

prof. zw. dr hab. Bolesław Niemierko

Elbląska Uczelnia Humanistyczno-Ekonomiczna

## Diagnostyka edukacyjna duża i mała

Edukacja ma dwojakie znaczenie. Jest zarówno systemem państwowym, działającym według potrzeb ekonomicznych, tradycji kulturowych i planów rozwojowych kraju, jak też programem osobistym jednostek – dzieci, młodzieży i dorosłych – wspomaganym przez otoczenie. Na obydwu tych szczeblach, społecznym i indywidualnym, uzyskiwanie informacji o warunkach, przebiegu i wynikach uczenia się, czyli dokonywanie diagnozy edukacyjnej, jest niezbędne.

Od szczebla edukacji zależy strategia diagnozy. Gdy górę bierze systemowy, instytucjonalny punkt widzenia, zbiory ujednoczonych danych są z reguły duże, co pozwala na ustalanie prawidłowości statystycznych o charakterze socjologicznym. Gdy zajmujemy się pojedynczymi osobami, zbieramy o każdej z nich rozmaite dane składające się na charakterystyki jakościowe z rozległą podstawą psychologiczną.

Te dwie strategie, grupowa i jednostkowa, będą w moim referacie porównane pod względem skuteczności, warsztatu poznawczego i przygotowania naukowego diagnosty. Moim postulatem będzie wiązanie badań statystycznych z badaniami jakościowymi oraz łączenie odpowiednich kompetencji w kształceniu diagnostów.

### Rozszczępienie się paradygmatu

**Paradygmat naukowy** diagnostyki edukacyjnej, to jest zbiór jej założeń, pojęć i metod, jest jeszcze płynny. Dla wielu pedagogów sprowadza się do *drobiazgowej analizy dziedziny osiągnięć jednostki. Jakie mikroskopijne umiejętności wstępne są silne lub słabe? Jakie błędy, złe nawyki i skojarzenia mogą utrudniać postępy w uczeniu się określonych umiejętności i wiadomości? Tylko niektóre z tych kwestii mogą znaleźć odpowiedź za pomocą właściwie zaprojektowanych testów* (Nitko, 1983: 15). Nic dziwnego, że tak pracochłonne zabiegi są podejmowane tylko w przypadkach **zaburzeń w uczeniu się**, to jest jawnie niezadowolających osiągnięć emocjonalno-motywacyjnych i poznawczych ucznia (Niemierko, 2009: 33n). Decyzja o ich podjęciu przypomina decyzję o skierowaniu chorego do lekarza, a więc należy do przykrych obowiązków pedagoga. Stwarza też pole do nadużyć, gdyż *domniemania [potrzeby diagnozy] wyrastają z konfliktów interpersonalnych, są zawsze niepewne i uruchamiają mechanizm samospełniającego się proroctwa* (Konarzewski, 2004: 55).

Na przeciwnym krańcu skali szczegółowości lokują się diagnozy dokonywane przez międzynarodowe organizacje, jak *IEA* – Międzynarodowe Stowarzyszenie Ewaluacji Osiągnięć Edukacyjnych (Husén, Poslethwaite, 1996) i *OECD* – Organizacja Współpracy Ekonomicznej i Rozwoju (Białecki i in., 2003), a także przez krajowe komitety ekspertów oświatowych powoływane do

przygotowywania reform (Szczępański, 1973; Kupisiewicz, 1985; Kupisiewicz, 1989). Ich zadaniem jest wykrywanie prawidłowości rozwoju edukacji, a w szczególności – czynników skuteczności systemów kształcenia.

Zagrożeniem **szerokich badań edukacyjnych** (*large-scale testing*), w tym analiz krajowych egzaminów zewnętrznych, jest traktowanie ich jako „wścigi konne” (Husén, 1987), w których najważniejsze jest miejsce krajów i szkół na mecie. „Tabele ligowe” wyników pomiaru osiągnięć uczniów znacznie bardziej interesują publicystów, rodziców i polityków niż złożone zależności pedagogiczne. Pytanie o jakość kształcenia w danej szkole jest zastępowane pytaniem łatwiejszym: o miejsce szkoły w ostatnim rankingu średnich wyników egzaminu zewnętrznego absolwentów.

Ze względu silny aspekt wartościujący, szerokie badania edukacyjne, w tym także egzaminy zewnętrzne, są częściej uważane za ewaluację niż diagnozę. **Ewaluacja w edukacji** jest zbieraniem informacji o warunkach przebiegu i wynikach kształcenia, dokonywanym w celu jego ulepszenia lub podjęcia decyzji o dalszym stosowaniu. Pole diagnozy edukacyjnej i ewaluacji w edukacji jest w przeważającej części wspólne, gdyż pedagodzy zawsze dążą do wysokich wyników kształcenia, a wnioski z badań muszą być oparte na szczegółowym rozpoznaniu procesu i kontekstu uczenia się (Niemierko, w druku). W ewaluacji jednak szczególną rolę odgrywa czynnik polityczny. Tak to określił Henry Brickell, dyrektor jednego z amerykańskich instytutów ewaluacyjnych (1976: 1):

*Służymy władzom państwowym, regionalnym, lokalnym. Jeżeli władza znaczy polityka, to nasze pole działania jest w 100% polityczne. [...] Urzędnicy państwowi, którzy zlecają nam zadania są uwikłani w zdobywanie i sprawowanie władzy – władzy decydowania o tym, co się komu należy. [...] Ich motywy zlecenia nam zadań są polityczne, to znaczy związane z władzą. [...] Wyniki ewaluacji, jakie dostarczamy, stają się kartą na politycznym stole używanym do gry zwanej sprawowanie władzy.*

Polityczna funkcja ewaluacji prowadzi do wynaturzeń obejmowanych nazwą **pseudoewaluacji**, rozumianej jako czynności wadliwe bądź pozorowane w ocenianiu jakości działania edukacyjnego i uzasadnianiu decyzji dotyczących tego działania (Niemierko, 2009: 323-325; Korporowicz, 2012). Takie praktyki oddalają ewaluację od diagnostyki edukacyjnej, w tym zwłaszcza od indywidualnej diagnozy uczenia się, zorientowanej na potrzeby ucznia, a nie administracji.

W łonie edukacyjnej teorii ewaluacji dokonał się rozłam, którego znaczenie wybiegło daleko poza początkowe intencje autora. Filozof Michael Scriven (1967) zauważył, że twórcy programów kształcenia, uwikłani w pracochłonne analizy szczegółowe, w miarę postępu tych prac coraz rzadziej rewidują pierwotne koncepcje i coraz oporniej wprowadzają zmiany w swoich wytworach. Scriven utworzył pojęcie **ewaluacji kształtującej** (formatywnej, konstruktywnej), jako wartościowania programu szkolnego w toku jego powstawania. Jego przeciwieństwo, wartościowanie gotowego programu, zostało nazwane **ewaluacją sumującą** (sumatywną, konkluzywną).

To rozróżnienie stało się w pedagogice bardzo popularne i wkrótce zostało rozciągnięte na wszelkie rodzaje ewaluacji – od jakości systemu kształcenia do

osiągnięć pojedynczego ucznia. Jako **ocenianie kształtujące** (*formative assessment*) przeszło dalszą ewolucję, bo zaczęto wymagać od niego dowodów na to, że istnieje zależność między ocenianiem a uczeniem się. Przed piętnastu laty dwaj Brytyjczycy, Paul Black i David Wiliam (1998), jako pierwsi wyrazili pogląd, że stosowanie terminu ocenianie „kształtujące” opiera się na założeniu, że między ocenianiem a przebiegiem uczenia się występuje **sprzężenie zwrotne**, wyrażające się wpływem informacji o skuteczności uczenia się na pracę ucznia.

