

prof. dr hab. Bolesław Niemierko

SWPS Uniwersytet Humanistycznospołeczny w Sopocie

Polskie Towarzystwo Diagnostyki Edukacyjnej

Ta para musi się pogodzić! **O potrzebie integracji egzaminów zewnętrznych z ocenianiem wewnątrzszkolnym**

Dokonująca się na szerokim świecie **zmiana funkcji oceniania szkolnego** z klasyfikacyjno-selekcyjnej na diagnostyczno-rozwojową wywołuje zainteresowanie teoretyków i niepokój praktyków. Referat jest próbą wytyczenia drogi, na której uda się pokonać opory, jakie tej zmianie stawiają zarówno znawcy psychometrii, jak i nauczyciele przedmiotów szkolnych.

Dwie miary osiągnięć uczniów

Dwie różne miary to wyniki testów standaryzowanych jako narzędzi egzaminu zewnętrznego (*large-scale assessment*) i stopnie szkolne uzyskiwane w izbie szkolnej (*class-room assessment*). Upraszczając, można je traktować jako swoiste dla dwu dobrze znanych w Polsce rodzajów oceniania: **sumującego**, zamykającego pewien etap kształcenia, i **kształtującego**, dostarczającego danych do indywidualizacji jego dalszego przebiegu (Stróżyński, 2011).

Tak głębokie rozróżnienie prowadzi do pytania o osobną i wspólną wartość edukacyjną tych dwu niezależnych miar (Black i in., 2011). Amerykańska Krajowa Rada Pomiaru Edukacyjnego (*National Council on Measurement in Education, NCME*) wzięła problem na warsztat i zorganizowała (we wrześniu 2017 roku) sympozjum o nazwie „Nadzwyczajna konferencja Krajowej Rady Pomiaru Edukacyjnego na temat oceniania wewnątrzszkolnego i masowej psychometrii: **Ta para musi się pogodzić**” (*NCME Special Conference on Classroom Assessment and Large-Scale Psychometrics: The Twain Shall Meet*). Przyjmijmy hasło tej konferencji za wytyczną dalszych rozważań.

Ocenianie wewnątrzszkolne i ocenianie zewnętrzne uznano za niedobrane małżeństwo, w którym każda z dwu osób ma inne zadania, inne obyczaje i poglądy, inny język. Nauczyciele traktują ocenianie jako składnik procesu kształcenia, kierują się w **nim własnym doświadczeniem i poczuciem racji**. Egzaminatorzy stosują teorię pomiaru psychologiczno-dydaktycznego, a czynniki kontekstowe osiągnięć uczniów zaliczają do źródeł błędu pomiaru. Nauczyciele wytykają egzaminatorom nieznaną lokalnych warunków kształcenia, a egzaminatorzy nauczycielom – brak obiektywizmu.

Przewagą nauczycieli jest znajomość danych uczniów, a przewagą egzaminatorów – znajomość psychometrii w jej szkolno-przedmiotowym wydaniu. Treść oceny nauczycielskiej jest wieloaspektowa, co nadaje jej charakter „hybrydowy” (Szyling, 2013), a treść testowania jest dokładnie wyznaczona planem zastosowanego narzędzia (Niemierko, 2011).

Trafność wewnętrzna testu standaryzowanego, polegająca na porównaniu jego treści z **wykazem umiejętności i wiadomości ucznia zawartym w podstawie programowej**, może być jednak edukacyjnie zwodnicza. Tak oto pisze Rick Stiggins w książce *Doskonały system oceniania* (Stiggins, 2017, s. 84):

W naszej wspólnocie pomiarowej myśleliśmy bardzo wąsko, widząc doskonałość oceniania w kategoriach jego rzetelności, trafności i bezstronności. Na tej liście kryteriów brak było wpływu oceniania na ucznia, czynnika o krytycznym znaczeniu emocjonalnym, a tym samym edukacyjnym. Moglibyście mi pokazać najtrafniejsze i najrzetelnwsze ocenianie na świecie, ale jeśli jego wyniki wiodą uczniów do popadania w beznadziejność, nie mogę go uznać za doskonałe ocenianie, skoro bardziej szkodzi niż pomaga. Innymi słowy, kwestia jakości oceniania nie może być zasadnie rozważana bez wnikania w jego kontekst i przebieg oraz wpływ, jaki faktycznie ma na uczenie się.

Wobec testu reprezentującego podstawę programową słabszy uczeń odczuwa **bezbardność intelektualną**, wyrażającą się upadkiem wiary w skuteczność uczenia się (Sędek, 1995). Ta **bezbardność** ma charakter wyuczony, jest pochodną porażek poniesionych w kolejnych próbach sprostania zbyt wysokim wymaganiom (Ciżkowicz, 2009).

Wskutek braku dopasowania między **odgórnie regulowanym** ocenianiem osiągnięć uczniów, wyznaczonym obowiązującym programem kształcenia, a indywidualnym rozwojem uczniów zawodzi **trafność prognostyczna** standaryzowanych testów osiągnięć szkolnych. Już trzydzieści lat temu amerykański kandydat na studenta czytał o Testach Uzdolnień Akademickich (*Scholastic Aptitude Tests*), najszerzej znanych w świecie bateriach testów dydaktycznych, następujące objaśnienia (*College...*, 1986, s. 8):

Test Uzdolnień Akademickich (SAT) jest testem wyboru wielokrotnego z częścią językową i matematyczną. **Zadania językowe sprawdzają twój słownik, rozumowania werbalne i rozumienie, co czytasz. Zadania matematyczne mierzą twoją zdolność do rozwiązywania problemów z arytmetyki oraz z elementarnej algebry i geometrii.** SAT nie mierzy innych czynników i zdolności – takich jak twórczość, szczególne talenty i motywacja – które mogą także ci pomóc na wyższej uczelni. [...] Twoje świadectwo szkolne jest zapewne najlepszym pojedynczym wskaźnikiem twoich przyszłych sukcesów na uczelni, ale kombinacja twoich stopni w szkole średniej i wyników testu jest jeszcze lepszym wskaźnikiem.

Uznanie dla **oceniania naturalnego**, dokonywanego własnymi środkami nauczyciela, wyrażone w oficjalnym wydawnictwie instytucji egzaminacyjnej, może zaskoczyć pedagogów nie dość biegłych w psychometrii, aby wyważyć zalety i ograniczenia testów. O tych ograniczeniach trzeba także przypominać psychologom dokonującym diagnoz rozwoju młodzieży.

Jerzy Brzeziński, nasz czołowy metodolog badań psychologicznych, ostrzega przed bezkrytycznym zaufaniem do wyniku testowania, zaleca, by takie „dane cząstkowe” **poddawać pogłębionej interpretacji psychometrycznej i konfrontacji z informacją kontekstową**, narzeka przy tym, że *kształcenie pedagogów w Polsce nie wyposaża absolwentów w kompetencje psychometryczne i odnoszące się do diagnozy psychologicznej* (s. 26). To ostatnie spostrzeżenie wydaje się trafiać w sedno obecnych potrzeb diagnostyki edukacyjnej.

Ocenianie jako motywacja

Głównym celem rekonstrukcji systemu oceniania szkolnego zapowiadanej hasłem *Ta para musi się pogodzić* jest zapewnienie mu korzystnego wpływu na wszystkich uczniów. Kluczowy (8) rozdział książki Stigginsa, z którego został zaczerpnięty pierwszy z cytatów w tym referacie, ma tytuł *Zastosowanie oceniania do motywowania **wszystkich** uczniów* (to „wszystkich” jest podkreślone w oryginale).

