

EGZAMIN

BIULETYN BADAWCZY
CENTRALNEJ KOMISJI EGZAMINACYJNEJ

Perspektywy badań
wspomagających system
egzaminowania zewnętrznego

1

2004

SPIS TREŚCI

DO CZYTELNIKÓW.....str. 3

WPROWADZENIE

■ System egzaminowania zewnętrznego.....str. 4

WOKÓŁ SYSTEMU EGZAMINOWANIA ZEWNĘTRZNEGO

■ Bolesław Niemierko: *Perspektywy badań psychologicznych, dydaktycznych, socjologicznych i politologicznych wspomagających system egzaminowania zewnętrznego*.....str. 7

TAJNIKI WARSZTATU BADAWCZEGO

■ Grażyna Szyling: *Trafność standardów wymagań egzaminacyjnych a jakość systemu egzaminowania zewnętrznego*.....str. 14

■ Przykłady przeprowadzonych w latach 2002 – 2003 badań dotyczących egzaminowania zewnętrznego.....str. 23

▶ Jan Herczyński: *Pierwsza odsłona. Społeczne i terytorialne zróżnicowanie wyników sprawdzianu szóstoklasistów i egzaminu gimnazjalnego przeprowadzonych wiosną 2002 r.*.....str. 23

▶ Zbigniew Powązka, Joanna Czaplińska: *Z badań nad przygotowaniem kandydatów do studiowania matematyki w Akademii Pedagogicznej w Krakowie*.....str. 23

▶ Ada Szylo: *Analiza motywacji do działania na przykładzie aplikacji kandydatów na egzaminatorów w nowym systemie oceniania zewnętrznego*.....str. 24

■ Projekty badań systemu egzaminowania zewnętrznego

▶ Jacek Rysiewicz: *Rzetelność egzaminu maturalnego z języka angielskiego w jego obecnej formie w oparciu o porównanie wyników kandydatów na studia anglistyczne w Instytucie Filologii Angielskiej UAM w Poznaniu*.....str. 25

▶ Wojciech Małecki: *Reforma systemu oceniania przez pryzmat osobistych teorii nauczycieli*.....str. 26

▶ Wojciech Małecki: *Modyfikacja szkolnych programów nauczania w gimnazjum*.....str. 27

■ Dorobek badawczy Zakładu Diagnostyki Edukacyjnej Instytutu Pedagogiki Uniwersytetu Gdańskiego.....str. 28

SŁOWNICZEK BADACZA EDUKACYJNEGO

■ Diagnostyka edukacyjna.....str. 30

PRZECZYTAJ, ZAJRZYJ, WYKORZYSTAJ.....str. 31

CZYTELNICY PISZA.....str. 32

Do Czytelników

Oddajemy do rąk Czytelników 1. numer *Egzaminu* z nadzieją. To ona pozwala nam bowiem sądzić, że zajądą do niego i badacze profesjonalnie zajmujący się polskim systemem edukacji, i troyący ciekawe tematy badawcze studenci, i nauczyciele poszukujący odpowiedzi na codziennie stawiane pytania, jak lepiej, skuteczniej, sprawiedliwiej oceniać, jak prowadzić ewaluację własnych działań, jakie metody stosować w procesie mierzenia jakości pracy szkoły, itp.

Nadzieja ta umacnia też redakcję w przekonaniu, że tematów nie zabraknie, bo badania związane z ocenianiem osiągnięć uczniów powinny towarzyszyć temu procesowi nieustannie, a coraz liczniej angażujący się w nie badacze zechcą podzielić się swoimi doświadczeniami z Czytelnikami biuletynu. Bardzo nam na tym zależy, by poznać Państwa projekty i przedsięwzięcia badawcze – zarówno te globalne, szeroko zakrojone, prowadzone na dużych populacjach, jak i te związane z np. ocenianiem wewnątrzszkolnym, prowadzone z użyciem metod właściwych dla studium przypadku.

Stąd apel – niech każdy, kogo interesuje system oceniania, egzaminowanie, ewaluacja osiągnięć, metodologia badań pedagogicznych, kto zamierza lub prowadzi badania zmierzające do wniosków związanych z jakością, trafnością, użytecznością, etc. systemu oświaty, podzieli się swą wiedzą i doświadczeniami z innymi. Nasze łamy stoją dla niego otworem. Ze swej strony obiecujemy z kolei spojrzenie na egzaminowanie zewnętrzne „od kuchni”, dowodzące, iż budowanie zasad ujednoczonego, porównywalnego, stabilnego systemu jest jednym z głównych celów Centralnej i okręgowych komisji egzaminacyjnych.

Pragniemy, by *Egzamin* był bezpłatnym kwartalnikiem, dostępnym w formie elektronicznej. Redagować go będzie zespół ekspertów zatrudnionych w Pracowni Ewaluacji CKE. W każdym z numerów znajdziecie Państwo m. in. działy poświęcone systemowi egzaminowania zewnętrznego, metodologii badań, projektom badawczym, literaturze przedmiotu.

Pierwszy numer biuletynu zdominowała problematyka związana z perspektywami badań systemu egzaminowania zewnętrznego, stanowiąca główny temat konferencji zorganizowanej przez CKE w październiku 2003 r. Prezentowane tu materiały – mamy nadzieję – pozwolą Państwu nie tylko na zorientowanie się w tej jakże ważnej dla nas problematyce, ale też pozwolą baczniej przyjrzeć się wybranym obszarom badawczym.

Prezentację otwiera referat Profesora Bolesława Niemierki, uczonego, który problematyką oceniania, ewaluacji osiągnięć szkolnych, diagnostyki edukacyjnej zajmuje się od wielu lat. Jest w tej dziedzinie niekwestionowanym autorytetem, a z Jego dorobku korzystają kolejne pokolenia nauczycieli i badaczy. W kierowanym przez Profesora Zakładzie Diagnostyki Edukacyjnej Uniwersytetu Gdańskiego (a wcześniej w Międzywydziałowym Studium Pedagogicznym) regularnie spotyka się coraz liczniejsza seminaryjna grupa nauczycieli-praktyków, planujących lub prowadzących badania związane z diagnozą osiągnięć uczniów. Dorobek kilkorga z nich prezentujemy w działach *Tajniki warsztatu badawczego* oraz *Doniesienia badawcze*.

Kończymy to wprowadzenie także z nadzieją – że nasz biuletyn przyjęty zostanie przez Państwa życzliwie, a jego lektura przybliży nieco tajniki systemu oceniania zewnętrznego, że rychło staniecie się Państwo współredaktorami wszystkich stałych działów, że zechcecie nas wesprzeć w poszukiwaniu dróg, jak oceniać i egzaminować lepiej, sprawiedliwiej, rzetelniej.

Pozostajemy z wyrazami szacunku i poważania –

Redakcja

WPROWADZENIE

System egzaminowania zewnętrznego

Reforma polskiej edukacji wprowadziła zasadnicze zmiany w procesie nauczania i oceniania osiągnięć uczniów. Zamiast centralnie wprowadzanych obowiązkowych programów nauczania dla poszczególnych przedmiotów, ustanowiono dla każdego typu szkoły podstawy kształcenia ogólnego, które wprawdzie ogólnie określają cele edukacyjne i zadania szkoły oraz sprawności i umiejętności uczniów kończących kolejne etapy kształcenia, ale już treści nauczania ograniczają do podstawowego kanonu. Daje to szkołom dużą autonomię w tworzeniu własnych autorskich koncepcji kształcenia i wychowania. Nauczyciele sami wybierają (lub tworzą autorskie) programy nauczania zgodne z podstawą programową i sami formułują wymagania edukacyjne oraz zasady oceniania. Tym samym ocen będących informacją o osiągnięciach i postępach uczniów poszczególnych szkół wystawionych na podstawie różnych wymagań nie można porównać bez uwzględnienia wszystkich kontekstów, w których przebiegał proces dydaktyczny.

W tej sytuacji, aby umożliwić bardziej zobiektywizowaną ocenę oddziaływań dydaktycznych nauczycieli i móc orzekać o jakości programów stosowanych w szkole oraz zapewnić spójność nauczania, konieczne stało się wprowadzenie do systemu edukacji **oceniania zewnętrznego**, opartego na jednolitych w całym kraju wymaganiach – standardach wymagań egzaminacyjnych.

Ocenianie zewnętrzne odbywa się po zakończeniu każdego z trzech etapów obowiązkowego kształcenia. Na zakończenie szkoły podstawowej – powszechny i obowiązkowy sprawdzian o charakterze ponadprzedmiotowym. W trzeciej klasie gimnazjum uczniowie przystępują do powszechnego, obowiązkowego egzaminu przeprowadzanego w części humanistycznej i matematyczno-przyrodniczej. Po ukończeniu liceum, liceum profilowanego i technikum można przystąpić do nieobowiązkowego egzaminu maturalnego, a po ukończeniu szkoły zawodowej do nieobowiązkowego egzaminu potwierdzającego uzyskanie kwalifikacji zawodowych, składającego się z części teoretycznej i praktycznej.

Pierwsze egzaminy zewnętrzne odbyły się w 2002 roku. Przystąpili do nich uczniowie kończący szkołę podstawową i gimnazjum oraz ci absolwenci liceum, którzy chcieli się sprawdzić na maturze przeprowadzanej wg nowej formuły.

Zadanie przygotowania zewnętrznego systemu oceniania oraz organizacji systemu egzaminowania zewnętrznego powierzono **Centralnej Komisji Egzaminacyjnej** z siedzibą w Warszawie, która koordynuje działania 8 okręgowych komisji egzaminacyjnych. Zarówno na Centralnej, jak i okręgowych komisjach w jednakowym stopniu spoczywa obowiązek zapewnienia jednolitości i wysokiej jakości przygotowywanych i przeprowadzanych przez nie egzaminów.

Centralna Komisja Egzaminacyjna, powołana *Ustawą o zmianie ustawy o systemie oświaty* z dnia 25 lipca 1998 r., swoją działalność rozpoczęła z dniem 1 stycznia 1999 roku. Jej ogólne zadania określa ustawa o systemie oświaty, natomiast szczegółowe – statut nadany przez właściwego ministra do spraw oświaty i wychowania. Zgodnie z zapisami we wspomnianych dokumentach Centralna Komisja Egzaminacyjna odpowiada za:

- opracowywanie propozycji standardów wymagań będących podstawą przeprowadzania sprawdzianów i egzaminów,
- przygotowywanie pytań, zadań i testów oraz ustalanie zestawów do przeprowadzenia sprawdzianu i egzaminów,
- dokonywanie analizy wyników sprawdzianów i egzaminów w skali całego kraju oraz składanie ministrowi corocznych sprawozdań dotyczących poziomu osiągnięć uczniów na poszczególnych etapach kształcenia,
- ustalanie jednakowych w całym kraju terminów sprawdzianu i egzaminów,
- we współpracy z komisjami okręgowymi opracowywanie i ogłaszanie informatorów zawierających opis zakresu sprawdzianu i egzaminów,
- koordynowanie działalności okręgowych komisji egzaminacyjnych oraz nadzorowanie ich prac związanych z opracowywaniem propozycji zestawów zadań, pytań i testów do sprawdzianu i egzaminów, a także przygotowywaniem i przeprowadzaniem sprawdzianu i egzaminów oraz ocenianiem arkuszy egzaminacyjnych,
- przygotowywanie i upowszechnianie programów doskonalenia nauczycieli w zakresie diagnozowania i oceniania,
- inspirowanie badań naukowych i innowacji w dziedzinie oceniania i egzaminowania.

Organem doradczym dyrektora CKE jest Rada Programowa, skupiająca uznane autorytety naukowe – profesorów reprezentujących wyższe uczelnie, konferencje rektorów, instytuty oraz placówki naukowo-badawcze. W obecnej kadencji funkcję przewodniczącego Rady Programowej pełni prof. dr hab. Bolesław Niemierko z Uniwersytetu Gdańskiego.

