

BOLESŁAW NIEMIERKO

Akademia Marynarki Wojennej, Gdynia

Czy diagnostyka edukacyjna i ewaluacja w edukacji zachowają odrębność?

Zacznę od paradygmatów ewaluacji i diagnozy, przejdę do porównania dwu typów procedur, a skończę na perspektywach wspólnych działań.

Słońca i meteory dydaktyczne

Słońca dydaktyczne są dwa: nauczyciel i uczeń. To pierwsze było wcześniej na pedagogicznym niebie i dla tradycjonalistów wciąż błyszczy w jego centrum. To drugie świeci większości pedagogów i nic nie wskazuje na to, by miało przygasnąć, choć trudności wychowawcze wywołują tęsknotę za dawną surową dyscypliną.

Te słońca ożywiają bowiem dwa różne paradygmaty dydaktyczne: autorytarny i humanistyczny. Paradygmat naukowy, jak wiadomo, to zbiór wspólnie uznawanych założeń, pojęć i metod (Kuhn 1962). Wywiedziony z fizyki (sukcesja systemów Arystotelesa, Newtona i Einsteina), zaowocował najpierw w naukach przyrodniczych, a potem i społecznych, w psychologii (Kozielecki 1997) i pedagogice. Pozwala porównywać sposoby myślenia.

Oprócz stale świecących słońc, tworzących w obecnej dobie swoistą gwiazdę podwójną, pojawiają się na dydaktycznym niebie jaskrawe meteory. Największym było nauczanie programowane, polegające na kierowaniu uczeniem się przez udzielanie uczniom małych dawek informacji wraz z zadaniami sprawdzającymi przyswojenie tej informacji. W latach sześćdziesiątych i siedemdziesiątych ubiegłego stulecia nauczanie programowane zapowiadało rewolucję w kształceniu szkolnym, a liczne ekspery-

menty dydaktyczne potwierdzały jego wyższość nad kształceniem „konwencjonalnym” w poszczególnych przedmiotach nauczania (Kupisiewicz 1966). Co ciekawe, kwitło bujnie nie tylko w Stanach Zjednoczonych, lecz także w NRD, w Polsce i nawet w Związku Radzieckim, gdzie, przemilczając behawiorystyczne podstawy, okrzyknięto je „naukowym systemem kształcenia”.

Programowanie sięgało śmiało w głąb dyscyplin naukowych, eksponując ich strukturę i metodę, co przyniosło mu uznanie wybitnych uczonych, nie wykluczając laureatów Nagrody Nobla z dziedzin przyrodniczych. Motorem postępu były publiczne fundusze inwestowane w wyposażenie (maszyny dydaktyczne) i studia (projekty badawcze, publikacje) na skalę nieznaną przedtem w naukach pedagogicznych.

Zróżnicowanie pracy uczniów pod względem tempa i ścieżki wiodącej poprzez program nie wystarczało jednak do samoregulacji uczenia się. W miarę poszerzania kręgu wdrożeń okazało się, że w aspekcie emocjonalno-motywacyjnym kształcenia nawet najdoskonalsze programy nie są na dłuższą metę w stanie zastąpić osoby nauczyciela. „Niemoc w dziedzinie wychowawczej” zdecydowała o degradacji nauczania programowanego do roli metody pomocniczej (Okoń 1987, s. 264–267). W kolejnych dekadach meteor rozsypał się całkowicie, wpadając w gęstą atmosferę konstruktywizmu, przyznającego dziecku prawo do tworzenia własnych znaczeń i chroniącego jego umysł przed schematami ściśle zaprojektowanej treści kształcenia (Klus-Stańska 2002).

W zakresie nadzoru pedagogicznego w Polsce meteorem kilkunastu niedawnych lat było mierzenie jakości pracy szkoły, polegające na gromadzeniu i analizowaniu wieloaspektowych danych o tej pracy, w tym zwłaszcza danych opartych na samoocenie nauczycieli i na poglądach uczniów. Jego główny teoretyk, S. Wlazło wyjaśniał, że *mierzenie – to przede wszystkim postawa poznawcza, humanistyczne i holistyczne spojrzenie na człowieka, a nie jakieś konkretne procedury. W mierzeniu dobre są wszystkie te czynności i działania, które służą dobru ucznia, jego satysfakcji z podejmowanych przez niego działań edukacyjnych w szkole* (1999, s. 24).

Powiało wolnością. Emisariusze nowej idei, solennie namaszczeni przez Ministerstwo Edukacji Narodowej, dotarli do dyrektorów szkół w najodleglejszych zakątkach kraju. Z zapałem zbierano dowody wypełniania ich misji edukacyjnej w postaci właściwych przekonań nauczycieli, uczniów i rodziców, sprowadzając pojęcie mierzenia do stosownych deklaracji.

Akcja „mierzenia jakości” zaowocowała głównie intensywnym ankietowaniem uczniów i nauczycieli, dostarczających dość powierzchownych, a często koniunkturalnych ocen przebiegu procesów dydaktycznych. Alarmujące wydarzenia (znieważanie nauczyciela, samobójstwo uczennicy) zarejestrowano w szkołach, których stan wychowawczy „zmierzone” jako wysoki. Tymczasem pomiar osiągnięć uczniów dokonywany przez egzaminy zewnętrzne biegł osobnym torem i dostarczał całkiem innej informacji o poziomie pracy szkoły.