Motywacja do uczenia się, najcenniejsza dyspozycja ucznia, była przedmiotem wielu eksperymentów psychologicznych, których wyniki można streścić następująco (Łukaszewski, 1972; Łukaszewski, 2002):

Informacja zwrotna a jakość działania

- A. Gdy uczący się określonej czynności nie uzyskuje informacji o jej opanowaniu, motywacja do jej wykonywania i poziom osiągnięć spadają.
- B. Gdy uczący się otrzymuje jedynie nagany (gorzej) lub jedynie pochwały (lepiej) oraz pochwały i nagany w różnych kombinacjach według skali stopni szkolnych, pojawia się **orientacja na sukces** (stopień), a jakość czynności nie wzrasta lub wzrasta niewiele.
- C. Gdy uczący się otrzymuje stopień i komentarz na temat osiągnięć, przeważa orientacja na sukces, a jakość osiągnięć wzrasta umiarkowanie.
- D. Gdy nauczyciel stosuje jedynie komentarz dydaktyczny, odkładając klasyfikację na późniejszy etap, pojawia się **orientacja na zadanie** i jakość czynności mocno wzrasta.

Informacja o wyniku własnego działania, czyli **informacja zwrotna**, jest niezbędna do podnoszenia jakości uczenia się. **Ocenianie szkolne** obejmuje komentarz do osiągnięć ucznia i **stopień, na jaki zasłużył. Komentarz jest bogatszy treściowo niż stopień i mniej podatny na zniekształcenia w odbiorze, a stopień jest tylko uproszczoną, stenograficzną odpowiedzią na pytanie, jak dobrze się uczył** (Brookhart, 2017).

Oto co napisała bohaterka moich książek, **pani Humańska, pod wypracowaniami czworga licealistów na temat *Czy warto być dorosłym/dorośli? List do rówieśników***:

(Andrzej) *Twój list jest żywy i konkretny. Na przyszłość pamiętaj jednak o dobrym uporządkowaniu treści: od faktów do wniosków. Unikaj wyrażen zbyt potocznych – jak „wygłupiać się”, „ubabrać”, „mieć ubaw” – zastępuj je bardziej starannymi, jak „ośmieszać się”, „wybrudzić”, „rozbawić”. Dbaj też o to, by kolejne zdania wiązały się z poprzednimi. Dostateczny.*

(Beata) *Pisziesz poprawnie i dobrze wykorzystujesz przykłady z lektur szkolnych, zwłaszcza z „Kamieni na szaniec” A. Kamińskiego. Mamy jednak już inne czasy i rówieśnicy chcieliby pewnie poznać Twoje własne obserwacje i oceny. Staraj się wykorzystać posiadaną wiedzę do śmiałego planowania swego życia i działania. Dobry.*

(Celina) *Twój list jest odważny i oryginalny. Jesteś krytyczna wobec otoczenia i zaangażowana w korzystne zmiany. Pisząc go, musisz jednak ograniczyć swoją żywiołowość na rzecz spokojnej, precyzyjnej argumentacji. Pamiętaj też o tym, by każdy swój tekst przeczytać po jego napisaniu i poprawić błędy. Dobry.*

(Dariusz) *Przeniosłeś zagadnienie w dziedzinę filozofii i przedyskutowałeś sens pytania o wartość kolejnych faz życia. Wprowadzając do listu takie pojęcia jak „alienacja” i „relacje społeczne”, powinieneś jednak zadbać o czytelnika i dostarczyć mu prostych definicji tych pojęć, a przykłady z życia ubarwiłyby tekst. Bardzo dobry.*

O tych czterech recenzjach mamy następujące spostrzeżenia:

1. Treść komentarza. **Komentarz do osiągnięć ucznia** dotyczy umiejętności wykazanej przez ucznia, odnosi się do jego pomysłu na odpowiedź i do formy realizacji tego pomysłu. Nie jest tylko uzasadnieniem stopnia szkolnego, lecz nieformalną **diagnozą rozwojową**, rozpoznaniem stylu i etapu zdobywania kompetencji oraz **strefy najbliższego rozwoju**, w której postęp jest możliwy w stosunkowo krótkim czasie (Wygotski, 1971). Taki komentarz niekoniecznie musi mieć postać pisemną, jest pożądanym także w formie ustnej po ustnych wypowiedziach uczniów. Ważne są: (1) intencja wsparcia pracy ucznia, (2) dostrzeżenie jego cech indywidualnych oraz (3) wskazanie mu drogi do wyższych osiągnięć.
2. Zróżnicowanie. Andrzej wciąż jeszcze się uczy posługiwać językiem literackim, Beata ma trudności z **formułowaniem własnego zdania**, Celina nie dba o argumentację, Dariusz zatapia się w dialektyce i zapomina, do kogo adresuje wypowiedź. Mamy tu wszystkie poziomy kompetencji licealistów: od opóźnienia rozwoju ogólnohumanistycznego do „klątwy wiedzy”, polegającej na tym, że *autorowi po prostu nie przychodzi do głowy, że czytelnicy nie wiedzą tego co on* (Pinker, 2016, s. 83). Nie byłoby pożytku z mylenia poziomów ani też traktowania wszystkich uczniów jednakowo, bo potrzeby rozwojowe – a zatem i **motywacyjne** – **poszczególne** uczniów są różne.
3. Przewaga aprobaty. W każdym przypadku pierwsze zdanie komentarza pani Humańskiej wyrażało uznanie dla włożonej pracy i osobistego stylu wypowiedzi ucznia. Ta kolejność jest ważna, bo utwierdza ucznia w przekonaniu o przyjaznym traktowaniu przez nauczyciela i otwiera go na przyjęcie uwag o niedociągnięciach. Uczeń nie powinien mieć wątpliwości co do tego, że nauczyciel ceni go jako osobę i **dostrzega wysiłek**, by podołał wymaganiom.
4. Stopień na koniec. Uczniowie i **ich rodzice z reguły przeceniają stopnie**, interpretując je chętnie po swojemu. Rola motywacyjna stopni jest w ich wyobrażeniach mocno uproszczona: mają „zmuszać” posiadacza do większego wysiłku lub do utrzymania poziomu. By jednak stały się istotnym elementem **samoregulacji** w postaci świadomych strategii poznawczych (Ledzińska, 2000), informacja o osiągnięciach musi być głębsza. Uczeń uzyskuje ją we wszelkich komentarzach do swoich wystąpień ustnych, pisemnych i ruchowo-sprawnościowych. Stopnie szkolne są klamrą spinającą komentarze i bez nich byłyby nazbyt zakodowanym przekazem. Gdyby zaczynać ocenę od stopnia, byłby tylko dodatkiem do głównej informacji, odbieranym jako jej uzupełnienie.
5. Rady, nie nakazy. Zauważmy, że w **zaleceniach pani Humańskiej nie pojawiło się słowo „musisz”**. Jej wskazówki zawierają wytyczne do poprawy jakości ich oryginalnych umiejętności, nie zaś przypomnienie wymagań i rejestracje błędów. Nauczycielka chce pomóc uczniom w dostępnym dla

nich sposób, wierząc, że znajdują w sobie energię na dokonanie postępu. Gdyby okazało się, że zalecenia są nieskuteczne, będzie szukać innych podejść w swym warsztacie metodycznym. Doradzanie, jak się uczyć, jest pośrednim zarządzaniem motywacją młodzieży do nauki. Gdy jest trafne, owocuje wzrostem motywacji.