Okręgowe komisje egzaminacyjne – z siedzibami w: Gdańsku, Łomży, Warszawie, Poznaniu, Łodzi, Krakowie, Jaworznie i we Wrocławiu – powołano Rozporządzeniem Ministra Edukacji Narodowej z dnia 18 lutego 1999 r. w *sprawie utworzenia okręgowych komisji egzaminacyjnych oraz określenia ich zasięgu terytorialnego*. Ogólne zadania komisji okręgowych, podobnie jak w przypadku Centralnej Komisji Egzaminacyjnej, określa *Ustawa o systemie oświaty*, a szczegółowe – statut nadany przez Ministra Edukacji Narodowej i Sportu.

Do zadań okręgowych komisji egzaminacyjnych należy w szczególności:

- przygotowywanie, w porozumieniu z CKE, propozycji pytań, zadań i testów oraz ich zestawów do przeprowadzenia sprawdzianów i egzaminów,
- organizacja i przeprowadzanie sprawdzianów i egzaminów, w tym porównywalne sprawdzenie i ocenianie arkuszy egzaminacyjnych,
- analizowanie wyników sprawdzianów i egzaminów przeprowadzonych na terenie OKE oraz formułowanie stosownych wniosków,
- opracowywanie i przekazywanie sprawozdań z przeprowadzonych sprawdzianów i egzaminów dyrektorom szkół, organom prowadzącym szkoły, kuratorom oświaty i Centralnej Komisji Egzaminacyjnej,
- szkolenie kandydatów na egzaminatorów oraz ustawiczne szkolenie egzaminatorów, według przygotowanych w CKE programów szkoleń,
- prowadzenie ewidencji egzaminatorów zamieszkujących na terenie objętym działaniem danej OKE,
- przygotowywanie, w porozumieniu z CKE, propozycji pytań, zadań i testów oraz kryteriów ich oceniania do informatorów,
- współpraca z kuratoriami oświaty właściwymi ze względu na zasięg terytorialny komisji w sprawach związanych z przeprowadzaniem sprawdzianu i egzaminów oraz dosko-

nalaniem nauczycieli w zakresie diagnozowania, oceniania, egzaminowania i badania osiągnięć edukacyjnych uczniów.

Pracami każdej okręgowej komisji egzaminacyjnej kierują dyrektorzy wyłonieni w drodze konkursu i powołani przez ministra właściwego do spraw oświaty i wychowania.



WOKÓŁ SYSTEMU EGZAMINOWANIA ZEWNĘTRZNEGO

prof. dr hab. Bolesław Niemierko

Zakład Diagnostyki Edukacyjnej

Uniwersytetu Gdańskiego

Przewodniczący Rady Programowej CKE

Perspektywy badań psychologicznych, dydaktycznych, socjologicznych i politologicznych wspomagających system egzaminowania zewnętrznego¹

Gdy nauka śpi...

...rosną w siłę populistyczne demony. Oburzamy się szczerze na nieuczciwe praktyki, prostackie interpretacje i pochopne decyzje wywołane *szkolnymi egzaminami zewnętrznymi*, lecz stajemy do nierównej walki. Rozpętane emocje krytyków (egzaminu, edukacji) bez trudu biorą górę nad naszymi argumentami i dyskusja prowadzi do *nieporozumienia* stron.

Jest głębszy powód, by zając się egzaminami. Otóż żadna dziedzina współczesnej cywilizacji nie może rozwijać się bez wsparcia naukowego. Wiedzą o tym potężne agencje egzaminacyjne, jak *ETS* w USA i *CITO* w Holandii, zatrudniając około 1/3 swych pracowników w pionie badawczym. Wie także Centralna Komisja Egzaminacyjna w Warszawie, utrzymując kilkusobową Pracownię Ewaluacji. W żadnym wszakże kraju własne badania profesjonalnego ośrodka nie wystarczą, by zapewnić *egzaminom doniosłym* pomyślny rozwój.

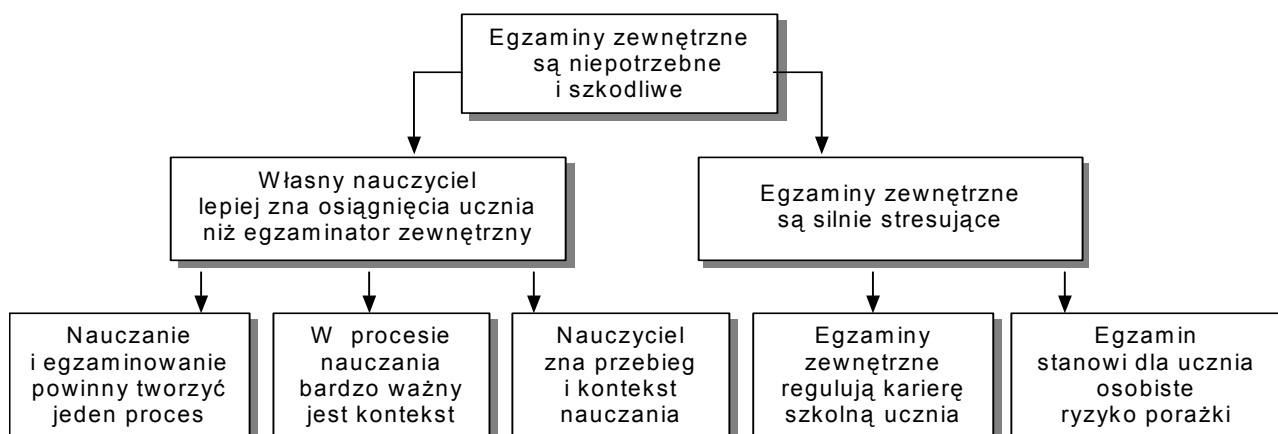
Złożoność problematyki egzaminu

Strategią niekompetencji jest koncentracja energii na jednym, intuicyjnie wybranym kierunku. Tak rodzą się uprzedzenia, stereotypy, znachorskie praktyki, posądzenia i medialne sądy czarownic. Spróbujmy przyrzeć się im, stosując metodę „rozpakowywania poglądów” (ang. *unpacking*), zaproponowaną przez Edwarda Haertela z Uniwersytetu Stanforda w USA².

¹ Referat został wygłoszony na konferencji *System egzaminowania zewnętrznego – perspektywy badawcze*, zorganizowanej przez Centralną Komisję Egzaminacyjną w Warszawie 29 października 2003 r.

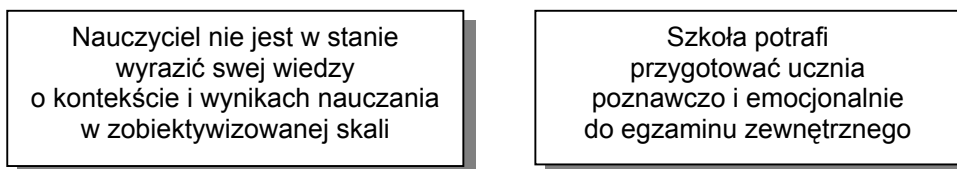
² E.H. Haertel, *Validity Arguments for High-Stakes Testing: In Search for Evidence*, „Educational Measurement: Issues and Practice” 1999 nr 4. Podane tu przykłady przedstawiłem wcześniej w referacie *Jaki pomiar dydaktyczny jest nam potrzebny?* na IX Krajowej Konferencji Diagnostyki Edukacyjnej w Łodzi (24 X 2003).

Rysunek 1. Rozpakowanie poglądu, że EGZAMINY ZEWNĘTRZNE SĄ NIEPOTRZEBNE I SZKODLIWE



Dekonstrukcja poglądu przez wykrycie jego założeń i „przedzałożeń”, tak modna w postmodernistycznym świecie, jest tu zastosowana pozytywnie, dla zrozumienia przekonań oponenta. Uogólnienie zawarte w górnym wierszu rysunku ma podstawę w dwu twierdzeniach środkowego wiersza, które możemy uznać za prawdziwe, a te wspierają się na przesłankach dolnego wiersza, które nie tylko są w pełni przekonujące, lecz także pedagogicznie fundamentalne. Zatem, jeżeli *redukcyjne* „rozpakowanie” poglądu zamienić na jego *dedukcyjne* uzasadnienie, to z prawdy (prawdziwych założeń) wynika fałsz (nieprawdziwe uogólnienie).

Gdzie zatem tkwi błąd rzeczowy? W *uproszczeniu*, to jest w ograniczeniu liczby przesłanek elementarnych. Gdyby w dolnym wierszu wykresu znalazły się i takie przesłanki:



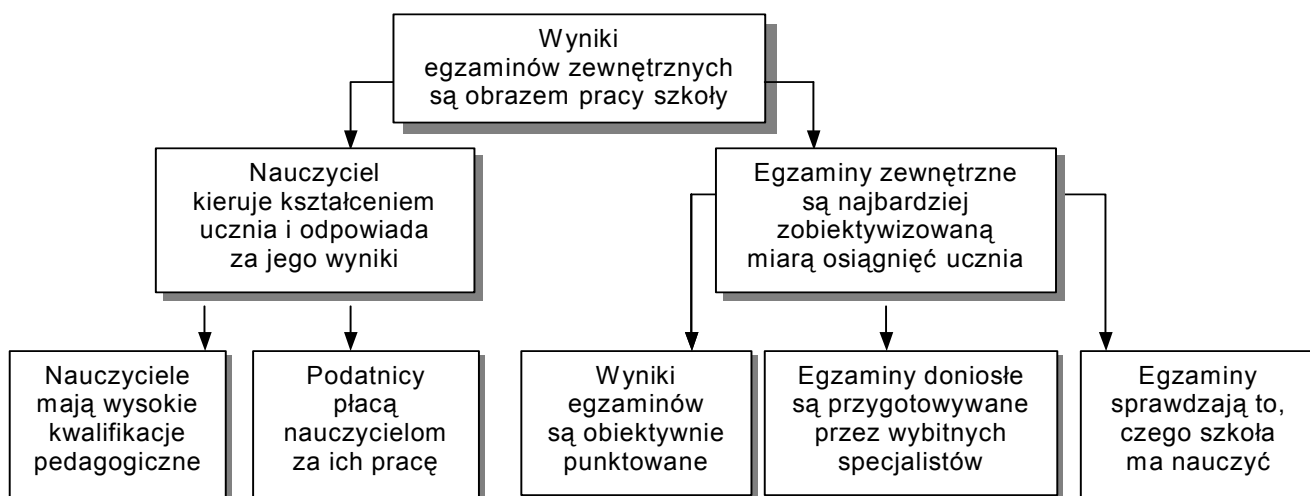
rozumowanie doprowadziłoby do ostrożniejszych wniosków.

Uproszczone rozwiązania złożonych zagadnień, nie tylko pedagogicznych, mogą doprowadzić do spustoszenia intelektualnego wśród odbiorców. Przypuśćmy, że utalentowany dziennikarz lub, co gorsza, literacko uzdolniony uczonej (politolog, ekonomista, pedagog) przedstawia swoją *dedukcję z emocjonalnie mu bliskich poglądów* w eleganckiej, quasi-indukcyjnej formie. Odbiorcy są zachwyceni prostotą wywodu, który wciąga ich i rzekomo rozwiązuje problem, z jakim przedtem się borykali. Tak tworzy się grono entuzjastów lub może sekta wyznawców proroka, przekonanych o bezwzględnej racji pewnej tezy.

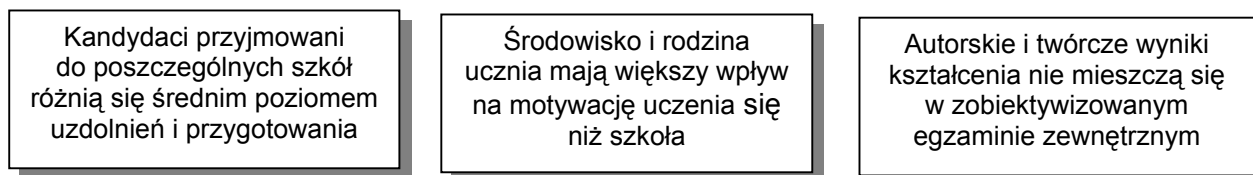
Największa herezja egzaminu zewnętrznego

Największą herezją egzaminacyjną, szerzącą się na całym bodaj świecie, jest przekonanie o pełnej i wyłącznej odpowiedzialności szkoły za osiągnięcia uczniów w egzaminach zewnętrznych. Oto jaka bywa argumentacja zwolenników tego poglądu.