Nie powinniśmy wizytatorów i dyrektorów szkół oskarżać o złą wolę, o świadome fałszowanie obrazu rzeczywistości. Jak dowodzą psychologowie, złudzenie poznaw-

cze pojawia się nieuchronnie w przypadku braku solidnych danych (Kahneman 2012, s. 340 i n.). Co więcej, wszelkie „spojrzenie od środka” musi prowadzić do pozytywnych złudzeń (myślenia życzeniowego), gdyż dodatnia samoocena jest warunkiem pełnej koncentracji na zadaniu. Tendencyjność optymistyczna przepełnia nasze działania, zarówno biznesowe, jak i naukowe.

Nie poznałem jeszcze odnoszącego sukcesy naukowca, który nie umiałby przesadnie ukazywać znaczenia własnych badań. Osoba nie mająca ani odrobiny urojeń na temat własnego znaczenia musi się załamać w obliczu licznych drobnych niepowodzeń i rzadkich sukcesów, które są zwykłym losem większości badaczy (tamże, s. 352).

Innym, bliższym kierowaniu szkołą, przykładem złudzenia poznawczego podanym przez D. Kahnemana jest rozkład odpowiedzi na pytania: *Czy jesteś dobrym kierownicą? Czy jesteś lepszym kierownicą od średniej?* (tamże s. 346). Brak ścisłych danych o średniej powoduje, że respondenci wykazują na ogół optymizm (większość kierowców odpowie, że „tak”), mocno odbiegający od rzeczywistości. W pedagogice amerykańskiej znany jest *Lake Wobegon effect*, który w latach osiemdziesiątych powodował, że wszystkie stany z wyjątkiem Hawajów w sprawozdaniach dla federalnego Departamentu Edukacji wykazywały poziom osiągnięć swoich uczniów jako wyższy od średniej krajowej (Cannel 1988).

Meteoryty (szczątki meteoru) „mierzenia jakości” zgasły w chwili pojawienia się rozporządzenia Ministra Edukacji Narodowej w sprawie nadzoru pedagogicznego (*Rozporządzenie...* 2009). Znowu dokonano dość pośpiesznych badań i analiz (Mazurkiewicz 2011), a następnie, przy dużym zakłopotaniu teoretyków (Korporowicz 2011; Mizerek 2011), rozpoczęto akcję usilnego wdrażania nowego systemu.

Krytyka wcześniejszej koncepcji nadzoru pedagogicznego była bezlitosna. Według L. Korporowicza, bardzo szybko faza ta popadła we wszystkie pułapki myślenia pozytywistycznego. Błędem było wygenerowanie obowiązującego przez znaczny czas systemu o metodologicznie bardzo eklektycznym charakterze, a nazwanego mierzaniem jakości pracy szkół (2011, s. 22). Nie poradzono sobie z „przełamywaniem redukcji w pojmowaniu badań ewaluacyjnych w edukacji” (Korporowicz 2012), sprowadzając je do schematycznej kontroli i biurokratycznej sprawozdawczości.

Kariera pojęcia „ewaluacja” w pedagogice

Można zaryzykować tezę, iż przyczyną porażki mierzenia jakości pracy szkoły była jej słabość teoretyczna, a dokładniej, zignorowanie dorobku teorii ewaluacji,

udostępnionego administracji oświatowej w Polsce w ramach unijnego programu *TERM (Training for Education Reform Management)* w postaci przedruku wybranych pozycji amerykańskich i brytyjskich (Korporowicz 1997; Mizerek 1997).

Ewaluacja jest zbieraniem informacji o warunkach, przebiegu i wynikach działania edukacyjnego w celu dokonania oceny wartości tego działania i uzasadnienia dotyczących go decyzji. Narodziła się jako teoria, nazywana obecnie podejściem zoperacjonalizowanych zadań (House 1997), w łonie ortodoksyjnego behawioryzmu (Tyler 1949), lecz – w przeciwieństwie do nauczania programowanego – nie zgasła wraz z jego zmierzchem, lecz pomyślnie ewoluowała. Prześledźmy więc jej karierę na świecie i w Polsce.

Ta kariera nabrała tempa z przyczyn politycznych. Ewaluatyka edukacyjna, jak będę – przez analogię do diagnostyki – nazywał teorię ewaluacji w edukacji, wspomagała bowiem rozliczanie federalnych programów podnoszenia jakości kształcenia, jakie wprowadzono w Stanach Zjednoczonych po „szoku sputnikowym” w 1957 roku, gdy okazało się, że ZSSR wyprzedził ten kraj w podboju kosmosu (Popham 1975, s. 7). W tym też czasie (1959) powstało Międzynarodowe Stowarzyszenie Badań Osiągnięć Pedagogicznych (*International Association for the Evaluation of Educational Achievement, IEA*), jako ośrodek badań porównawczych nad edukacją. Wprawdzie Związku Radzieckiego nie udało się pozyskać do współpracy, ale w pionierskich badaniach *IEA* wzięły udział Węgry i Polska, zdobywając cenne doświadczenie metodologiczne (Niemierko 1990).

