prof. dr hab. **Bolesław Niemierko**
Międzywydziałowe Studium Pedagogiczne
Uniwersytet Gdański

DIAGNOSTYKA EDUKACYJNA
Cele, metody i szanse rozwoju

Mimo zwołania kilku konferencji i zredagowania dwóch książek z „diagnostyką edukacyjną" w tytule[1] mam poczucie, że tu, w Legnicy, zaczynam coś nowego. Jest tak dlatego, że nazwa Ośrodka i pomysł obecnej konferencji mają charakter oryginalnej inicjatywy lokalnej, rzucającej zdrowe ziarno na wyjałowioną glebę krajową.

Przy takim założeniu, moje wystąpienie, jako pierwsze, musi zawierać przegląd dziedzin, modeli i praktycznych osiągnięć diagnostyki edukacyjnej w Polsce. Dokonam go w ograniczonym zakresie, licząc bardzo na pozostałych, starannie dobranych referentów. Będzie to nadto przegląd dość stronniczy, ześrodkowany na dydaktyce i zabarwiony własnymi pomysłami autora, bo czyż warto byłoby organizować konferencje, gdyby referenci ograniczali się do przedstawienia wiedzy ogólnej i szeroko dostępnej?

Humanizacja pedagogiki a diagnoza

Jesteśmy w Polsce świadkami i uczestnikami silnego nurtu humanizacji pedagogiki. Minął już, na szczęście, czas *autorytarnej* pedagogiki instrumentalnej, produkującej „budowniczych socjalizmu". Współcześnie na szerokim świecie walczą o lepsze dwie wielkie szkoły: *humanistyczna*, skoncentrowana na potrzebach i samorozwoju ucznia, w tym głównie na przeżywaniu wartości, oraz *technologiczna*, wskazująca jak wykorzystać teorię i narzędzia komunikacji międzyludzkiej do podniesienia skuteczności kształcenia.

Stosunki między przedstawicielami tych dwu szkół nie są najlepsze, zwłaszcza humaniści nie czują się dobrze w towarzystwie technologów i dają temu wyraz, ale w dobie zachęt do pojednania zwróćmy raczej uwagę na „komplementarność", czyli wzajemne uzupełnianie się tych szkół.

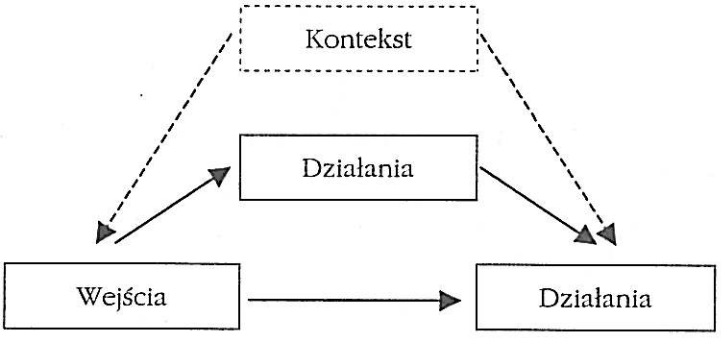
Otóż obydwie grupy pragną usamodzielnić ucznia w pewien sposób: przez wyposażenie go w wartości duchowe lub przez uzbrojenie w strategie osiągania celów. Obydwie walczą - z nie-pełnym sukcesem - z autorytarnymi reliktami w pedagogice: z przymuszaniem do uznawania przebrzmiałych wartości poprzednich pokoleń i ze zbyt silnym wpływem charyzmatycznych nauczycieli, a także nihilizmu moralnego kultury masowej i środków jej przekazu.

Koncentracja na uczniu wywołuje głębokie zainteresowanie jego rozwojem i podnosi rangę *diagnozy edukacyjnej*: ze swego rodzaju „materiałoznawstwa" dydaktycznego do rozpoznawania głębokich wartości, jakimi kieruje się partner w procesie edukacyjnym.

Nie sposób przecenić znaczenia humanizacji pedagogiki dla rozwoju diagnostyki edukacyjnej. Wybór słusznego kierunku jednak jej nie wystarcza. Potrzeba odpowiedniej technologii, by rozwinąć program uzyskiwania pożytecznej informacji.

Diagnoza edukacyjna a uczenie się

Najprostszy i najczęściej spotykany w literaturze metodologicznej model działania nauczycielskiego systemu kształcenia ma postać następującą:



Rys. 1. Składniki systemu dydaktycznego

Składnikami systemu dydaktycznego są:

1. *Wejścia systemu*, to jest uzdolnienia, zainteresowania, aspiracje, wiadomości i umiejętności młodzieży oraz ich nauczycieli, a także właściwości wcześniejszych systemów dydaktycznych, jakie poznali uczniowie.
2. *Działania systemowe*, to jest te czynności uczniów, nauczycieli i administracji szkolnej, które zmierzają do osiągnięcia celów kształcenia.
3. *Wyjścia systemu*, to jest wiadomości, umiejętności, postawy wobec uczenia się, wartości i cechy osobowości uczniów oraz osiągnięcia metodyczno-organizacyjne nauczycieli i szkół, w tym także nowe kompetencje nauczycieli.
4. *Kontekst systemu*, to jest okoliczności, które nie podlegają zmianie w toku działań, jak stałe urządzenia szkoły, sieć szkolnictwa i jego struktura administracyjna, sposób finansowania i pozycja społeczna edukacji, przekonania rodziców i pracodawców, rynek pracy itp.

Zauważmy, że wyjścia systemu są równolegle kształtowane przez wszystkie pozostałe jego składniki: wejścia, działania i kontekst systemu. Żaden z nich nie może być więc uznany za nieistotny dla wyników kształcenia.

Diagnoza edukacyjna jest rozpoznaniem przebiegu uczenia się. Obejmuje wszystkie cztery grupy zmiennych, przedstawione na rys. 1, ze szczególnym uwzględnieniem kontekstu i jego oddziaływania na poszczególne fazy procesu dydaktycznego: wejścia, działania i wyjścia. Diagnosta stara się wykryć, jakie warunki uczenia się wpływają na przebieg i wyniki

uczenia się, a jego celem w tym zakresie jest uświadomienie uczniom i nauczycielom odpowiednich zależności po, to by skuteczniej szukać drogi do spełnienia wymagań programowych.