Rysunek 2. Rozpakowanie poglądu, że WYNIKI EGZAMINÓW SĄ OBRAZEM PRACY SZKOŁY



Także i tym razem nietrudno jest wskazać przesłanki relatywizujących prawdziwość naczelną tezę zwolenników *administracyjnych sankcji* wobec szkół, których uczniowie uzyskują zbyt niskie wyniki egzaminów zewnętrznych:

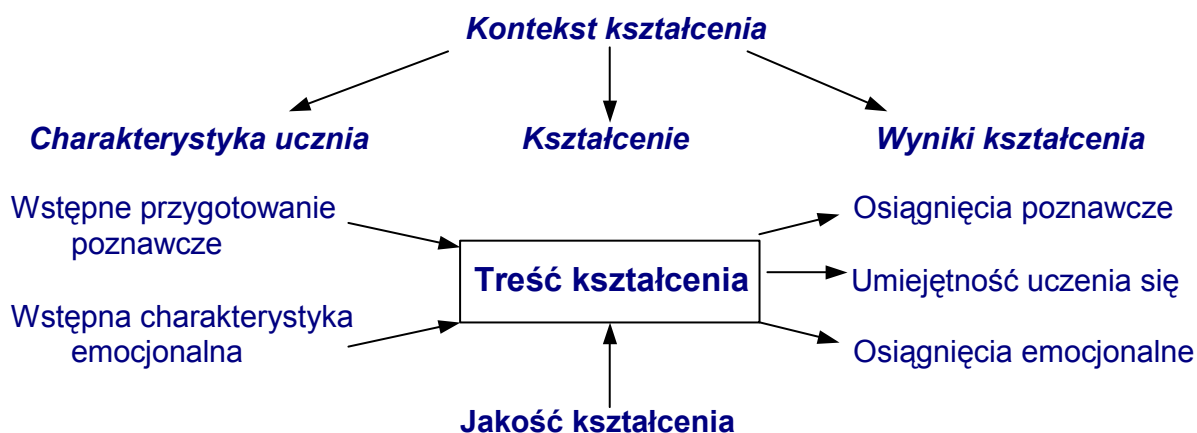


Metoda *rozpakowywania poglądów* uzupełniona o poszukiwanie *pominiętych przesłanek* niechybnie ujawnia słabość wąskich, osobistych i środowiskowych „teorii”, lansowanych – w dobrej lub nieco gorszej wierze – na różnych szczeblach kompetencji pedagogicznej.

Najprostszy model diagnostyki edukacyjnej

Nawet najprostszy model diagnostyki edukacyjnej wyznacza szerokie pole naukowym *badaniom interdyscyplinarnym* nad warunkami, przebiegiem i wynikami kształcenia. Oto taki model, zbudowany według wzoru dostarczonego przez Beniamina Blooma³:

Rysunek 3. Model kształcenia jako podstawa DIAGNOSTYKI EDUKACYJNEJ



³ B.S. Bloom, *Human Characteristics and School Learning*. New York 1976, McGraw-Hill. Ten model przedstawiłem w książce *Ocenianie szkolne bez tajemnic*, Warszawa 2002, WSiP, s. 72.

Model uzależnia *wyniki kształcenia*, których część – a mianowicie osiągnięcia poznawcze uczniów i, w coraz większym stopniu, umiejętność uczenia się – są objęte egzaminami wewnątrzszkolnymi i zewnętrznymi, od:

1. *wstępnej charakterystyki uczniów* oraz zjawisk rozwojowych, będących domeną analiz psychologicznych,
2. doboru *treści* i uzyskanej *jakości* kształcenia, stanowiących zagadnienia w przewadze dydaktyczne,
3. szeroko rozumianego, ekonomiczno-społecznego *kontekstu kształcenia*, stanowiącego dziedzinę studiów socjologicznych i politologicznych.

Model przestrzega przed wnioskowaniem o jakości kształcenia wyłącznie na podstawie jego wyników. Ma charakter *regresyjny*. Badania pokazują, że w pojedynkę (na podstawie korelacji dwu zmiennych) jakość kształcenia tłumaczy tylko $\frac{1}{4}$ końcowych osiągnięć ucznia, podczas gdy wstępna charakterystyka poznawcza ucznia tłumaczy około $\frac{1}{2}$ osiągnięć, a więc znacznie więcej⁴.

Zagrożenie edukacji spowodowane wycinkową interpretacją uwarunkowań wyników kształcenia jest realne. Wyniki egzaminu zewnętrznego zwykle rozczarowują, gdyż – w przeciwieństwie do oceniania wewnątrzszkolnego – nie obejmuje on osiągnięć emocjonalnych i demonstracja pilności przez ucznia nie podnosi jego ocen⁵. Zaczyna się szukanie „winnych”, a najłatwiej ich znaleźć wśród osób zatrudnionych w systemie edukacji. Skutki przedsięwziętych akcji porównuje się z totalitarną polityką Mao: wytępiono ptaki, które zjadały część ziarna, a w konsekwencji owady tak się rozmnożyły, że zjadły całe ziarno i zapanował głód⁶.

Najpilniejsza problematyka badawcza

Wracamy z Chin do Polski. Według mego rozeznania, na czoło listy palących *problemów egzaminowania zewnętrznego* wysuwa się siedem problemów, które zestawilem w Tabeli 1. wraz ze swoją interpretacją ich charakteru badawczego.

Tabela 1. Najpilniejsza problematyka badawcza egzaminu zewnętrznego

T e m a t y k a	D y s c y p l i n a w i o d a c a			
	psychologia	dydaktyka	socjologia	politologia
	<i>rozwój ucznia</i>	<i>treść kształcenia</i>	<i>grupy społeczne</i>	<i>stosunki władzy</i>
<i>Stres egzaminacyjny</i>	X		X	
<i>Standardy wymagań</i>		X	X	
<i>Wielopoziomowość</i>		X		X
<i>Uczciwość egzaminacyjna</i>	X		X	
<i>Ocenianie wewnątrzszkolne</i>	X	X		
<i>Odpowiedzialność szkoły</i>			X	X
<i>Efekt zwrotny</i>		X		X

⁴ B. S. Bloom, op. cit., s. 192. Chodzi o stosunek wariancji wyjaśnionej do wariancji całkowitej wyników kształcenia.

⁵ B. Niemierko, *Ocenianie szkolne ...*, ss. 212 – 215.

⁶ W. Rich, *Historical High-Stakes Policies Relating to Unintended Consequences of High-Stakes Testing. Response to Cizek*, „Educational Measurement: Issues and Practice” 2003 nr 1. Jest to głos w szerokiej dyskusji o systemowych konsekwencjach egzaminowania zewnętrznego, jaka toczy się na łamach dydaktycznopomiarowych czasopism amerykańskich.

A oto krótki komentarz do treści tej tabeli:

1. *Stres egzaminacyjny* jest od wielu dziesięcioleci przedmiotem badań psychologów w wielu krajach⁷. *Niepokój testowy* (ang. *test anxiety*) podlega w pewnym stopniu oddziaływaniom edukacyjnym („obyście testowe”), ale jest trudny do całkowitego wyeliminowania, a jego zależność od czynników psychologicznych (osobowość ucznia, aspekty sytuacyjne) i socjologicznych (aspiracje do wykształcenia, kulturowe wzory zachowań) jest silna i historycznie zmienna.
2. *Standardy wymagań egzaminacyjnych* dojrzewają w Polsce do starannie przygotowanej rewizji. Należałoby ją oprzeć na dostatecznie szerokich badaniach odbioru tych tekstów przez różne grupy użytkowników (uczniów, rodziców, nauczycieli, administrację szkolną, metodyków przedmiotowych, wydawców oświatowych, badaczy i dziennikarzy oświatowych, wykładowców uczelni pedagogicznych).
3. *Wielopoziomowość wymagań* jest osobnym problemem egzaminacyjnym. Standardy jednego poziomu nie wystarczą dla sprawdzenia osiągnięć żadnej realnej populacji uczniów: dla jednych są za wysokie, a dla innych za niskie. Wytwarzają ogromne różnice ilościowe (zdobytych punktów). Wprowadzenie kilku poziomów wymagań podnosi *znaczenie treściowe* standardów i umożliwia (jak w angielskim *National Curriculum*) zbudowanie *norm podłużnych*. Takie rozwiązania jednak, dokumentujące i utrwalające środowiskowe różnice między uczniami, nie są obojętne politycznie. Czy lepsza jest zatem fasadowa „szkoła jednolita”?
4. *Uczciwość egzaminacyjna* z dnia na dzień stała się międzynarodowym problemem badawczym. Dużą rolę odegrała tu głośna książka Gregory Cizka, pokazująca rozmiary i konsekwencje *oszustwa egzaminacyjnego*.⁸ W Polsce zorganizowano już kilka konferencji, których uczestnicy wyrażali potępienie dla „podpowiadania” i „ściągnięcia”. Realny, choć powolny postęp można tu jednak uzyskać tylko przez badanie okoliczności i natury oszustwa oraz przez eksperymentowanie z programami zaradczymi, licząc się przy tym z faktem, że normy uczciwości egzaminacyjnej są silnie uwarunkowane środowiskowo⁹.
5. *Ocenianie wewnętrzne* stało się w Polsce problemem badawczym z dwu powodów: (1) wyzwolenia *szkolnych i przedmiotowych systemów oceniania* z gorsetu ministerialnych regulaminów i (2) „konkurencji” ze strony egzaminowania zewnętrznego. Jak odbierany przez uczniów i rodziców jest taki pluralizm?
6. *Odpowiedzialność szkoły za wyniki kształcenia* była już charakteryzowana jako dziedzina rażących uproszczeń i niesprawiedliwych decyzji. Potrzeba interdyscyplinarnych badań i upowszechniania wyników tych badań jest szczególnie pilna.
7. *Efekt zwrotny* (ang. *backwash effect*) egzaminów doniosłych w postaci redukcji programów kształcenia do tego, „co będzie na egzaminie”, jest również szeroko znanym zjawiskiem.¹⁰ Występuje we wszystkich krajach stosujących egzaminy zewnętrzne lub choćby tylko testy standaryzowane. Ani żądanie, by egzaminy „bezszywowo przylegały” (ang. *seamlessness*) do podstawy programowej i „sprawdzały wszystko”, ani zakaz koniunkturalnej redukcji programu kształcenia nie jest skutecznym rozwiązaniem. Polityka programowo-egzaminacyjna kraju musi polegać na dokładnym przebadaniu zależności: „naprzód” (od programu do egzaminu) i „wstecz” (od egzaminu do programu).

⁷ Zob. np. S. Szuman, *Badania nad wzruszeniem egzaminowym i jego wpływem na sprawność umysłową w czasie egzaminu*, „Polskie Archiwum Psychologii” 1932, tom X, nr 4.

⁸ G. J. Cizek, *Cheating. How to Do It, Detect It, and Prevent It*, Mahwah 1999, Erlbaum.

⁹ Tamże, ss. 85-88 i 95. Kultura grupowej pomocy słabszym bywa nazywana „marksistowską teorią oszustwa egzaminacyjnego”!

¹⁰ „System egzaminacyjny zabija wszystko” – mówi dyrektor szkoły. „Gazeta Wyborcza” 2003-10-23.

Pomiar dydaktyczny

Jako procedura sprawdzania i oceniania osiągnięć szkolnych, egzaminy szkolne podlegają interpretacji pomiarowej. Teoria *pomiaru dydaktycznego* może być więc podstawą oceny ich jakości. Odwołanie się do hierarchii niezbędnych właściwości pomiaru dydaktycznego¹¹ pozwala nam uporządkować problematykę badawczą egzaminów od koncepcji i konstrukcji sytuacji egzaminacyjnej do interpretacji uzyskanych wyników jako wskaźników skuteczności edukacji narodowej. Przykład takiego uporządkowania zawiera Tabela 2.

Tabela 2. Problematyka badawcza egzaminu uporządkowana według teorii pomiaru sprawdzającego

Właściwość egzaminu	Główne zmienne niezależne i kryteria
<i>Bezstronność egzaminu</i>	Warstwy społeczne. Płeć ucznia. Dyscyplina organizacyjna.
<i>Dokładność punktowania</i>	Forma zadań: otwarta, zamknięta. Egzaminator.
<i>Rzetelność egzaminu</i>	Jednorodność testu. Błąd pomiaru. Funkcja informacyjna.
<i>Trafność egzaminu</i>	Jakość pracy szkoły. Przewidywanie przyszłych osiągnięć.
<i>Obiektywizm egzaminu</i>	Podstawy programowe. Standardy. Wskaźniki OECD.