Przełomowe dla ewaluatyki okazało się rozróżnienie, którego dokonał M. Scriven (1967). Zauważył on, że twórcy programów kształcenia, uwikłani w pracochłonne analizy szczegółowe, w miarę postępu tych prac coraz rzadziej rewidują pierwotne koncepcje i coraz bardziej opornie wprowadzają zmiany w swoich wytworach. Najlepszym okresem ewaluacji programów kształcenia jest więc wczesny etap ich budowania, gdy nie trzeba cofać się daleko, a metody wytwarzania tych dokumentów są jeszcze elastyczne. M. Scriven stworzył pojęcie ewaluacji kształtującej (formatywnej, konstruktywnej), jako wartościowania programu szkolnego w toku jego powstawania. Jego przeciwieństwo, wartościowanie gotowego programu, zostało nazwane ewaluacją sumującą (sumatywną, konkluzywną).

Wkrótce ewaluacja kształtująca zdominowała ocenianie osiągnięć nie tylko szkół i nauczycieli, lecz także pojedynczych uczniów, jako „ocenianie kształtujące” (Kingston, Nash 2011). Pedagodzy uznali, że wprawdzie sumowanie osiągnięć jest potrzebne administracji szkolnej i społeczeństwu, ale kształtowanie osiągnięć ma większe znaczenie dla uczniów i ich nauczycieli (Stróżyński 2011). *Gdy kucharz próbuje zupę, jest to ocenianie kształtujące, a gdy klient próbuje zupę, jest to ocenianie sumujące* – zauważa S. Brookhart (2004, s. 45). Uczeń i nauczyciel są w roli kucharzy, mających bezpośredni wpływ na uczenie się w szkole.

Piętnaście lat temu dwaj Brytyjczycy, P. Black i D. Wiliam (1998), wyrazili pogląd, że stosowanie terminu „ocenianie kształtujące” (*formative assessment*) wymaga

dowodu, że między ocenianiem a przebiegiem uczenia się ucznia występuje sprzężenie zwrotne (*feedback*), wyrażające się zależnością uczenia się od informacji o jego przebiegu. Obecnie mocno wzrosło zainteresowanie tym podejściem i pojawiła się definicja: ocenianie kształtujące to informacja o procesie uczenia się, którą nauczyciele mogą wykorzystać do decyzji dydaktycznych, a uczniowie mogą wykorzystać do podniesienia osiągnięć, motywując uczniów (Brookhart 2009). Ocenianie niegwarantujące sprzężenia, rozumiane jako *wartościowanie wyników uczenia się ze względu na cele, do których osiągnięcia jednostka lub grupa zbliża się różnymi drogami i w różnym tempie* (Niemierko 2010, s. 41), może być zaś nazwane ocenianiem orientującym (*interim assessment*), etapowo sumującym.

Na szczeblu systemów szkolnych ewaluacja przybiera obecnie postać systemów odpowiedzialności edukacyjnej (*educational accountability systems*). Składają się na nie regulacje prawne i organizacyjne, według których instytucje edukacyjne i ich kadry są rozliczane z osiągnięć uczniów (Koretz, Hamilton 2006). Najbardziej wyrazistym przykładem takiej regulacji jest amerykańska ustawa *Żadne dziecko nie zostaje w tyle* (*No Child Left Behind* 2002), nakładająca na wszystkie szkoły ogólnokształcące w tym kraju obowiązek systematycznego zmniejszania liczby uczniów o wynikach niezadowolających i dostatecznych (poziom 1), tak by w ciągu 12 lat 100 procent uczniów w normie psychologicznej (z bardzo niewielkimi wyłączeniami) uzyskało w tych szkołach biegłość (poziom 2) w zakresie treści objętej standardami edukacyjnymi języka angielskiego, matematyki i nauk przyrodniczych.

Coroczne badanie osiągnięć uczniów zostało nakazane w każdym ze stanów, a poszczególne szkoły są oceniane według normy „właściwego postępu rocznego”, ustalonego za pomocą specjalnego algorytmu (Koretz, Hamilton 2006, s. 531). Szkoły niepełniające tych oczekiwań, wypadające poza trajektorię nakazanego postępu, ponoszą surowe konsekwencje, do likwidacji placówki włącznie, przy rozpaczliwych protestach związków nauczycielskich.

Ustawa *Żadne dziecko nie zostaje w tyle* spowodowała wzrost zainteresowania personelu amerykańskich szkół standardami edukacyjnymi i pomiarem sprawdzającym (tamże, s. 536 i n). Etapowe badania wskazują jednak na to, że odgórnie inspirowana interwencja pedagogiczna wobec dzieci „pozostających w tyle” nie w pełni jest pomyslna. Wśród jej niepożądanych skutków wymienia się zwięźenie programów kształcenia i zwiększenie liczby uczniów rezygnujących ze szkoły (Kane 2007, s. 53). Doświadczenia systemów edukacyjnych w wielu krajach skłaniają do ostrożności w przewidywaniu skuteczności wszelkich akcji wyrównywania różnic rozwojowych między uczniami.

Skojarzenie ewaluacji z nadzorem pedagogicznym, jakie dokonało się w Polsce, też budzi poważne obawy teoretyków (Korporowicz 2011; Mizerek 2011). Preferują oni brytyjską ewaluację dialogiczną (*responsive evaluation*), opartą na partnerskiej wymianie informacji między dwiema stronami procesu, a niechętnie nawiązują do wyników pomiaru osiągnięć uczniów dokonywanego w egzaminach zewnętrznych.