Możemy to zapisać następująco:

Cel pierwszy diagnostyki edukacyjnej:	dostarczanie uczniom i nauczycielom informacji wyjaśniającej przebieg i wyniki uczenia się.

Przykłady warunków uczenia się, które mogą być określone metodami diagnostycznymi są następujące:

a. wysokie/niskie aspiracje do wykształcenia w rodzinie dziecka,
b. wysokie/niskie aspiracje własne dziecka do wykształcenia,
c. dobra/słaba umiejętność rozumienia cicho czytanego tekstu określonego typu,
d. dobre/słabe rozumienie tego, co nauczyciel mówi podczas lekcji,
e. opanowanie/brak nawyków samodzielnej pracy z książką,
f. wysokie/niskie zainteresowanie przedmiotem,
g. wysoka/niska umiejętność zapamiętywania wiadomości,
h. wysoka/niska umiejętność rozwiązywania problemów,
i. wysoka/niska pozycja w grupie klasowej,
j. dodatnia/ujemna postawa wobec przyrody i społeczeństwa.

Diagnoza edukacyjna a ocenianie osiągnięć uczniów

Przyjmiemy tutaj, że *ocena szkolna* jest informacją o wyniku kształcenia wraz z komentarzem, a *ocenianie osiągnięć uczniów* jest ustalaniem i komunikowaniem oceny szkolnej. Komentowanie informacji o wyniku kształcenia i komunikowanie oceny szkolnej są - w zwykłych, to jest codziennych, warunkach - umiejętnościami tak samo ważnymi i trudnymi, jak pozostałe składniki oceny i oceniania. Będąc nauczycielem, trzeba specjalnie się ich uczyć.[2]

Komentarz do informacji o wyniku kształcenia dotyczy zwykle:

a. warunków uczenia się,
b. sposobu uzyskania informacji o wyniku oraz poprawnej interpretacji wyniku,
c. sposobu wykorzystania tej informacji w toku dalszego uczenia się.[3]

Wszystkie te składniki komentarza muszą być w pewnej mierze oparte na diagnozie edukacyjnej, a najsilniej i bezpośrednio - pierwszy z nich, podstawowy. Dzięki rozpoznaniu warunków (kontekstu) kształcenia możemy:

1. uogólnić wynik sprawdzania osiągnięć na sytuacje podobne do tej, w której uczeń opanował wybraną treść kształcenia, a także wskazać drogę do wyższych osiągnięć,
2. zdać sobie i uczniowi sprawę z względności uzyskanego wyniku,

3. zmodyfikować informację o wyniku kształcenia,
4. uczynić warunki uczenia się podstawą oceny szkolnej,
5. zastąpić ocenę szkolną diagnozą warunków uczenia się.

Te sposoby wykorzystania diagnozy nie są dla dydaktyka równorzędne. Najważniejsze dla niego jest *znaczenie treściowe*, to jest programowe, oceny szkolnej. Według znaczenia treściowego i roli diagnozy można zbudować następującą skalę wartości oceny szkolnej:

Tabela 1. Skala wartości oceny szkolnej

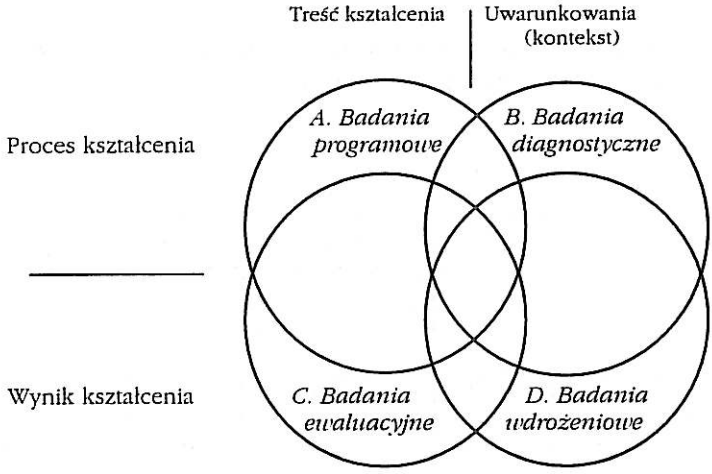
Znaczenie treściowe	*Typ osiągnięć*	*Rola diagnozy*	*Typ komentarza*
Bardzo wysokie	Wynik pomiaru dydaktycznego sprawdzającego wybrany zakres treści kształcenia	Generalizuje wynik kształcenia	*Twój wynik pozwala przypuszczać, że...*
Wysokie	Wykonanie pojedynczego zadania dobranego dla danego ucznia z wybranego zakresu treści	Relatywizuje wynik kształcenia	*Za wykonania tego trudnego (łatwego) zadania...*
Umiarkowane	Spełnienie wymagań skorygowane oceną postępu, motywacji lub zachowania ucznia	Modyfikuje informację o wyniku kształcenia	*Stawiam ci czwórkę, ponieważ ty...*
Niskie	Intuicyjna ocena „całego ucznia" uzupełniona wynikiem sprawdzania osiągnięć lub obserwacją zachowania	Ogranicza informację o wyniku kształcenia	*Jesteś, jak zwykle, dobrze (źle)...*
Bardzo niskie	Stałe lub chwilowe zadowolenie lub niezadowolenie oceniającego z zachowania ucznia	Zastępuje informację o wyniku kształcenia	*Ja dobrze wiem, że ty zawsze...*

Jak widać, informacja o warunkach uczenia się może być wykorzystana zarówno do wzbogacenia oceny szkolnej, jak i do osłabienia jej znaczenia treściowego i podważenia sensu. Posługując się zhumanizowanym, rozszerzonym pojęciem „trafności", obejmującym zarówno zgodność procedury z założeniami, jak i długofalowe skutki jej zastosowania[4], możemy powiedzieć, że chodzi nam o *trafność oceniania szkolnego*. Drugi z celów diagnozy edukacyjnej jest zatem następujący:

Cel drugi diagnostyki edukacyjnej: podnoszenie trafności oceniania szkolnego.