Badania nad *bezstronnością* egzaminu i *dokładnością punktowania*, związane z przygotowaniem i corocznym prowadzeniem akcji, najłatwiej prowadzić wewnątrz komisji egzaminacyjnych, co wcale nie wyklucza udziału niezależnych badaczy, reprezentujących różnorodne dyscypliny naukowe. Chodzi o to, by egzaminy jednakowo sprzyjały młodzieży z miasta i ze wsi, chłopcom i dziewczynkom, humanistom i „ściśłowcom”, a ich wynik był jak najmniej zależny od wybranej formy zadań (zgadywania, sprytu testowego, łatwości pisania) i od osobowości egzaminatora, który będzie punktował wyniki.

Badania nad *rzetelnością* egzaminu zewnętrznego wymagają zaawansowanego aparatu analiz statystycznych, włączając w to znajomość probabilistycznej teorii zadania testowego (ang. *item response theory, IRT*). Są niezbędne jako przestroga przed naiwnym traktowaniem wyniku ucznia jako dokładnej miary osiągnięć mierzonych testem, podczas gdy w grę wchodzi tylko przybliżone oszacowania przedziałowe. Osobne zagadnienie stanowi *zrównywanie* (ang. *equating*) wyników wersji testu stosowanych równolegle i w kolejnych latach.

Dwie najwyższe właściwości pomiaru, trafność i obiektywizm, dotyczą interpretacji jego wyników. Określanie wartości dodanej (ang. *value added*) poszczególnych kompetencji uczniów przez szkołę, porównywanie zakresu egzaminu z podstawą programową i programami autorskimi (*trafność programowa*) oraz z ustalonymi standardami wymagań (*obiektywizm sprawdzania*), porównywanie wyników egzaminu z wynikami badań osiągnięć szkolnych (*trafność diagnostyczna*) i przewidywanie osiągnięć ucznia w szkole wyższego szczebla (*trafność prognostyczna*) – oto najważniejsze rodzaje analiz wymagających wiedzy psychologicznej (*rozwój ucznia*), dydaktycznej (*treść kształcenia i egzaminowania*), socjologicznej (*grupy społeczne*) i politologicznej (*kierowanie edukacją*).

Czego oczekujemy?

Rada Programowa Centralnej Komisji Egzaminacyjnej pozwala sobie wyrazić pod adresem uczestników konferencji *System egzaminowania zewnętrznego – perspektywy badawcze* następującą nadzieję:

¹¹ B. Niemierko, *Pomiar sprawdzający w dydaktyce. Teoria i zastosowania*, Warszawa 1990, PWN, ss. 112-121.

1. że problematyka badań nad założeniami, przebiegiem, wynikami i konsekwencjami egzaminów zewnętrznych zostanie uznana za dostatecznie dojrzałą i atrakcyjną do podjęcia przez *interdyscyplinarne zespoły badawcze* spoza komisji egzaminacyjnych,
2. że uda się na ten cel pozyskać *fundusze celowe i granty*, w czym Rada Programowa może pomóc, korzystnie opiniując odpowiednio dojrzałe koncepcje,
3. że Centralna Komisja Egzaminacyjna będzie w stanie udostępnić – z poszanowaniem ochrony danych osobowych – *informację* niezbędną do właściwych analiz,
4. że badania będą podjęte w dobrej wierze („bona fide”) *uniknięcia jednostronności i sensacji*, wielce szkodliwych dla wszelkich procesów edukacyjnych,
5. że badania dostarczą wiedzy umożliwiającej stopniowe *podnoszenie jakości* procedur, narzędzi i interpretacji wyników egzaminów zewnętrznych na trzech szczeblach kształcenia ogólnego i zawodowego, a więc nie wyłączając matury,
6. że uda się *syntetyzować i popularyzować* wyniki badań (przez biuletyny, warsztaty, konferencje) w taki sposób, by diagnostyka edukacyjna zyskała w Polsce pozycję rozwiniętej, użytecznej nauki stosowanej.

Niech się tak stanie!

TAJNIKI WARSZTATU BADAWCZEGO

Grażyna Szyling

Zakład Diagnostyki Edukacyjnej
Uniwersytetu Gdańskiego

Trafność standardów wymagań egzaminacyjnych a jakość egzaminowania zewnętrznego¹²

Wprowadzenie

Standardy wymagań egzaminacyjnych określają wynik edukacji szkolnej, trudno więc przecenić ich znaczenie nie tylko dla procesu kształcenia. Z tego też powodu budzą ogromne emocje, czego dowodem jest chociażby szereg wystąpień i burzliwe dyskusje podczas IX Krajowej Konferencji z cyklu *Diagnostyka Edukacyjna* (2003).

Zawarte w standardach opisy oczekiwanych osiągnięć uczniów są także jakościowymi normami testów egzaminacyjnych (Niemierko, 2003), należy więc troszczyć się o ich szeroko rozumianą trafność. Jej podnoszeniu niewątpliwie mogą służyć doświadczenia nabywane w efekcie zastosowania standardów w praktyce, czyli podczas egzaminów zewnętrznych. Zapisem tych swoistych doświadczeń i związanych z nimi refleksji są dokumenty przygotowane na zlecenie Centralnej Komisji Egzaminacyjnej przez jej pracowników bądź ekspertów zewnętrznych.

Celem badania, którego wyniki są prezentowane w niniejszym wystąpieniu, było zrekonstruowanie kryteriów trafności standardów egzaminacyjnych przyjętych przez autorów tych dokumentów i spojrzenie na nie z perspektywy jakości egzaminowania zewnętrznego.

Założenia metodologiczne

Analizowane dokumenty różnią się od siebie: czasem powstania, formą, podejściem metodologicznym, stopniem uszczegółowienia, zakresem badania, doбором próby. Łączy je natomiast problem oceny standardów wymagań egzaminacyjnych dokonanej po ich zastosowaniu jako podstawy egzaminu gimnazjalnego.

Dążąc do zrekonstruowania w miarę jednolitego układu kategorii myślenia o standardach jako podstawowym kryterium jakości egzaminowania zewnętrznego, założono celową redukcję materiału pod kątem cech treściowych tekstu. Zaś jako podstawę porządkowania materiału zawartego w dokumentach przyjęto poszerzone pojęcie trafności S. Messicka (1989), przedstawione postaci macierzy postępującej aspektów trafności (Tabela 1.).

¹² Wystąpienie na konferencji *System egzaminowania zewnętrznego – perspektywy badawcze*, zorganizowanej przez Centralną Komisję Egzaminacyjną w Warszawie 29 października 2003 r.

Tabela 1. Macierz postępująca aspektów trafności S. Messicka zastosowana do standardów wymagań egzaminacyjnych

Funkcje oceny trafności Uzasadnienie Trafności	INTERPRETACJA	ZASTOSOWANIE
NA PODSTAWIE DOWODÓW	<p>TRAFNOŚĆ TEORETYCZNA</p> <p><i>Interpretacja potwierdzona empirycznymi dowodami</i></p> <p>Dlaczego ten, a nie inny zbiór umiejętności ucznia jest podstawą przeprowadzania egzaminu gimnazjalnego?</p> <p>Czy standardy gimnazjalne są dokumentem spójnym wewnątrznie i jednoznacznym pod względem pojęć i terminów?</p>	<p>TRAFNOŚĆ TEORETYCZNA + STOSOWNOŚĆ I UŻYTECZNOŚĆ</p> <p><i>Wykorzystanie standardów do jakiegoś celu</i></p> <p>Czy sformułowania standardów umożliwiają konstruktorom stosowne opracowanie kartoteki i zadań testu egzaminacyjnego?</p> <p>Czy wyniki egzaminu gimnazjalnego odniesione do standardów egzaminacyjnych pozwalają na dostarczenie uczniom użytecznej informacji o ich osiągnięciach?</p>
NA PODSTAWIE KONSEKWENCJI	<p>TRAFNOŚĆ TEORETYCZNA + STOSOWNOŚĆ I UŻYTECZNOŚĆ + KSZTAŁTOWANIE WARTOŚCI</p> <p><i>Oszacowanie wartości wywołanych interpretacją standardów</i></p>	<p>TRAFNOŚĆ TEORETYCZNA + STOSOWNOŚĆ I UŻYTECZNOŚĆ + KSZTAŁTOWANIE WARTOŚCI + SKUTKI SPOŁECZNE</p> <p><i>Oszacowanie konsekwencji społecznych</i></p>

Kluczowe dla teorii trafności pytanie brzmi, *w jakim stopniu – jeśli w ogóle – na podstawie dowodów i rozumowych uzasadnień [standardy wymagań] powinny być interpretowane i wykorzystywane w zaproponowany sposób* (Messick, 1989).

Cztery pola macierzy odpowiadają czterem podstawowym zagadnieniom trafności. W takim ujęciu trafność jest jednolitym pojęciem, którego podstawę stanowi trafność teoretyczna potwierdzona empirycznymi dowodami. Kolejne kategorie macierzy rozwijają ten fundamentalny aspekt o jego zewnętrzne zastosowania służące różnym celom.

Messick zakłada, że ustalanie trafności (walidacja) jest procesem polegającym na gromadzeniu i porządkowaniu danych empirycznych oraz argumentów (racjonalnych uzasadnień), traktowanych jako równoprawne. Co więcej, uznaje, że szacowanie trafności konsekwencyjnej (dolny wiersz tabeli) nie wymaga oparcia w jakimkolwiek zbiorze danych empirycznych, ale może być przewidywane na podstawie przesłanek.

W niniejszym opracowaniu uwaga skupiona zostanie na oszacowaniu trafności standardów gimnazjalnych na podstawie dowodów zawartych w dokumentach poddanych analizie. Tak więc interesować nas będzie głównie górny wiersz macierzy postępującej aspektów trafności, chociaż nie zostaną pominięte przewidywane konsekwencje spełniania bądź niespełniania przez standardy egzaminacyjne określonych kryteriów trafności.

Ważne wydaje się także przypomnienie, że punktem wyjścia rozważań w analizowanych dokumentach nie jest (jak zakłada Messick) interpretacja pojęcia standardów egzaminacyjnych (czyli ich znaczenie założone, uzasadnione teoretycznie), ale spojrzenie na trafność standardów z perspektywy ich zastosowania jako podstawy egzaminów zewnętrznych. To z kolei pociąga za sobą bardziej uytylitarne podejście do zagadnienia.

Zrekonstruowane w drodze analiz kryteria trafności dotyczą głównie części humanistycznej standardów gimnazjalnych. Trudno więc je uznać za zbiór wyczerpujący. Mogą jednak służyć różnym celom, mniej lub bardziej pragmatycznym. Są propozycją, którą można wykorzystać zarówno do wieloaspektowej oceny jakości standardów jak i do budowania modeli teoretycznych.

Trafność teoretyczna

Trafność teoretyczna związana jest z interpretacją pojęcia standardów egzaminu gimnazjalnego. Jej zasadnicze znaczenie wynika nie tylko z tego, że jasne określenie pojęcia (*konstruktu*) stanowi podstawę rozważań o kolejnych aspektach trafności. Od uzyskania odpowiedzi na pytanie: Dlaczego ten, a nie inny zbiór umiejętności ucznia ma być podstawą przeprowadzenia egzaminu zewnętrznego? – zależy uprawomocnienie (K. Konarzewski, 2003) testów budowanych w oparciu o standardy egzaminacyjne.

Autorzy analizowanych dokumentów, nastawiając się głównie na realizację celów użytecznych, nie rozważali, jakie *dyspozycje wchodzą w skład pojęcia wykształcenia* (tamże, s. 42), nie podejmowali też dyskursu na temat sensu teoretycznego pojęć, które są reprezentowane w standardach. Zagadnienia te bowiem, choć kluczowe dla trafności teoretycznej, z trudem poddają się ocenie dokonywanej z perspektywy zastosowania standardów, a ponadto wymagają uzasadnień uzgodnionych społecznie. (Taka szeroko zakrojona dyskusja społeczna miała miejsce przed wprowadzeniem ostatecznej wersji standardów.)