Przecenianie tych wyników, nagminne wśród niespecjalistów, grozi redukcją analiz kulturowego tła edukacji do ministerialnych rozporządzeń (Korporowicz 2012).

Na świecie i w Polsce ewaluacja w edukacji jest zorientowana głównie na systemy i placówki, a dopiero pośrednio na ucznia i jego problemy. Kontynuując astronomiczną metaforę, można uznać ją za planetę krążącą wokół nauczycielsko-kierowniczego słońca, oświetlaną jego blaskiem, wrażliwą na plamy. Polityczne uzależnienie wywołuje podatność na wypaczenia, których długie listy można znaleźć w literaturze (Niemierko 2009, s. 324; Korporowicz 2012, s. 44-46).

Teoretycy usilnie pragną oczyścić ewaluację z naleciałości, akcentując różnicę między gromadzeniem danych a podejmowaniem decyzji, lecz oddziaływanie decydentów i ich politycznych potrzeb na metody i postawy ewaluatorów jest dokumentowane od zarania dziejów instytucji tej dziedziny. *Czasem siły polityczne wskazują populację, z której możemy pobrać próbę. Czasem ograniczają dane, jakie możemy zbierać. (...) Czasem kształtują nasze wnioski. Czasem dotyczą stylu naszych sprawozdań. I zawsze wpływają na skutki naszych zaleceń* – napisał H. Brickell, dyrektor pedagogicznego ośrodka ewaluacyjnego w Stanach Zjednoczonych (Brickell 1976, s. 5).

Uczenie się jako przedmiot diagnostyki

Na tle rozmachu działań ewaluacyjnych dorobek diagnostyki edukacyjnej przedstawia się skromnie. Jej funkcja jest w Stanach Zjednoczonych pojmowana jako rozpoznawanie psychologicznych, fizycznych lub środowiskowych symptomów wykazywanych przez uczniów o nietypowych lub powtarzających się problemach uczenia się i zachowania. Jej metody są

głównie obserwacyjne, ale dla pewnych symptomów ogólne techniki przesiewowe potrzebne do sprawdzania hipotez mogą być dostępne dla nauczycieli. Zwykle po odnotowaniu symptomów nauczyciel przekazuje wyniki obserwacji właściwym specjalistom, np. doradcom, pielęgniarcom szkolnym, psychologom (Nitko 1989, s. 451).

Nauczyciel-diagnosta jest tam zatem „lekarzem pierwszego kontaktu”, wstępnie rozpoznającym chorobę i kierującym ucznia do dalszych analiz.

Pomiar dydaktyczny ma stać się metodą diagnozowania braków w opanowaniu treści kształcenia (Glaser 1988). Jednak *sztuka testowania diagnostycznego nie jest dobrze rozwinięta*, jak pisał ćwierć wieku temu jej wybitny znawca A. Nitko (1983, s. 15). Te stanowiska i oceny dominują do dzisiaj (Leighton, Gierl 2007).

Pomiar diagnostyczny jest wylaniającym się polem psychometrii zogniskowanym na uzyskiwaniu roboczej informacji na podstawie testów i ocen (...), by wytyczyć ścieżki naprawcze uczniom wymagającym pomocy w przedmiocie szkolnym – czytamy w programie tegorocznej konferencji *National Council for Measurement in Education* (Templin, Bradshaw 2013).

W Polsce diagnostyka edukacyjna rozwija się dwoma torami. Pierwszy to teoria diagnozy zaburzeń w uczeniu się, rozumianej jako rozpoznawanie przyczyn niezadawalających osiągnięć emocjonalno-motywacyjnych i poznawczych wybranych uczniów (Jarosz, Wysocka 2006). Podobnie jak w Stanach Zjednoczonych, diagnoza ma być podejmowana na skutek wydarzeń sygnalizujących nieprawidłowości rozwojowe: niewłaściwych zachowań w szkole i poza nią, trudności przystosowawczych, rażącego braku motywacji do uczenia się, trudności w opanowaniu podstawowych umiejętności poznawczych (np. w czytaniu), niskich osiągnięć poznawczych ucznia lub grupy uczniów (oddziału szkolnego, szkoły, sieci szkół).

Ten nurt ma w dydaktyce polskiej tradycję w postaci badań niepowodzeń szkolnych (Konopnicki 1957; Konopnicki 1961; Kupisiewicz 1964), ale w epoce przewidywania autorytarnych tendencji i postępującej humanizacji kształcenia szkolnego budzi sporo zastrzeżeń. Zdaniem K. Konarzewskiego, *zadaniem nauczyciela jest kształcić, a nie diagnozować*, zaś nauczyciela *na profesjonalną diagnozę psychologiczną nie stać*, więc wyręcza go *mądrość pokoju nauczycielskiego*, w którym *kolektywne doświadczenie skłania nauczycieli do patrzenia na uczniów przez pryzmat typowych wad*. Wskutek tego diagnozowanie edukacyjne staje się *bezwstydnie stronnicze, systematycznie gwałcące zasadę bezstronnej nadziei* (Konarzewski 1999). Znacznie mniej wątpliwości budzi diagnostyka resocjalizacyjna, zogniskowana na przypadkach ostrego niedostosowania społecznego (Wysocka 2008; Wojnarska 2011).