Diagnostyka edukacyjna jako rodzaj badań dydaktycznych

W kilku niedawnych publikacjach przedstawiłem następującą klasyfikację stosowanych badań dydaktycznych:



Rys. 2. Uproszczona typologia badań dydaktycznych

Rys. 2 pokazuje, że badania diagnostyczne różnią się od programowych i ewaluacyjnych akcentowaniem uwarunkowań kształcenia, a nie jego treści, a od badań ewaluacyjnych i wdrożeniowych - akcentowaniem procesu, a nie wyniku kształcenia. *Badania diagnostyczne* zostały tu zatem określone jako systematyczne rejestrowanie i wyjaśnianie wpływu kontekstu materialnego i społecznego na przebieg uczenia się, dokonywane niezależnie lub jako istotny składnik badań programowych, ewaluacyjnych lub wdrożeniowych.

Gdy chodzi o metody badań diagnostyki edukacyjnej, to za oryginalny uważam nie tyle ich dobór, ile *układ preferencyjny*, czyli porządek co do znaczenia. Cztery metody są, kolejno, najbardziej zalecane:

1. **Obserwacja bezpośrednia**. Teza o konieczności przyjrzenia się zjawiskom, o których się wypowiada i wnioskuje, ma w przypadku diagnostyki edukacyjnej pełne uzasadnienie. Obserwacja jest najlepszym sposobem gromadzenia hipotez, które mogą być potem sprawdzane bardziej obiektywnymi metodami. W dobie humanizacji pedagogiki rozsądnie regulowana obecność badacza podczas zajęć szkolnych i w szerzej pojmowanym środowisku edukacyjnym zyskuje na znaczeniu.

2. **Analiza dokumentów**. Chodzi o wszelkie przedmioty materialne, nie tylko pisane, które mogą być świadectwem wybranych procesów pedagogicznych, jak wyposażenie ucznia i pracowni szkolnych, stan posiadania rodziny, a także zeszyty ucznia, dzienniki lekcyjne, protokoły zebrań rady pedagogicznej, arkusze ocen. Dokumenty te mogą, choć nie muszą, rejestrować fakty, które uczestnicy procesów pedagogicznych potrafią w pewien sposób zinterpretować.

3. **Metody kwestionariuszowe**. Ankieta i wywiad są ogromnie nadużywane w badaniach dydaktycznych i najbardziej powierzchownie interpretowane. Najczęściej zadaje się uczniom i nauczycielom pojedyncze pytania o fakty i opinie, a następnie porównuje się odpowiedzi ze zdroworozsądkowym modelem zjawisk. Rzadko tworzy się skale ocen i próbuje analiz wielozmiennowych, docierających do zmiennych ukrytych.[5] Nawet jednak w najskromniejszej metodologicznie wersji, ankieta i wywiad mogą dostarczyć nam informacji niezwykle cennej: *obrazu* procesów pedagogicznych w oczach objętych nimi uczniów i nauczycieli.

4. **Pomiar dydaktyczny**. Bez pomiaru dydaktycznego, przynoszącego zobiektywizowaną informację o osiągnięciach uczniów, diagnoza edukacyjna traci naturalne kryterium trafności działania szkoły i łatwo popada w - niekoniecznie miłe - złudzenia. Z kolei pojedynczy pomiar nie mówi nam bezpośrednio niczego o kontekście osiągnięć uczniów. Testy mogą mieć jednak charakter diagnostyczny, gdy zawierają zadania sięgające czynnościowych podstaw potrzebnych wiadomości i umiejętności. Diagnoza tego rodzaju nie jest na ogół dokładna, ale w badaniach masowych może być użyteczna.

Dziedzictwo międzynarodowych badań osiągnięć uczniów w Polsce

W latach 1968 - 1992 Polska uczestniczyła w kilku badaniach Międzynarodowego Stowarzyszenia Badań Osiągnięć Pedagogicznych jako jeden z dwu najodważniejszych krajów Europy Środkowej. Angielska nazwa stowarzyszenia - *International Association for the Evaluation of Educational Achievement* (*IEA*) - oraz model metodologiczny prowadzonych badań[6] dowodzą, że - mimo zapewnień o unikaniu „wyścigów konnych między krajami" - miały one charakter głównie ewaluacyjny, ale składnik diagnostyczny był tam także obecny i to w bardzo rozbudowanym zakresie.

Dla przykładu przedstawię listę zmiennych kontekstowych Drugiego Studium Matematyki IEA[7] (tab. 2)

Jak widać, liczba zmiennych kontekstowych rośnie w miarę „schodzenia" ze szczebla kraju, przez szczeble lokalny i szkolny, na szczeble nauczycielski i uczniowski kształcenia matematycznego. Większość z tych zmiennych wykazuje pewien związek z osiągnięciami uczniów w różnych krajach, ale ten związek jest słaby, zawikłany i mało podobny między krajami. Nic więc dziwnego, że autorzy dzieła, z którego zaczerpnięto dane do tab. 2, w jego zakończeniu napisali, iż „obraz, jaki ukazał się dotychczas, jest złożony i wieloaspektowy. Mimo iż stwierdzono wiele istotnych podobieństw i różnic w poglądach nauczycieli i uczniów z różnych środowisk „nie udało się w tym sprawozdaniu wznieść się w analizach tych podobieństw i różnic ponad odnotowanie ich istnienia i, w paru przypadkach, próbę badania możliwych związków z osiągnięciami uczniów. Ogromna większość dalszej pracy musi być wykonana wewnątrz systemów dydaktycznych, a także międzynarodowo, jeśli pełne korzyści z udziału w tych badaniach mają być uzyskane."