Z tych też powodów za decydującą o uprawomocnieniu praktycznym uznano w dokumentach zgodność standardów egzaminacyjnych z podstawą programową. To fundamentalne z pragmatycznego punktu widzenia założenie otwiera zestawienie kryteriów trafności teoretycznej prezentowane w Tabeli 2.

Tabela 2. Kryteria trafności teoretycznej standardów będących podstawą egzaminu gimnazjalnego

TRAFNOŚĆ TEORETYCZNA	KRYTERIA TRAFNOŚCI TEORETYCZNEJ STANDARDÓW będących podstawą egzaminu gimnazjalnego
Trafność dydaktyczna (Niemierko, 1999) Uprawomocnienie praktyczne (Konarzewski, 2003)	A. Standardy wymagań egzaminu gimnazjalnego obejmują umiejętności wpisane w Podstawę programową. Wybór tych umiejętności determinują następujące kryteria: A.1. Zgodność (spójność) z podstawą programową. A.2. Istotne znaczenie umiejętności w procesie kształcenia humanistycznego. A.3. Zgodność z koncepcją egzaminu gimnazjalnego: A.3.1. Zapewnienie ponadprzedmiotowości egzaminu. A.3.2. Łączenie wiadomości i umiejętności. A.3.3. Możliwość sprawdzenia pisemną formą egzaminu.
Trafność wewnętrzna	B. Standardy wymagań egzaminu gimnazjalnego są przyporządkowane obszarom umiejętności. Przyporządkowanie to określają następujące zasady: B.1. Obszary standardów określają najistotniejsze kompetencje uczniów gimnazjum. B.2. Przypisanie umiejętności uczniów do danych standardów jest uzasadnione i konsekwentne. B.3. Obszary standardów są rozłączne.

Trafność redakcyjna	<p>C. Standardy wymagań egzaminu gimnazjalnego są poprawnie sformułowane.</p> <p>Poprawność tę wyznaczają następujące kryteria:</p> <p>C.1. Stopień uszczegółowienia standardów pozwala na ich zrozumienie i jednoznaczną interpretację.</p> <p>C.2. Standardy są jednoznaczne pod względem pojęć i terminów, które są wykorzystane przy ich formułowaniu.</p> <p>C.3. Standardy są poprawnie sformułowane pod względem językowym i stylistycznym (proste składniowo).</p>
------------------------	---

By zapewnić wysoką jakość egzaminowania zewnętrznego, standardy powinny obejmować umiejętności najistotniejsze w procesie kształcenia humanistycznego i matematyczno-przyrodniczego, co bezpośrednio wiąże się także z założeniem ponadprzedmiotowości egzaminu gimnazjalnego. Koncepcja egzaminu narzuca także dwa pozostałe wyznaczniki uprawomocnienia praktycznego: dobór umiejętności możliwych do sprawdzenia pisemną formą egzaminu oraz zasadę łącznego sprawdzania wiadomości i umiejętności uczniów. To ostatnie kryterium silnie powiązane jest także z wysokimi kompetencjami twórców arkuszy egzaminacyjnych.

Warto zastanowić się, co może zagrażać trafności dydaktycznej standardów, tym samym obniżając jakość egzaminowania zewnętrznego. W badanych dokumentach pojawiają się przede wszystkim tendencje do uzupełnienia sprawdzanych egzaminem treści kształcenia o konkretne wymagania programowe z danego przedmiotu. Ich uwzględnienie prawdopodobnie przybliżyłoby egzamin do szkolnego nauczania, ale też mogłoby doprowadzić do ograniczenia jego ponadprzedmiotowego charakteru. Z drugiej strony pojawia się obawa, że zawężenie zakresu umiejętności objętych standardami może stymulować proces ich marginalizacji w procesie kształcenia w szkole.

Te ostatnie uwagi dotyczą nie tyle samej trafności dydaktycznej, ile przewidywanych konsekwencji jej obniżenia. Natomiast eksperci, których opinie znajdują się w analizowanych dokumentach, są skłonni - choć nie bez zastrzeżeń - uznać, że standardy wymagań obejmują umiejętności najważniejsze dla kształcenia humanistycznego i matematyczno-przyrodniczego w gimnazjum.

Ten aspekt nie wyczerpuje jednak obszaru rozważań o trafności teoretycznej standardów egzaminacyjnych. Dla jej uzasadnienia na podstawie dowodów ważna jest odpowiedź na pytanie, czy standardy gimnazjalne są dokumentem spójnym wewnątrz i jednoznacznym pod względem wykorzystanych pojęć i terminów? Zagadnienie to wyrasta z przekonania, że pojęcia reprezentowane pomiarem *powinny być starannie przedyskutowane, nim powstanie zbiór zadań je reprezentujący* (Lee Cronbach, za: Niemierko, 2003, s. 20).

Trafność wewnętrzna i redakcyjna standardów (patrz: dwa kolejne pola Tabeli nr 2) ma szczególne znaczenie dla twórców arkuszy egzaminacyjnych i konstruktorów zadań. Idealem w tym zakresie byłby *komplementarny charakter wydzielonych obszarów standardów, w których poszczególne standardy wpływają na siebie logicznie, ale równocześnie są rozłączne* (Raport CKE, s. 102). Wyraźnie widoczne jest także oczekiwanie, by stopień uszczegółowienia standardów oraz terminy wykorzystywane przy ich formułowaniu nie budziły wątpliwości interpretacyjnych, a przede wszystkim pozwalały precyzyjnie określić te umiejętności ucznia, których opanowanie mają sprawdzać egzaminy.

Jeśli zatem przyjmiemy jako dowód empiryczny krytyczne stwierdzenie prof. K. Konarzewskiego (2003, s. 43), że standardy sformułowane *zostały w sposób niewiążący, tak by mogły uprawomocnić każdy test osiągnięć*, to poddamy w wątpliwość ich trafność wewnętrzną. W tym przeświadczeniu utrzymują nas także uwagi pochodzące z innych źródeł,

dotyczące chociażby nakładania się na siebie zakresów znaczeniowych poszczególnych standardów. W wielu opiniach pojawia się spostrzeżenie, że np. standard: *Uczeń czyta teksty kultury [...] na poziomie dosłownym, przenośnym i symbolicznym* jest – ze względu na swą pojemność znaczeniową – rodzajem „nadstandardu”, obejmującego swym zakresem pozostałe standardy związane z czytaniem i odbiorem tekstów kultury.

Czy zatem, wskazując na ograniczenia trafności wewnętrznej i redakcyjnej standardów gimnazjalnych, podważamy tym samym także jakość egzaminowania zewnętrznego?

Odpowiedź na to pytanie wcale nie jest jednoznaczna. Trzeba bowiem pamiętać, że negatywne skutki niepełnej rozłączności standardów i ograniczonej precyzji ich zapisu mogą zostać zminimalizowane dzięki *starannej nad nimi dyskusji*, o potrzebie której mówił jeden z twórców teorii trafności, Lee Cronbach. Do takiej dyskusji dojrzeli już także ci użytkownicy standardów egzaminacyjnych, którzy najsilniej odczuwają skutki ich niepełnej trafności wewnętrznej i redakcyjnej, czyli konstruktorzy arkuszy egzaminacyjnych i osoby odpowiedzialne za komentowanie wyników egzaminów. Nie chcą się oni pogodzić z wieloznacznością standardów, która nadmiernie komplikuje ich pracę i nie ułatwia komunikacji. Z drugiej jednak strony zdają sobie sprawę, że pomysł rozpisania standardów na szereg prostych czynności może (choć nie musi) grozić nadmierną atomizacją, która z kolei będzie utrudniała syntetyczne interpretowanie osiągnięć ucznia. Wielu z nich za rozwiązanie najlepsze dla doraźnego podniesienia trafności wewnętrznej standardów egzaminacyjnych uznaje ustalenie przez sędziów kompetentnych takiej interpretacji standardów, która uwzględni ich wzajemny logiczny związek, a zarazem podniesie ich rozłączność. W tym celu należałoby także jednoznacznie zdefiniować pojęcia i terminy wykorzystane w sformułowaniach standardów, zwłaszcza że pewne składniki ich znaczenia mogą się różnić w zależności od przedmiotu, którego dotyczą, jak dzieje się na przykład z rozumieniem przez historyka i polonistę pojęcia *interpretacja tekstu*.

Wszystkie te działania, o których mowa powyżej, były zapewne podjęte także w procesie konstruowania standardów. Jednak opublikowany dokument, stanowiący podstawę egzaminów zewnętrznych, oderwał się od intencji swoich twórców i zaczął żyć własnym życiem, obrastając w nowe interpretacje, budząc wątpliwości i domagając się kolejnych uzasadnień trafności teoretycznej. Im wyższa bowiem trafność standardów, tym bardziej zasadne może być przekonanie, że oparty na nich egzamin nie tylko jest prawomocny, ale i *sprawdza to, co miał sprawdzać*.

Stosowność i użyteczność

Kolejnym aspektem rozszerzonego pojęcia trafności standardów gimnazjalnych jest ich stosowność i użyteczność. Przy jej szacowaniu rozważa się wykorzystanie standardów do jakiegoś celu, czyli odpowiada na pytanie, *czy ich użytkownicy czerpią korzyści z posługiwania się nimi* (Niemierko, maszynopis, 2003). To zagadnienie można rozpatrywać z perspektywy uczniów, nauczycieli, konstruktorów narzędzi egzaminacyjnych czy wreszcie innych zainteresowanych osób. Należy przy tym pamiętać, że trafność teoretyczna pozostaje ciągle obecnym punktem odniesienia rozważań o stosowności i użyteczności standardów.

Zrekonstruowane kryteria tego aspektu trafności przedstawia Tabela 3.

Tabela 3. Kryteria stosowności i użyteczności standardów będących podstawą egzaminu gimnazjalnego

STOSOWNOŚĆ I UŻYTECZNOŚĆ	KRYTERIA STOSOWNOŚCI I UŻYTECZNOŚCI STANDARDÓW będących podstawą egzaminu gimnazjalnego
Standardy jako podstawa konstruowania kartoteki testu egzaminacyjnego	<p>A. Standardy wyznaczają czynności ucznia uwzględniane w kartotece testu egzaminacyjnego</p> <p>A.1. Umiejętności ucznia określone w standardach są uwzględniane w kartotekach testów gimnazjalnych.</p> <p>A.2. Standardy umożliwiają dokonanie precyzyjnego i konkretnego zapisu (nazwania) czynności ucznia w kartotece testu.</p>
Standardy jako podstawa konstruowania zadań egzaminacyjnych	<p>B. Opanowanie przez ucznia umiejętności określonych w standardach gimnazjalnych można w sposób jednoznaczny sprawdzać za pomocą zadań egzaminacyjnych.</p> <p>B.1 Zakwalifikowanie zadania do standardu nie budzi wątpliwości.</p> <p>B.2. Wszystkie standardy gimnazjalne są sprawdzane za pomocą zadań egzaminacyjnych.</p> <p>B.3. Standardy umożliwiają konstruowanie zadań stosownych do określonych w nich umiejętności.</p>
Standardy jako podstawa do wnioskowania o osiągnięciach uczniów	<p>C. Wyniki egzaminów odniesione do standardów gimnazjalnych pozwalają na dostarczenie uczniom użytecznej informacji o ich osiągnięciach w badanym zakresie.</p> <p>C.1. Podział gimnazjalnych standardów obszary pozwala zasadnie wnioskować o osiągnięciach ucznia w zakresie umiejętności kluczowych.</p> <p>C.2. Wyniki testów odniesione do standardów gimnazjalnych pozwalają zasadnie wnioskować o umiejętnościach ucznia na koniec danego etapu kształcenia.</p>

Pierwsza grupa zrekonstruowanych w Tabeli 3. kryteriów stosowności i użyteczności standardów gimnazjalnych związana jest z oceną ich **przydatności do budowania kartoteki testu**, którą możemy traktować jako ogniwo pośredniczące między standardami a zadaniami egzaminacyjnymi. Z punktu widzenia konstruktora sposób sformułowania standardów powinien ułatwiać operacjonalizację czynności sprawdzanych testem. Na pewno zadanie to jest utrudnione z powodu wspomnianego już wcześniej nakładania się na siebie obszarów standardów.