Zainteresowanie tym podejściem na świecie i w Polsce (Stróżyński, 2011) nadal rośnie. Jego współczesną wersję ujmuje definicja: *Ocenianie kształtujące to informacja o procesie uczenia się, którą nauczyciele mogą wykorzystać do decyzji dydaktycznych, a uczniowie mogą wykorzystać do podniesienia osiągnięć, motywująca uczniów* (Brookhart, 2009). Tak kreślone kształtowanie osiągnięć uczniów odbiegło tak daleko od sumowania tych osiągnięć, że powstała potrzeba wypełnienia luki (Perie i in., 2009).

Ocenianie niezakładające 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: 41), może być teraz nazwane **ocenianiem orientującym** (*interim assessment*), etapowo sumującym. Z kształtowaniem ma ono wspólny potencjalny efekt regulacyjny, a z sumowaniem – kryterium końcowego celu. Ma łączyć trzy główne funkcje oceniania: dydaktyczną, ewaluacyjną i prognostyczną (Perie i in. 2009: 7-8).

Modelu oceniania orientującego można dopatrzeć się w nowopowstałych **systemach 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 i 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 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** w postaci redukcji liczby uczniów nieuzyskujących należytej biegłości (Koretz i Hamilton, 2006: 531). Szkoły niespełniające 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, pomiarem sprawdzającym i diagnozą edukacyjną (tamże, 536n). Etapowe badania wskazują jednak na to, że ogólnie inspirowana interwencja pedagogiczna wobec „dzieci pozostających w tyle” nie w pełni jest pomyślna. Wśród jej niepożądanych skutków wymienia się zwężenie programów kształcenia i zwiększenie liczby uczniów rezygnujących ze szkoły (Kane, 2007: 53).

Doświadczenia systemów edukacyjnych w wielu krajach skłaniają do ostrożności w przewidywaniu skuteczności wszelkich akcji wyrównywania różnic rozwojowych między uczniami. Wzrost różnic w miarę uczenia się stanowi naturalny **efekt wachlarzowy** w postaci wzrostu wariancji osiągnięć wraz ze wzrostem średniej tych osiągnięć, gdyż przy wyższym poziomie osiągnięć czas potrzebny na nauczenie się nowej czynności jest krótszy niż przy niższym poziomie osiągnięć bez względu na metodę kształcenia (Cronbach i Snow, 1977). Wiązanie ewaluacji pracy szkoły z diagnozą osiągnięć pojedynczych uczniów budzi jednak nadzieję na harmonijny rozwój dziedziny, którą obejmujemy nazwą diagnostyki edukacyjnej.

W Polsce diagnostyka edukacyjna poczyniła pierwsze kroki organizacyjne w 1993 roku w toku konferencji naukowej w Instytucie Pedagogiki Uniwersytetu Gdańskiego (Niemierko, Kowalik, 1994). Dziesięć lat później w Krakowie, w Oddziale Doskonalenia Nauczycieli przodującym w studiach nad systemem egzaminów zewnętrznych, narodziło się Polskie Stowarzyszenie Diagnostyki Edukacyjnej. Na kolejnych krajowych konferencjach Towarzystwa zagadnienia egzaminów zewnętrznych zdominowały problematykę obrad, pozostawiając diagnostykę indywidualnej pracy ucznia w tle szerszych badań i analiz. Tak więc perspektywa dalsza rozwoju ucznia, wyznaczona przez kolejne szczeble systemu szkolnictwa, wzięła górę nad jego perspektywą bliższą, określoną przez bieżące potrzeby edukacyjne.

## Dwie perspektywy diagnozy

Perspektywa to pole widzenia wyznaczone kompetencją i pozycją obserwatora. *Istnieje tylko to, co widzisz (What you see is all there is)* jest aforystycznym założeniem teorii perspektywy wyłożonej w *Pułapkach myślenia* noblisty Daniela Kahnemana (2012). Ktoś, kto zwiedził kilkadziesiąt krajów świata ma inne doświadczenia i „wartości bazowe” niż ktoś, kto nigdy nie opuścił swojej wioski. Starzy ludzie myślą inaczej niż młodzi, ministrowie edukacji inaczej niż rodzice uczniów, a tym bardziej – niż sami uczniowie, filozofowie inaczej niż technicy laboratoryjni, a pedagodzy ogólni inaczej niż nauczyciele poszczególnych przedmiotów szkolnych.

Do różnic układów odniesienia dołączają się różnice procedur poznawczych, określonych przez Kahnemana jako *System I* i *System II*. **System 1** rozstrzyga problemy przez analogie, działa intuicyjnie i natychmiastowo, jak w rutynowych reakcjach rzemieślnika, sportowca, kierowcy pojazdu lub nauczyciela podczas lekcji, wchodzącego codziennie w ponad 1000 osobistych interakcji z uczniami (Good i Brophy, 1987). Wiele z nich to interakcje nieoczekiwane, nagłe, bogate w podteksty, znaczące dla przyszłej współpracy. W okamgnieniu trzeba odczytać intencje ucznia, wybrać zachowanie i przewidzieć reakcje klasy (Janowski, 1989).

*Klasa szkolna to złożone środowisko, w którym nauczyciele często muszą podjąć szybką decyzję, posługując się niepełną informacją. [...] Świadomość wszystkiego, co się zdarza, jest niemożliwa* – stwierdzili Good i Brophy (1987: 1, 43). Ta świadomość, niekiedy gorzka, pojawia się później, tuż po lekcjach lub podczas źle przespanych nocy. U niespecjalistów wszakże, ze względu na brak fachowej wiedzy, pożyteczna autorefleksja znacznie rzadziej występuje.

Większość z nich będzie tylko umacniać się w przekonaniach i bronić swych emocjonalnych rozstrzygnięć.

Gdy są na to warunki zewnętrzne i wewnętrzne, do akcji wkracza inny system przetwarzania informacji. **System 2** myślenia działa wolno i pod pełną kontrolą świadomości. Posługujący się nim rozważa jak najszerszy krąg doświadczeń, w tym podstawy teorii wyjaśniających zjawiska. Tu „kadrowanie” jest szerokie, spojrzenie dokonywane z dystansu, bez emocjonalnego napięcia, ale z dużym wysiłkiem umysłowym. Ten system polega na mozolnym wyszukiwaniu informacji, ich starannej analizie i syntezie, skrupulatnym sprawdzaniu wniosków, ciągłej rewizji doświadczeń. Jest trybem działania refleksyjnego praktyka (Schön, 1983) i każdego, w zasadzie, teoretyka.

Jak spowodować, by diagnostyka edukacyjna była częścią profesjonalnej kompetencji pedagoga na wszystkich szczeblach organizacji procesu edukacyjnego? Jak powiązać systemowe i osobiste obszary działań, postaw, wiedzy i umiejętności potrzebnych do rozpoznawania sytuacji dydaktyczno-wychowawczych? Wydają się konieczne cztery etapy akcji:

1. Trzeba jasno odróżnić **perspektywę instytucjonalną** diagnostyki, prowadzącą do interpretacji przypadków według prawidłowości ogólnych i do formułowania wytycznych systemowych dla nauczycieli-wychowawców, administracji oświatowej i nadzoru pedagogicznego, od **perspektywy personalnej**, zorientowanej na indywidualność uczniów i nauczycieli, niepowtarzalność sytuacji, oryginalną typologię procesów edukacyjnych. Tę pierwszą możemy nazwać diagnozą „z daleka”, gdyż obejmuje rzeczywistość jakby z lotu ptaka, rysując topografię rozległych terenów, a tę drugą – diagnozą „z bliska”, bo liczy się w niej każdy szczegół, a nadto pomieć zdarzeń z dłuższego okresu.
2. Trzeba wyważyć **trafność diagnoz** dokonywanych z tych dwu perspektyw, to jest ich użyteczność do określonych celów edukacyjnych. Żadna z nich nie okaże się bezwzględnie trafna, ale też żadna nie okaże się bezużyteczna.