Tabela 2. Skrócony wykaz zmiennych kontekstowych międzynarodowych badań osiągnięć matematycznych

Szczebel	*Zmienna*	*Przykładowy wskaźnik*
Krajowy	Stopień centralizacji	Instancja decydująca o programach i standardach
	Elastyczność	Reformy szkolne w ostatnim ćwierćwieczu
Lokalny	Typ regionu	Zatrudnienie w przemyśle i rolnictwie
	Typ szkoły	Selekcyjność wstępu do szkoły
	Język kształcenia	Procent uczniów z domów innojęzycznych
	Środowisko lokalne	Średni poziom wykształcenia rodziców w regionie
Szkolny	Koedukacja	Procent uczniów według płci w szkole
	Sprawność szkoły	Drugoroczność spowodowana przez dany przedmiot
	Wielkość szkoły	Liczba uczniów
	Kadra pedagogiczna	Stosunek liczby uczniów do liczby nauczycieli
	Kierowanie szkołą	Częstość narad nauczycieli przedmiotu w danej szkole
	Znaczenie przedmiotu	Tygodniowa liczba godzin nauczania przedmiotu
Nauczycielski	Wykształcenie nauczyciela	Studia przedmiotowe i pedagogiczne w latach/godzinach
	Doświadczenie pedagogiczne	Staż pracy pedagogicznej w latach
	Obciążenie dydaktyczne	Pensum godzinowe nauczyciela
	Wielkość oddziału szkolnego	Średnia liczba uczniów w klasie
	Metody kształcenia	Procent zajęć zbiorowych, grupowych, indywidualnych
	Indywidualizacja kształcenia	Procent prac zróżnicowanych co do trudności
	Wyobrażenie przedmiotu	Skale *Matematyka jako proces, Użyteczność matematyki*
	Postawa wobec przedmiotu	Skale: *Ważne / nieważne, Łatwe / trudne, Lubię / nie lubię*
Uczniowski	Wiek uczniów	Średni wiek ucznia w latach
	Płeć uczniów	Procent uczniów według płci w oddziale szkolnym
	Zawód rodziców	Od pracy niewykwalifikowanej do profesjonalnej
	Wykształcenie rodziców	Brak wykształcenia / podstawowe / średnie / wyższe
	Opinia rodziców o przedmiocie	Skala *Moi rodzice uważają, że matematyka ...*
	Aspiracje edukacyjne	Przewidywana liczba lat dalszego kształcenia
	Praca domowa	Średnia liczba godzin pracy domowej tygodniowo
	Korepetycje	Procent korzystających z pomocy korepetytorów
	Wyobrażenie przedmiotu	Skale: *Matematyka jako proces, Użyteczność matematyki*
	Postawa wobec przedmiotu	Skale: *Ważne / nieważne, Łatwe / trudne, Lubię / nie lubię*
	Motywacja uczenia się	Skala *Matematyka i ja*

Znajdujemy się zatem na początku długiej drogi naukowego poznania systemów dydaktycznych, obejmującego ich główne składniki oraz zależności między nimi (por. rys. 1). Dotyczy to zarówno wielkich systemów - krajowych i regionalnych (skala *makro*), jak też i małych systemów - szkolnych i nauczycielskich (skala *mikro*). Żmudne podążanie tą drogą stanowi trzeci z celów diagnostyki edukacyjnej:

Cel trzeci diagnostyki edukacyjnej: poznawanie zależności między kontekstem, wejściami i działaniami a wynikami kształcenia w różnych systemach dydaktycznych.

To, co stanowi wielki dorobek międzynarodowych i ogólnopolskich badań osiągnięć uczniów z lat osiemdziesiątych i dziewięćdziesiątych, to liczne wydawnictwa metodyczne i stosunkowo liczne kadry przedmiotowych specjalistów pomiaru sprawdzającego. Zyskaliśmy znacznie więcej w zakresie metodologii pomiaru niż diagnozy systemu oświatowego.[8] Ani Ministerstwo Edukacji Narodowej, ani Instytut Badań Edukacyjnych w Warszawie nie wykazały większego zainteresowania systemem powtarzalnych badań osiągnięć uczniów w Polsce.

Dlaczego diagnostyka edukacyjna nie rozwija się w Polsce?

Na to pytanie próbowano odpowiedzieć na konferencji naukowej w Elblągu, którą można uznać za poprzedniczkę naszego spotkania. Oto wypowiedzi trzech znakomitych pedagogów:

Krzysztof Kruszewski w błyskotliwym eseju *Test wiadomości dobrego i złego*[9] potwierdza: „Rzeczywiście, masowe badania wyników nauczania w Polsce przeprowadzono kilkakrotnie, zwykle w ramach badań międzynarodowych, przy czynnym, choć niekonsekwentnym oporze władz, co tłumaczy ich nikłą liczbę oraz brak na nie środków, i przy małym zainteresowaniu środowiska pedagogicznego, co tłumaczy ich mały rezonans." Dalej wymienia następujące przyczyny nikłego zainteresowania pedagogów tymi badaniami:

1. niedojrzałość a zarazem hermetyczność empirycznej wersji dydaktyki,
2. konserwatyzm pedagogów wyrażający się lekceważeniem wyników badań,
3. zbyt wielka - dla niespecjalisty - trudność pojęciowa pomiaru dydaktycznego,
4. brak odpowiedniej propagandy, to jest „sprzedania" wyników badań.

Heliodor Muszyński w obszernej rozprawie *Diagnoza edukacyjna i jej miejsce w praktyce i teorii pedagogicznej*[10] zauważa, że „decydent edukacyjny ma wiele powodów, aby wyniki edukacyjnych diagnoz uważać za niebezpieczne, a więc niepożądane, lub też aby być zainteresowanym

wynikami wyłącznie pomyślnymi dla siebie, aby odnosić się do nich z nieufnością i rezerwą, a wreszcie, aby traktować je w sposób upraszczający, pozorujący jedynie ich uwzględnianie w swoich decyzjach." Zdaniem Muszyńskiego, „uznajemy za niezbędne możliwie dokładne diagnozy dotyczące naszego stanu zdrowia, zamożności, wydajności pracy, spożycia czy przestrzegania prawa. Jedynie w edukacji informacje takie uważane są dość często za zbędne." Przyczyną tego jest niechęć do zmiany oświatowej, która byłaby oparta na racjonalnych, nie zaś politycznych podstawach.