Reżim stopni szkolnych

Peter Kroepe, wieloletni zagraniczny członek PTDE, już przed ćwierćwieczem sygnalizował nam niepożądaną wszechwładzę skali stopni szkolnych. Oto skrót jego opowieści (Kroepe, 1994):

W 1989 roku w Szlezewiku-Holsztynie, jednym z niemieckich landów, wprowadzono ustawą parlamentarną ocenę opisową jako obowiązkową w trzech początkowych latach nauki i zalecaną w klasach wyższych. Po pięciu latach doświadczeń 90% rodziców zażądało tam przywrócenia stopni szkolnych jako bardziej zrozumiałych. Podjęto więc próby instruowania rodziców i nauczycieli w zakresie interpretacji oceny opisowej, ale nie przyniosły one zadowalających wyników.

To wydarzenie nie powstrzymało ekspansji **oceny opisowej** jako nieformalnej informacji o wyniku kształcenia z rozbudowanym komentarzem pisemnym. Taki sposób oceniania jest obecnie stosowany w wielu krajach dla ułatwienia najmłodszym uczniom wstępu do sformalizowanych systemów ocen, takich jak skala stopni szkolnych. Protesty rodziców przeciw ocenie opisowej zdarzały się i w Polsce, ale wkrótce ucichły w obliczu innych problemów, w tym wyznaczenia wieku podjęcia obowiązku szkolnego. Ujawniły jednak słabości oceny opisowej i potrzeby pedagogiczne, których ona nie zaspokaja, w tym przede wszystkim trudność językową i brak klucza do prostych porównań.

Prześledźmy zalety **stopni szkolnych** jako miar osiągnięć uczniów o znaczeniu przyjętym w danym systemie edukacyjnym.

1. Łatwość porównań. Stopnie szkolne tworzą skalę porządkową, opartą na hierarchii wartości. Porównywanie jakości osiągnięć uczniów jest fundamentalnym założeniem tej skali. Wyższy stopień oznacza wyższą jakość, a ten sam stopień – jakość zbliżoną. Ponieważ skale stopni szkolnych – ze względu na małą dokładność dokonywanych diagnoz – są zazwyczaj krótkie, wielu uczniów uzyskuje ten sam stopień, co w statystyce nazywamy „słabym uporządkowaniem” zbioru danych. w edukacji najważniejsze jest odróżnienie zaliczenia od niezaliczenia przedmiotu, wytwarzające szczególne napięcie w tym punkcie skali. To zjawisko jest psychologicznie niekorzystne, ale trudno mu zapobiec.
2. Każdy adresat. Prostota skali pozwala na jej zastosowanie w **niezmienionej** formie wobec każdego odbiorcy komunikatu o osiągnięciach ucznia. Wyjątkiem są najmłodsi uczniowie, skłonni do przeceniania trafności stopni i ich **znaczenia wychowawczego**, toteż zaleca się **ostrożne** wprowadzanie sformalizowanego oceniania w **szkołach podstawowych**. Zwykle kolejne pokolenia w kraju posługują się tą samą skalą, a jej modyfikacja lub zmiana wywołuje protesty, jak w **przykładzie** zaczerpniętym z doświadczeń naszych zachodnich sąsiadów.

3. Stały schemat. Liniowa konstrukcja skali nie może już być uproszczona. Gdy nauczyciel ocenia określone czynności ucznia, pozostaje mu tylko wędrówka w górę i w dół hierarchii ich wartości, tak by zatrzymać się na poziomie odpowiedniego stopnia. Można natomiast wyróżnić nauczycieli surowych, którzy doszukują się błędów, by stopień obniżyć, i nauczycieli pobłażliwych, wypatrujących zalet pracy ucznia, by stopień podwyższyć.
4. Wielowymiarowość stopnia. Po latach prób uwolnienia treści ocen szkolnych od emocjonalnego aspektu uczenia się psychologowie uznali, że ich zaletą jest **wielowymiarowość**, w tym przede wszystkich związek poznania z **motywacją** (Brookhart i in., 2016). **Przyjmujemy, że motywacja do uczenia się jest też wynikiem uczenia się i stanowi zadatek na przyszłe osiągnięcia.** Składnikami oceny szkolnej są więc oszacowania osiągnięć poznawczych i emocjonalno-motywacyjnych w proporcjach stosownych do potrzeb rozwojowych uczniów.
5. Rzeczowość uzasadnień. Komentarz do stopnia jest potrzebny, ale powinien być równie prosty jak konstrukcja skali. Polega na porównaniu osiągnięć ucznia ze **standardami wymagań programowych**, to jest normami osiągnięć uznanych za niezbędne na danym poziomie kształcenia. Standardy mają zwykle zasięg krajowy, ale mogą być rozwinięte i sprecyzowane przez szkołę i nauczyciela w porozumieniu z władzami oświatowymi. Bardziej szczegółowy opis osiągnięć ucznia może być potrzebny w rozmowach z rodzicami uczniów, którzy nie znają programu kształcenia i nie mają możliwości obserwowania zajęć szkolnych.
6. Uogólniona opinia. Wydarzenia szkolne znikają w **treści stopnia szkolnego, z czego niektórzy uczniowie mogą być zadowoleni i tylko bardziej zasłużeni w pracy nad sobą chętnie prowadzą teczkę prac, w których zbierają dowody swoich sukcesów.** Uproszczona, obiegowa opinia w postaci najczęściej uzyskiwanego stopnia łatwo staje się **stereotypem**, ujednoliconą kategorią zbiorową. Stereotyp „ucznia dostatecznego”, bez ambicji i **polotu, oraz „ucznia bardzo dobrego”, zawsze w pełni przygotowanego, może stać się etykietą wygodną w masowej edukacji, lecz także źródłem szkodliwych uprzedzeń** (Konarzewski, 1991).
7. Subiektywna pewność. Ocenianie osiągnięć uczniów w **stopniach szkolnych jest mocno oparte na intuicji**, przeświadczeniu pojawiającym się w świadomości nauczyciela nagle i z dużą mocą bez pełnego dowodu prawdziwości (Bruner, 1964). Jest pozbawione szacowania błędu pomiaru, jakie obowiązuje w **statystycznych analizach wyników testowania osiągnięć.** Odbiorca oceny musi więc odbierać ją także intuicyjnie, opierając się na zaufaniu do nauczyciela, a powoływanie się na fakty odgrywa tu tylko rolę pomocniczą. Nie da się w pełni uzasadnić stopnia szkolnego, choć można kwestionować jego trafność.

Często zdarza się, niestety, że ocenianie wewnątrzszkolne ogranicza samodzielność uczniów. Ilustruje to wyznaczenie pewnej australijskiej nauczycielki (Tripp, 1996):

Mam taką uczennicę, Kasię, która nigdy nie zacznie pracy nad tym, co poleciłam klasie, dopóki nie podejść do niej, aby wyjaśnić parę błahostek. Jeszcze nie skończę tłumaczyć czegoś klasie, a już ręka Kasi jest w górze. To mnie naprawdę czasami irytuje.

Objaśnienie tego przypadku jest następujące:

1. Czynności przymusowe. w **obliczu nowych zadań Kasia odczuwa niepokój**, przykre napięcie, które może być usunięte tylko osobistym kontaktem z nauczycielką. Jej stan może być zaliczony do kategorii uzależnień, utraty kontroli nad swoimi potrzebami wskutek nawyku ich zaspokajania w pewien sposób. Ludzie wpadają w uzależnienia nieświadomie, stosując **mechanizmy obronne**, rzekome usprawiedliwienia swoich zachowań. Swoje natrętne pytania Kasia tłumaczy chęcią **wzorowego wykonania** nowych zadań. Mechanizmy obronne umacniają uzależnienia i utrudniają ich terapię (Woronowicz, 2009).