Drugi tor rozwoju diagnostyki edukacyjnej w Polsce to diagnostyka rozwojowa, nastawiona na rozpoznawanie zmian długotrwałych i względnie nieodwracalnych w uczeniu i wychowanku. Jej podstawą jest psychologia różnic indywidualnych, dyscyplina zajmująca się porównywaniem osobowości jako względnie stałych regulatorów działania kształtujących się na podłożu genetycznym i środowiskowym (Strelau 2002). Dotyczy wszystkich uczniów, a nie tylko sprawiających kłopoty.

W zakresie wychowawczym edukacyjną diagnostykę rozwojową zapoczątkował A. Janowski (1975). W zakresie dydaktycznym pionierem tego rodzaju diagnostyki był W. Okoń (1967), który rozróżnił cztery drogi uczenia się: przez przyswajanie, działanie, odkrywanie i przeżywanie. Ten podział pozwolił autorowi artykułu na rozróżnienie czterech równoległych modeli uczenia się, nazwanych „alfa”, „beta”, „gamma” i „delta”. Ich występowanie zostało potwierdzone w ogólnopolskich reprezentatywnych badaniach gimnazjalistów w proporcjach 31% – 13% – 43% – 13% (Niemierko 2009, s. 149–153). Dwa modele, przez przyswajanie (*alfa*) i przez odkrywanie (*gamma*), wyróżniły się skutecznością mierzoną wynikami egzaminu gim-

nazjalnego. W jednym z podręczników dydaktyki (Niemierko 2007) została przedstawiona próba uprawomocnienia wszystkich czterech modeli uczenia się jako zindywidualizowanych dróg rozwoju uczniów.

Diagnostykę rozwojową reprezentuje także system egzaminów zewnętrznych w Polsce, na który składają się trzy kolejne pomiary: po 6, 9 i 12 latach kształcenia. Jego funkcjonowanie stało się głównym przedmiotem analiz powstałego w 2004 roku Polskiego Towarzystwa Diagnostyki Edukacyjnej, skupiającego około 200 uczestników dorocznych konferencji. Na tych konferencjach systematycznie prezentowane są analizy edukacyjnej wartości dodanej (EWD), przyrostu osiągnięć uczniów w wybranym zakresie programowym w określonym czasie, np. między sprawdzianem po szkole podstawowej a egzaminem gimnazjalnym (Dolata 2007; Pokropek 2009). Trzyletni wskaźnik EWD, pozwalający na wykorzystanie danych o rozwoju osiągnięć poszczególnych gimnazjalistów do rozróżnienia „szkół sukcesu”, „szkół wspierających rozwój”, „szkół niewykorzystanych możliwości” i „szkół wymagających pomocy”, zyskał uznanie jako sposób wykorzystania diagnostyki do ewaluacji kształtującej działalność dydaktyczną szkół średnich (Szmigel 2010).

Zadaniem najbliższych lat w dziedzinie diagnostyki rozwojowej jest wzmocnienie jej zastosowań nieformalnych, będących rozpoznawaniem postępów w uczeniu się, dokonywanym środkami własnymi nauczyciela-wychowawcy, w toku działania edukacyjnego, jako systematyczna czynność pomocnicza w stosunku do kształcenia. T. Good i J. Brophy (1987) nazwali to „patrzeniem w klasie szkolnej”. Chodziło im, oczywiście, nie o to, by tylko patrzeć, lecz by widzieć, jak przebiega uczenie się uczniów. Według nich, klasa szkolna jest środowiskiem, w którym nauczyciel często (około tysiąc razy dziennie) musi podejmować szybkie decyzje przy niepełnej informacji.

Fundamentalnym czynnikiem trudności, jaką napotykają nauczyciele – piszą (s. 43) – oceniając zachowanie w klasie szkolnej, jest to, że tak wiele dzieje się tak szybko, iż nie są w stanie uświadomić sobie tego wszystkiego, co robią. Ten problem można częściowo rozwiązać przez ćwiczenie, gdyż dobre nauczanie wymaga przygotowania naukowego, doświadczenia i zaangażowania (s. 49).

Diagnostyka edukacyjna powinna zaoferować zarówno przyszłemu, jak i czynnemu nauczycielowi-wychowawcy program kształcenia niezbędnych umiejętności.

O ile diagnostyka zaburzeń może być traktowana jak wsparcie edukacyjnej roli nauczyciela, o tyle planeta diagnostyki rozwojowej niewątpliwie krąży wokół ucznia. To może sprawić, że nie okaże się meteorem, mało znaczącą odmianą działania pedagogicznego. Żadne nowe rozporządzenia i centralne akcje nie wyszłyby jej zapewne na dobre. Przyszłość zależeć będzie od potrzeb placówek edukacyjnych, które mogą wykorzystać określone rozwiązania metodyczne, pozorować je tylko lub odrzucić. Taki wybór dokonuje się już obecnie, przy rosnącym poziomie świadomości zysków i strat.