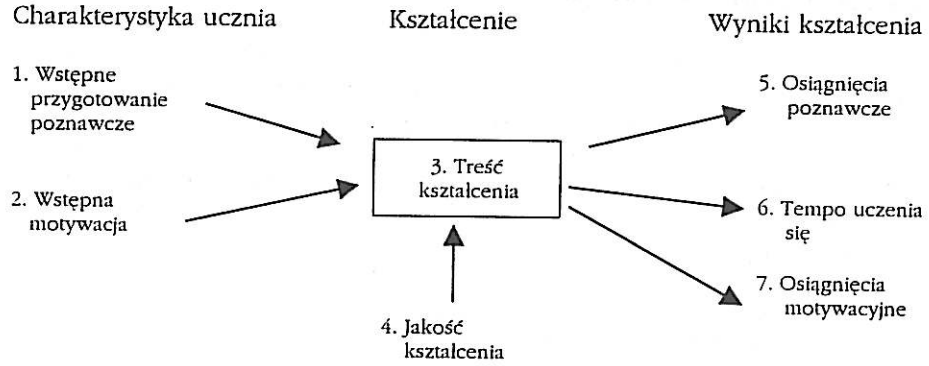
Tadeusz Lewowicki w referacie *Współczesne wyzwania wobec diagnostyki edukacyjnej - przemiany teleologii oświatowej i pojmowania pedagogiki a szanse rozwoju diagnostyki* [11] formułuje postulat diagnozowania sfery emocjonalnej ucznia. Pisze: „dawniejsze spostrzeganie głównych funkcji, celów i zadań oświaty (przede wszystkim oświaty szkolnej) uzasadniało koncentrację poczynań diagnostycznych na pomiarach wiedzy, diagnozowaniu wiadomości i umiejętności. (...) Dzisiejsze widzenie funkcji i celów edukacji prowadzi (prowadzić po-winno?) do prób lepszego diagnozowania rozwoju osobowości i sfery wartości." Diagnostyka edukacyjna, zdaniem Lewowickiego, „z powodzeniem może być subdyscypliną nie tylko technologiczną, ale również dyscypliną *rozumiejącą*, ogólnopedagogiczną, humanistyczną." Nalega na taką reorientację, stwierdzając na koniec: „Poszerzenie obszaru zainteresowań i zastosowań diagnostyki jest szansą zarówno dla rozwoju tej subdyscypliny, jak i ulepszenia oświaty. I szansy tej nie wolno zmarnować."

O diagnostykę procesów motywacyjnych

W związku z szybko postępującą humanizacją pedagogiki zalecenia Tadeusza Lewowickiego wydają się zyskiwać na aktualności. Jego uwagi o przewadze zainteresowania dziedziną poznawczą są trafne. Na przykład znany niemiecki podręcznik diagnostyki edukacyjnej [12] poświęca 72 strony teorii pomiaru dydaktycznego, kolejne 60 stron - testom osiągnięć szkolnych, 44 strony - inteligencji i zdolnościom specjalnym, a tylko 27 stron „diagnostyce właściwości społecznych i emocjonalnych" - dynamice małej grupy społecznej (socjometrii), motywacji uczenia się i lękowi przed egzaminami. Można więc uznać, że dziedziny pozapoznawcze są reprezentowane w tym podręczniku śladowo.

Dokładnie przeciwne założenia przyjął Andrzej Janowski, autor najlepszego, moim zdaniem, i najczęściej wznawianego polskiego podręcznika diagnostyki edukacyjnej, zatytułowanego *Poznawanie uczniów.* [13] Wśród „dziedzin rozpoznawania", jakimi się zajął, nie ma osiągnięć poznawczych, a wśród „metod zdobywania informacji" nie ma pomiaru dydaktycznego. W ujęciu Janowskiego sprawdzanie i ocenianie wyników kształcenia nie należą do diagnostyki.

Istnieją próby równoległego traktowania dziedzin poznawczej i motywacyjnej. Ich przykładem jest Beniamina Blooma *model uczenia się.*[14] Ma on postać następującą:



Rys. 3. Blooma model uczenia się jako model diagnozy edukacyjnej

Bloom był jednym z doradców Międzynarodowego Stowarzyszenia Badań Osiągnięć Pedagogicznych, a jego model był w tych badaniach stosowany, ale bez większego powodzenia. Potraktowanie postaw wobec poszczególnych przedmiotów jako *zmiennych zależnych*, to jest jako wyników kształcenia[15], dało obraz zależności tak pogmatwany, że pożytek okazał się niewielki. Uczniowie deklarowali swój stosunek do przyrody, społeczeństwa, wiedzy naukowej, przedmiotu i szkoły (por. tab. 2) niekonsekwentnie wobec pozostałych zmiennych.

W Polsce pewien dorobek w tym zakresie wniosły prace doktorskie napisane na moim seminarium przez Grażynę Penkowską, Barbarę Ciżkowicz i Rytę Suską-Wróbel.[16] Dowiodły one, że motywacja raczej maleje w procesie uczenia się przedmiotów matematyczno-przyrod-niczych, co może być spowodowane kumulowaniem się trudności merytorycznych. Dziedzina zainteresowania zamienia się w dziedzinę ciężkiej pracy. Dla większości uczniów i studentów osiągnięcia motywacyjne (różnica między wartościami zmiennych 7 i 2 na rys. 3) są ujemne! Ponieważ osiągnięcia poznawcze (różnica między wartościami zmiennych 5 i 1) niemal zawsze, nawet u najsłabszych uczniów i studentów, rosną, równoległe traktowanie tych dwu dziedzin jest bezpłodne.

Istnieją także głębsze powody odrębnego traktowania celów poznawczych i celów emocjonalnych kształcenia. Należę do tych pedagogów, którzy sądzą, że *nie powinniśmy wymagać* od uczniów określonych zainteresowań, aspiracji, wartości i postaw, aczkolwiek powinniśmy tworzyć sytuacje pedagogiczne kształtujące te właściwości. Nie wymagać, to znaczy nie poddawać egzaminom i nie oceniać formalnie. W tych zakresach liczy się tylko samodzielna, wewnętrzna decyzja wychowanka, nie zaś dostosowanie do wymagań szkoły i umiejętność zademonstrowania odpowiednich zachowań. Takie założenie zasadniczo odróżnia diagnozę procesów emocjonalnych od ewaluacji poznawczych wyników kształcenia.