2. Źródło uzależnienia. Przypadek Kasi stał się przedmiotem zespołowej analizy pedagogicznej z następującym wynikiem:

Gdy omawialiśmy zdarzenie, nauczycielka (Kasi) wyraźnie zdała sobie sprawę, że najprawdopodobniej zbyt mocno akcentowała wobec uczniów żądanie, by pracowali dokładnie tak, jak sobie życzyła. Postanowiła więc uczyć tak, by uczniowie poczuli się mniej od niej zależni, żeby ich prac były bardziej samodzielne i zróżnicowane (Tripp, 1996, s. 57).

3. Skutki oceniania. Ocenianie osiągnięć uczniów, najgłębiej spersonalizowana czynność nauczyciela w **edukacji masowej, należy do najbardziej uzależniających. Uzależnieniem od oceny szkolnej** będziemy nazywać przemożną potrzebę potwierdzeń obrazu siebie jako ucznia. Może przejawiać się niską motywacją do uczenia się tego, co „nie liczy się do oceny”, stałym oczekiwaniem wysokich i **coraz wyższych ocen** (Turska, 2006) lub **rezygnacją z ambitnych celów rozwojowych** (Kwieciński, 1990; Kwieciński, 2002a). w **każdej** z tych postaci jest szkodliwe, bo zwęża horyzonty i osłabia motywację wewnętrzną na rzecz zachowania uzyskanej pozycji.

Oswajanie egzaminów

Oto wspomnienia sprzed sprawdzianu (A) i opinie szóstoklasistów po jego zakończeniu (B):

(A)

Całą noc nie mogłam spać. Bolała mnie głowa, byłam cała w szoku. Kiedy wstałam rano, wszystko wypadło mi z rąk. Rozbiłam talerz. z tego zdenerwowania nic nie zjadłam. Trzęsły mi się ręce i nogi, a serce stanęło w gardle. Wszystko, czego nauczyłam się w ciągu ostatnich trzech lat, gdzieś uciekło. w głowie miałam zupełną pustkę. Prosiłam mamę, żeby nie iść do szkoły, ale zdałam sobie sprawę, jakie mogą być konsekwencje (Groenwald, 2004).

(B)

- *Sprawdzian, który pisaliśmy, był ważnym wydarzeniem w szkole, jak i moim życiu.*
- *Sprawdzian podobał mi się bardzo, ponieważ u nas na wsi [też] piecze się chleb. Najtrudniejsze zadanie to gdy trzeba było obliczyć, że piekarz się pomylił.*
- *Uważamy, że ten test był dostosowany do możliwości ucznia zdolnego. Szkoda! Przecież wszyscy wiemy, że większość klasy to uczniowie przeciętni.*

- *Nauka w szkole podstawowej to przecież nie tylko język polski i matematyka! Są też inne przedmioty, którymi uczniowie się interesują! A może trzeba się uczyć tylko języka polskiego i matematyki? Prosimy o więcej zadań z przyrody.*
- *Proszę egzaminatorów, by nie oceniali zbyt surowo sprawdzianów. Proszę o wyrozumiałość dla uczniów (Szmigel, 2004).*

Te dwa fragmenty ilustrują znaczenie egzaminu zewnętrznego dla ucznia, jego uczestnika. A oto interpretacja ich treści:

1. Siła oddziaływania egzaminu. Egzamin zewnętrzny jako **egzamin doniosły**, o wynikach wpływających na dalsze losy uczniów, stanowi jedną z podstaw odpowiedzialności edukacyjnej nauczycieli i szkół. Jego **efekt zwrotny**, wpływ na treść i formę uczenia się i kształcenia, jest uważany za przemożny, ale koszt psychologiczny uczniów nie zawsze jest doceniany. Fragment A go wyolbrzymia, bo dotyczy apogeum przeżyć w przededniu życiowej próby. Fragment B, zawierający oceny już odbytego egzaminu, można uznać za bardziej miarodajny. Uczniowie, których wypowiedzi ankietowe przytoczono, proszą o dopasowanie treści egzaminu do ich zainteresowań i o wyrozumiałość w interpretacji uzyskanych wyników.
2. **Wzbudzenie lęku**. Fragment A ilustruje atak lęku egzaminacyjnego, niepokoju o sprawność psychiczną w czasie egzaminu i o wynik, jaki uda się uzyskać. w myślach uczniów pojawia się widmo porażki, które może, ale nie musi zniknąć po przeczytaniu zadań. Potrzebują życzliwości w toku przygotowań, tymczasem w intencji mobilizacji do pracy bywają – w domu rodzinnym i w szkole – straszeni egzaminem. Obycie uzyskane w egzaminach powszednich zmniejsza lęk, ale go nie usuwa, bo egzamin doniosły ma znacznie większą wagę. Tylko uczynienie go przyjaznym dla przeciętnego i słabszego ucznia może zapobiec utrwalaniu się lęku jako zjawiska towarzyszącego przyszłemu zdobywaniu kwalifikacji.
3. Narzędzie selekcji. Sprawdzian w VI klasie szkoły podstawowej był pomyślany jako zobiektywizowana diagnoza, ale dość szybko przybrał postać egzaminu doniosłego. Wprawdzie wszyscy uczestnicy egzaminu mieli prawo do nauki w gimnazjach, ale szybko wytworzyła się konkurencja nierejonowych placówek o najlepszych uczniów, a także konkurencja uczniów o najlepsze gimnazjum. *Dla szóstoklasisty, który nie dostał się do wymarzonego gimnazjum, to jest mały koniec świata* – skarży się jeden z nich (za: Niemierko, 2009, s. 267). Pojawiła się zatem **selekcja**, to jest wybór jednych, a odrzucenie innych kandydatów do określonych form kształcenia (Dolata, 2008). Na wyższych poziomach systemu edukacji selekcja zaostrza się i najlepsi mają prostą drogę do uprzywilejowanych zawodów, a najsłabsi są skazani na marginalizację (Kwieciński, 2002a, s. 32). Selekcja ma charakter sumujący, bo w większości przypadków jej wyniki są nieodwracalne (Kwieciński, 2002b).
4. Redukcja programu kształcenia. Zwężanie się zakresu kształcenia do treści obejmowanej egzaminami zewnętrznymi jest ich najsilniejszym efektem (Smith i Rottenberg, 1991). To zjawisko wynika z „połączenia bezszwowego” kształcenia z egzaminem (Nitko, 1998), dającego prze-

wagę egzaminowi jako czynnikowi wspólnemu systemowi oświatowemu i w pełni obowiązkowemu, nad ocenianiem naturalnym. Wśród nauczycieli anglosaskich krążą anegdoty o tym, jak po każdej próbie wprowadzenia na lekcję oryginalnego wątku podnoszą się ręce pilnych uczniów: Przepraszam, czy to będzie na egzaminie? Również rodzice domagają się, by „nie tracić czasu” na jakiegokolwiek kwestie uboczne, więc nauczyciele wykraczający poza podstawę programową muszą się z tym zawsze liczyć.

Za redukcją treści kształcenia podąża ujednoclenie jego metod. Skoro pisemne formy egzaminów są sprawniejsze i **częściej stosowane niż formy laboratoryjne**, to pokazy i eksperymenty ustępują metodom wąsko intelektualnym z ewentualnym wykorzystaniem symulacji komputerowych. Przestrzeń twórczości własnej uczniów i nauczycieli kurczy się pod wpływem egzaminów doniosłych, zwłaszcza w jej aspekcie emocjonalnym.

