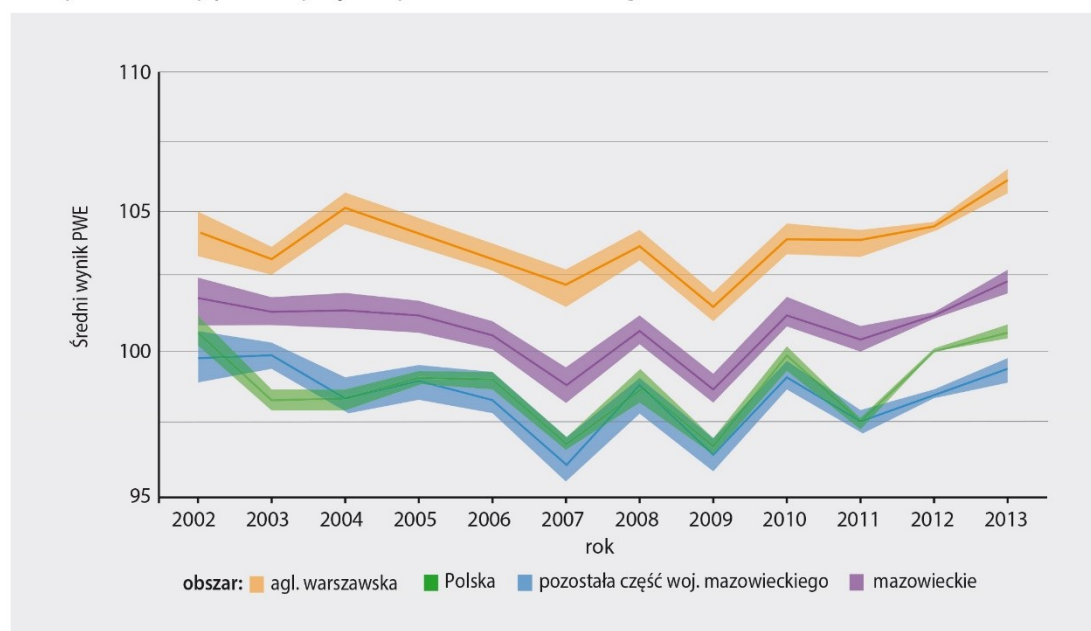
bezwzględną porównywalność wyników i możliwość wnioskowania o zmianie umiejętności w populacji. W drugim przypadku uzyskujemy względną porównywalność w stosunku do średniej w danym roku.

W latach 2011-2014 Zespół Analiz Osiągnięć Uczniów przeprowadził badania zrównujące w różnych typach szkół, co po zastosowaniu złożonej procedury zrównywania wyników pozwoliło uzyskać porównywalność wyników sprawdzianu i egzaminu gimnazjalnego od 2002 roku, a matury z j. polskiego, angielskiego oraz matematyki od 2010 roku. Za rok bazowy dla wszystkich egzaminów przyjęto rok 2012 (krajowa średnia dla 2012 roku wynosi 100), a wyniki danego egzaminu w pozostałych latach należy rozumieć jako wzrost/spadek poziomu umiejętności uczniów w porównaniu do 2012 roku.

Zrównane wyniki egzaminacyjne zapewniają porównywalność rezultatów egzaminacyjnych pomiędzy latami, dlatego też nazwaliśmy je Porównywalnymi Wynikami Egzaminacyjnymi (PWE). Mogą być one wykorzystywane przede wszystkim do monitorowania poziomu osiągnięć uczniów (sprawdzanych w trakcie egzaminu) w skali kraju. Mogą też być przydatne do monitorowania w wieloletnich przedziałach czasowych procesów edukacyjnych w skali województwa i powiatu.

Rys. 1. Porównywalne wyniki egzaminacyjne dla Warszawy i Mazowsza

Średnie PWE z części humanistycznej egzaminu gimnazjalnego wraz z 95% przedziałami ufności dla aglomeracji warszawskiej, pozostałej części województwa mazowieckiego oraz Polski



Edukacyjna wartość dodana

Gdyby wszystkie szkoły w Polsce pracowały z uczniami o tym samym poziomie osiągnięć szkolnych „na wejściu”, to faktycznie wyniki egzaminu końcowego byłyby dobrą miarą efektywności. Ale tak nie jest. Metoda EWD pozwala uwzględnić zróżnicowanie szkół ze względu na zasoby „na wejściu”.