Dylematy dyrektora szkoły

Najciekawiej opowiada o tym książka czwórki wykonawców unijnego programu oświatowego (MacBeath i in. 2003). Jej tłumacz, K. Kruszewski, nie zdecydował się na użycie terminu „ewaluacja” (w oryginale: *self-evaluation*), stwierdzając, że *niepotrzebnie używamy „ewaluacji”, skoro mamy „ocenianie”* (s. 9) i uzupełnił jej wymowę mottem: *Polityka oświatowa ma chronić ucznia przed złym nauczycielem i złą szkołą, a dobrego nauczyciela i dobrą szkołę – przed administracją oświatową*. Narzędziem tej obrony jest samoocena (autoewaluacja, ewaluacja wewnętrzna) placówki, dokonywana przez pracowników danej placówki. Placówka edukacyjna, która systematycznie dokonuje ulepszeń własnej pracy na podstawie jej diagnozy i ewaluacji, jest nazywana uczącą się szkołą.

Pokonując rozliczne wątpliwości i przeszkody, bohaterowie omawianej książki posługują się systematycznie arkuszem obserwacji nazwanym „profilem samooceny szkoły” (s. 24). Na ten profil składają się wymiary:

1. Wyniki, obejmujące „osiągnięcia w nauce” (osiągnięcia poznawcze), „rozwój osobisty i społeczny” (osiągnięcia emocjonalne) i „losy uczniów”.
2. Nauczanie w klasie, obejmujące wykorzystanie czasu, jakość kształcenia i „pomoc w przypadku trudności” (elastyczność kształcenia).
3. Szkoła jako środowisko poznawcze i wychowawcze dla uczniów.
4. Środowisko lokalne – współpraca z domem rodzinnym, ze społecznością lokalną i z zakładami pracy.

Te wymiary są podobne do „obszarów ewaluacji zewnętrznej” wymienionych w aktualnym rozporządzeniu o nadzorze pedagogicznym w Polsce (*Rozporządzenie...* 2009):

1. Efekty działalności dydaktycznej, wychowawczej i opiekuńczej.
2. Procesy zachodzące w szkole lub placówce.
3. Funkcjonowanie szkoły lub placówki w środowisku lokalnym.
4. Zarządzanie szkołą lub placówką.

Dyrektor szkoły poinformowany o rozporządzeniu i wstępnie przygotowany do udziału w jego realizacji musi rozdysonować czas, siły i środki na trzy dziedziny wspomagania rozwoju placówki:

1. Kompetencje podstawowe kadry (studia podyplomowe, szkoły dydaktyki, usprawniania dydaktyczno-wychowawcze, współpraca z uczelniami).
2. Diagnoza edukacyjna (rozpoznawanie i ocenianie dróg uczenia się i osiągnięć uczniów, wspieranie samoregulacji uczenia się).
3. Ewaluacja wewnętrzna (działania o charakterze „uczącej się szkoły”, przygotowanie do egzaminów zewnętrznych i do wizytacji).

Pominiemy tu inwestycje w kompetencje podstawowe kadry, będące osobnym zagadnieniem, a zajmiemy się dwiema pozostałymi dziedzinami: diagnozą i ewaluacją.

Uznając, że większa część ich pola jest wspólna, porównamy i rozróżnimy paradygmaty tych dziedzin, to jest zasady i metody w ich ramach stosowane.

Zasady wspólne współczesnej ewaluatyki i diagnostyki wydają się następujące:

1. Staranne planowanie działań optymalizujących proces edukacji.
2. Wielość źródeł informacji, określana w metodologii badań jako triangulacja.
3. Równouprawnienie jakościowych i ilościowych metod badawczych.
4. Udział ocenianego w interpretacji wyników.

Zasady różne ewaluatyki i diagnostyki są przede wszystkim dwojakie:

1. Decyzje edukacyjne (zmiany, kontynuacje, promocje), podejmowane przez uprawnione instancje, są celem ewaluacji, a samoregulacja uczniów/wychowanków jest perspektywicznym celem diagnozy edukacyjnej.
2. Różnice systemowe (statutowe, kulturowe) mają pierwszeństwo w ewaluacji, a różnice indywidualne (genetyczne, środowiskowe, osobowościowe) – w diagnozie uczenia się.

Jako metody wspólne uzyskiwania informacji w zakresie ewaluatyki i diagnostyki można wymienić:

1. Studium przypadku, polegające na uzyskiwaniu wielu informacji z wielu źródeł.
2. Analizy jakościowe w zakresie kulturowo zróżnicowanych zjawisk.
3. Wywiad kwestionariuszowy, preferowany w stosunku do anonimowej ankiety.
4. Pomiar dydaktyczny z wykorzystaniem standaryzowanych baterii testów.

Pewne metody jednak są wyraźnie zróżnicowane między dwoma paradygmatami:

1. Obserwacja uczestnicząca dominuje w diagnostyce edukacyjnej, podczas gdy dla ewaluacji typowa jest obserwacja kontrolowana (rejestrowana obiektywnie).
2. Nauczyciel może podjąć wobec ucznia lub grupy uczniów eksperyment naturalny, przebiegający w ramach procesu kształcenia, podczas gdy ewaluator może tylko zastosować skale opisowe (szacunkowe), odzwierciedlające ten proces w opinii udziałowców edukacji.

Można dostrzec, że diagnostyka i ewaluatyka edukacyjna ujmują te same zjawiska z dwu przeciwnych krańców czterech kontynuów. Są to:

1. Przyczynowość: diagnoza szuka źródeł, a ewaluacja skutków. Bezzałożeniowość analiz (Scriven 1973) jest nie do przyjęcia dla pedagoga kierującego rozwojem ucznia.
2. Stałość zjawisk: diagnoza dotyczy głównie tego, co powtarzalne, podczas gdy ewaluacja ma zwykle na celu potwierdzić korzystną zmianę.
3. Skala obiektu: diagnoza jest zorientowana na pojedynczego ucznia/wychowanka, a ewaluacja na instytucję (sieć placówek, szkołę lub zakład, oddział szkolny).
4. Proces edukacyjny: diagnoza obejmuje uczenie się, a ewaluacja kształcenie – odpowiednio zrównoważone działanie edukacyjne (Niemierko 2007, s. 34–36).

Podobieństw i różnic jest wiele. Nic dziwnego, że pani dyrektor Barre, bohaterka dzieła *Czy nasza szkoła jest dobra?*, ma zamęt w głowie i *nie rzuciła pracy w tej szkole (...) tylko ze względu na grupę naprawdę oddanych nauczycieli* (MacBeath i in. 2003, s. 57). Nie każdy jej więc wybierać między diagnozą a ewaluacją!

W pedagogice częste jest krzyżowanie się zakresów podstawowych pojęć, jak „oświata” i „edukacja”, „nauczanie” i „wychowanie”, „sprawdzanie” i „ocenianie”, „diagnoza” i „ewaluacja”. Próby rozgraniczenia takich pojęć, podejmowane przez teoretyków, nie ułatwiają praktyki pedagogicznej. Lepiej dążyć do uzgodnienia zasad i do wytyczenia skutecznych działań, pozostawiając kwestię nazewnictwa otwartą.

Bibliografia

- BLACK P., WILIAM D., 1998, *Assessment and classroom learning*, Educational Assessment: Principles, Policy and Practice”, 5.
- BRICKELL H., 1976, *The influence of external political factors on the role and methodology of evaluation*, Evaluation Comment, 2.
- BROOKHART S.M., 2004, *Grading*, Pearson, Upper Saddle River.
- BROOKHART S.M., 2009, *Editorial*, Educational Measurement: Issues and Practice, Special Issue: The Validity of Formative and Interim Assessment, 3.
- CANNELL J.J., 1988, *Nationally normed achievement testing in America's public schools. How all 50 states are above the national average*, Educational Measurement, Issues and Practice, No. 2.
- DOLATA R. (red.), 2007, *Edukacyjna wartość dodana jako metoda oceny efektywności nauczania na podstawie wyników egzaminów zewnętrznych*, CKE, Warszawa.
- GLASER R., 1988, *Przyszłość pomiaru dydaktycznego*, Edukacja, 1.
- GOOD T.L., BROPHY J.E., 1987, *Looking in classroom*, Harper & Row, New York.
- HOUSE E.R., 1997, *Ewaluacja i jej uprawomocnianie. Główne podejścia*, [w:] L. Korporowicz (red.), *Ewaluacja w edukacji*, Oficyna Naukowa, Warszawa.
- JANOWSKI A., 1975, *Poznawanie uczniów. Zdobywanie informacji w pracy wychowawczej*, WSiP, Warszawa.
- JAROSZ E., WYSOCKA E., 2006, *Diagnoza psychopedagogiczna. Podstawowe problemy i rozwiązania*, Żak, Warszawa.
- KAHNEMAN D., 2012, *Pułapki myślenia. O myśleniu szybkim i wolnym*, Media Rodzina, Poznań.
- KANE M.D., 2007, *Validation*, [in:] R.L. Brennan (ed.), *Educational measurement*, Fourth edition, American Council on Education – Praeger, Westport.
- KINGSTON N., NASH B., 2011, *Formative assessment: A meta-analysis and a call for research*, Educational Measurement: Issues and Practice, No. 4.
- KLUS-STĄŃSKA D., 2002, *Konstruowanie wiedzy w szkole*, Wyd. Uniwersytetu Warmińsko-Mazurskiego, Olsztyn.
- KONARZEWSKI K., 1999, *Komu jest potrzebna diagnostyka oświatowa?*, [w:] B. Niemierko, B. Machowska (red.), *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*, Ośrodek Diagnostyki, Egzaminów Szkolnych i Informacji Pedagogicznej, Legnica.
- KONOPNICKI J., 1957, *Zaburzenia w zachowaniu się dzieci szkolnych i środowisko*, PWN, Warszawa.
- KONOPNICKI J., 1961, *Problem opóźnienia w nauce szkolnej*, Ossolineum, Wrocław.
- KORETZ D.M., HAMILTON L.S., 2006, *Testing for accountability in K-12*, [in:] R.L. Brennan (ed.), *Educational measurement*, Fourth edition, American Council on Education – Praeger, Westport.