Z drugiej jednak strony to poziom kompetencji konstruktora w dużej mierze decyduje o jakości tego swoistego „przełożenia” sformułowania danego standardu na nazwę czynności ucznia. Nie można zatem na podstawie analizy kartotek testów egzaminacyjnych wnioskować wprost o trafności standardów, zwłaszcza że brak dowodów empirycznych na to, że kartoteka rzeczywiście pełni funkcję ogniwa pośredniczącego między standardem a zadaniem go reprezentującym.

To zagadnienie łączy się bezpośrednio z rozważeniem użyteczności standardów jako **podstawy konstruowania zadań egzaminacyjnych**. Właściwie we wszystkich analizowanych dokumentach można znaleźć stwierdzenie, że zakwalifikowanie zadania do standardu bywa umowne, ale z reguły można je racjonalnie uzasadnić. Omawiane zjawisko, podobnie jak poprzednie, ma swoje źródło w niepełnej trafności wewnętrznej i redakcyjnej standardów. Nie można też polemizować z poglądem, że *proces „przekładu” standardów na język zadań testowych jest niejednoznaczny i wymaga trudnych decyzji interpretacyj-*

nych (Szyłko, 2003). To stwierdzenie ponownie odsyła nas nie tylko do standardów, ale i do kompetencji autorów zadań i arkuszy egzaminacyjnych jako podstawowych wyznaczników jakości egzaminowania zewnętrznego.

Wszystkie analizowane dokumenty rozpatrują zagadnienie *reprezentatywności standardów w arkuszach egzaminacyjnych*, jako jedno z kryteriów ich użyteczności. Dostyc powszechny jest pogląd, że *opanowanie standardów da się sprawdzać za pomocą zadań egzaminacyjnych*, ale nie wszystkimi typami zadań i nie z równą łatwością. Wśród standardów humanistycznych, które konstruktorzy uznali za trudne do *przełożenia* na zadania, znajduje się między innymi taki: *Uczeń odnajduje i interpretuje związki przyczynowo-skutkowe w rozwoju cywilizacyjnym Polski i świata – odnajduje i interpretuje związki przyczynowo-skutkowe w polityce, gospodarce, kulturze i życiu społecznym (I/5)*. Wydaje się on niezbyt stosowny nie tylko z punktu widzenia konstruktora. Jego opanowanie może przekraczać także możliwości przeciętnego ucznia gimnazjum. Obydwa te aspekty prawdopodobnie sprawiają, że rzadko jest reprezentowany zadaniami egzaminacyjnymi.

Czy można jednak kusić się o jednoznaczna ocenę użyteczności i stosowności standardów biorąc pod uwagę zestawienia liczbowe ilustrujące częstość ich wybierania przez autorów arkuszy egzaminacyjnych? Takie podejście wydaje się być nadużyciem. Należy bowiem pamiętać o omawianych wcześniej zjawiskach obniżających trafność teoretyczną, takich jak: różny zakres znaczeniowy poszczególnych standardów czy ich niepełna rozłączność. Te zakłócenia trafności wewnętrznej nie mogą nie znaleźć swojego odzwierciedlenia w nierównej liczbie zadań, którymi są sprawdzane poszczególne standardy podczas egzaminu.

Ta kwestia wymaga jednak dalszych badań, zwłaszcza że „przyzwyczajenie” konstruktorów do wykorzystywania określonych standardów (i związanych z nimi rodzajów zadań) może w konsekwencji doprowadzić do schematyzmu w budowaniu testów i *rutynowego postrzegania treści objętych badaniem*. To zaś z kolei może wywoływać niekorzystne zmiany w procesie kształcenia, zwłaszcza w szkołach, w których świadomie ogranicza się treść kształcenia i uczy „pod egzamin”.

Wszystkie wskazane uprzednio silne i słabe strony standardów gimnazjalnych znajdują swoje odzwierciedlenie w szacowaniu ich użyteczności do zasadnego **wnioskowania o osiągnięciach ucznia**. Jest to bardzo ważny aspekt trafności standardów, ponieważ *brak informacji zwrotnych o wyniku działania prawie zawsze upośledza działania następne* (Łukaszewski, s. 87). Im więcej w ocenianiu osiągnięć ucznia kryteriów incydentalnych, niejawnych i niezależnych od jego postawy, tym trudniej mu samodzielnie podejmować decyzje o procesie uczenia się.

Analizowane dokumenty wskazują, że wnioskowanie o osiągnięciach ucznia na podstawie wyników testów odniesionych do standardów egzaminacyjnych ma charakter bardzo ogólny i nie zawsze jednoznaczny, co związane jest z niepełną trafnością wewnętrzną standardów oraz uchybieniami w budowaniu kartotek testów.

W tym kontekście pod rozwagę można wziąć propozycję doskonalenia **operacjonalizacji standardów**, której celem byłoby doprowadzenie do jednoznacznej interpretacji tego, *co to znaczy opanować w całości (lub nie) takie czynności, jak: wyodrębnić, określić, przekształcić, odczytuje [...] (Raport CKE, 2003)*. Informacja dla ucznia zbudowana w oparciu o tak precyzyjne kryteria byłaby niewątpliwie szczegółowa, ale mogłaby pójść w kierunku rozbicia złożonych umiejętności na szereg prostych czynności.

Warto także pamiętać, że nie można zasadnie wnioskować o osiągnięciach ucznia określonych w danym standardzie, jeśli jest on reprezentowany w teście egzaminacyjnym jednym zadaniem zamkniętym lub zadaniem krótkiej odpowiedzi.

Mimo tych wszystkich ograniczeń, wielu badanych nauczycieli (66%) jest przekonanych, że standardy gimnazjalnych wymagań egzaminacyjnych *ułatwiają ocenę osiągnięć ucznia z perspektywy zewnętrznej czy umożliwiają porównywanie wyników uzyskanych w wewnątrzszkolnych badaniach osiągnięć* (Raport CKE, 2003). Takie deklaracje zdają się sugerować, że liczna grupa nauczycieli uznaje standardy za stosowne i użyteczne narzędzie, co implikuje także pozytywną ocenę jakości egzaminowania zewnętrznego.

O trafności konsekwencyjnej

Zamiast zakończenia

Bardzo niewiele wiemy jednak o tym, czy uczniowie rzeczywiście czerpią korzyści z odniesionych do standardów egzaminacyjnych informacji o swoich osiągnięciach. Podobnie nikt nie ma naszej wiedzy o postawach i wartościach, które u uczniów wywołało zastosowanie standardów jako podstawy egzaminu zewnętrznego. Bez zbadania tych aspektów trudno będzie zasadnie ocenić i trafność standardów, i jakość egzaminowania zewnętrznego.

Ważnym obszarem dociekań powinno stać się też poszukiwanie odpowiedzi na pytanie: *Czy standardy wymagań egzaminacyjnych wpłynęły korzystnie na jakość pracy szkoły?* Pierwsze próby udzielenia odpowiedzi zostały podjęte przez Pracownię Ewaluacji CKE w raporcie *Wykorzystywanie przez szkoły wyników sprawdzianu i egzaminu 2002*, ale to zbyt mało, jak na obszar problemów, które wiążą się z tym zagadnieniem. Warto przy tym pamiętać, że w szerokim odbiorze społecznym standardy wymagań, testy egzaminacyjne i wyniki osiągnięte przez uczniów są składnikami pewnej całości o nazwie „szkoła”, w której standardy nie muszą wcale odgrywać roli najważniejszej.

Samuel Messick wymagał, by rozważając zastosowanie jakiegoś pojęcia, oszacować tego wielorakie konsekwencje, do społecznych włącznie. My znajdujemy się w takim momencie, w którym skutki te możemy już badać i to nie tylko po to, by zgromadzić dowody przemawiające za lub przeciw trafności standardów egzaminacyjnych. Istotnym celem powinna stać się dyskusja nad podnoszeniem jakości standardów i zarazem jakości egzaminowania zewnętrznego, które stało się już trwałym elementem systemu edukacji, a także ważnym czynnikiem regulującym kształcenie szkolne.

Literatura

- K. Konarzewski, *Jak uprawiać badania oświatowe. Metodologia praktyczna*, WSiP, Warszawa 2000.
- W. Łukaszeński, *Zwrotne informacje o wyniku czynności*, [w:] Kurcz I., Kądziałowa D. (red.), *Psychologia czynności. Nowe perspektywy*, WN Scholar, Warszawa 2002.
- S. Messick, *Validity*, [w:] B. Niemierko (red.) *Pedagogika, pomiar dydaktyczny i ocenianie szkolne. Skrypt dla uczestników II PSED UG*, Tłumaczenie: Maciąg A. i Ludwikowska E., Gdańsk 1999, ss. 38-42.
- S. Messick, *Meaning and Values in Test Validation: The Science and Ethics of Assessment*, "Educational Researcher" 1989, Vol. 18, No. 2, ss. 5-17.
- B. Niemierko, *Ocenianie szkolne bez tajemnic*, WSiP, Warszawa 2002.
- B. Niemierko, *Kształcenie według wymagań*, [w:] K. Kruszewski (red.), *Pedagogika w pokoju nauczycielskim*, WSiP, Warszawa 2000.
- B. Niemierko, *Pomiar wyników kształcenia*, WSiP, Warszawa 1999.
- Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych, Red. B. Niemierko, WSHE w Łodzi, Łódź 2003. [w tym: K. Konarzewski, *Egzamin 2002: rzetelność i trafność testu matematyczno-przyrodniczego*; B. Niemierko, *Jaki pomiar dydaktyczny jest nam potrzebny?*].

Dokumenty poddane analizie

1. Pisma Okręgowych Komisji Egzaminacyjnych dotyczące propozycji weryfikacji standardów (OKE w Jarorzynie, OKE w Krakowie, OKE w Łomży, OKE w Łodzi, OKE w Poznaniu, OKE we Wrocławiu), 2002.
2. Herczyński J., Małecki W., *Standardy po pierwszej próbie (standardy wymagań egzaminacyjnych po przeprowadzeniu pierwszego sprawdzianu i egzaminu gimnazjalnego)*. Raport przygotowany na zlecenie

Centralnej Komisji Egzaminacyjnej – dokument poufny, nie przeznaczony do rozpowszechniania, listopad 2002.

3. *Standardy wymagań egzaminacyjnych. Raport z badania ewaluacyjnego.* Opracował zespół ekspertów Pracowni Ewaluacji pod kierunkiem Zofii Lisieckiej, Centralna Komisja Egzaminacyjna, Warszawa 2003.
4. Szyłko M., *Część humanistyczna egzaminu gimnazjalnego w świetle standardów. Realizacja standardów będących podstawą egzaminu gimnazjalnego w części humanistycznej,* Opracowanie przygotowane na zlecenie Centralnej Komisji Egzaminacyjnej, Gdańsk 2003.
5. Zestawienia: *Wykorzystanie standardów egzaminacyjnych w testach z egzaminów gimnazjalnych* (także próbnych) z lat 2002 i 2003 wykonane przez Wydz. Egzaminów dla Uczniów Gimnazjum CKE.

DONIESIENIA BADAWCZE**Przykłady przeprowadzonych w latach 2002 – 2003
badań dotyczących egzaminowania zewnętrznego¹³**

dr Jan Herczyński

Interdyscyplinarne Centrum Modelowania
Uniwersytetu Warszawskiego

Pierwsza odsłona***Społeczne i terytorialne zróżnicowanie wyników sprawdzianu szóstoklasistów
i egzaminu gimnazjalnego przeprowadzonych wiosną 2002 r.***

Podstawę wniosków stanowi tu analiza zagregowanych na poziomie gminy wyników odbytych wiosną 2002 roku sprawdzianu dla uczniów klasy szóstej oraz egzaminu gimnazjalnego. Głównym celem badawczym stało się opisanie podstawowych prawidłowości terytorialnego zróżnicowania wyników i określenie ich podłoża. Interpretacja dotyczy cech rozkładu wyników testów egzaminacyjnych w świetle dostępnych informacji o uczniach, sieci osadniczej i szkolnej, historycznych uwarunkowaniach rozwoju i bieżącej sytuacji społeczno-gospodarczej gmin. Dokonano bezpośredniego porównania umiejętności uczniów w całej Polsce w odniesieniu do zróżnicowań terytorialnych oraz poszukiwano cech charakterystycznych gminy wpływających (dodatnio bądź ujemnie) na wyniki testów. Analizy stwierdzonych empirycznych współzależności społecznych oraz historycznych dokonano przy użyciu narzędzi statystyczno-ekonometrycznych i map.