Cytowany już tu Kahneman przypomniał zażarte spory w innym niż edukacja „obszarze niskiej trafności”: medycynie. Wywołała je mała książeczka Paula Meehla (1954) na temat prognozowania przebiegu choroby według równania regresji empirycznych wskaźników statystycznych oraz według osądu klinicystów. Te i inne badania porównawcze, w tym także dotyczące przewidywania osiągnięć studentów kolegium na podstawie pomiaru zdolności oraz ocen nauczycielskich ze szkoły średniej, doprowadziły do eskalacji wysoce emocjonalnych określeń, które warto tu przytoczyć (Meehl, 1968, cyt. za: Kahneman, 2012: 305):

[Metoda statystyczno-algorytmiczna] – *mechaniczna, atomistyczna, matematyczna, nieprawdziwa, arbitralna, niepełna, martwa, pedantyczna, wyrzywkowa, trywialna, wymuszona, statyczna, powierzchowna, sztywna, jałowa, akademicka, pseudonaukowa oraz ślepa.*

[Metoda kliniczno-intuicyjna] – *dynamiczna, całościowa, otwarta na głębszy sens, holistyczna, subtelna, zakorzeniona we współczuciu, konfiguralna, czuła na prawidłowości, zorganizowana, bogata, głęboka, autentyczna, wrażliwa, wyrafinowana, prawdziwa, żywa, konkretna, naturalna, życiowa oraz rozumiejąca.*



Jak widać, klinicyści łatwo uzyskali werbalną przewagę nad statystykami. Badania pokazały jednak, że *ich trafne osądy dotyczą krótkoterminowych prognoz, a więc umiejętności, którym klinicyści poświęcili lata praktyki*, ale w długim terminie te umiejętności zawodzą i „wzory matematyczne” uzyskują przewagę, co oczywiście nie znaczy, że są wolne od błędu losowego oszacowań (tamże, 304n).

3. Trzeba kształcić kadry pedagogów, którzy będą zdolni stosować obydwa rodzaje perspektywy diagnostycznej, instytucjonalną i personalną. To pozwoli im na wybór strategii edukacyjnej możliwej do realizacji na odpowiednim szczeblu edukacji.

**System 1** podejmowania decyzji wymaga zbierania osobistych doświadczeń, które mogą stanowić układ odniesienia dla nowych sytuacji decyzyjnych. Teoretycy, którzy nigdy nie byli profesjonalnymi nauczycielami i wychowawcami, powinni uzyskać takie kompetencje lub przynajmniej działać w zespołach, w których praktycy będą mieli wiele do powiedzenia, a ich zdanie będzie się liczyć.

**System 2** podejmowania decyzji wymaga zrozumienia zależności probabilistycznych i wnioskowania statystycznego. Praktycy, którzy nie rozumieją i „boją się” statystyki, będą skłonni przyjmować wyniki szerokich diagnoz bezkrytycznie lub, częściej, odrzucać je w całości. (Niemierko, 2012). Nie jest im łatwo wyjść poza jednostkowe doświadczenie, będące ich najcenniejszą kompetencją zawodową.

4. Trzeba wreszcie nauczyć się **przenoszenia problematyki diagnostycznej** z układu instytucjonalno-szkolnego na układ personalno-dydaktyczny i odwrotnie. To bardzo odpowiedzialne zadanie. Problemy pedagogiczne zyskują nową postać przy zmianie perspektywy – z reguły stają się trudniejsze do rozwiązania. To, co teoretykowi wydaje się właściwe jako statutowa powinność każdego nauczyciela, bywa dla praktyka mgliste i uciążliwe, a niekiedy wręcz sprzeczne z dorobkiem zawodowym. By zaś uogólnić to, co dla praktyka jest naturalne i oczywiste, możemy potrzebować skomplikowanych modeli teoretycznych, które i tak zawodzą w nietypowych sytuacjach.

Istnieje silna tendencja do zamykania się w swoim gronie i zostawiania zadań zmiany perspektywy drugiej stronie. Niech biorą od nas to, co dla nich przydatne, a my będziemy mówić o naszych sprawach naszym językiem! Taka postawa niewątpliwie hamuje postęp diagnostyki edukacyjnej. Poprawne przełożenie wiedzy na działania w innej skali może udać się tylko jej pełnym posiadaczom, więc takie zadania należy traktować jako istotną część roli twórczego teoretyka i praktyka. Fascynującą część ich roli!

Przedstawię teraz po dwa przykłady zadań diagnostycznych, które zyskują na atrakcyjności po zmianie perspektywy: z instytucjonalnej na personalną i personalnej na instytucjonalną, z dalekiej na bliską i z bliskiej na daleką.

## Od instytucji do jednostki

I. Do najogólniejszych pojęć, jakimi posługują się współczesne nauki społeczne, należą kapitał ludzki i kapitał społeczny. Edukacja jest dość zgodnie uznawana za główny czynnik wzrostu kapitału niematerialnego, jak nazywamy kapitał ludzki i kapitał społeczny łącznie.

Potężny program operacyjny Unii Europejskiej zatytułowany *Kapitał Ludzki. Narodowa strategia spójności*, coraz bardziej pomyślne wyniki międzynarodowych badań umiejętności uczniów (Białecki i in., 2003), a także niebywały rozwój ilościowy szkolnictwa wyższego, zaszczepiły w nas przekonanie, że wskutek transformacji ustrojowej kapitał ludzki w Polsce rośnie. Znacznie niżej wypada ocena kapitału społecznego. Alarmujące przykłady jego niedostatku w szkole (Dudzikowa, 2008) i w środowisku (Marzec-Holka, 2009) wywołują zaniepokojenie opinii publicznej. Obydwie oceny dotyczą jednak krajowego systemu edukacji i instytucji takich jak szkoły różnego szczebla, a nie są pochodną i nie prowadzą do metodyki diagnozy kapitału zdobywanego przez uczniów i studentów.

W ekonomii **kapitał** to *zasób, który daje lub potencjalnie może dać zysk, czyli nadwyżkę przychodów nad kosztami.* (Walukiewicz, 2010: 24). Przeważa w nim zysk niepewny, zależny od gry rynkowej, określony słowem „może”. *Rynek jest wszędzie: w zarządzaniu, produkcji, zaopatrzeniu, sporcie, rekreacji, służbie zdrowia, nauce, sztuce, edukacji* (Walukiewicz, 2012: 45n). Każdemu, z wyjątkiem klientów opieki społecznej, przyjdzie stawić czoła ostrej konkurencji. Ani w skali masowej, ani w skali jednostki, nie ma gwarancji, że okoliczności pozwolą nam na zysk z kapitału, w tym z wybranego elementu wykształcenia. To ograniczenie nie przekłada się jeszcze na pedagogikę szkolną, ale zaczyna już być dostrzegane przez publicystów. Znany gdański pisarz, Stefan Chwin, ostrzeża (2013):

*„Nauka to potęgi klucz. Jak się będziesz dobrze uczył, to osiągniesz sukces.” Pięknie, tylko co to ma wspólnego z rzeczywistością? [...] Oni dowiadują się z telewizji, kolorowych gazet, że mają obowiązek być szczęśliwymi ludźmi sukcesu, którymi w większości nie będą. [...] Powinni czytać więcej książek i oglądać więcej filmów, w których byłaby mowa na przykład o tym, jakie kłopoty psychologiczne pojawiają się w relacjach między matką a córką, co to jest związek toksyczny z partnerem, jak rodzi się kryzys przyjaźni między ojcem a synem, czym jest dla pracownika utrata zajęcia, dla dziecka śmierć rodziców czy przewlekła choroba kogoś bliskiego. Mądry nauczyciel powinien rozmawiać z młodymi ludźmi także o takich sprawach, bo prawdziwa szkoła powinna ich przygotować na czołowe zderzenie z życiem.*

Trzeba odróżnić statystyczną prawidłowość sukcesu związanego z wykształceniem, w tym zwłaszcza z sukcesem związanym z wykształceniem rejestrowanym przez wyniki egzaminów zewnętrznych, od obecnej i przyszłej sytuacji jednostki.