Bardzo istotną dla ucznia różnicę między testami standaryzowanymi a sytuacją oceniania naturalnego stwarza obecność w tych pierwszych **zadań zamkniętych**, z gotowymi odpowiedziami do wyboru. Ich właściwości przedyskutujemy na przykładzie jednego zadania, wybranego z cytowanego już amerykańskiego zbioru ze względu na swojską dla nas tematykę:

Wskaż najlepsze sformułowanie podkreślonej części zdania: Joseph Conrad urodził się i wychował w Polsce i napisał wszystkie swoje powieści po angielsku.

- (A) Joseph Conrad urodził się i wychował w Polsce i
- (B) Joseph Conrad, będąc urodzonym i wychowanym w Polsce,
- (C) Mimo że będąc urodzony i wychowany w Polsce, Joseph Conrad
- (D) Joseph Conrad urodził się i wychował w Polsce, ale
- (E) Będąc z Polski, gdzie urodził się i wychował, Joseph Conrad

(College..., 1986, s. 49)

To zadanie **wyboru wielokrotnego** z jedną **odповідzią prawidłową**, uznaną za najlepszą, daje nam okazję do następujących komentarzy:

1. Treść zadań. Łatwo zaznaczyć jedną z pięciu liter na karcie odpowiedzi, ale nie to jest istotą zadania wyboru wielokrotnego. **Treść zadania testowego** stanowią czynności, które powinny być wykonane w celu jego rozwiązania. Przytoczone zadanie sprawdza umiejętność redagowania tekstu pisemnego, a więc mieści się w **wysokiej kategorii celów kształcenia**: stosowania wiadomości w **sytuacjach problemowych, dla ucznia nowych**. Obydwa mocno wykraczają poza poziom pamiętania wiadomości. Pytania o fakty z życia Konrada Korzeniowskiego byłyby czynnościowo uboższe, a ich zamknięcie służyłoby tylko ułatwieniu punktowania.
2. Walory sytuacyjne. Najlepsze zadania zamknięte są głęboko osadzone w **potrzebach i doświadczeniach ucznia**. **Odpowiednią informację wprowadza trzon zadania**, objaśniający czynności i sytuację, w jakiej mają być wykonane. W zadaniach wyboru wielokrotnego trzon zadania jest często dość rozbudowany, czasem wspólny dla pewnej grupy zadań, by skrócić czas ich czytania. Przytoczone zadanie przypomina opowiadanie życiorysów w szkole i w rodzinie. Jest umiarkowanie trudne, bo rozwiązało go

- poprawnie tylko 44% absolwentów amerykańskiej szkoły średniej. Ten wynik dowodzi, że elementy sytuacyjne nie ułatwiają rozwiązania zadań wyboru wielokrotnego, a często wręcz przeciwnie – utrudniają je.
3. Wybór odpowiedzi. Typowe zadanie wyboru wielokrotnego składa się z trzonu i odpowiedzi do wyboru, wśród których mamy odpowiedź prawidłową, zwykle jedną, i 2–4 **dystraktory** wykazujące pozory prawidłowości lub stanowiące alternatywne, lecz słabsze rozwiązanie. Dystraktory atrakcyjne, wybierane przez dużą liczbę badanych, uzupełniają treść trzonu i zaostrzają problem. w przytoczonym zadaniu atrakcyjna jest – w Stanach Zjednoczonych i w Polsce – imiesłowowa konstrukcja zdania złożonego (dystraktory B, C i D), a wielu uczniów nie dostrzega większej wyrazistości przeciwstawienia zdań prostych (... , ale ...) niż ich następstwa wyrażonego spójnikiem „i” (dystraktor A). Atrakcyjność dystraktorów mogą zwiększać zawarte w nich **ukryte wskazówki**, charakterystyczne cechy formalne, jak np. długość, zgodność z trzonem lub sprzeczność w pewnej parze, ale tego nie ma w dyskutowanym zadaniu.
 4. Wpływ doświadczeń ucznia. Ceną, jaką płacimy za zbliżenie treści zadań do życia, jest uzależnienie ich odbioru od pozaszkolnych i szkolnych doświadczeń uczniów. Dzieci imigrantów mogą silniej odczuwać kontrast kultur będący założeniem zadania o naszym rodaku w Anglii. Życiorys autora *Lorda Jima* mógł pojawić się lub nie pojawić na lekcjach literatury angielskiej. Duży błąd losowy pomiaru umiejętności pisarskich ucznia jest nieunikniony, ale to nie powód, by w testach ograniczać się do reprodukcji faktów i teorii naukowych. Uwarunkowanie rozwiązań doświadczeniem życiowym występuje także w zadaniach otwartych, ale w zadaniach zamkniętych jest uważane za **stronniczość** (Hornowska, 1999). w **każdym przypadku należy sprawdzić, czy bezstronność** oceniania osiągnięć uczniów nie jest zbyt silnie zagrożona przez związek sytuacji zadaniowej z ich życiem.
 5. Niezbędne obycie testowe. Zadania zamknięte są konstrukcją sztuczną w tym sensie, że w **życiu codziennym nieczęsto mamy do czynienia z wyborem** między gotowymi rozwiązaniami. Jeszcze bardziej nietypowa jest presja czasu: w **teście językowym SAT uczniowie mieli tylko po 36 sekund** na 50 zadań (łącznie 30 minut)! Gospodarka czasem jest głównym składnikiem **obycia testowego**, obejmującego także znajomość głównych strategii rozwiązywania zadań w **danej formie i trafną interpretację** intencji ich autorów. Obycie uzyskuje się przez rozwiązywanie testów ćwiczebnych, zwykle łatwiejszych i z mniejszą dyscypliną czasu.
 6. Podatność na zgadywanie. Obecność prawidłowej odpowiedzi w **tekście** zadania zamkniętego umożliwia jej **zgadywanie**, losowe wskazywanie przez ucznia. w przypadku pięciu odpowiedzi do wyboru szansa jej zgadnięcia wynosi 20%, znacznie mniej niż w **zadaniach typu prawda-falsz**, gdzie wynosi aż 50%. Zakładając, że żaden uczeń nie opuścił cytowanego zadania językowego, 56% podało jedną z nieprawidłowych odpowiedzi, średnio po 14% wybrało każdy dystraktor, co pokazuje nam przypuszczalny rozmiar zgadywania. Gdy odejmiemy tę wielkość od 44% podających prawidłową odpowiedź, otrzymujemy 30% jako oszacowanie rzeczywistej łatwości zadania. Tego rodzaju poprawka jest stosowa-

wana w wielu krajach, ale teoretycy uważają ją za zbyt słabą (Niemierko, 1999, s. 263–265), bo uczniowie wykazują **spryt testowy**, polegający na posługiwaniu się wiedzą częściową i **ukrytymi wskazówkami w dystraktorach**, co pozwala im eliminować niektóre z nich i zwiększać szansę na „trafienie” prawidłowej odpowiedzi. Poprawka nie uwzględnia indywidualnego „szczęścia” w **zgadywaniu i nie podnosi trafności ani rzetelności** oceniania, może natomiast zmniejszyć liczbę uczniów, których wynik punktowy uznano za spełnienie wymagań egzaminacyjnych.