Rola diagnozy postaw uczniów

W ustroju demokratycznym uczeń i nauczyciel mają prawo do autonomicznych postaw wobec procesów pedagogicznych, w których uczestniczą, a w miarę postępów demokratyzacji w Polsce coraz śmielej te postawy wyrażają. Jako przykład przytoczę wybrane opinie uczniów i nauczycieli o sprawdzianie kompetencji w Wałbrzychu, Słupsku, Pile i Elblągu. Opinie te dotyczą egzaminu opartego na rozwiązywaniu zadań wyboru wielokrotnego.

Tabela 3. Wybrane opinie uczniów w Wałbrzychu o sprawdzianie kompetencji (n = 1054)[17]

Wybrane pozycje kwestionariusza B. Niemierki *Test kompetencji szkolnych. Opinia o egzaminie*	Procentowy rozkład odpowiedzi					Średni wskaźnik aprobaty	
	zdec	raczej	nie mam zdania	raczej	zdec	uczniowie	nauczyciele
	(+2)	(+1)	(0)	(-1)	(-2)		
Każdy uczeń dobrze wiedział, co ma robić	42	45	3	7	1	1,2	1,6
Obyło się bez podpowiadania	30	24	12	21	13	0,4	0,2
Zadania były interesujące dla uczniów	16	38	19	17	10	0,3	0,4
Zadania były związane z życiem uczniów 6	6	16	20	31	27	-0,6	0,0
Test kompetencji pokazał uczniom, co umieją	32	39	10	13	6	0,8	0,4
Test kompetencji zachęcił uczniów do nauki	17	26	24	20	13	0,1	0,1

Opinie zestawione w tab. 3 są mocno rozproszone, ale ogólne tendencje są czytelne. Uczniowie na ogół doceniają jasność instrukcji testowania i wartość poznawczą wyników, mają wszakże poważne wątpliwości co do uczciwości swojej pracy nad testem, atrakcyjności intelektualnej zadań i motywacyjnego znaczenia sprawdzianu, zaś wyraźnie ujemnie oceniają „związek zadań z życiem", co musi rozczarować konstruktorów, którzy dwoją się i troją, by zadania ubarwić przygodami dzieci i ozdobić rysunkami o tematyce ornitologiczno-szkolnej... Nauczyciele są bardziej przekonani o samodzielności pracy uczniów, a mniej o trafności testu („pokazał, co umieją"). Niemal tak samo jak uczniowie wątpią w emocjonalno-motywacyjną wartość tej formy egzaminu.

Średni wskaźnik aprobaty uczniów dla trzydziestu podobnych twierdzeń wyniósł w Wałbrzychu 0,61, w Słupsku 0,51, w Elblągu 0,43, w Pile 0,41. Te różnice mogą być spowodowane znacznie większym obyciem z podobnymi zadaniami w byłych województwach wałbrzyskim i pilskim. Nauczyciele aprobują sprawdzian nieco wyżej (ogólna średnia wynosi około 0,6) niż uczniowie, a dyrektorzy szkół - wyraźnie wyżej (średnia około 0,9).

Porównanie testowego egzaminu „wałbrzyskiego" z próbną nową maturą wypada dla niego niekorzystnie. Wskaźniki aprobaty sześciu twierdzeń z tab. 3 przez nauczycieli wynoszą dla tej matury kolejno: 1,6; 1,2; 1,0; 0,3; 1,2; 1,2, są więc (z wyjątkiem pierwszego) wyraźnie wyższe,

zwłaszcza w zakresie motywacyjnym. Ogólna ocena próbnej matury przez nauczycieli wyraża się wskaźnikiem aprobaty około 0,9.

Ocena sprawdzianu kompetencji nie była ograniczona do obliczania wskaźników ilościowych. Wszyscy respondenci byli zachęcani do otwartych uwag na ten temat. Oto charakterystyczne wypowiedzi związane z twierdzeniami zebranymi w tab. 3:[18]

Ławki były za blisko i można było odpatrywać od innych (uczeń, Piła)
Wielu uczniów podpowiadało sobie nawzajem, co bardzo przeszkadzało (uczeń, Piła)
Nauczyciele w komisji rozmawiali i nie można było się nieraz skupić (uczeń, Piła)
Nie ściągaliśmy i był rygor (uczeń, Elbląg)
Korzystali uczniowie sprytni, zaradni, bez kompleksów (nauczyciel, Piła)
Niestety, w ściąganiu i podpowiadaniu uczniowie nie widzą niczego nagannego. Dla nich jest to zwyczajna „koleżeńska pomoc" (nauczyciel, Piła)
Uczniowie przejęci, skupieni (nauczyciel, Słupsk)
Duża część uczniów nie rozwiązywała zadań, lecz zakreślała odpowiedzi na „chybił trafił" (nauczyciel, Piła)
Na pewno nie był nudny jak poprzednia forma egzaminów! (uczeń, Piła)
Moim zdaniem te zadania były nieodpowiednie. Jeżeli mają być takie pytania, to po co uczymy się w szkole podstawowej aż osiem lat (zupełnie o czym innym)? My nie uczymy się myśleć, a jedynie wkuwamy na pamięć! Zróbcie coś! (uczeń, Słupsk)
Podoba mi się to, iż ten sprawdzian był na myślenie, ponieważ nie daje szans „kujonom", którzy całymi dniami wertują encyklopedie (uczeń, Słupsk)
Treść bardzo odbiega od tego, czego uczymy się w szkole (uczeń, Słupsk)
Zadania były trudne, o tematyce różnej. Niektóre z nich nie wiążą się wcale z życiem uczniów.
Uczniowie nie byli przygotowani do takiej formy egzaminów, a jeden rok szkolny na pewno nie wystarczy do przygotowania się do niego (uczeń, Słupsk)
Treść sprawdzianu była zachęcająca dla ucznia do myślenia (uczeń, Elbląg)
Takie coś ma mnie zachęcić do nauki? Nigdy! (uczeń, Piła)
Pytania „rozstrzelone" tematycznie, brak logicznego układu, część znacznie wykraczających poza ramy życia, błędy rzeczowe. Wiele pytań wymagało tylko odtwórczej wiedzy, z życiem uczniów niewiele miały wspólnego (nauczyciel, Słupsk)
Zadania postawione przed uczniami spłycały wartości literackie, nie prowokowały do przemyśleń i podejmowania własnego stosunku do omawianych treści.
Nie pozwalały na indywidualne opracowanie zagadnień (nauczyciel, Słupsk)
Zadania były ciekawe, o różnorodnej tematyce, były „życiowe" (nauczyciel, Piła)
Sprawdzian kompetencji pobudza uczniów do samodzielnego myślenia i zachęca do zdobywania wszechstronnej wiedzy. (nauczyciel, Słupsk)
Zadania były mało interesujące dla uczniów, mało związane z ich życiem (nauczyciel, Piła)
Absolutnie nie daje obrazu myślenia matematycznego (nauczyciel, Piła)
Treść testu była dobra, odpowiadała temu, czego uczyłam (nauczyciel, Elbląg)