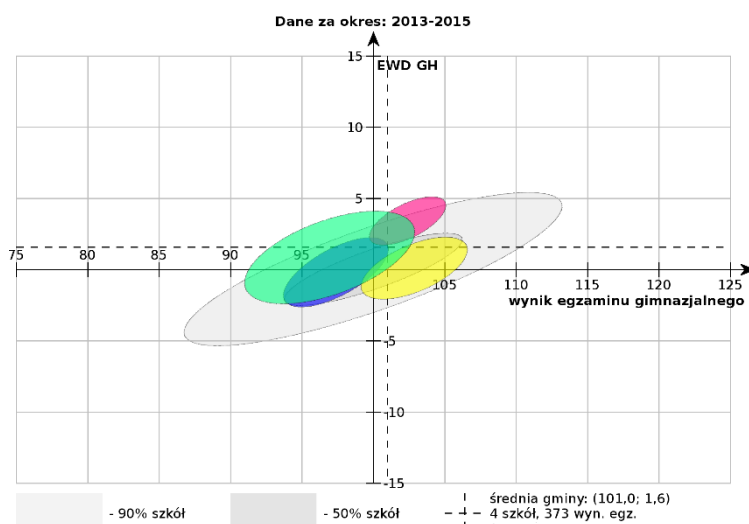


Pozwala zobaczyć wkład szkoły w osiągnięcia ucznia. Na czym polega? W skrócie na zbadaniu i porównaniu, z jakimi umiejętnościami uczeń zaczyna kolejny etap edukacji, a z jakimi go kończy. EWD pozwala określić, czy efektywność nauczania w danej szkole była wyższa (EWD dodatnie) lub niższa (EWD ujemne) od przeciętnej w Polsce dla danego typu szkół (EWD zerowe).

Polski system egzaminacyjny dostarcza danych do wyliczenia EWD dla gimnazjów (sprawdzian w klasie VI szkoły podstawowej jako miara na wejściu, egzamin gimnazjalny jako miara na wyjściu) oraz szkół maturalnych (egzamin gimnazjalny jako miara na wejściu, egzamin maturalny jako miara na wyjściu). Natomiast budowane od 2014 roku modele EWD dla szkół podstawowych wykorzystują jako pomiar na progu edukacyjnym wyniki Ogólnopolskiego badania umiejętności trzecioklasistów (OBUT) z lat 2012 i 2013 odpowiednio, a wyniki sprawdzianu w VI klasie szkoły podstawowej jako miarę wyników końcowych.

Ogólnodostępne są **trzyletnie** wskaźniki EWD dla gimnazjów i szkół maturalnych. Wskaźniki te wyliczane są na podstawie danych z kolejnych trzech edycji egzaminów i przedstawiane są w graficznej formie jako elipsa, która pokazuje pozycję szkoły na tle populacji ze względu na wyniki egzaminacyjne i efektywność nauczania. Wielkość elipsy zależy od liczby uczniów w szkole i zróżnicowania ich wyników: im więcej uczniów, tym mniejsza elipsa. Dla gimnazjów początkowo konstruowane były dwa typy wskaźników: humanistyczny i matematyczno-przyrodniczy, ale zmiana formuły egzaminu gimnazjalnego umożliwiła wyliczenie od 2014 roku dodatkowych czterech typów: z języka polskiego, z historii i WOSu oraz z matematyki i z przedmiotów przyrodniczych. Dla szkół maturalnych są to cztery wskaźniki: polonistyczny, humanistyczny, matematyczny oraz matematyczno-przyrodniczy.