7. Wartość dydaktyczna zadań zamkniętych. Zadania zamknięte, a w szczególności zadania wyboru wielokrotnego, wnoszą do procesu kształcenia sprawność przeglądania dużych zakresów treści. Ich wyniki tworzą bardziej szczegółową mapę osiągnięć uczniów niż wyniki zadań otwartych. By jednak korzystnie wpływać na te osiągnięcia, musimy zwolnić tempo; 2–3 wybrane zadania mogą wypełnić całą lekcję typu problemowego przynoszącą odpowiedzi na następujące pytania: (a) dlaczego określone rozwiązanie jest najlepsze? (b) jaka część prawdy tkwi w pozostałych rozwiązaniach? (c) jak przeredagować zadanie w celu **większej jasności, użyteczności i atrakcyjności**? (d) jakie pomysły mają uczniowie na podobne zadania? Takie ćwiczenia pogłębiają myślenie analityczne, wdrażają do syntez wiedzy przedmiotowej i **uzasadniania decyzji, rozwijają umiejętności twórcze, oduczają powierzchowności i zgadywania**.

Po pogłębionej analizie wybranych zadań testowych uczniowie powinni otrzymać indywidualne komentarze do swoich osiągnięć, najlepiej w ujęciu rozwojowym, przez porównanie ze swoimi wcześniejszymi wynikami. Oto przykłady takich komentarzy do osiągnięć egzaminacyjnych gimnazjalistów (Szmigel i Rappe, 2005, s. 152):

A (wynik niski): *Adamie, masz powody do zadowolenia, bo stwierdzono u ciebie postęp w zakresie wiedzy i umiejętności w ciągu trzech lat gimnazjum. Jednak w stosunku do rówieśników twój postęp oceniono jako mały. Sam odpowiedz sobie na pytanie, czy nie poświęcasz zbyt mało czasu na naukę w gimnazjum. Sądzę, że stać cię na korzystniejsze wyniki kolejnych egzaminów. Zachęcam cię, byś zaczął poważniej traktować swoje obowiązki szkolne.*

B (wynik przeciętny): *Gratuluję ci, Zosiu, że w licznej gromadzie uczniów gimnazjalnych uzyskałaś wynik równy przewidywanemu, a więc w stosunku do swoich rówieśników zachowałaś taką samą pozycję. Czy ten wynik ci wystarcza? Czy jest zgodny z twoimi dalszymi planami edukacyjnymi? Warto, żebyś wiedziała, że są w twojej szkole i klasie uczniowie, którzy w tym samym czasie uczynili większy postęp. Jeśli masz ambicję dorównania czołówce, nie trać czasu!*

C (wynik wysoki): *Heniu, jestem ciekawa, jak bardzo wynik egzaminu gimnazjalnego cię satysfakcjonuje. Chyba bardzo. Podnieś głowę wyżej i bądź dumny z siebie! Mam dla ciebie dobrą wiadomość: jesteś w grupie uczniów, którzy w ciągu trzech lat nauki w gimnazjum uczynili największy postęp w stosunku do wyniku sprawdzianu w szkole podstawowej. Brawo! Twoje zdolności oraz chęć uczenia się mogą być źródłem wielu sukcesów w przyszłości. Życzę ci wytrwałości! Stawiaj przed sobą coraz trudniejsze zadania!*

O tych komentarzach poczynimy parę spostrzeżeń:

1. Przyjazna zachęta. Każdy komentarz zaczyna się uznaniem wysiłku, jaki uczeń włożył w dotychczasowe osiągnięcia. Potem pojawia się w nim zachęta do refleksji nad poziomem osiągnięć i **do zaplanowania** dalszej pracy nad sobą. Podkreślana jest samodzielność decyzji co do tego i **własna odpowiedzialność za postępy w liceum. Uczeń powinien odczuć, że nauczyciel w niego wierzy i że on powinien wierzyć w siebie.**
2. Edukacyjna wartość dodana. Przyrost osiągnięć w wybranym zakresie programowym w określonym czasie nazywamy **edukacyjną wartością dodaną** (EWD). Koncepcja EWD jest pochodną **psychologii poznawczej**, akcentującej indywidualne ścieżki rozwoju człowieka, bardziej rozmaite i bogatsze niż przyswajanie skodyfikowanej wiedzy (Kohlberg i Mayer, 1993). Uczniowie wędrują po tych ścieżkach szybszym lub wolniejszym tempem, co plasuje ich w różnych etapach marszruty. Adam jest bliższy startu niż reszta rówieśników, Zosia jest w grupie środkowej, a Henio zdecydowanie wyprzedza tę grupę. Wszyscy mogą przyspieszyć swój rozwój, co wykazałaby kolejna wartość dodana.
3. Wzrost osiągnięć. w edukacji masowej rozróżnianie indywidualnych ścieżek uczenia się jest trudne, a operowanie statystykami wyników jest stosunkowo łatwe. Przytoczone tu komentarze są oparte na pomiarze osiągnięć w dwu ujednoczonych egzaminach: sprawdzianie w szkole podstawowej i **egzaminie gimnazjalnym. Te egzaminy obejmowały różne zakresy treści, więc ustalana była wartość dodana względna** osiągnięć uczniów, wyrażana zmianą pozycji ucznia w populacji badanych uczniów. Adam obniżył swoją pozycję w stosunku do rówieśników, Zosia ją utrzymała, a Henio wyraźnie poprawił. w **każdym** przypadku układem odniesienia był dla nich wynik koleżanek i kolegów. **Zapewne cała trójka nauczyła się wiele przez trzy lata** pozostawania w murach szkoły, ale o tym się nie dowiadujemy, bo wartość dodana rejestruje tylko **zmianę pozycji** ucznia w rozkładzie wyników populacji. Utrzymanie średniej pozycji – za co Zosia otrzymała gratulacje – jest w tym układzie sukcesem, choć sugeruje nastawienie ucznia na trwałość, a nie na rozwój osiągnięć.
4. Etapy osiągnięć. Rozwój to nie tylko gromadzenie nowych doświadczeń, ale także, i to przede wszystkim, podnoszenie jakości działania. By to wykazać, potrzebne jest ocenianie wartości dodanej oparte na dwukrotnym szacowaniu jednego zakresu osiągnięć ucznia. Gdy umiejętności, a nie wiadomości uznamy za zasadnicze dla rozwoju ucznia, możliwe jest projektowanie **hierarchii uczenia się**, to jest takiego układu treści kształcenia, w którym opanowanie pewnych czynności jest warunkiem opanowania innych czynności (Resnick, 1973). Przykładami mogą być: czytanie tekstów o rosnącej długości i trudności, rozwiązywanie zadań matematycznych o rosnącej złożoności i oryginalności, dokonywanie obserwacji z **zastosowaniem aparatury i procedur o rosnącym zaawansowaniu**. Odkrywanie i stosowanie naturalnych hierarchii uczenia się jest niezbędną kompetencją nauczyciela, ale oparcie na nich oceniania osiągnięć uczniów (Popham, 2008) jest obarczone ryzykiem niedoce-