Uczniowie już po testach próbnych nabrali przekonania, że rzetelna wiedza z języka polskiego nie jest potrzebna, wystarczy umiejętność szybkiego czytania i inteligencja (nauczyciel, Elbląg)

Przytoczone przykłady ilustrują dodatnie i ujemne poglądy i odczucia uczniów i nauczycieli biorących udział w regulaminowych oraz próbnych sprawdzianach kompetencji. Poznanie ich postaw wobec nieoczekiwanej zamiany tradycyjnych egzaminów ustnych i pisemnych na testowanie osiągnięć *wyłącznie* zadaniami wyboru wielokrotnego powinno wpłynąć na strategie egzaminacyjne w Polsce w najbliższych latach. Władze oświatowe muszą jednak mieć dowody na to, że postawy uczestników egzaminu wobec formy egzaminu wpływają na uczenie się.[19]

Obserwacja zachowań oraz ankietowanie uczniów i nauczycieli pozwala na pogłębioną interpretację wszelkich procesów pedagogicznych, nie tylko egzaminów szkolnych. W każdym przypadku *obraz* tych procesów w oczach uczniów i nauczycieli jest równie ważny jak faktyczny przebieg tych procesów, a najcenniejsze byłoby dokładne zbadanie związku między postawami uczestników a przebiegiem i wynikami działań edukacyjnych.

Czwarty, ostatni już z głównych celów diagnostyki edukacyjnej możemy sformułować następująco:

Cel czwarty diagnostyki edukacyjnej:	poznawanie wpływu postaw uczestników procesów pedagogicznych wobec tych procesów na przebieg procesów.

Szanse postępu diagnostyki edukacyjnej w Polsce

Wpadanie w euforię z okazji kolejnej ogólnopolskiej konferencji na temat diagnostyki edukacyjnej uważam za niewskazane. Raczej należy zdać sobie sprawę z niejasności, trudności i ograniczeń w tym zakresie oraz poszukać możliwości lepszego zaspokojenia potrzeb.

Żaden program dydaktycznych badań empirycznych na miarę reformy systemu edukacji w Polsce nie został podjęty i nie zanosi się na jego organizację. Brak nam także instytucji centralnej, która mogłaby koordynować, tu i tam podejmowane, programy lokalne. Sekcja Nauk Społecznych Komitetu Badań Naukowych ogranicza się do recenzowania projektów, przyznawania i rozliczania grantów. Komitet Nauk Pedagogicznych PAN prawie nie interesuje się badaniami empirycznymi. Instytut Badań Pedagogicznych nie ma ani pieniędzy, ani kadr wykonawców takich badań. Nowo powstała Centralna Komisja Egzaminacyjna zbyt jest zajęta normowaniem prawnym i organizowaniem przyszłych egzaminów, by w najbliższej przyszłości poważnie zająć się diagnozą chociażby tylko skutków własnych działań.

Szansę rozwoju diagnostyki edukacyjnej dostrzegam w ośrodkach regionalnych, takich jak Ośrodek Diagnozy, Egzaminów Szkolnych i Informacji Pedagogicznej w Legnicy. Będąc znacznie bliżej praktyki pedagogicznej niż instytucje centralne, zechcą zapewne podjąć problematykę, która już teraz jest bardzo ważna, a po wdrożeniu elementów reformy systemu edukacji w Polsce, a więc za kilka lat, stanie się pierwszoplanowa. W ślad za rozpoznaniem po-trzeb pójdą zapewne różne formy kształcenia kursowego i podyplomowego w zakresie diagnostyki edukacyjnej. Jak uczy doświadczenie, wiele z tych form kształcenia nie wzniesie się ponad pozorowanie naukowości i praktycznej użyteczności, ale przynajmniej niektóre dostarczą kadr wartościowych specjalistów. Tak właśnie było z pomiarem dydaktycznym, który dopiero teraz, i to nie bez oporów, okazuje się niezbędny jako podstawa działalności systemu egzaminacyjnego w Polsce.

Podsumowanie

Diagnostyka edukacyjna została tu określona jako rozpoznawanie przebiegu uczenia się ze szczególnym uwzględnieniem jego *kontekstu*, to jest ogółu czynników nie poddających się lub mało podatnych na oddziaływania pedagogiczne.

Tak pojęta diagnostyka ma dwa główne *cele praktyczne*:

1. dostarczanie uczniom i nauczycielom informacji wyjaśniającej przebieg i wyniki uczenia się,
2. podnoszenie trafności oceniania szkolnego.

Cele naukowe diagnostyki edukacyjnej stanowią natomiast:

3. poznawanie zależności między kontekstem, wejściami i działaniami a wynikami kształcenia w różnych systemach dydaktycznych,
4. poznawanie wpływu postaw uczestników procesów pedagogicznych wobec tych procesów na przebieg procesów.