Rys. 2. Wyniki czterech gimnazjów w jednej gminie



Wykorzystanie analizy egzaminów do pracy szkoły

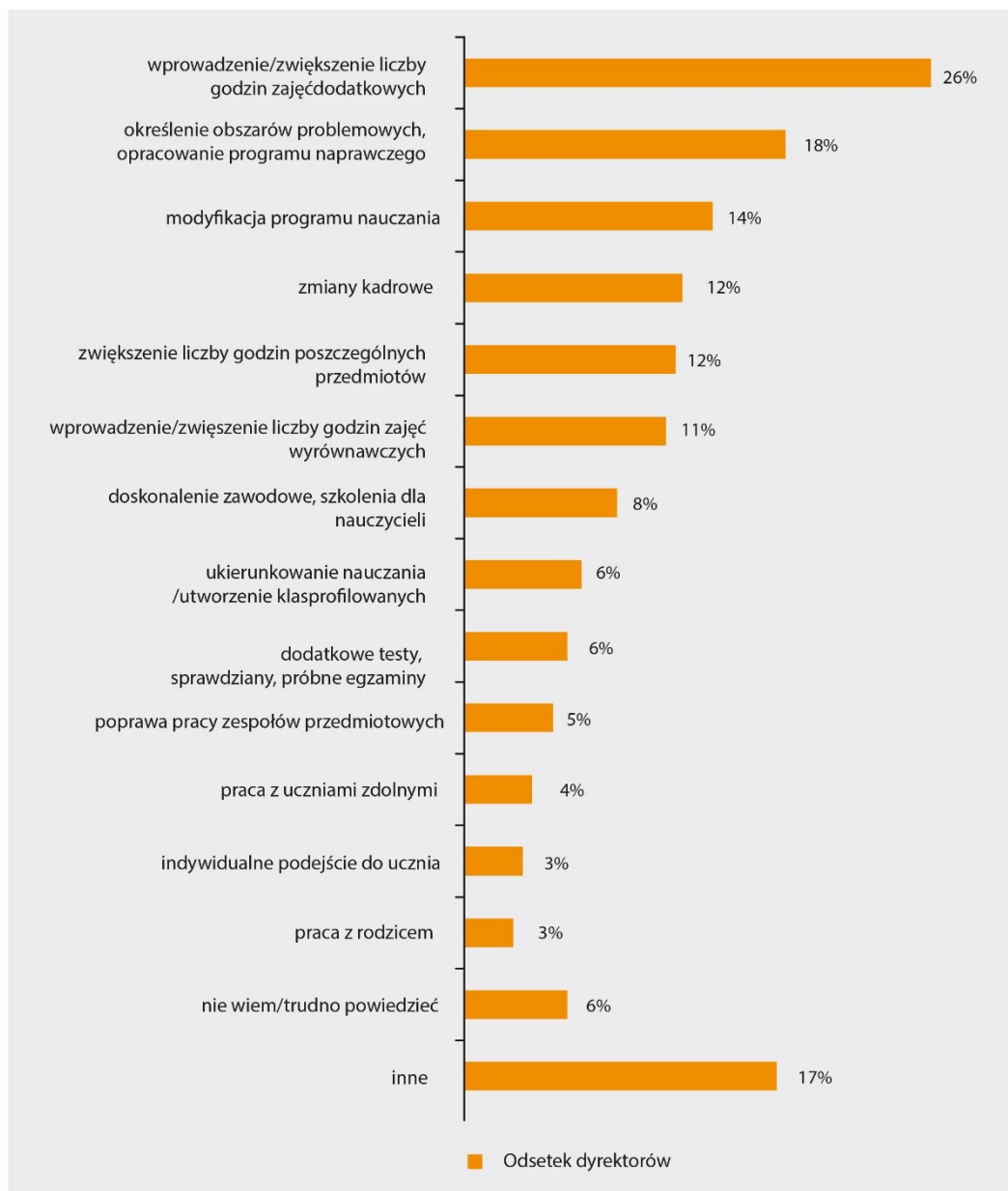
Badania międzynarodowe pokazują, że dyrektorzy i nauczyciele z polskich szkół znajdują się zazwyczaj wśród państw, w których najczęściej deklarowane jest wykorzystywanie wyników egzaminacyjnych w różnych aspektach. Z pogłębionych analiz z kolei wynika, że szkoły mają liczne problemy z prowadzonymi analizami, a większość z nich dopiero się uczy jak i w jakim celu analizować i wykorzystywać dane z systemu egzaminacyjnego. Zebrany materiał pokazuje, że motywacja do gromadzenia i przetwarzania danych w szkołach ma zazwyczaj charakter zewnętrzny – związany z regulacjami, które obligują szkoły do wykorzystywania wyników egzaminacyjnych, a nie z poszukiwaniem odpowiedzi na pojawiające się wyzwania, chęcią wprowadzania zmian w szkole.



W tej chwili wśród działań podejmowanych po analizach wyników egzaminacyjnych dominuje raczej intensyfikacja pracy nad konkretnymi umiejętnościami/przedmiotami, przede wszystkim poprzez zwiększenie liczby konkretnych ćwiczeń i liczby zajęć, na których pracuje się nad umiejętnościami uznanymi za słabiej opanowane, a nie bardziej ogólne doskonalenie sposobów nauczania, czy modyfikacje w warsztacie pracy.