- niania indywidualnych ścieżek uczenia się. Pewne umiejętności Zosi są z pewnością wyższe niż Adama, a Henio opanował je na jeszcze wyższym poziomie, ale nie wiadomo, czy podążali tą samą drogą.
5. Wyprzedzanie i zostawanie w tyle. Zarówno EWD względna, jak i EWD bezwzględna pokazują różnice w **osiągnięciach uczniów, to jak jedni wyprzedzają większość grupy, a inni pozostają w tyle**. W procesie uczenia się te różnice rosną, bo w grę wchodzi **efekt wachlarzowy** rozwoju, wiążący EWD z wyjściowym poziomem osiągnięć: przy wyższym poziomie osiągnięć czas potrzebny na nauczenie się nowej czynności jest krótszy niż przy niższym poziomie osiągnięć i w tym sensie wachlarz się otwiera (Niemierko, 2007, s. 348–350). To zjawisko zaprzecza tradycyjnej koncepcji **szkoły jednolitej**, w której wszyscy uczniowie uczą się tego samego i spełniają te same wymagania. Taka szkoła jest marzeniem polityków, ale w praktyce okazuje się rygorystyczna, selekcyjna i daleka od demokracji.
 6. Interpretacje podłużne. Umiejętności rozwijane na wszystkich poziomach szkoły – takie jak rozumienie czytanego tekstu, posługiwanie się liczbami, obserwacje przyrodnicze, wnioskowanie – wyznaczają naturalne skale ciągłe rozwoju uczniów. Można ustalić średnie punktów zdobywane w kolejnych klasach systemu szkolnego jako **równoważniki klasy** i z nimi porównywać wynik pojedynczego ucznia. Może okazać się, że w czytaniu poziom Adama to 5,2, Zosi to 8,5, a Henia aż 12,0, co odpowiada początkowi piątej klasy szkoły podstawowej, połowie trzeciej klasy gimnazjum i zakończeniu liceum. Ta urokliwa prostota miara, chętnie stosowana w amerykańskich testach szerokiego użytku, jest powszechnie nadinterpretowana przez laików, którzy skłonni byłiby przesunąć niektórych uczniów o parę szczebli drabiny szkolnej (Nitko, 1983; Niemierko, 2009, s. 254–259). Także uczniom może to zakłócać wyobrażenie kariery szkolnej.

Alfabetyzm funkcjonalny w ocenianiu szkolnym

Pedagogika intensywnie korzysta z dorobku psychologii w swych zasadach i metodach, ale bywa, że z pewnym opóźnieniem. Przednaukowa, wywiedziona z filozofii „mechanika umysłu” Jana Fryderyka Herbarta (1835), sprowadzająca uczenie się do przyswajania, porządkowania i **stosowania wiadomości zaowocowała autorytarnym systemem kształcenia**, którego *triumfalny pochód* zaczął się w połowie XIX wieku (Nawroczyński, 1987, s. 372) i nie zakończył do dzisiaj. Jego ślady, utrwalone w dydaktyce socjalistycznej (Okoń, 1987), znajdujemy w obiegowych modelach lekcji. Psychologię dwudziestowieczną zdominował **behawioryzm**, zbudowany na fizjologicznych prawidłowościach bodźca i reakcji, który oddziałwał na przekonania pedagogów o wynikach uczenia się i testowych metodach ich rejestrowania. Dopiero współczesny **konstruktywizm**, badający systemy znaczeń powstające w **umysłach jednostek, daje solidną podstawę indywidualizacji kształcenia i oceniania** (Shepard, 2000; Mietzel, 2002, rozdz. 7).

Ten referat otwierało hasło *Ta para musi się pogodzić*, skierowane do nauczycieli i do **egzaminatorów zaopatrzonych w standaryzowane narzędzia pomiaru dydaktycznego**. Nadzieję na zbliżenie stanowisk dwu obozów Krajowa Rada Pomiaru Edukacyjnego (NCME) oparła na kształceniu i doksztalceniu

nauczycieli. Na ubiegłorocznej konferencji kluczowy był nowy dla nas termin **assessment literacy**, który można przetłumaczyć jako **alfabetyzm funkcjonalny w ocenianiu**, umiejętność formułowania i komunikowania ocen według potrzeb rozwojowych ucznia. Taki termin jest neologizmem, ale „alfabetyzacja” pojawiła się w pedagogice już wcześniej (Kwieciński, 2002b). Aspekt czynnościowy, a nie erudycyjny, alfabetyzmu akcentuje poniższy cytat z wystąpienia konferencyjnego (Nitko, 2017):

Alfabetyzm w ocenianiu powinien być funkcjonalny: nie chodzi o znajomość pojęć, lecz o **skuteczne ich stosowanie w życiu codziennym i pracy. Role pedagogów zmieniają się**, w miarę uzyskiwania doświadczeń i podejmowania odpowiedzialności, tak że skupiają się oni na różnych aspektach oceniania i głębia alfabetyzmu rośnie.

Kandydat na nauczyciela mógłby zdobyć odpowiednie kompetencje poprzez moduł (przedmiot) jego wykształcenia obejmujący wykład, ćwiczenia i **praktykę w ocenianiu szkolnym**. Taki moduł pojawił się już w **kształceniu nauczycieli chemii** na Uniwersytecie Łódzkim (Zakrzewski i Wypych-Stasiewicz, 2017), o czym dowiedzieli się uczestnicy poprzedniej konferencji PTDE.

Przykładowy zbiór pojęć niezbędnych w **alfabetyzacji funkcjonalnej kandydatów do zawodu nauczyciela** mógłby być następujący:

ALFABETYZM FUNKCJONALNY W OCENIANIU SZKOLNYM

Zbiór pojęć podstawowych (uporządkowany alfabetycznie)

Alfabetyzm funkcjonalny w ocenianiu	Motywacja wewnętrzna i zewnętrzna
Bezradność intelektualna i motywacyjna	Nastawienie na trwałość i na rozwój
Diagnoza rozwojowa	Ocenianie kształtujące i sumujące
Dysonans poznawczy	Ocenianie naturalne i regulowane
Edukacyjna wartość dodana	Pomiar sprawdzający i różnicujący
Efekt wachlarzowy	Rozwój wszechstronny i ukierunkowany
Egzamin powszedni i doniosły	Rzetelność i trafność oceniania
Hierarchia uczenia się	Samoregulacja uczenia się
Informacja zwrotna w uczeniu się	Uczenie się – cztery modele
Komunikacja obronna i kompromisowa	Zadanie otwarte i zadanie zamknięte

Te dwadzieścia pojęć wyznacza zakres części wykładowo-konwersatoryjnej modułu. Ich objaśnienie można znaleźć w książkach autora tego wykazu (Niemierko, 1975; Niemierko, 1990; Niemierko, 2002; Niemierko, 2007), ale wykładowcy z pewnością będą chcieli dobrać własny zestaw pojęć i lektur, najkorzystniejszy dla specjalizacji przedmiotowej kandydatów. Jest ważne, by dyskusje za punkt wyjścia miały realne sytuacje i **wydarzenia** – albo **zaobserwowanych** w szkołach, albo **wybranych z lektur** jak w niniejszym referacie. Zaliczeniowe eseje, także oparte na odpowiednich epizodach edukacyjnych, mogą sprawdzać, czy studenci potrafią te pojęcia twórczo interpretować, ale polem ich użyteczności jest praktyka.

Podsumowanie referatu

1. w psychologii i pedagogice pojawiły się tendencje do integrowania oceniania wewnątrzszkolnego z pomiarem dokonywanym standaryzowanymi testami osiągnięć szkolnych.
2. Podstawą integracji ma być psychologia poznawcza, akcentująca rolę motywacji do uczenia się i wartość indywidualnych konstrukcji wiedzy.
3. Stopnie szkolne muszą ustąpić pierwszeństwa jakościowemu komentowaniu rozwoju emocjonalno-motywacyjnego i poznawczego ucznia.
4. Wyniki egzaminów zewnętrznych muszą być traktowane jako pomocniczy wskaźnik rozwoju ucznia, a nie jako cel kształcenia.
5. Drogą postępu integracji dwu odmian oceniania szkolnego jest konsekwentna **alfabetyzacja funkcjonalna** nauczycieli w tej dziedzinie.