Więcej w: J. Herczyński, M. Herbst, *Pierwsza odsłona. Społeczne i terytorialne zróżnicowanie wyników sprawdzianu szóstoklasistów i egzaminu gimnazjalnego przeprowadzonych wiosną 2002 roku*. Raport przygotowany na zlecenie Fundacji Klub Obywatelski, Warszawa 2002.



dr Zbigniew Powązka, Joanna Czaplińska

Akademia Pedagogiczna w Krakowie

**Z badań nad przygotowaniem kandydatów do studiowania matematyki
w Akademii Pedagogicznej w Krakowie**

Badania nad zestawem zadań, który pozwoliłby ocenić kompetencje kandydatów do studiowania matematyki prowadzone są w Akademii Pedagogicznej od r. 2002. Przygotowując te zestawy kierowano się projektami standardów wymagań egzaminacyjnych, które docierały do uczelni z komisji egzaminacyjnej w Krakowie. W OKE znaleziono też partne-

¹³ Abstrakty wystąpień na konferencji CKE *System egzaminowania zewnętrznego – perspektywy badawcze*, Warszawa, październik 2003 r.

rów do dyskusji o jakości i trafności narzędzi badawczych. W wyniku zastosowania w r. 2003 jednego z przygotowanych narzędzi, uzyskano rozkład punktów za rozwiązanie zadań zbliżony do rozkładu normalnego. Analiza rozwiązań zadań ujawniła jednak szereg niedostatków w wiadomościach i umiejętnościach opisanych w standardach wymagań maturalnych. Są one związane m. in. z: wyznaczaniem dziedziny funkcji, punktów przecięcia wykresu funkcji z osiami układu współrzędnych, przekształcaniem i szkicowaniem wykresu funkcji, itd. Zaobserwowano również dużą nieporadność kandydatów związaną z posługiwaniem się definicjami i twierdzeniami oraz umiejętnością argumentowania.

Więcej w: J. Czaplińska, *Trudności w stosowaniu pojęć analitycznych przez absolwentów szkół średnich podczas rozwiązywania zadania matematycznego*, *Disputationes Scientifcae Universitatis Catholicae in Ružomberok*, Ružomberok 2003, roč 3, č. 3, ss. 3 -10.



Ada Szyłko

absolwentka Wydziału Psychologii
Uniwersytetu Gdańskiego

Analiza motywacji do działania na przykładzie aplikacji kandydatów na egzaminatorów w nowym systemie oceniania zewnętrznego

W prezentowanych badaniach podjęto próbę analizy motywacji leżącej u podstaw podjęcia decyzji o zaangażowaniu się w działanie przez osoby ubiegające się o nabycie konkretnego statusu społecznego: statusu egzaminatora w nowym, zewnętrznym systemie oceniania, wprowadzanym do systemu edukacji w związku z reformą oświaty mającą miejsce w Polsce. W toku analizy 1733 aplikacji nadesłanych przez kandydatów na egzaminatorów z dwóch województw (pomorskiego oraz kujawsko-pomorskiego) skategoryzowano ich sposób uzasadniania swojej decyzji. Aplikacje nadsyłane były przede wszystkim przez osoby zajmujące się wyłącznie dydaktyką w różnego typu szkołach. W próbie znalazły się też osoby pełniące funkcje kierownicze w placówkach oświatowych oraz jeszcze w inny sposób związane z edukacją. Oprócz analizy motywacji osób badanych dokonano też opisu socjodemograficznego próby.

Więcej w: A. Szyłko, *Analiza motywacji do działania na przykładzie aplikacji kandydatów na egzaminatorów w nowym systemie oceniania zewnętrznego w Polsce*. Projekt badawczy zrealizowany pod kierunkiem prof. dr hab. B. Wojciszke w Instytucie Psychologii Uniwersytetu Gdańskiego w okresie grudzień 2001 – kwiecień 2003, Gdańsk 2003, maszynopis.

Projekty badań systemu egzaminowania zewnętrznego¹⁴

Jacek Rysiewicz

Instytut Filologii Angielskiej
Uniwersytetu Adama Mickiewicza w Poznaniu

Rzetelność egzaminu maturalnego z języka angielskiego w jego obecnej formie w oparciu o porównanie wyników kandydatów na studia anglistyczne w Instytucie Filologii Angielskiej UAM w Poznaniu

Pomysł przeprowadzenia badania nasunął się autorowi pod wpływem kilku czynników. Są nimi:

1. obserwacje autora dotyczące poziomu przygotowania kandydatów na studia anglistyczne w IFA UAM w Poznaniu, poczynione podczas corocznych egzaminów wstępnych;
2. potrzeba stworzenia systemu corocznej analizy wyników egzaminów wstępnych, który, obejmując wszelkie możliwe dane dostępne o kandydacie, pozwoliłby na porównawcze badania zmieniających się tendencji w grupie kandydatów;
3. chęć i potrzeba zweryfikowania rzetelności ocen uzyskiwanych przez kandydatów na egzaminie maturalnym z języka angielskiego przed wprowadzeniem i po wprowadzeniu tzw. *Nowej Matury*;
4. przeświadczenie autora o szerokiej skali *patologii* dotychczasowego wewnętrznego oceniania egzaminów maturalnych.

Cele projektu

1. sprawdzenie, czy po wprowadzeniu nowego zewnętrznego systemu oceniania w 2005 r. wzrośnie rzetelność ocen z języka angielskiego uzyskiwanych na egzaminie maturalnym (zarówno pisemnym, jak i ustnym);
2. wykorzystanie uzyskanych danych porównawczych sprzed i po 2005 r. do opracowania przyszłego systemu rekrutacji na studia;
3. sporządzenie profilu przeciętnego kandydata przyjętego na studia anglistyczne w celu uzyskania informacji służących modyfikacji systemu kształcenia lub weryfikacji decyzji dotyczących wyboru kierunków studiów.

Realizacja badania

Do celów badawczych zgromadzono dotychczas dane z 4 kolejnych lat rekrutacji na filologię angielską UAM (lata 2001 – 2003). Przyjmuje się, że łącznie z rokiem 2004 próbę badawczą stanowić będzie szacunkowo 3000 kandydatów. Ponadto zakłada się, że po wprowadzeniu tzw. *Nowej Matury* w 2005 r. dane o kandydatach będą gromadzone przez dwa kolejne lata. Tak uzyskany materiał badawczy zostanie poddany analizie porównawczej.

Z dokumentów dostępnych instytutowej komisji rekrutacyjnej w UM systematycznie do bazy danych wprowadzane są następujące informacje o kandydatach: imię i nazwisko, wiek, płeć, szkoła (wraz z nazwą miejscowości), w której kandydat zdawał egzamin maturalny, oceny z pisemnego i ustnego egzaminu maturalnego z języka angielskiego, oceny z pisemnego i ustnego egzaminu wstępnego, ocena łączna z egzaminu wstępnego, informacja czy kandydat został przyjęty na studia.

¹⁴ zgłoszone przez różne podmioty badawcze podczas konferencji *System egzaminów zewnętrznych – perspektywy badawcze*

Hipotezy badawcze

Oczekuje się, że wprowadzenie zewnętrznego systemu oceniania zmniejszy dotychczasowy rozdzźwięk pomiędzy ocenami uzyskiwanymi na pisemnym i ustnym egzaminie maturalnym a ocenami uzyskiwanymi na egzaminie wstępnym (pisemnym i ustnym) przez tych samych kandydatów. Zakłada się więc, że zgodność (korelacja) pomiędzy ocenami uzyskiwanymi na maturze i na egzaminie wstępnym będzie większa.

Zakłada się też, że w wyniku wprowadzenia oceniania zewnętrznego abiturienti ze szkół zlokalizowanych w małych miejscowościach będą rzetelniej oceniani niż dotychczas. W tych placówkach bowiem zaobserwowano tendencję do zawyżania ocen na egzaminie dojrzałości. W rezultacie kandydaci na studia z tych ośrodków, mimo że na maturze uzyskali wysokie noty, na egzaminie wstępnym otrzymują oceny niedostateczne.



Wojciech Małecki

Dolnośląska Szkoła Wyższa Edukacji TWP
we Wrocławiu

Reforma systemu oceniania przez pryzmat osobistych teorii nauczycieli

Informacje ogólne	Realizatorzy	Bogusława Dorota Gołębnik, Wojciech Małecki
	Kierownik projektu	Bogusława Dorota Gołębnik
	Termin realizacji	grudzień 2004
	Zasięg terytorialny badania	województwo dolnośląskie
Obszar i przedmiot badań	Relacje między wprowadzaniem systemem egzaminów zewnętrznych a zmianą strategii oceniania w szkołach	
Problematyka	Egzaminy zewnętrzne, ocenianie arkuszy egzaminacyjnych, rozwój umiejętności egzaminatorów, przenoszenie kryteriów oceniania stosowanych w egzaminach zewnętrznych do szkół.	
Cele badania	Opisanie wymuszanej odgórnie reformy systemu oceniania w szkole polskiej przez pryzmat analiz narracji nauczycieli w różnym stopniu zaangażowanych w reformę	
Metodologia	Wywiady indywidualne (dotyczące biografii zawodowej), dyskusje grupowe, kwerenda i analiza dokumentów źródłowych	
Próba badawcza	Dobrana w sposób celowy (tzw. próbkowanie teoretyczne) - około 100 osób	
Forma dokumentacji	Transkrypcje wywiadów, analizy dokumentów i źródeł, raporty badawcze	
Sposoby upowszechniania wyników badań	Referaty na konferencjach, publikacje artykułów	

Wojciech Małecki
Dolnośląska Szkoła Wyższa Edukacji TWP
we Wrocławiu

Modyfikowanie szkolnych programów nauczania w gimnazjum

Informacje ogólne	Realizatorzy	Beata Krywuł-Szczudło, Wojciech Małecki
	Kierownik projektu	Wojciech Małecki
	Termin realizacji	listopad 2005 r.
	Zasięg terytorialny badania	Gimnazjum we Wrocławiu
Obszar i przedmiot badań	Możliwości korzystania z wyników egzaminów zewnętrznych do modyfikowania szkolnych programów nauczania – studium przypadku.	
Problematyka	Analizowanie i interpretowanie wyników egzaminów zewnętrznych, wnioskowanie o umiejętnościach uczniów na podstawie wyników egzaminów, modyfikowanie szkolnych programów nauczania	
Cele badania	Opracowanie procedur korzystania z wyników egzaminów zewnętrznych w gimnazjum	
Metodologia	Analiza wyników egzaminów zewnętrznych uczniów gimnazjum, analiza procedur i wyników oceniania wewnętrznego, obserwacje, wywiady indywidualne i grupowe	
Próba badawcza	Zespoły nauczycieli gimnazjum uczący przedmiotów „egzaminacyjnych”	
Forma dokumentacji	Szkolne raporty diagnoz i egzaminów zewnętrznych, transkrypcje wywiadów, protokoły obserwacji, opracowania dokumentów	
Sposoby upowszechniania wyników badań	Referaty na konferencjach, publikacje artykułów, przewodnik dla nauczycieli gimnazjum	

Dorobek badawczy Zakładu Diagnostyki Edukacyjnej Instytutu Pedagogiki Uniwersytetu Gdańskiego

W nowej strukturze Zakład funkcjonuje od 1 października 2003 r. Wcześniej kierujący nim prof. zw. dr hab. Bolesław Niemierko stał na czele Międzywydziałowego Studium Pedagogicznego. Ponieważ od wielu lat zainteresowania Profesora skupiają się wokół szeroko pojętej problematyki oceniania (zob. np. publikacje: *Między oceną szkolną a dydaktyką (bliżej dydaktyki)*, 1997, *Pomiar wyników kształcenia*, 1999, *Ocenianie szkolne bez tajemnic*, 2002, liczne publikacje związane z dziewięcioma konferencjami naukowymi *Diagnostyka edukacyjna*), Zakład realizuje zadania związane z kształceniem przyszłych i pracujących już zawodowo nauczycieli w zakresie pomiaru dydaktycznego, oceniania, egzaminowania, diagnozy edukacyjnej, ewaluacji w szkole, prowadząc zajęcia na studiach magisterskich i podyplomowych oraz organizując seminaria doktoranckie. Tematyka badań, analiz i opracowań podejmowanych w Zakładzie odpowiada potrzebom reformy systemu edukacji. Obejmuje ona m.in.: standardy wymagań egzaminacyjnych (hierarchie celów i materiału kształcenia, normowanie testów i normy podłużne, kształcenie według wymagań, jako rodzącą się teorię dydaktyczną), egzaminy szkolne (wewnętrzne i zewnętrzne, powszednie i doniosłe, ich pozytywny i negatywny efekt zwrotny), a wreszcie ocenianie szkolne, rozumiane jako ustalanie i komunikowanie oceny wyników uczenia się w wybranych nauczycielskich systemach kształcenia, z uwzględnieniem uczniów o specjalnych potrzebach edukacyjnych.