Rys. 3. Wykorzystywanie wyników egzaminów do pracy szkoły

Decyzje przy których pomocne było wykorzystanie wskaźników EWD



Źródło: Na podstawie badania CATI przeprowadzonego przez IBE wśród dyrektorów w 2017 r. w podstawie procentowania: dyrektorzy, dla których EWD pomogło w podjęciu decyzji w dużym lub bardzo dużym stopniu (N = 776).



Możliwe kierunki rozwoju systemu egzaminacyjnego

Główne kierunki rozwoju systemu egzaminacyjnego powinny dotyczyć przede wszystkim lepszego przygotowywania zadań egzaminacyjnych i monitorowania trafności arkuszy egzaminacyjnych w różnych jej aspektach. Bardzo ważnym obszarem wymagającym poprawy jest proces oceniania zadań rozszerzonej odpowiedzi przez egzaminatorów (w czym pomocne będzie wdrażanie e-oceniania), a także udoskonalanie i uspojnianie rozwiązań organizacyjnych i informatycznych w zakresie przetwarzania i komunikowania wyników.

Aby diametralnie podnieść jakość arkuszy egzaminacyjnych konieczne jest przejście od autorsko tworzonych testów do budowania narzędzi egzaminacyjnych bazując na wykalibrowanych bankach zadań. Przygotowywanie egzaminów powinno przebiegać podobnie jak przygotowanie rzetelnego badania naukowego. Do tego potrzebni są nie tylko eksperci znający się na dydaktyce poszczególnych obszarów edukacyjnych ale także eksperci w dziedzinie psychometrii i sztuki tworzenia egzaminacyjnych testów. O pilnej potrzebie przyspieszenia prac nad bankami zadań może świadczyć fakt, że podczas analizy zadań z egzaminu gimnazjalnego w części humanistycznej i matematyczno-przyrodniczej w latach 2002-2010 Zespół Analiz Osiągnięć Uczniów IBE wytypował 66 zadań o słabych właściwościach psychometrycznych, które nie powinny znaleźć się w testach egzaminacyjnych.

Ważne jest też wprowadzenie do systemu egzaminacyjnego mechanizmów zapewniających porównywalność wyników kolejnych edycji egzaminów. Zwróciła na to uwagę Najwyższa Izba Kontroli, która w raporcie *System egzaminów zewnętrznych* wśród rekomendacji wnioskuje wprowadzenie na etapie tworzenia testów egzaminacyjnych metod, które nadadzą wynikom egzaminacyjnym walor międzyedycyjnej porównywalności.

Warto rozważyć wdrożenie jednego z trzech opisanych poniżej rozwiązań. Warianty te w mniejszym lub większym stopniu wiążą się z koniecznością zmian w systemie egzaminów zewnętrznych i związanych z tym zmian w prawie. Trzy warianty są przedstawione w kolejności pożądaných własności metodologicznych.

Wariant A – włączenie do egzaminu niejawnych zadań kotwiczących (rekomendowany, ale najtrudniejszy legislacyjnie)

Z metodologicznego i pomiarowego punktu widzenia jest to rozwiązanie najlepsze, dające najbardziej precyzyjne oszacowanie funkcji zrównującej i najmniej narażone na potencjalne czynniki zakłócające. Rozwiązanie to wymaga jednak istotnych zmian w sposobie przeprowadzania egzaminów.

Wariant B – jednoczesna standaryzacja zadań z wielu edycji egzaminu

Rozwiązanie jest mniej efektywne ze względu na wyznaczanie funkcji zrównującej. Wyniki mogą być obciążone wpływem niższego poziomu motywacji testowej uczniów na sesji standaryzacyjnej w porównaniu do sesji egzaminacyjnej.

Wariant C – dodatkowe badania na reprezentatywnej próbie

Jest to rozwiązanie nie wprowadzające żadnych zmian w systemie egzaminacyjnym, przez co byłoby najprostsze we wdrożeniu. Jednak wariant C rodzi największą liczbą potencjalnych problemów prowadzących do zakłócenia procesu zrównywania, ponadto odznacza się najniższą efektywnością. Ze względu na znaczną niepewność statystyczną wyników, jakie przynosi ta metoda, jej przydatność do zrównywania wyników pojedynczych uczniów jest ograniczona. Zaletą tego rozwiązania jest



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

IBE  *entuzjaści
edukacji*

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY



możliwość śledzenia zmian w poziomie umiejętności populacji między latami oraz zbieranie informacji o stabilności trudności testów egzaminacyjnych między latami.