Wielu pedagogów wzbrania się przed nazywaniem osiągnięć ucznia kapitałem, a już na pewno – przed porównywaniem kapitału ludzkiego między uczniami, dopatrując się w tym dyskryminacji nieuprzywilejowanych jednostek i warstw społecznych. Ci, którzy pokonają aksjologiczną barierę humanizmu, od razu

zauważą, że egzaminy szkolne obejmują tylko część zakresu pojęcia, którym chcemy się posłużyć. **Kapitał ludzki** bowiem to *wszystkie, ale to absolutnie wszystkie zasoby niematerialne (przymioty) kojarzone z człowiekiem traktowanym jako samodzielna istota ludzka* (Walukiewicz, 2010: 25). Należą do nich (tamże, 25-26):

1. kompetencje i doświadczenie,
2. wiedza i zdolności (intelektualne, artystyczne, przywódcze itp.),
3. zdrowie i wydolność organizmu,
4. nastawienie życiowe (optymizm – pesymizm, otwartość – skrytość, prawdomówność – hipokryzja itp.),
5. zadatki na kapitał społeczny (zaufanie – podejrzliwość, tolerancja – kłótność, lojalność – nielojalność itp.).

Te przymioty są wysoko cenione przez pracodawców, większość z tych cech sprzyja jakości i radości życia, a żadna z nich nie jest obojętna młodzieży. Jednak wprowadzenie ich do oceny szkolnej, w tym do dominującej – oceny przedmiotowej, aczkolwiek teoretycznie możliwe (Niemierko, Walukiewicz, 2012), napawa nas przerażeniem. Czy polonista będzie oceniał kondycję fizyczną, fizyk – samoocenę, a matematyk – uspołecznienie ucznia?

Z drugiej strony, **ocena wąskodydaktyczna**, ustalana wyłącznie na podstawie wyników sprawdzania wiadomości i umiejętności programowych (Niemierko, 2007: 302), także pedagogów nie cieszy. Badania pokazują, że społeczno-wychowawczy składnik ocen wewnątrzszkolnych, choć nie w pełni legalny, jest wszechobecny (Brookhart, 1993). Co więcej, w przewidywaniu przyszłych osiągnięć akademickich ucznia zapewnia on ocenom wewnątrzszkolnym przewagę nad wynikami egzaminów zewnętrznych (Niemierko, 2002: 269; Smith, 2003). Nauczyciele umiejętnie oceniają przygotowanie akademickie uczniów, ale brak badań, które wykazałyby trafność oceny ich kapitału ludzkiego ze względu na wartości inne niż dalsze kształcenie naukowe, a tym bardziej brak nam metodyki diagnozy tego kapitału.

W szkolnej ławce niełatwo zaobserwować zachowania wybiegające poza taksonomię celów kształcenia dziedziny poznawczej. Osiągnięcia emocjonalne, praktyczne i światopoglądowe, stanowiące dowody rozwoju inteligencji wielorakich (Gardner, 2002), mogą zarysować się w toku rozmowy diagnostycznej (Suchańska, 2007), a na pewno zabłyszczą na wycieczce szkolnej, ulubionej przez uczniów metodzie kształcenia. To spostrzeżenie można przyjąć za wytyczną szkolnej metodyki rozwoju i diagnozy kapitału ludzkiego uczniów (Niemierko, 2007: 251-253):

Organizuj zajęcia szkolne na wzór wycieczki w nieznanne:

1. Przygotuj uczniów do wyprawy odkrywczej.
2. Utwórz heterogeniczne zespoły zadaniowe.
3. Pozwól zespołom na samodzielność organizacyjną i poznawczą.
4. Zbierz, oceń i wykorzystaj dorobek uczestników wyprawy.

Te zalecenia brzmią mało rewelacyjnie dla dydaktyków, którym bliski jest dorobek amerykańskiego pragmatyzmu (Dewey, 1972; Dewey, 1988). Nowością natomiast w Polsce byłoby ich planowe zastosowanie do indywidualnej diagnozy edukacyjnej uczniów.



II. Badania edukacyjnej wartości dodanej, prowadzone pod kierunkiem prof. Romana Dolaty, są niewątpliwie najbardziej oryginalnym, zaawansowanym psychometrycznie i konsekwentnie wdrażanym projektem diagnostyki edukacyjnej w naszym kraju.

**Edukacyjna wartość dodana** to przyrost osiągnięć ucznia w wybranym zakresie programowym w pewnym czasie. W badaniach zespołu Dolaty przyjęto jej węższe określenie, ograniczone do porównania średnich wyników grup uczniów w kolejnych egzaminach zewnętrznych. Oto stosowne objaśnienie (Żółtak, w druku):

*Idea wskaźników edukacyjnej wartości dodanej (EWD) opiera się na porównywaniu ze sobą wyników dwóch (lub więcej) egzaminów w celu określenia względnego przyrostu wiedzy poszczególnych uczniów. Wskaźniki te mogą być traktowane jako szczególnie sposób komunikowania wyników egzaminacyjnych dla grup uczniów, w odniesieniu do ich wcześniejszych osiągnięć. Jednym z najistotniejszych, a często wręcz podstawowym, zastosowaniem wskaźników EWD jest ewaluacja pracy szkół (a czasem również nauczycieli), jednak mogą one być wykorzystywane również w innych celach, np. w celu wsparcia rodziców przy wyborze szkoły, do której posłą oni swoje dzieci.*

Artur Pokropek (2009) zbudował model trzyletniego wskaźnika tak rozumianej edukacyjnej wartości dodanej w gimnazjum. Jego model pozwala klasyfikować placówki edukacyjne na szkoły sukcesu (wysoki poziom, wysoka EWD), wspierające rozwój (niski poziom, wysoka EWD), niewykorzystanych możliwości (wysoki poziom, niska EWD) i wymagające pomocy (niski poziom, niska EWD). Tego rodzaju klasyfikacja wnosi do ewaluacji pracy szkoły czynnik rozwojowy, priorytetowy we współczesnej pedagogice (Kohlberg, Mayer, 1993), doceniany także przez władze terenowe (Stożek, 2009).

By administracja oświatowa opanowała nową miarę, potrafiła wykorzystać informację, jaką o pracy szkół przynoszą wskaźniki EWD, a zarazem ustrzec się ich nadinterpretacji, trzeba napisać i wydać podręczniki, w których będą przystępnie objaśnione podstawy teoretyczne tych wskaźników, zasady ich stosowania i odpowiednie ćwiczenia. Przydatne byłyby wydawnictwa popularnonaukowe, zawierające barwne przykłady poprawnej i niepoprawnej interpretacji wskaźników, a nawet wątki fabularne, jak w bardzo przystępnej książce unijnego zespołu autorów propagującej autoewaluację placówki edukacyjnej (MacBeath i in., 2003). Trudno bowiem liczyć na masowy odbiór tekstów pedagogicznych, w których przeważają wzory matematyczne i analizy komputerowe.

Działania Zespołu EWD mogą ożywić nadzór pedagogiczny i ułatwić dyrekcjom szkół argumentację za zwiększaniem wysiłków nauczycieli i uczniów w procesie dydaktycznym. Niewiele natomiast pomogą w budowaniu strategii podnoszenia wartości dodanej wyników uczenia się, bo na to trzeba jej nadać **znaczenie treściowe** (*content meaning*) w postaci rejestru czynności opanowanych przez ucznia uzyskującego określoną liczbę punktów.

Postulat znaczenia treściowego wyniku punktowego pojawił się już w zaraniu pomiaru sprawdzającego (Ebel, 1962) i pozostał koniecznym, choć niełatwym do urzeczywistnienia warunkiem jego trafności. W przypadku wskaźnika

edukacyjnej wartości dodanej, uczeń i nauczyciel muszą dobrze wiedzieć, co jest w obu pomiarach stałe (dziedzina czynności), a co zmienne (poziom biegłości), by móc śledzić postępy nie tylko jako zmianę pozycji w krajowej populacji rówieśników i nie tylko od egzaminu doniosłego do egzaminu doniosłego, mającego, jak wiemy, sumujący, a nie kształtujący charakter.

**Diagnostyka indywidualna edukacyjnej wartości dodanej**, polegająca na oszacowaniach różnicy poziomu umiejętności ucznia w wybranych fazach procesu uczenia się (Niemierko, 2009: 254n), pozostaje wyzwaniem dla teoretyków i praktyków. By ją uprawiać, trzeba pokonać cztery wysokie bariery:

1. **Bariera programowa.** Nauczyciele i uczniowie są na ogół przekonani, że powinno się sprawdzać i oceniać wszystko, co jest kolejno nauczane w szkole, a tymczasem, by oceniać postęp, musimy co najmniej dwukrotnie *mierzyć ten sam konstrukt* (Pokropek, w druku). Na to musi być zachowana ciągłość ćwiczenia odpowiednich czynności, co ogranicza treść EWD do pewnej części zakresu podstawy programowej i materiału wybranych lekcji, a więc odróżnia ten wskaźnik od nauczycielskiej praktyki oceniania osiągnięć uczniów i może obniżać jej wartość .

2. **Bariera komunikacyjna.** Skale standardowe, oparte na normalizacji rozkładu mierzonej właściwości w populacji, są stosowane tylko w pomiarze różnicującym, z założenia odrębnym od praktyki oceniania szkolnego. W ostatnim dziesięcioleciu administracja szkolna w Polsce w pewnym stopniu przywykła do skali staninowej, ale stosowana przez Zespół EWD **skala standardowa IQ** (średnia 100, odchylenie standardowe 15) jest w pedagogice unikatowa, bo stwarza sugestię związku EWD z inteligencją, czyli z cechą z definicji stałą. Z kolei **skala równoważników klasy**, odmierzająca wartość dodaną w latach i miesiącach średniego postępu w populacji (Niemierko, 2009: 255-256), mimo swojskości symboliki („zrobiłeś postęp 3 miesięcy”, „czytasz na poziomie 5. klasy”), często budzi (w USA) nieporozumienia, gdyż wydaje się naruszać system klasowo-lekcyjny (Nitko, 293: 372-380). Jak widać, popularny zwrot „postępy ucznia” opornie poddaje się skalowaniu, które byłoby szerzej użyteczne.

3. **Bariera technologiczna.** By oszacować indywidualną EWD, musimy dysponować co najmniej dwiema bateriami testów unormowanych dla całego okresu jej zastosowań (Niemierko, 1975: 152:158). Takie wydawnictwa są ogólnie dostępne w wielu krajach zachodnich, ale w Polsce uczyniono tylko pewne próby w kierunku ich przygotowania (Mulawa i in., 2007). Rynek testowych wydawnictw egzaminopodobnych jest w Polsce bardzo duży, ale rzadko znajdujemy w nich coś więcej niż teksty zadań i klucze punktowania, prawie nigdy zaś nie zawierają norm empirycznych i nigdy – norm podłużnych. Autorzy tych wydawnictw milcząco zakładają, że użytkownicy testów będą zainteresowani tylko bezbłędnymi wynikami pomiaru i nie potrzebują porównań z innymi uczniami ani z własnym stanem wcześniejszym.

4. **Bariera probabilistyczna.** Błąd pomiaru różnicy między dwoma niezależnymi oszacowaniami osiągnięć ucznia jest tak duży, że wartość dodana w czasie krótszym od jednego semestru ma znikome znaczenie orientujące. Przy typowym współczynniku rzetelności  $r_{tt} = 0,80$ , istotna ( $p = 0,05$ ) różnica

między dwoma wynikami wersji równoległych testu wynosi aż 2,6 stanina. Przy wysokim współczynniku rzetelności  $r_{tt} = 0,90$ , taka różnica wynosi 1,8 stanina (Niemierko, 1975: 241-243). Nawet coroczny pomiar edukacyjnej wartości dodanej miałby więcej z zawodów sportowych, w których dyspozycja danego dnia i łut szczęścia odgrywają wielką rolę, niż z naukowych badań laboratoryjnych. Młodzież w większości chętnie zaakceptowałaby tę rozgrywkę, ale trudniej spodziewać się jej aprobaty ze strony nauczycieli i rodziców. Należy przypomnieć jednak, że półprzedziały ufności dla pojedynczych wyników pomiaru wynoszą odpowiednio 1,8 stanina i 1,3 stanina, a więc także są oszałamiające dla niespecjalistów pomiaru (Niemierko, 2012). Podwójna niepewność zwiększa błąd standardowy tylko o około 40 procent.

Te cztery bariery łącznie mogą zniechęcić teoretyków i praktyków do diagnozowania indywidualnej EWD. Szansa na podniesienie samoregulacji uczenia się i motywacji wewnętrznej uczniów do rozwoju powinna skłonić ich jednak do podjęcia prób w tym kierunku.

## Od jednostki do instytucji

I. Wszyscy znamy taksonomie celów kształcenia, pierwowzór (Bloom, 1956) i jej pochodne. Gdy powstawały kilkadziesiąt lat temu, w psychologii królował **behawioryzm**, oparty na obserwacji pojedynczych zachowań podmiotu jako reakcji na bodźce otoczenia. Mimo zmasowanej krytyki podstaw filozoficzno-psychologicznych taksonomii Blooma (Furst, 1981) długo nie znajdowano bardziej użytecznej hierarchii celów i wyników kształcenia. Dopiero sukces **psychologii poznawczej**, skoncentrowanej na organizacji wiedzy w umyśle człowieka, zmienił wizerunek uczenia i doprowadził do rewizji taksonomii dziedziny poznawczej (Snow, Lohman, 1989; Mislevy, 2007).

Rewizja taksonomii Blooma nie naruszyła układu pamiętanie – rozumienie – stosowanie – rozwiązywanie problemów (analiza, synteza i ocena), ale dodała na szczycie nowy rodzaj wyższych procesów umysłowych: twórczość (Anderson, Krathwohl, 2002). **Twórczość** jest działaniem przynoszącym nowe, a zarazem użyteczne wytwory. Ponieważ narusza autorytety i wygodne wzory, wymaga zwykle pokonania oporu zewnętrznego (Nęcka, 1999). Tak pisze pedagog twórczości Krzysztof Szmidt (2007: 183):

*Uczniowie twórczy są powszechnie myleni z uczniami krnąbrnymi, nieuspołecznionymi, niegrzecznymi lub sprawiającymi nauczycielom wiele trudności wychowawczych. Dzieje się tak dlatego, że niektóre ich zachowania są postrzegane przez nauczycieli nie jako wyraz kreatywności, lecz przejaw nonkonformizmu, skłonności buntowniczych, braku uwagi i szacunku dla dyscypliny klasowej oraz wychowawcy.*

Podobne spostrzeżenia przedstawił Maciej Karwowski (2010). Przeprowadził on solidne reprezentatywne badania inteligencji akademickiej, inteligencji emocjonalnej i zdolności twórczych polskich uczniów, w których stwierdził niskie, ale pozytywne interkorelacje między tymi właściwościami oraz między zdolnościami twórczymi i wewnątrzszkolnymi ocenami uczniów (Karwowski, 2005). Ta ostatnia zależność pozostaje jednak niepewna, gdyż inne badania

pokazały, że szkolni prymusi uzyskują niższe wyniki psychologicznych testów twórczości niż przeciętni uczniowie gimnazjów i liceów (Turska, 2006: 83-84).

**Kreatywność**, jako zespół uzdolnień i zainteresowań twórczych, jest jednym z najwyższej cenionych walorów kandydatów do pracy w korporacjach. Szkoła jednak, szczególnie w krajach posocjalistycznych, jest mocno konserwatywna. Występuje w niej duża rozbieżność między sprawnym myśleniem odtwórczym, cechującym najlepszych uczniów, a myśleniem twórczym, promowanym w gospodarce. Wciąż zdarzają się nauczyciele uprawiający encyklopedyzm i promujący pamiętanie, a więc najniższą kategorię taksonomii celów kształcenia.

**Rozszerzona taksonomia dziedziny poznawczej w postaci:**

1. pamiętanie
2. rozumienie
3. stosowanie
4. analiza
5. ewaluacja
6. twórczość

jest na ogół chętnie przyjmowana za model osiągnięć uczniów. Pedagodzy podkreślają jego szczególną przydatność w **społeczeństwie wiedzy**, wykorzystującym technologię informacyjno-komunikacyjną. Uważa się, że obecne „pokolenie cyfrowe”, dysponujące taką technologią (Prensky, 2001; Jukes, Dosay, 2006), jest predysponowane do działania na nowym szczycie dziedziny poznawczej. Dzięki Internetowi, współczesna młodzież łatwo nawiązuje kontakty i łączy się w zespoły wytwarzające oryginalne pomysły i bardziej niż jednostka zdolne do pokonywania **inhibitorów** twórczości, czyli barier psychicznych wstrzymujących naturalny proces twórczy.

Wciąż brak nam badań i działań systemowych, które pozwoliłyby wdrożyć rozszerzoną taksonomię Blooma do naszej edukacji. Psychologowie skarżą się na niskie kompetencje nauczycieli w zakresie odróżniania zdolności twórczych ucznia od jego inteligencji i motywacji, ale porównują nauczycielskie oceny twórczości z wynikami testów psychologicznych pozbawionych związku z materiałem przedmiotowym, niemierzących twórczości w określonym przedmiocie szkolnym (Karwowski, 2009a; Karwowski, 2009b). Potrzebne są zatem: (1) koncepcje **uczniowskiej twórczości przedmiotowej**, wykraczającej poza poziom programowy danego szczebla szkoły, ale możliwej do uznania w określonej dyscyplinie naukowej, (2) procedury pomiarowe twórczości uczniowskiej w zakresie poszczególnych przedmiotów i szczebli kształcenia, (3) mocne dowody liniowości rozszerzonej taksonomii dziedziny poznawczej i na koniec (4) warsztaty ćwiczące nauczycieli w stosowaniu rozszerzonej taksonomii do projektowania zajęć edukacyjnych i sprawdzania osiągnięć uczniów.

Nie ma przeszkód w posługiwaniu się już teraz rozszerzoną taksonomią Blooma w indywidualnej diagnozie osiągnięć ucznia dokonywanej przez nauczyciela przedmiotu szkolnego. By jednak przejść w tym działaniu na poziom diagnozy instytucjonalnej, potrzebne są badania, które mogłyby odpowiedzieć na następujące pytania:

1. Jak zachęcać uczniów do myślenia twórczego i usuwać jego inhibitory?
2. Jak mierzyć przedmiotową twórczość ucznia na kolejnych szczeblach kształcenia?
3. Jak kształcić nauczycieli promujących myślenie twórcze?
4. Jak myślenie twórcze wpływa na osiągnięcia w zakresie niższych procesów umysłowych?
5. Jak budować zadania twórcze do egzaminu zewnętrznego?
6. Jak podnieść średni wynik szkoły w rozwiązywaniu zadań twórczych?
7. Jak przedmiotowe osiągnięcia twórcze w szkole potwierdzają się w pracy zawodowej?

Bez rozwinięcia frontu podobnych badań metodyka i diagnoza twórczych osiągnięć uczniów pozostanie zaledwie hobbystycznym zajęciem zapaleńców. A to stanowczo za mało.

II. Za najcenniejszy dorobek dydaktyki ogólnej w powojennej Polsce uważam okoniowską **teorię kształcenia wielostronnego**, wiążącą cztery drogi uczenia się: przez przyswajanie, działanie, odkrywanie i przeżywanie. Jej autor uznał nie tylko herbartowską tradycję apercpcji gotowej wiedzy, lecz także kerschensteinerowską szkołę pracy, deweyowkie poszukiwanie wiedzy i nawet sporadyczne koncepcje szkoły twórczej za kierunki jednostronne, a więc zaproponował ich integrację (Okoń, 1967: 81n).

Zwolennicy nowej teorii próbowali przełamać *gnostyczną tradycję w dydaktyce polskiej* (Zaczyński, 1988: 49), wprowadzając zróżnicowane wzory lekcji, zalecając indywidualizację nauczania (Lewowicki, 1977) i definiując **modele uczenia się** (Niemierko, 2009):

1. model alfa (przyswajanie), zorientowany na fakty naukowe, pojęcia i prawa,
2. model beta (działanie), zorientowany na wiadomości i czynności praktyczne,
3. model gamma (odkrywanie), zorientowany na badanie otaczającej rzeczywistości,
4. model delta (przeżywanie), zorientowany na pożytek społeczny.

Te cztery modele zostały oprzyrządowane skalami opisowymi (Niemierko, 2009: 102-104), wstępnie zastosowanymi w kilku pracach licencjackich i w ogólnopolskich badaniach gimnazjalistów ( $n = 1500$ ).

Nasza wiedza ogólnosystemowa na temat modeli uczenia się młodzieży jest jeszcze bardzo skromna. W szkole podstawowej wydają się przeważać modele alfa i gamma (przyswajanie i odkrywanie) w mieście, a beta i gamma (działanie i przeżywanie) na wsi (Roman, 2009; Domżańska, 2009). W gimnazjum wyraźnie przeważa odkrywanie, zwłaszcza wśród dziewcząt i na wsi, ale najlepsze wyniki egzaminu zewnętrznego osiągają uczniowie uczący się przez przyswajanie (Niemierko, 2006). W liceach ogólnokształcących i technicach chłopcy opowiadają się najczęściej za uczeniem się przez działanie, zaś wśród dziewcząt duża grupa wskazuje na przeżywanie (Jaszewski, 2009). Wśród studentów preferencje modelowe wyrównują się, tylko na kierunkach ścisłych utrzymuje się przewaga uczących się przez działanie (Szczepański, 2009).



Z dotychczasowych zastosowań teorii kształcenia wielostronnego niewiele wynika dla diagnostyki edukacyjnej na szczeblu polityki oświatowej i zarządzania szkołą. Lista problemów do rozwiązania w drodze reprezentatywnych badań krajowych mogłaby zawierać następujące pozycje:

1. Jak diagnoza kwestionariuszowa uczenia się potwierdza się w obserwacjach nauczycieli?
2. Jak drogi uczenia się zależą od przedmiotu kształcenia?
3. Jakie drogi uczenia się preferują nauczyciele kolejnych szczebli systemu edukacji?
4. Jaki jest związek dróg uczenia się z motywacją ucznia do osiągnięć szkolnych?
5. Jak osiągnięcia uczniów zależą od wielostronności kształcenia?
6. Jakie różnice modeli uczenia się występują w skali międzynarodowej?
7. Jakie modele uczenia się zapowiadają sukces w pracy zawodowej?

Analizy wyników egzaminów zewnętrznych nie wystarczą. Objęcie instytucji edukacyjnych diagnozą modeli uczenia się i kształcenia mogłoby doprowadzić do wykrycia ważnych prawidłowości dydaktycznych i sformułowania wytycznych do lepszego wykorzystania różnorodności psychologicznej i kulturowej młodzieży.

### Podsumowanie

1. Pół wieku temu ewaluacja w edukacji rozpadła się na **sumującą i kształtującą**, co wyodrębniło dwie dziedziny badań pedagogicznych.
2. Diagnostyka edukacyjna zaczęła od badania indywidualnych **zaburzeń w uczeniu się**, a obecnie sięga do **odpowiedzialności edukacyjnej** instytucji pedagogicznych.
3. Nadzór pedagogiczny przyjmuje **perspektywę instytucjonalną** diagnozy, natomiast nauczyciel – jej **perspektywę personalną**, indywidualną.
4. Diagnoza instytucjonalna jest dokonywana w psychologicznym **Systemie 2**, refleksyjnym, zaś diagnoza indywidualna – głównie w **Systemie 1**, intuicyjnym.
5. Najtrudniejsze, a zarazem najcenniejsze w diagnostyce jest **przechodzenie** od diagnozy instytucjonalnej do indywidualnej i przeciwnie.
6. Z poziomu instytucji na poziom indywidualny warto przenieść problematykę **kapitału ludzkiego** i **edukacyjnej wartości dodanej**.
7. Z poziomu indywidualnego na poziom instytucji warto przenieść problematykę **osiągnięć twórczych** ucznia i **kształcenia wielostronnego**.

### Bibliografia

1. Anderson, L.W., Krathwohl, D. (eds.) (2002) *A taxonomy for learning. Teaching and assessing a revision of Bloom's taxonomy of educational objectives*. New York: Longman.
2. Białecki, I., Blumsztajn, A., Cyngot, D. (2003) *PISA – Program Międzynarodowej Oceny Umiejętności Ucznia*. Warszawa: ZNP.

3. Black, P., Wiliam, D. (1998). Assessment and classroom learning. *Educational Assessment: Principles, Policy, and Practice*, 5.
4. Bloom, B.S. (red.) (1956) *Taxonomy of educational objectives. The classification of educational goals. Handbook I: Cognitive domain*. New York: McKay.
5. Brickell, H. (1976) *The influence of external political factors on the role and methodology of evaluation*. Evaluation Comment, 2.
6. Brookhart S. (1993) *Teachers' grading practices: Meaning and values*. Journal of Educational Measurement, 2.
7. Brookhart, S.M. (2009) *Editorial. Educational Measurement: Issues and Practice* (Special Issue: The Validity of Formative and Interim Assessment), 3.
8. Chwin, S. (2013) *Szkoła nie przygotowuje na zderzenie z życiem*. „Gazeta Wyborcza. Trójmiasto, 21 czerwca.
9. Cronbach, L.J., Snow, R.E. (1977) *Aptitude and instructional methods. A handbook for research on interactions*. New York: Irvington.
10. Dewey J. (1972, wyd. amer. 1916) *Demokracja i wychowanie*. Ossolineum.
11. Dewey J. (1988, wyd. amer. 1933) *Jak myślimy?* Warszawa: PWN.
12. Dolata, R. (red.) (2007) *Edukacyjna wartość dodana jako metoda oceny efektywności nauczania na podstawie egzaminów zewnętrznych*. Warszawa: CKE.
13. Domżańska, P. (2009) *Modele uczenia się w szkole podstawowej i gimnazjum*. AMW, nieopublikowana praca licencjacka.
14. Dudzikowa, M. (2008) *Erozja kapitału społecznego w szkole w kulturze nieufności*. [w:] M. Dudzikowa, M. Czerepaniak-Walczak (red.) *Wychowanie. Pojęcia, procesy, konteksty*. Gdańsk: GWP.
15. Ebel. R.L. (1962) *Content standard test scores*. Educational and Psychological Measurement, 22.
16. Furst, E.J. (1981) *Bloom's taxonomy of educational objectives for the cognitive domain: Philosophical and educational issues*. Review of Educational Research, 441-453.
17. Gardner, H. (2002) *Inteligencje wielorakie. Teoria w praktyce*. Poznań: Media Rodzina.
18. Good, T.L., Brophy, J.E. (1987) *Looking in classroom*. New York: Harper & Row.
19. Husén, T. (1987) *Policy impact of IEA research*. Comparative Education Review, 31.
20. Husén, T., Postlethwaite, T.N. (1996) *A brief history of the IEA*. Assessment in Education: Principles, Policy & Practice, 2.
21. Janowski A. (1989) *Uczeń w teatrze życia szkolnego*. Warszawa: WSiP.
22. Jaszewski, M. (2009) *Uczenie się w technikum i liceum ogólnokształcącym*. AMW, nieopublikowana praca licencjacka.
23. Jukes, I.J., Dosay, A (2006) *Understanding digital children. Teaching and learning in the new digital landscape*. The InfoSsavvy Group.
24. Kahneman, D. (2012) *Pułapki myślenia. O myśleniu szybkim i wolnym*. Poznań: Media Rodzina.
25. Kane, M.D. (2007) *Validation*. [w:] R.L. Brennan (red.) *Educational measurement. Fourth edition*. Westport: American Council on Education – Praeger.
26. Karwowski, M. (2005) *Konstelacje zdolności. Typy inteligencji a kreatywność*. Kraków, 2005.
27. Karwowski, M. (2009a) *Zgłębianie kreatywności. Studia nad pomiarem poziomu i stylu twórczości*. Warszawa: Wyd. Akademii Pedagogiki Specjalnej.
28. Karwowski, M. (2010) *Kreatywny czyli niegrzeczny?* Psychologia w szkole, 5.
29. Karwowski, M. (red.) (2009b) *Identyfikacja potencjału twórczego. Teoria - metodologia - diagnostyka*. Warszawa: Wyd. Akademii Pedagogiki Specjalnej.
30. Karwowski, M. (w druku) *Ścieżki rozwoju edukacyjnego młodzieży – szkoły pogimnazjalne. Trafność wskaźników edukacyjnej wartości dodanej dla szkół maturalnych*.

31. Kohlberg, L., Mayer R. (1993) *Rozwój jako cel wychowania*. [w:] Z. Kwieciński, L. Witkowski (red.). *Spory o edukację. Dylematy i kontrowersje we współczesnych pedagogiach*. Toruń: Edytor.
32. Konarzewski, K. (2004) *Komu jest potrzebna diagnostyka oświatowa?* [w:] B. Niemierko (red.). *Diagnostyka edukacyjna. Teoria i praktyka. Wybór tekstów*. Kraków: PTDE.
33. Koretz, D.M., Hamilton, L.S. (2006). *Testing for accountability in K – 12*. [w:] R.L. Brennan (red.) *Educational measurement. Fourth edition*. Westport: American Council on Education – Praeger.
34. Korporowicz, L. (2012) *Refleksja jako działanie. O stałej potrzebie przełamywania redukcji w pojmowaniu badań ewaluacyjnych w edukacji*, [w:] B. Niemierko, M.K. Szmigel *Regionalne i lokalne diagnozy edukacyjne*. Kraków: PTDE.
35. Kupisiewicz, C. (1985) *Paradygmaty i wizje reform oświatowych*. Warszawa: PWN.
36. Kupisiewicz, C. (red.) (1989) *Raport o kierunkach i stanie edukacji narodowej w Polskiej Rzeczypospolitej Ludowej*. Warszawa: PWN.
37. Lewowicki, T. (1977) *Indywidualizacja kształcenia. Dydaktyka różnicowa*. Warszawa: PWN.
38. MacBeath, J., Schratz, M., Meuret, D., Jakobsen, L. (2003) *Czy nasza szkoła jest dobra?* Warszawa: WSiP.
39. Marzec-Holka, K. (2009) (red.). *Kapitał społeczny a nierówności – kumulacja i redystrybucja*. Bydgoszcz: Wyd. UKW.
40. Meehl, P.E. (1954) *Clinical versus statistical prediction*. Minneapolis: University of Minnesota Press.
41. Meehl, P.E. (1968) *Causes and effects of my disturbing little book*. Journal of Personality Assessment, 50.
42. Mulawa, J., Sokołowska, J., Walewska, J., Sroka, D. (2007) *Bank testów osiągnięć szkolnych. Szkoła podstawowa. Klasa 0-3; Szkoła podstawowa. Klasa 4-6; Gimnazjum. Przedmioty humanistyczne; Gimnazjum. Przedmioty matematyczno-przyrodnicze* Wałbrzych: Instytut Badań Kompetencji.
43. Nęcka, E. (1999) *Proces twórczy i jego ograniczenia*. Kraków: Impuls.
44. Niemierko, B. (1975) *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki obliczeniowe*. Warszawa: WSiP.
45. Niemierko, B., Kowalik, E. (1994) *Diagnostyka edukacyjna*. Gdańsk: Wyd. UG.
46. Niemierko, B. (2002) *Ocenianie szkolne bez tajemnic*. Warszawa: WSiP.
47. Niemierko, B. (2006) *Obszar 3: Uwarunkowania psychologiczne zróżnicowania wyników egzaminów zewnętrznych. Temat 3: Poznawcze hierarchie osiągnięć ucznia (modele alfa, beta, gamma, delta) a wyniki egzaminów zewnętrznych*. Pentor, maszynopis.
48. Niemierko, B. (2007) *Kształcenie szkolne. Podręcznik skutecznej dydaktyki*. Warszawa: WAiP.
49. Niemierko, B. (2009) *Diagnostyka edukacyjna. Podręcznik akademicki*. Warszawa: PWN.
50. Niemierko, B. (2010). *Między prawdą a skutecznością – perspektywy oceniania szkolnego*. [w:] B. Niemierko, M.K. Szmigel (red.) *Teraźniejszość i przyszłość oceniania szkolnego*. Kraków: PTDE.
51. Niemierko, B. (2012) *Jawne i ukryte błędy pomiaru dydaktycznego*. [w:] B. Niemierko i M.K. Szmigel *Regionalne i lokalne diagnozy edukacyjne*. Kraków: PTDE.
52. Niemierko, B. (w druku) *Czy diagnostyka edukacyjna i ewaluacja w edukacji zachowują odrębność?*
53. Niemierko, B., Wałukiewicz, S. (2011) *Jak mierzyć kapitał ludzki? Ewaluacyjne perspektywy operacjonalizacji pojęć ekonomicznych*. [w:] B. Niemierko, M.K. Szmigel (red.) *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Kraków: PTDE.

54. Nitko, A.J. (1983) *Educational tests and measurement. An introduction*. New York: Harcourt.
55. No Child Left Behind (2002). *Pub Law Act No. 107-110*. USA.
56. Okoń, W. (1967) *Podstawy wykształcenia ogólnego*. Warszawa: Nasza Księgarnia.
57. Perie, M., Scott, M., Gong, B. (2009) *Moving toward a comprehensive assessment system: A framework for considering interim assessments*. Educational Measurement: Issues and Practice (Special Issue: The Validity of Formative and Interim Assessment), 3.
58. Pokropek A. (2009) *Metody statystyczne wykorzystywane w szacowaniu trzyletnich wskaźników egzaminacyjnych*. [w:] B. Niemierko, M. K. Szmigel (red.) *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Kraków: PTDE.
59. Pokropek, A. (w druku) *Trafność testów egzaminacyjnych*. [w:] M. Karwowski (red.) *Ścieżki rozwoju edukacyjnego młodzieży – szkoły pogimnazjalne. Trafność wskaźników edukacyjnej wartości dodanej dla szkół maturalnych*.
60. Prensley, M. (2001) *Digital natives, digital immigrants*. NCB University Press, 5.
61. Roman, M. (2009). *Uczenie się w szkole podstawowej na wsi i w mieście*. AMW, nieopublikowana praca licencjacka.
62. Schmidt, K.J. (2007) *Pedagogika twórczości*. Gdańsk: GWP.
63. Schön D.A. (1983) *The reflexive practitioner. How professionals think in action*, New York: Basic Books.
64. Scriven, M. (1967) *The methodology of evaluation*. [w:] R.E. Stake (red.) *Perspectives of curriculum evaluation*. Chicago: Rand McNally.
65. Smith J.K. (2003) *Reconsidering reliability in classroom assessment and grading*. Educational Measurement: Issues and Practice, 4.
66. Snow, R.E. Lohman, D.F. *Implications of cognitive psychology for educational measurement*. [w:] R.L. Brennan (red.) *Educational measurement. Fourth edition*. Westport: American Council on Education – Praeger.
67. Stożek, E. (2000) *Z EWD wśród samorządowców*. [w:] B. Niemierko, M. K. Szmigel (red.) *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Kraków: PTDE.
68. Stróżyński, K. (red.) (2011). *Ocenianie kształtujące po polsku. Kurs dla doradców metodycznych. Scenariusze zajęć*. Warszawa: CODN.
69. Suchańska, A. (2007) *Rozmowa i obserwacja w diagnozie psychologicznej*. Warszawa: WAIp.
70. Szczepański, J. (1973) *Założenia i tezy raportu o stanie oświaty*. Warszawa: Ministerstwo Oświaty i Wychowania.
71. Szczepański, P. (2009) *Trafność teoretyczna modeli uczenia się*. AMW, nieopublikowana praca licencjacka.
72. Turska, D. (2006) *Skuteczność ucznia. Od czego zależy udana realizacja wymogów edukacyjnych*. Lublin: Wyd. UMCS.
73. Walukiewicz, S. (2010) *Kapitał ludzki. Skrypt akademicki*. Warszawa: IBS PAN.
74. Walukiewicz, S. (2012) *Kapitał społeczny. Skrypt akademicki*. Warszawa: IBS PAN.
75. Zaczyński, W.P. (1988) *Metodologiczna tożsamość dydaktyki*. Warszawa: WSiP.
76. Żółtak, T. (w druku) *EWD jako sposób badania efektywności szkół*. [w:] M. Karwowski (red.) *Ścieżki rozwoju edukacyjnego młodzieży – szkoły pogimnazjalne. Trafność wskaźników edukacyjnej wartości dodanej dla szkół maturalnych*.