Bibliografia

- Black P., Wilson M., Shih Ying Yao (2011) Road Maps for Learning: A Guide for the Navigation of Learning Progressions. *Measurement*, 9, 71–123.
- Brookhart S.M., Guskey T.R., Bowers A.J., McMillan J.H., Smith J.K., Smith L.F., Stevens M.T., Welsh M.E. (2016) A Century of Grading Research: Meaning and Value in the Most Common Educational Measure. *Review of Educational Research*, 86, No. 4, 803–848.
- Brookhart S.M. (2017) *How to Use Grading to Improve Learning*. Alexandria: ASCD.
- Bruner J.S. (1964) *Proces kształcenia*. Warszawa: PWN.
- Brzeziński J. (2017) Diagnostyka psychologiczna – za całościowym ujęciem, a przeciw testowemu redukcjonizmowi. W: W.J. Paluchowski (red.) *Diagnozowanie. Wyzwania i konteksty*. Poznań: Wydawnictwo Naukowe WNS UAM, 11–30.
- Cizkowicz B. (2009) *Wyczerpana bezradność młodzieży*. Bydgoszcz: Wydawnictwo UKW.
- College Entrance Examination Board (1986) *10 SATs. The Actual and Complete Scholastic Aptitude Test Plus Advice from the College Board on How to Prepare for It*. New York: CEEB.
- Dolata R. (2008) *Szkoła – segregacje – nierówności*. Warszawa: Wydawnictwa UW.
- Groenwald M. (2004) Uczniowie wobec doniosłego charakteru egzaminów zewnętrznych. W: K. Denek, T. Koszczyk, M. Lewandowski *Edukacja jutra*. Wrocław: Wydawnictwo WTN.
- Herbart J.F. (1835) *Umriss paedagogischer Vorlesungen*. Dietrich. (1912 wyd. pol. *Wykłady pedagogiczne w zarysie*).
- Hornowska E. (1999) *Stronniczość testów psychologicznych. Problemy – kierunki – kontrowersje*. Poznań: Humaniora.
- 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.
- Konarzewski K. (1991) *Problemy i schematy. Pierwszy rok nauki dziecka*. Poznań: Akademos.
- Krope P. (1994) Ocena opisowa w pedagogice jako źródło nieporozumień. W: B. Niemierko (red.) *Diagnostyka edukacyjna*. Gdańsk: Wydawnictwo UG.
- Kwieciński Z. (1990) Błędne koło ubóstwa kulturowego. W: Z. Kwieciński, L. Witkowski (red.) *Ku pedagogii pogranicza*. Toruń: Wydawnictwo UMK.

- Kwieciński Z. (2002a) *Wykluczanie*. Toruń: Wydawnictwo UMK.
- Kwieciński Z. (2002b) *Nieuniknione? Funkcje alfabetyzacji w dorosłości*. Toruń, Wydawnictwo UMK.
- Ledzińska M. (2000) *Uczenie się wykraczające poza warunkowanie*. W: J. Strelau (red.) *Psychologia. Podręcznik akademicki*. T. 2. *Psychologia ogólna*. Gdańsk: GWP.
- Łukaszewski W. (1972) *Informacje dotyczące osiągniętych wyników a poziom wykonania zadań umysłowych*. Wrocław: Ossolineum.
- Łukaszewski W. (2002) *Zwrotne informacje o wyniku czynności*. W: I. Kurcz, D. Kądziaława (red.) *Psychologia czynności. Nowe perspektywy*. Warszawa: Scholar.
- Mietzel G. (2002) *Psychologia kształcenia. Praktyczny podręcznik dla pedagogów i nauczycieli*. Gdańsk: GWP.
- Nawroczyński (1987) Jan Fryderyk Herbart. Pisma pedagogiczne. Wstęp. W: *Dzieła wybrane. Tom I*. Warszawa: WSiP.
- Niemierko B. (1975) *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki obliczeniowe*. Warszawa 1975b, WSiP.
- Niemierko B. (1990) *Pomiar sprawdzający w dydaktyce. Teoria i zastosowania*. Warszawa: Wydawnictwo Naukowe PWN.
- Niemierko B. (1999) *Pomiar wyników kształcenia*. Warszawa: WSiP.
- Niemierko B. (2002) *Ocenianie szkolne bez tajemnic*. Warszawa: WSiP.
- Niemierko B. (2007) *Kształcenie szkolne. Podręcznik skutecznej dydaktyki*. Warszawa: WAiP.
- Niemierko B. (2009) *Diagnostyka edukacyjna. Podręcznik akademicki*. Warszawa: Wydawnictwo Naukowe PWN.
- Nitko A.J. (1983) *Educational Tests and Measurements. An Introduction*. New York: Harcourt.
- Nitko A.J. (1998) Model egzaminów państwowych opartych na programie nauczania, sprawdzających i różnicujących, przeznaczonych do dyplomowania i selekcji uczniów. W: B. Niemierko, E. Kowalik (red.) *Perspektywy diagnostyki edukacyjnej*. Gdańsk: Wydawnictwo UG.
- Nitko A.J. (2017) *What a Classic Resource Can Tell Us about the Future of Assessment Literacy*. Referat wygłoszony na konferencji NCME: „Ta para musi się pogodzić”.
- Okoń W. (1987) *Wprowadzenie do dydaktyki ogólnej*. Warszawa: PWN.
- Popham W.J. (2008) *Transformative Assessment*. Alexandria, VA: ASCD.
- Resnick L.B. (1973) Hierarchies in Children's Learning. A Symposium. *Instructional Science*, 2.
- Sędek G. (1995) *Bezradność intelektualna w szkole*. Warszawa: Wydawnictwo Instytutu Psychologii PAN.
- Shepard L.A. (2000) The Role of Assessment in a Learning Culture. *Educational Researcher*, 27, 7, 4–14.
- Smith M.L., Rottenberg C. (1991) Unintended Consequences of External Testing in Elementary Schools. *Educational Measurement: Issues and Practice* 4, 11, 7–11.
- Stiggins R. (2017) *The Perfect Assessment System*. Alexandria, VA: ASCD.
- Stróżyński K. (red.) (2011) *Ocenianie kształtujące po polsku. Kurs dla doradców metodycznych. Scenariusze zajęć*. Warszawa: CODN.
- Szmigel M.K. (2004) Uczniowie o sprawdzianie 2004 i jego wyniku. W: B. Niemierko, H. Szalaniec (red.) *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Kraków: PTDE.
- Szmigel M.K., Rappe A. (2005) Komunikowanie wartości dodanej osiągnięć szkolnych uczniom, nauczycielom i dyrektorom szkół. W: B. Niemierko, G. Szyling (red.) *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Gdańsk: Fundacja Rozwoju Uniwersytetu Gdańskiego.
- Szyling G. (2013) Hybrydowy charakter oceny szkolnej we wczesniej edukacji: stan i wyzwania. *Studia Edukacyjne*, 28, 303–323.

- Tripp D. (1996) *Zdarzenia krytyczne w nauczaniu. Kształtowanie profesjonalnego osądu*. Warszawa: WSiP.
- Turska D. (2006) *Skuteczność ucznia. Od czego zależy udana realizacja wymogów edukacyjnych*. Lublin: Wydawnictwo UMCS.
- Woronowicz B.T. (2009) *Uzależnienia. Geneza, terapia, powrót do zdrowia*. Poznań: Media Rodzina.
- Wygotski L.S. (1971) *Wybrane prace psychologiczne*. Warszawa: PWN.
- Zakrzewski R., Wypych-Stasiewicz A. (2017) Zwiększenie kompetencji studentów Wydziału Chemii Uniwersytetu Łódzkiego w zakresie sprawdzania arkuszy egzaminu maturalnego z chemii. W: B. Niemierko, M.K. Szmigel *Diagnozowanie umiejętności praktycznych w toku kształcenia i egzaminowania*. Kraków: PTDE, 65–74.