W wyniku wieloletniej pracy dydaktycznej Profesora, pod Jego kierunkiem powstało kilkadziesiąt prac naukowych. Poniżej prezentujemy sylwetki niektórych badaczy skupionych wokół seminarium prof. Niemierki oraz krótkie informacje o ich dorobku naukowym.

Jerzy Chodnicki

nauczyciel matematyki w gimnazjum w Piątnicy k. Łomży

Interesuje się problemami związanymi z ocenianiem osiągnięć edukacyjnych uczniów. Obecnie prowadzi badania związane z trafnością i rzetelnością modeli oceniania zadań otwartych stosowanych w egzaminowaniu powszednim i doniosłym.

dr Maria Greonwald

nauczyciel akademicki w Pracowni Dydaktyki Geografii i Biologii Wydziału Geografii, Biologii i Oceanologii Uniwersytetu Gdańskiego

Zajmuje się etyką oceniania. W kilku publikacjach konferencyjnych z cyklu *Diagnostyka edukacyjna* porusza kwestie związane z uczciwością egzaminacyjną uczniów, nauczycieli nadzorujących przebieg egzaminu oraz egzaminatorów, którzy sprawdzają prace egzaminacyjne.

dr Arleta Poręba-Konopczyńska

nauczyciel-konsultant w Pracowni Organizacji Doskonalenia Ośrodka Doskonalenia Nauczycieli w Zielonej Górze

Interesuje ją miejsce oceny w procesie dydaktycznym. W rozprawie doktorskiej zajęła się dydaktyczną funkcją testów zespołowych z wyposażeniem.

Piotr Skorupiński

nauczyciel języka polskiego w liceum ogólnokształcącym w Warszawie

Zajmuje się trafnością teoretyczną egzaminów. W swojej, będącej na ukończeniu, pracy doktorskiej bada trafność działań programowych związanych z tzw. *Nową Maturą* z języka polskiego.

dr Henryk Szaleniec

wicedyrektor Okręgowej Komisji Egzaminacyjnej w Krakowie

Interesuje się probabilistyczną teorią testu. Bada możliwości zastosowania modelu Rascha do analizy wyników egzaminu zewnętrznego.

Grażyna Szyling

asystent w Zakładzie Diagnostyki Edukacyjnej Uniwersytetu Gdańskiego

Bada drugi układ wymagań programowych, zwany także *szarą strefą oceniania szkolnego*. W kręgu jej zainteresowań znajdują się nauczycielskie strategie oceniania osiągnięć uczniów, którzy nie są w stanie spełnić podstawowych wymagań programowych. Jej badania wskazują na istnienie rozbieżności między zalecanymi procedurami a nauczycielskimi praktykami oceniania. Trudna do oszacowania wielkość kompromisów podejmowanych w tym obszarze – nierzadko wymuszana warunkami pracy szkoły – będzie coraz wyraźniej ujawniana przez egzaminy zewnętrzne.

s. Leokadia Wojciechowska

dyrektor Zespołu Szkół Salezjańskich w Suwałkach

Interesuje ją stosowanie założeń programowych koncepcji kształcenia wg wymagań. M. in. bada relacje pomiędzy wymaganiami programowymi a egzaminacyjnymi.

SŁOWNICZEK BADACZA EDUKACYJNEGO

Diagnostyka edukacyjna

Diagnoza edukacyjna wyłoniła się w świecie (Ingenkamp, 1975; Bloom, 1976) i w Polsce (Janowski, 1985; Niemierko, 1994) w procesie humanizacji szkoły.

Jako teoria naukowa, diagnostyka edukacyjna zajmuje się opisem i wyjaśnianiem sposobów rozpoznawania sytuacji emocjonalnej i poznawczej ucznia, w tym zwłaszcza zmian, jakie w nim zachodzą pod wpływem działania edukacyjnego. Jako dyscyplina stosowana rejestruje stan i wyjaśnia kontekst osiągnięć uczniów.

W praktyce diagnostyka edukacyjna najczęściej służy rozpoznawaniu jakości procesu wychowania i kształcenia. Ewaluacja efektywności wszelkich działań związanych z budowaniem jakości tego procesu, polegająca na rozpoznawaniu jego przebiegu, wyników i uwarunkowań, powinna obejmować znaczące sytuacje wynikające m.in. z warunków uczenia się, stosowania programów i metod nauczania czy kształtowania wymagań, uwzględniając szeroki kontekst psychologiczny, społeczny i pedagogiczny oddziaływań szkoły. Diagnoza pedagogiczna nie ogranicza się bowiem tylko do sprawdzania i oceniania stanu osiągnięć uczniów, lecz obejmuje także wyjaśnianie przyczyn (genezy) tego stanu i przewidywanie (prognozę) rozwoju ich osiągnięć. B. Niemierko proponuje planowanie diagnozy edukacyjnej z uwzględnieniem modelu uczenia się B. Blooma (p. rysunek 3. na str. 7. naszego *Biuletynu*).

Diagnozę edukacyjną prowadzi się z wykorzystaniem metod ilościowych i jakościowych, np. metody pomiaru dydaktycznego. Celem diagnostyki edukacyjnej jest rozpoznawanie uwarunkowań procesu kształcenia w skali indywidualnej, szkolnej, regionalnej, krajowej, a nawet międzynarodowej. Celem badaczy edukacyjnych może stać się więc zarówno poznawanie zależności między kontekstem, stanem początkowym i działaniami a wynikami kształcenia w różnych systemach dydaktycznych czy poznawanie wpływu postaw uczestników procesów pedagogicznych na przebieg tych procesów, jak i podnoszenie trafności oceniania szkolnego poprzez dostarczanie uczniom i nauczycielom informacji wyjaśniającej przebieg i wyniki procesu kształcenia.

Opracowanie Zofia Lisiecka (Pracownia Ewaluacji CKE) na podstawie:

B. Niemierko, *Diagnostyka dydaktyczno-wychowawcza*, [w:] *Encyklopedia pedagogiczna*, red. W. Pomykało, Warszawa 1993, Fundacja Innowacja

Perspektywy diagnostyki edukacyjnej, red. nauk. Bolesław Niemierko, Gdańsk 1998, Wyd. Uniwersytetu Gdańskiego

B. Niemierko, *Pomiar wyników kształcenia*, Warszawa 1999, WSiP

B. Niemierko, *Ocenianie szkolne bez tajemnic*, Warszawa 2002, WSiP

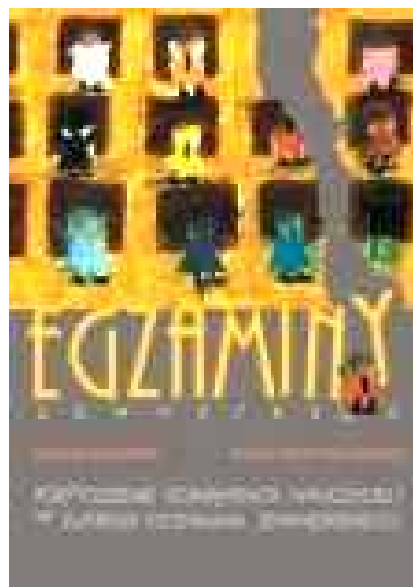
Encyklopedia pedagogiczna XXI wieku, red. T. Pilch, tom 1., Warszawa 2003, Wyd. Akademickie „Żak”

PRZECZYTAJ, ZAJRZYJ, WYKORZYSTAJ

Henryk Szaleniec, Maria K. Szmigel: *Egzaminy zewnętrzne. Podnoszenie kompetencji nauczycieli w zakresie oceniania zewnętrznego.* Wyd. „Zamiast Korepetycji”, Kraków 2001

Książka adresowana jest przede wszystkim do nauczycieli przygotowujących uczniów do egzaminów, dyrektorów szkół, konsultantów i doradców metodycznych, egzaminatorów i kandydatów na egzaminatorów. Jej autorzy starali się przede wszystkim odpowiedzieć na następujące pytania:

- Kim są pośredni i bezpośredni uczestnicy systemu egzaminowania i jakie są ich oczekiwania wobec egzaminów?
- Jak tworzyć koncepcje arkusza egzaminacyjnego i jak oceniać jakość zadań egzaminacyjnych i całego testu?
- Jak prowadzić analizę ilościową wyników egzaminów?
- W jaki sposób prezentować i komunikować wyniki egzaminów zewnętrznych?
- Jak wyniki egzaminów mogą wpłynąć na proces kształcenia?
- W jaki stopniu egzaminy zewnętrzne mogą być wykorzystane do ewaluacji w edukacji?



Opracowanie to może także odgrywać rolę pomocniczego podręcznika na kursach pomiaru dydaktycznego, ponieważ autorzy, oprócz teoretycznych rozważań na temat zastosowania teorii pomiaru dydaktycznego w egzaminowaniu zewnętrznym jak i procesie dydaktycznym, proponują także wiele praktycznych rozwiązań (w formie ćwiczeń) związanych np. z prezentowaniem uczniom i rodzicom wymagań egzaminacyjnych. W książce znaleźć można ponadto projekty szkoleń dla rad pedagogicznych czy zespołów samokształceniowych.

Jest to wartościowy podręcznik, dobrze wykorzystujący praktyczny dorobek egzaminacyjno-pomiarowy autorów, którzy na co dzień są pracownikami Okręgowej Komisji Egzaminacyjnej w Krakowie.

CZYTELNICY PISZA

Czekając na listy od Państwa, w których podzielicie się z Czytelnikami *Egzaminu* swoją wiedzą i doświadczeniami badawczymi, zamieszczamy dziś projekt przygotowanej w Pracowni Ewaluacji CKE karty informacyjnej o badaniach, eksperymentach, innowacjach pedagogicznych, etc. Interesują nas badania w każdej ich fazie – mile widziane zarówno plany badawcze, jak też opisy badań już wdrożonych czy zakończonych. Obszar badawczy powinien koncentrować się na kwestiach związanych z diagnozą edukacyjną, ocenianiem wewnątrzszkolnym i zewnętrznym, egzaminowaniem doniosłym, ewaluacją procesu nauczania-uczenia się, itp. Karta jest propozycją i oczywiście może być modyfikowana przez osoby ją wypełniające tak, by jak najpełniej odzwierciedlała projekt badawczy.

KARTA INFORMACYJNA O BADANIU NAUKOWYM / EKSPERYMENCIE / BADANIU EWALUACYJNYM / INNOWACJI / INNYM BADANIU

(niepotrzebne skreślić)

Temat/tytuł projektu badawczego:

Informacje ogólne	Zleceniodawca/y	
	Realizator/zy	
	Kierownik projektu / Osoba/y odpowiedzialna/e	
	Termin realizacji	
	Zasięg terytorialny badania Miejsce/a realizacji	
Obszar i przedmiot badań		
Problematyka		
Cele badania		
Zagadnienia badawcze Hipotezy Pytania ewaluacyjne		
Metodologia		
Próba badawcza		
Forma/y dokumentacji		
Sposoby upowszechnienia wyników badań		
Wnioski ogólne i rekomendacje wynikające z badań		
Recenzje i rekomendacje projektu badawczego		
Załączniki		

Sporządził: