

O JUBILEUSZU, JUBILATACH I TREŚCI TOMU

Dziesiąta Konferencja Diagnostyki Edukacyjnej wyróżnia się od poprzednich nie tylko „okrągłą” liczbą porządkową! Jest znamienna miejscem, którego nam użyła zasłużone Towarzystwo Salezjańskie, a także nadziejami, jakie budzi powołanie Polskiego Towarzystwa Diagnostyki Edukacyjnej i działalność w nim krakowskiej grupy założycielskiej. Kolorystykę konferencji wzbogaca Wieliczka, nasuwająca pamięć o beczce soli...

Przeszkód było dużo. Gdy ustąpiły **doktrynalne**, dające się we znaki w zaraniu naszej współpracy, pojawiły się **administracyjne** — skłonność władzy do skrywania i nadużywania informacji oraz **metodologiczne**, uzasadniające opór wobec testów. A jednak rosła grupa zwolenników pomiaru dydaktycznego. W jaki sposób? Nad tym warto chwilę się zastanowić.

Bardzo szybko okazało się, że wiedza pomiarowa jest chętnie przyjmowana przez dydaktyków przedmiotowych i to nie tylko nauk ścisłych. Wkrótce kursy i seminaria pomiarowe, a potem studia podyplomowe w Gdańsku, wysunęły się na czoło **diagnostyki edukacyjnej**, w tym sensie, w jakim ją dziś uprawiamy. Wiele osób uzyskało drugi zawód — „pomiarowca”, zdobyty na podstawie swej specjalności przedmiotowej, dając przykład nowoczesnej elastyczności wykształcenia. Duża, wciąż odmładzana kadra pionierów dziedziny stanowi główny dorobek ponad trzydziestu lat jej rozwoju w II i III Rzeczypospolitej.

Kiedyś były to seminaria teorii i konstrukcji testów osiągnięć szkolnych (łącznie ponad 50, w latach siedemdziesiątych i osiemdziesiątych ubiegłego wieku), a od jedenastu lat (w różnym rytmie — stąd kłopoty z numeracją) odbywają się krajowe konferencje diagnostyki edukacyjnej. Ich dorobek reprezentuje osobna antologia, „Diagnostyka edukacyjna. Teoria i praktyka”, przygotowywana równolegle przez dzielny ośrodek krakowski. Niniejszy tom natomiast zbiera materiały nadesłane na Dziesiątą Konferencję, której tematem są „Standardy wymagań i normy testowe w diagnostyce edukacyjnej”. Ten temat i jego rozwinięcie (w 22 punktach, rozsyłane wraz z zaproszeniami) zaowocowały oryginalnymi opracowaniami.

Pierwszą część tomu (IA i IB) wypełniają referaty plenarne i panelowo-plenarne, wnoszące problematykę teoretyczną i metodologię dla nas nową, ale często już znaną na Zachodzie, głównie w Stanach Zjednoczonych i w Anglii. Drugą część (IIA i IIB) stanowią doniesienia z sekcji „Standardy wymagań i skale pomiarowe” oraz z sekcji „Ocenianie osiągnięć i komunikowanie wyników oceniania”.

IA. METODOLOGIA DIAGNOSTYKI EDUKACYJNEJ

Tekst „Założone i uzyskane, sprawdzające i różnicujące znaczenie treściowe egzaminacyjnej skali pomiarowej” **Bolesława Niemierki** został pomyślany jako wprowadzenie do tematu konferencji. Zawiera przegląd najważniejszych skal pomiarowych, stosowanych w egzaminach szkolnych i w badaniach osiągnięć uczniów, prowadzący do wniosku, że najbardziej użyteczne są te, których pozycje mają znaczenie treściowe, co znaczy, że opisują jakość osiągnięć uczniów w kategoriach opanowanej treści kształcenia. Znaczenie treściowe uzyskane w pomiarze może odbiegać od założonego w standardach wymagań egzaminacyjnych i w planie testu, bowiem łatwe zadania (np.

z czytania) silniej wpływają na surowy wynik ucznia, a dobrze różnicujące zadania (np. z matematyki) lepiej wykazują zmiany jakości osiągnięć. W edukacji stale rośnie zainteresowanie zmianą, rozumianą jako postęp ucznia i szkoły – jako „wartość dodana” osiągnięć w procesie kształcenia. Przyszłość egzaminów szkolnych, zdaniem autora, stanowi skomputeryzowane testowanie adaptacyjne (STA), w którym sekwencja zadań jest tak dobierana do ucznia, by zoptymalizować pomiar jego postępów.

Referat gościa z Niemiec, **Petera Kroepe**, „Od życiowych doświadczeń do naukowych umiejętności. Pomiar wyników kształcenia wzdłuż ścieżki uczenia się”, oparty na danych empirycznych uzyskanych w Polsce (w Gorzowie Wielkopolskim), pokazuje jakościowe różnice wiedzy fizycznej w trzech etapach: 1) opis zjawisk w języku potocznym, 2) odtwarzanie opisu naukowego i 3) samodzielne stosowanie pojęć naukowych. Pomiar sprawdzający powinien pokazywać postęp osiągnięć ucznia wzdłuż tej ścieżki, przypominającej piązetańskie modele stadialnego rozwoju dziecka.

Roman Konarski, reprezentujący Uniwersytet Gdański i Sopotką Pracownię Badań Społecznych, w referacie „Model cechy latentnej w analizie psychometrycznej testów i pozycji testowych” przedstawia interpretację danych pomiarowych (na przykładzie przejawów dziecięcej agresji zaobserwowanej w Kanadzie) oparte na założeniu cechy ukrytej. Pokonanie trudności, jakie tekst sprawia niespecjalistom psychometrii, zaowocuje lepszym zrozumieniem podstaw analizy czynnikowej i teorii wyniku zadania testowego (IRT).

Starannie udokumentowane studium „Wybrane problemy rzetelności i trafności nowej matury” przedstawił **Roman Dolata** z Uniwersytetu Warszawskiego. Dotyczy ono głównie rzetelności, a właściwie — nierzetelności punktowania wypracowań maturalnych z języka polskiego. Efekt prosty egzaminatora oraz interakcja egzaminatora z wypracowaniem mogą sięgać połowy wariancji wyników tej części egzaminu (tab. 6), co jest wielkością alarmującą! Ponieważ przekracza ona wartości podawane w literaturze zagranicznej i krajowej (autor nie próbuje się do niej ustosunkować), należy poddać krytyce procedurę eksperymentu, co jednak nie osłabia siły wniosków: albo trzeba będzie korygować oceny egzaminatorów, albo (jak zaleca literatura, patrz „Pomiar wyników kształcenia”, s. 201n) zwiększyć liczbę sędziów. To doniosłe ostrzeżenie (Instytutu Spraw Publicznych) na rok przed „nową maturą”!

W starannie udokumentowanym eseju „Krytyka standardów wymagań egzaminacyjnych ze stanowiska radykalnego konstrukttywizmu” **Piotr Skorupiński** śledzi drogę, jaką te standardy powstawały i przeciwstawia jej postulaty współczesnych teorii nauk społecznych. Konstrukttywizm, określony jako „nowy paradygmat pedagogiczny” i niewątpliwie awansujący w dydaktyce, jest tu dobrze objaśniony i zinterpretowany, co podnosi wartość kształcącą tego pogłębianego studium.

W studium „Zarządzanie wiedzą w systemie egzaminów zewnętrznych — szkic do autodiagnozy” **Henryk Szaleniec** dowodzi pożytku zastosowania aparatu pojęciowego teorii zarządzania wiedzą (G. Probst i in. „Zarządzanie wiedzą w organizacji”, 2002) do informacji uzyskiwanej w systemie egzaminów zewnętrznych. Ten system rozrasta się, gdyż przybywają egzaminy potwierdzające kwalifikacje zawodowe i, co zwróci na niego powszechną uwagę, nowa matura. „Wiedza” jest najwyższą kategorią informacji, a „znaki” i „dane” są jej kategoriami podrzędnymi (rys. 1). Ponieważ liczba danych z corocznych egzaminów narasta lawinowo, zaniedbanie odpowiedniej wiedzy i właściwego zarządzania nią byłoby dużym błędem.

W kierunku przeciwnym psychometrii podąża sprawozdanie z monograficznego numeru czasopisma „Educational Measurement: Issues and Practice”, jakie przedstawia

Michał Daszkiewicz pod tytułem „Pierwsze kroki klasometrii”. Pod roboczą nazwą „klasometrii” ukrywają się pomysły pedagogów amerykańskich zmierzające do zbudowania teorii pomiaru przydatnej dla nauczyciela w klasie szkolnej. Nowe modele trafności, rzetelności i zastosowań pomiaru są tu rozważane swobodnie, jeszcze bez operacjonalizacji i weryfikacji, jako podejście różne od stosowanego w egzaminach zewnętrznych (patrz tab. 1) i w psychologii.

„Prosto z pieca” uzyskała swój materiał **Aleksandra Bartmańska**, bowiem z Anglii, gdzie – po smutnych doświadczeniach z „tabelami ligowymi” — po raz pierwszy spróbowano przedstawiania wartości dodanej wyniku egzaminu. Tytułem referatu jest „Wartość dodana w komunikowaniu wyników egzaminu a przeciwdziałanie niewłaściwej interpretacji wyników egzaminów zewnętrznych”, a jego słuszną troską — wzbogacenie informacji o wynikach egzaminów zewnętrznych miarami postępu ucznia i szkoły, tak by wkład pracujących w trudniejszych warunkach mógł być doceniony.

Na koniec mamy teorię cokolwiek fantazyjną, kosmologiczną i metafizyczną, w której szuka oparcia **Arleta Poręba-Konopczyńska**. Jej studium nosi nazwę „Miejsce teorii pomiaru dydaktycznego w naukach społecznych na podstawie przemyśleń Kena Wilbera”. Okazuje się, że w „teorii wszystkiego”, a dokładniej — w teorii „holonów”, zbiegają się kręte ścieżki współczesnej pedagogiki i może nam ona doradzić, jak pomiar dydaktyczny ma „wynurzyć się ze swego kąta kosmosu”! Takie teksty ubarwiają konferencje naukowe, a ich język ożywi luźniejsze dyskusje (na przykład w scenarii Wieliczki).

Wiele wystąpień i referatów tej konferencji związanych jest ze zmianami zewnętrznego systemu oceniania — nasi koledzy z Ukrainy są na początku tej drogi. Liliya Hrynevyh, w artykule „Reforma systemu egzaminacyjnego na Ukrainie — stan i problemy” dzieli się pierwszymi doświadczeniami z wprowadzania egzaminów zewnętrznych na Ukrainie.

IB. W STRONĘ UCZNIA

Tę tematykę otwiera tekst **s. Leokadii Wojciechowskiej** „System wychowawczy ks. Bosko jako podstawa wewnątrzszkolnego oceniania osiągnięć uczniów”. Patron — goszczącej nas w Krakowie — Federacji Szkół Salezjańskich, nazwany przez Ojca Świętego Jana Pawła II „fenomenem pedagogicznym”, wyrażał w XIX wieku poglądy zadziwiająco współczesne. Przedstawione są kompetentnie, a w komentarzach — żarliwie, co spaja podstawowe racje religijne z racjami ucznia i szkoły.

W obszernym opracowaniu „Uczniowie o Sprawdzianie 2004 i jego wyniku” **Maria Krystyna Szmigel** dokonała wyboru materiałów uzyskanych w ankiecie przeprowadzonej w losowej próbie systematycznej ponad 3800 uczniów z krakowskiego okręgu egzaminacyjnego. Wyłania się z nich sympatyczny obraz ucznia, walczącego ze stresem („Gdy zobaczyłem tytuł (chleb) zatkało mnie. Lecz pytania do testu były dość proste.”), niezdolnego do oszacowania uzyskanego wyniku (tab. 2), lecz dającego zadziwiająco dojrzałe rady swoim następcom. Aspiracje są zróżnicowane (zaznacza się silny wpływ wykształcenia rodziców), nadspodziewanie dużą rolę odgrywa czynnik zdrowia. Ten referat jest, mam nadzieję, zaczątkiem monografii egzaminowanego ucznia, bardzo potrzebnej adeptom diagnostyki edukacyjnej.

Równie cenną problematykę, a mianowicie — etyczną, kontynuuje na kolejnych konferencjach **Maria Groenwald**. W referacie „Meandry sprawiedliwego oceniania” przedstawia sześć wersji aksjomatu sprawiedliwości według Chaima Perelmana, zna-

komicie je ilustrując wypowiedziami uczniów. Konkluzja jest złożona, bo obok jasności zasad (deklarowanych w wewnętrznych i przedmiotowych systemach oceniania) zostawia duże pole kompetencji, uczciwości i dojrzałości (inteligencji) emocjonalnej nauczyciela. Także i tutaj widzę wyłanianie się monografii — podręcznika etyki egzaminacyjnej.

Grażyna Szyling konsekwentnie rozwija swoje badania nad ocenianiem szkolnym. We wzorowo uporządkowanej referacie „To wszystko kwestia strategii? Szara strefa oceniania w oczach uczniów” metoda epizodów (Susan Brookhart) została uzupełniona o wypowiedzi 100 uczniów oceniających decyzje nauczycieli. Uczniowskie poczucie sprawiedliwości okazuje się dość rygorystyczne w kwestii wymagań, co jednak nie wyklucza ulg dla ciężko pracujących. Potępiane są natomiast takie przejawy „dobrego serca” wobec słabego ucznia, które zniechęcają innych uczniów do wysiłku.

Komunikowaniem wyników sprawdzania zajęli się **Krzysztof Bednarek** i **Cezary Lempa** w artykule zatytułowanym „W drodze do skutecznego komentarza”. Podłożem ich analiz były doświadczenia zdobyte „w trakcie całorocznej pracy z Radami Pedagogicznymi, nauczycielami, doradcami metodycznymi i nadzorem pedagogicznym”. Metodą „metaplanu” (systematyzacji burzy mózgów) zbudowali kwestionariusz ankiety korespondencyjnej, której najistotniejszym wynikiem jest potwierdzenie znaczenia uczniowskiego planu rozwoju, zapewniającego sprzężenie zwrotne między wynikami oceniania a uczeniem się.

Problematykę dydaktycznej roli wymagań programowo-egzaminacyjnych podjęła **Małgorzata Lipska** w referacie „Pokaż uczniowi cel, czyli o wybranych aspektach planowania pracy nauczyciela w kształceniu według wymagań”. Chce, by doceniano rolę planowania kierunkowego i planowania wynikowego. Jej postulaty są mocno zbieżne z dorobkiem Jere Brophy’ego („Motywowanie uczniów do nauki”, wydane w Polsce przed dwoma laty), choć ich źródła są rodzime: własne doświadczenia i krajowe lektury.

IIA. STANDARDY I SKALE

Materiały tej sekcji otwiera artykuł **Elżbiety Kowalik** „Standardy edukacyjne w kształceniu chemicznym”, ilustrujący spiralną konstrukcję programu kształcenia ogólnego w tym przedmiocie na przykładzie hierarchii kilku teorii atomu. Taki układ jest naturalną podstawą wielostopniowej budowy testu. Opracowanie **Janusza Mula-wy**, **Jolanty Sokółowskiej** i **Dariusza Sroki** pod prowokacyjnym tytułem „Testowanie w szkole: zmora czy metoda?” prezentuje wykorzystanie banku danych Instytutu Badań Kompetencji w Wałbrzychu do sporządzania staninowego profilu osiągnięć ucznia i *programów rozwoju* (nazwanych, trochę deprymująco, „programami naprawczymi”; zmieniłbym też „wskaźnik zgodności” na *iloraz osiągnięć klasy*), stałość parametrów zadań kotwiczących oraz porównanie osiągnięć klasy VIII z II klasą gimnazjum, nastawiające nas krytycznie do reformy edukacji (tu należałoby sprawdzić, czy motywacja gimnazjalistów do pracy nad testem była równie wielka jak motywacja absolwentów dawnej szkoły podstawowej). Pomysłowość, pozytywna fascynacja statystyką, nauczycielski, żywy język — oto niewątpliwe zalety tej kolejnej publikacji Instytutu.

Henryk Palkij w tekście „Wpływ standardów wymagań z historii i wiedzy o społeczeństwie na kształcenie przedmiotowe”, opisuje i dokumentuje ewolucję kształcenia przedmiotowego pod wpływem wprowadzenia standardów wymagań egzaminacyjnych, korzystnie ocenia rolę standardów, wprowadza nowy wątek — spowodowane

zmiany podręcznikowe. **Maria Gawrysiuk i Małgorzata Szurlej-Gielen** są autorkami „Refleksji nad rozumieniem maturalnych standardów wymagań egzaminacyjnych z języków obcych nowożytnych”, w której dokładnie przedstawiły Europejski System Opisu Kształcenia Językowego, ilustrowany przykładowymi zadaniami testowymi w różnej formie. **Ewa Chorąży** zastanawia się „Jaką rolę mogą odegrać standardy wymagań egzaminacyjnych z historii w kształceniu według wymagań”, a jej rozważania prowadzą do standardów podłużnych. Takimi właśnie standardami zajmuje się **Anna Dubiecka** w doniesieniu „O jednolitości standardów wymagań egzaminacyjnych”. Pozostałe doniesienia w tym zakresie tematyki są dziełem **Bożeny Śniadek, Teresy Sadoń-Osowieckiej, Tomasza Maciejewskiego i Marka Zapieraczyńskiego**.

Zagadnienia konstrukcji i jakościowej interpretacji skal pomiarowych zilustrowała danymi egzaminacyjnymi **Maria Krystyna Szmigel** w doniesieniu „Treściowe znaczenie wyniku egzaminacyjnego ucznia i szkoły na przykładzie sprawdzianu w 2004 roku”, uzupełniającym wiele wątków referatu otwierającego konferencję. Tabele i charakterystyki uczniów zawarte w tym bogatym tekście pokazują prawidłowości warte dalszych systematycznych analiz. Zastosowanie teorii wyniku zadania (IRT) do eliminacji zadań obciążonych błędem kulturowym znajdzie Czytelnik w opracowaniu **Barbary Ciżkowicz** „Stronniczość pozycji testowych” na przykładzie — co ciekawe — skali wyuczony bezradności. Na przykładach wybranych zadań sprawdzianu osiągnięć po szkole podstawowej **Maria Sobczak** sygnalizuje możliwość „Wykorzystania wskaźnika łatwości w badaniach podłużnych”, zbliżając się w swych analizach do problematyki zrównywania wyników egzaminów zewnętrznych w kolejnych latach. O skali staninowej piszą **Barbara Hankus i Anna Stelmach**, a stosowną problematykę kształcenia zawodowego przedstawiają, w osobnych opracowaniach, **Janusz Figurski i Mariola Przekopiak** oraz **Izabela Górecka i Witold Woźniak**.

IIB. OCENIANIE I KOMUNIKOWANIE OCENY

Ocenianie szkolne to wciąż zapalny temat. W polemicznie (wobec działań administracyjnych w tym zakresie) nastawionym, krótkim tekście „Zmierzyć wychowanie” **Czesław Ficner** chce „włożyć kij w mrowisko” problematyki oceniania osiągnięć ucznia i szkoły. **Siostra Leokadia Wojciechowska** jest autorką studium „Po co nauczycielom i uczniom przedmiotowy system oceniania — refleksje nauczycielki matematyki w szkole zawodowej”, w którym zawarła spostrzeżenia o żmudnym, lecz dla wychowawcy — radosnym, procesie budowania systemu. Natomiast **Małgorzata Jagodzińska** w doniesieniu „Wewnątrzszkolne systemy oceniania — założenia i rzeczywistość szkolna” stara się wydobyć racje z opinii nauczycieli na ten temat i pokazać nieuchronną złożoność realnych systemów. Autorami pozostałych tekstów na temat oceniania szkolnego są **Bożena Żuk, Teresa Wejner, Danuta Hyżak, Elżbieta i Marek Jasiński, Maria Sarnecka, Ewa Czaja**.

Komunikowanie wyniku egzaminu i oceny szkolnej jest tematem ostatniej, bardzo ważnej dydaktycznie, serii doniesień. Otwiera ją tekst **Barbary Przychodzeń i Marii Kutajczyk** „Komunikowanie wyników egzaminu na przykładzie próbnych egzaminów gimnazjalnych organizowanych przez Okręgową Komisję Egzaminacyjną w Gdańsku”. To solidne sprawozdanie pokazuje szeroki zakres i różnorodność kontaktów OKE z udziałowcami edukacji, wymagającymi swoistej informacji o treści, formie i wynikach egzaminów zewnętrznych. **Bożena Kotarska-Lewandowska** przedstawia „Komunikowanie wyników egzaminu wstępnego na wydziały architektury”, rozciąga-

jąc — na tym przykładzie — postulaty pomiarowo-diagnostyczne ze szkół niższego szczebla na szkołę wyższą. „Ryzyko błędu w pracy egzaminatora” **Bogumiły Frączak** jest krótkim zbiorem warsztatowych obserwacji i wniosków na temat punktowania zadań testowych. Podobną tematykę podjęły **Małgorzata Boba** i **Maria Michłowicz** w swym studium „Egzaminator części humanistycznej egzaminu gimnazjalnego wobec schematu punktowania i odpowiedzi ucznia”. Materiały tej sekcji dopełniają opracowania **Ireny Marii Nyckowskiej**, **Janiny Różanowskiej** i **Jerzego Paczkowskiego**.

III. PRACOWNIA METODY EKWICENTYLOWEJ

Osobne miejsce w wydawnictwie — ze względu na inną rolę — zajmuje roboczy raport **Bolesława Niemierki** „Zrównywanie wyników Sprawdzianu 2004 do wyników Sprawdziany 2003”. Raport pokazuje, że trudniejszy sprawdzian tegoroczny nie może być interpretowany jako wskaźnik obniżenia osiągnięć absolwentów szkoły podstawowej w skali kraju (rys. 9). Z powodu dużego zainteresowania metodą zastosowaną w tym raporcie w toku konferencji odbędą się ćwiczenia w kalibrowaniu testów (bezpośredniego przenoszenia wyników jednej wersji na inną wersję), jako przygotowanie do procedur zrównywania wyników.

* * *

Konferencji towarzyszy wystawa zorganizowana przez Pedagogiczną Bibliotekę Wojewódzką w Krakowie na temat egzaminów szkolnych. Możemy na niej znaleźć pierwsze wzmianki na temat „examen maturitatis” w szkołach polskich. Zebrane i wyeksponowane przez **Weronikę Smoleń** dokumenty, książki, patenty uczniowskie pozwalają uczestnikom konferencji wsiąść do wehikułu czasu i pożegłować do bardziej odległych zakątków historii egzaminów. Proszę wczuć się w zapisy ustawy zatytułowanej „Wewnętrzne urządzenie szkół departamentowych 1812 roku”. (**J. Lipiński** „Wewnętrzne urządzenie szkół departamentowych”, 1812), gdzie w paragrafie 41 czytamy: *Nikt do Uniwersytetu przyjętym nie będzie, kto całego kursu nauk w szkole Departamentowej nie ukończy, niżej opisanego egzaminu przed najwyższą Magistraturą edukacyjną miejscową nie odprawi, i nie otrzyma patentu, czyli zaświadczenia maturitatis od teyże Magistratury, od Rektora i Profesorów examinujących podpisanego.* Czy jest jasne, dlaczego spotkaliśmy się w Krakowie?

* * *

Jako autor tego wstępu (z wyjątkiem akapitu pomiędzy gwiazdkami, który zredagował **H. Szaleniec**), zdaję sobie sprawę, że o powodzeniu konferencji decydują nie tylko, a bywa — że nie głównie, nadesłane teksty, lecz wyteżona, drobiazgowo praca organizacyjna. Pragnę więc podziękować Komitetowi Organizacyjnemu działającemu pod kierunkiem **Henryka Szaleńca** za niezwykle dzielne pokonywanie rozmaitych przeszkód. **Siostrze Leokadii Wojciechowskiej** i **Federacji Szkół Salezjańskich** dziękuję za przyjęcie nas do siebie w Krakowie. Niech się mnoży to, co sprzyja edukacji młodzieży!

Bolesław Niemierko