- KORPOROWICZ L. (red.), 1997, *Ewaluacja w edukacji*, Oficyna Naukowa, Warszawa.
- KORPOROWICZ L., 2011, *Zmienne losy polskiej ewaluacji. Pomiędzy nadzieją, animacją i konfuzją*, [w:] B. Niemierko, M.K. Szmigel (red.), *Ewaluacja w edukacji: koncepcje, metody, perspektywy*, PTDE, Kraków.
- KORPOROWICZ L., 2012, *Refleksja jako działanie. O stałej potrzebie przelamywania redukcji w pojmowaniu badań ewaluacyjnych w edukacji*, [w:] B. Niemierko, M.K. Szmigel (red.), *Regionalne i lokalne diagnozy edukacyjne*, PTDE, Kraków.
- KOZIELECKI J., 1997, *Koncepcje psychologiczne człowieka*, Żak, Warszawa.
- KUHN T.S., 1962, *The structure of scientific revolutions*, wyd. pol. 1968, *Struktura rewolucji naukowych*, PWN, Warszawa.
- KUPISIEWICZ C., 1964, *Niepowodzenia dydaktyczne. Przyczyny i niektóre środki zaradcze*, PWN, Warszawa.
- LEIGHTON J.P., GIERL M.J. (eds.), 2009, *Cognitive diagnostics assessment for education. Theory and applications*, Cambridge University Press, New York.
- MACBEATH J., SCHRATZ M., MEURET D., JAKOBSEN L., 2003, *Czy nasza szkoła jest dobra?*, WSiP, Warszawa.
- MAZURKIEWICZ G. (red.), 2011, *Ewaluacja w nadzorze pedagogicznym. Refleksje*, Wyd. UJ, Kraków.
- MIZEREK H. (red.), 1997, *Ewaluacja w szkole. Wybór tekstów*, M/G, Olsztyn.
- MIZEREK H., 2011, *Tam, gdzie kończą się zakłęcia, a zaczyna ewaluacja*, [w:] B. Niemierko, M.K. Szmigel (red.), *Ewaluacja w edukacji: koncepcje, metody, perspektywy*, PTDE, Kraków.
- NIEMIERKO B., 1990, *Ogólnopolskie badania osiągnięć uczniów, nauczycieli i szkół. Synteza badań*, Centrum Doskonalenia Nauczycieli, Warszawa.
- NIEMIERKO B., 2007, *Kształcenie szkolne. Podręcznik skutecznej dydaktyki*, WAiP, Warszawa.
- NIEMIERKO B., 2009, *Diagnostyka edukacyjna. Podręcznik akademicki*, PWN, Warszawa.
- NIEMIERKO B., 2010, *Między prawdą a skutecznością – perspektywy oceniania szkolnego*, [w:] B. Niemierko, M.K. Szmigel (red.), *Teraźniejszość i przyszłość oceniania szkolnego*, PTDE, Kraków.
- NITKO A.J., 1983, *Educational tests and measurements. An introduction*, Harcourt, New York.
- NITKO A.J., 1989, *Designing tests that are integrated with instruction*, [in:] R.L. Linn (ed.), *Educational measurement. Third edition*, American Council on Education, New York.
- OKOŃ W., 1987, *Wprowadzenie do dydaktyki ogólnej*, PWN, Warszawa.
- POKROPEK A., 2009, *Metody statystyczne wykorzystywane w szacowaniu trzyletnich wskaźników egzaminacyjnych*, [w:] B. Niemierko, M.K. Szmigel (red.), *Teraźniejszość i przyszłość oceniania szkolnego*, PTDE, Kraków.
- POPHAM W.J., 1975, *Educational evaluation*, Prentice-Hall, Englewood Cliffs.
- Rozporządzenie Ministra Edukacji Narodowej z dnia 7 października 2009 r. w sprawie nadzoru pedagogicznego, Dziennik Urzędowy, 09/168/1324 z dnia 9 października 2009 r.
- SCRIVEN M., 1967, *The methodology of evaluation*, [in:] R.E. Stake (ed.), *Perspectives of curriculum evaluation*, Rand McNally, Chicago.
- SCRIVEN M., 1973, *Goal-free evaluation*, [in:] E.R. House (ed.), *School evaluation: The politics and process*, McCutchan, Berkeley.
- STRELAU J., 2002, *Psychologia różnic indywidualnych*, Scholar, Warszawa.
- STRÓŻYŃSKI K. (red.), 2011, *Ocenianie kształtujące po polsku. Kurs dla doradców metodycznych. Scenariusze zajęć*, CODN, Warszawa.
- SZMIGEL M.K., 2010, *Planowanie rozwoju szkoły z wykorzystaniem trzyletnich wskaźników egzaminacyjnych*, [w:] B. Niemierko, M.K. Szmigel (red.), *Teraźniejszość i przyszłość oceniania szkolnego*, PTDE, Kraków.
- TEMPLIN J., BRADSHAW L., 2013, *Diagnostic measurement theory, methods and applications*, NCME Annual Meeting Program.

- TYLER R., 1949, *Basic principles of curriculum and instruction*, The University of Chicago Press, Chicago.
- WLAZŁO S., 1999, *Mierzenie jakości pracy szkoły. Część trzecia*, MarMar, Wrocław.
- WOJNARSKA A. (red.), 2011, *Diagnostyka resocjalizacyjna. Wybrane zagadnienia*, Wyd. UMCS, Lublin.
- WYSOCKA E., 2008, *Diagnoza w resocjalizacji. Obszary problemowe i modele rozwiązań w ujęciu psychopedagogicznym*, PWN, Warszawa.

Will educational diagnostics and evaluation in education keep separate?

Education abounds with fleeting fascinations like programmed learning or “measuring the school-work quality” (in Poland); however, the interest in teachers and students is permanent. It may provide stability for, accordingly, evaluation and diagnostics. The author compared paradigms of the domains and found that formative evaluation and developmental diagnosis have more similarities than differences. Therefore, practitioners can unite the operations according to their needs and leave possible distinctions to theorists.