Mimo szerokiej użyteczności, zwłaszcza w dobie reformy systemu edukacji, diagnostyka edukacyjna jest w Polsce zaniedbana. Szanse jej rozwoju autor tego tekstu dostrzega w spontanicznej działalności lokalnych ośrodków metodycznych oraz w intensywnym kształceniu naukowym kadr terenowych.

PRZYPISY

[1] *Diagnostyka edukacyjna*, Gdańsk 1994, Wyd. UG; *Perspektywy diagnostyki edukacyjnej*, Gdańsk 1998, Wyd. UG.

[2] S. M. Brookhart *Teaching About Communicating Assessment Results and Grading*, „Educational Measurement: Issues and Practice" 1/1999.

[3] B. Niemierko *Pomiar wyników kształcenia. Podręcznik pomiaru dydaktycznego dla nauczycieli i studentów*. Rozdział XIII: *Ocenianie osiągnięć uczniów*. Warszawa 1999, WSiP (w druku).

[4] Poszerzenie pojęcia trafności pomiaru przypisuje się Samuelowi Messickowi, patrz: *Validity* (w:) R. L. Linn *Educational Measurement. Third Edition*, Washington 1989, ACE. Objaśnienia zamieszczam w podręczniku wymienionym w poprzednim przypisie (Rozdział X: *Jakościowa analiza testu. Rozszerzone pojęcie trafności teoretycznej*)

[5] W książce Krystyny Lutyńskiej *Wywiad kwestionariuszowy. Przygotowanie i sprawdzanie narzędzia badawczego* (Wrocław 1984, Ossolineum) znajdujemy następującą charakterystykę Jana Lutyńskiego trzech koncepcji pytania kwestionariuszowego: 1) koncepcja *informacyjno-tradycyjna*, polegająca na tym, iż badacz pyta bezpośrednio o to, czego chce się dowiedzieć (np. jakie wykształcenie ma ojciec lub czy uczeń lubi czytać książki), a każde pytanie stanowi samodzielną informację, 2) koncepcja *wskaźnikowa*, opierająca się na założeniu, że odpowiedź respondenta jest wskaźnikiem dla danego pojęcia (np. poziomu intelektualnego rodziny), a w istocie - dla danego zjawiska, które interesuje badacza, przy czym zjawisko to nie zawsze bywa przed badaniami definiowane i 3) koncepcja *testowa*, w której celem zbioru pytań jest przypisanie respondentowi wyniku liczbowego, określającego jego miejsce na skali (np. skali poziomu intelektualnego rodziny), a istotne znaczenie ma tylko ten wynik.

[6] Ten model dokładnie przedstawia J. P. Keeves (red.) *Methodology and Measurement in International Surveys*, Hague 1992, IEA.

[7] D. F. Robitaille i R. A. Garden *The IEA Study of Mathematics II: Contteżts and Outcomes of School Mathematics*, Oxford 1989, Pergamon.

[8] B. Niemierko *Ogólnopolskie badania osiągnięć uczniów, nauczycieli i szkół. Synteza badań*, Warszawa 1990, Centrum Doskonalenia Nauczycieli.

[9] B. Niemierko (red.) *Perspektywy...* , s. 75 - 81. Ten tekst nie był wygłaszany i dyskutowany na konferencji, lecz został nadesłany przez autora do wydawnictwa.

[10] Jw., s. 15 - 42.

[11] Jw., s. 43 - 53.

[12] K. Ingenkamp *Lehrbuch der Padagogische Diagnostik. Studiensausgabe*, 2 Auflage, Weinheim 1992, Beltz.

[13] Warszawa 1975, WSiP.

[14] B. S. Bloom *Human Characteristics and School Learning*, New York 1976, McGraw-Hill. Model ten cytowałem już w *Diagnostyce edukacyjnej*.

[15] Zob. B. Niemierko *Osiągnięcia szkolne dzieci i młodzieży - koncepcja badań*, „Badania Oświatowe" 1982 nr 3.

[16] G. Penkowska *Badanie osiągnięć motywacyjnych w kształceniu informatycznym studentów pedagogiki* (w:) B. Niemierko i E. Kowalik (red.) *Perspektywy diagnostyki edukacyjnej*, Gdańsk 1998, Wyd. UG; B. Ciżkowicz *Motywacja w uczeniu się statystyki*, Bydgoszcz 1999, WSP; R. Suska-Wróbel *Proces zmiany motywacji w toku uczenia się przedmiotu szkolnego na przykładzie botaniki w liceum ogólnokształcącym* (maszynopis).

[17] Dane zostały zebrane i dostarczone mi przez Ośrodek Badań Kompetencji Ucznia w Wałbrzychu.

[18] Przykłady te zostały wybrane spośród znacznie większej liczby z cytowanych w moim opracowaniu *Ewaluacja sprawdzianu kompetencji przeprowadzonego w szkołach podstawowych województw słupskiego i pilskiego* oraz Ryszarda Wodzisławskiego *Ewaluacja sprawdzianu osiągnięć uczniów przeprowadzonego w województwie warmińsko-mazurskim 25 i 26 maja 1999 roku*.

[19] Zagadnienie *oddziaływania zwrotnego* egzaminów na treść i formy uczenia się jest obecnie przedmiotem analiz teoretycznych i badań empirycznych, zob. A. Nitko *Model egzaminów państwowych opartych na programie nauczania, sprawdzających i różniczujących, przeznaczonych do dyplomowania i selekcji uczniów* (w:) B. Niemierko i E. Kowalik (red.) *Perspektywy diagnostyki edukacyjnej*, Gdańsk 1998, Wyd. UG. Rośnie zainteresowanie *trafnością konsekwencyjną* pomiaru dydaktycznego, to jest wpływem jego stosowania na hierarchię wartości uczniów i nauczycieli, jednym z aspektów rozszerzonej trafności teoretycznej (por. przypis[4]). Poświęcono jej monograficzny numer kwartalnika *Educational Measurement: Issues and Practice* 1998 nr 17.