

RUCH

PEDAGOGICZNY

2



WYŻSZA SZKOŁA PEDAGOGICZNA ZNP

Rok LXXXV Warszawa 2014

Skład redakcji:

Andrzej Hankała (psychologia), Elżbieta Hoffmann (sekretarz redakcji; język polski),
Bogumiła Kwiatkowska-Kowal (pedagogika), Timothy Lomas (język angielski),
Ludmiła Marnej (język rosyjski), Stefan Mieszalski (redaktor naczelny),
Swietłana Sysojewa (pedagogika), Bartłomiej Walczak (socjologia)

Redaktor statystyczny:

Bolesław Niemierko

Rada naukowa:

Przewodniczący: Tadeusz Lewowicki (Wyższa Szkoła Pedagogiczna ZNP),
członkowie: Wojciech Kojs (Wyższa Szkoła Biznesu w Dąbrowie Górniczej),
Wasył G. Kremień (Narodowa Akademia Nauk Pedagogicznych Ukrainy),
Jacek Kurzępa (Wyższa Szkoła Psychologii Społecznej we Wrocławiu),
Jerzy Nikitorowicz (Uniwersytet w Białymstoku),
Wiktor O. Ogniewjuk (Kijowski Uniwersytet im. Borysa Grinczenki),
Andrzej Radziejewicz-Winnicki (Uniwersytet Zielonogórski),
Andrzej Sadowski (Uniwersytet w Białymstoku),
Petro Sauch (Uniwersytet im. Iwona Franki w Żytomierzu),
Bronisław Siemieniecki (Uniwersytet Mikołaja Kopernika)

Recenzenci:

Tadeusz Gałkowski, Maria Machalova, Bogusław Milerski, Ewa Ogrodzka-Mazur,
Rafał Piwowski, Julius Sekera, Mirosław J. Szymański, Frank Tosch

Redaktorzy naukowcy numeru 2/2014:

dr Grażyna Szyling, prof. UG dr hab. Maria Groenwald, Uniwersytet Gdański

Recenzenci numeru:

Ewa Jarosz, Alicja Jurgiel-Aleksander, Dorota Klus-Stańska

Redakcja tekstów w języku polskim:

Grażyna Szyling (UG)

Redakcja tekstów w języku angielskim:

Michał Daszkiewicz (UG)

Adres redakcji:

„Ruch Pedagogiczny”, Wyższa Szkoła Pedagogiczna ZNP
ul. Smulikowskiego 6/8, 00-389 Warszawa
tel. 22 330 57 46
e-mail: rektor@wsp.edu.pl

Uwagi i informacje dla autorów znajdują się na stronie internetowej „Ruchu Pedagogicznego”:
www.rp.edu.pl

Skład, łamanie, druk i oprawa:

Sowa – Druk na życzenie

www.sowadruk.pl

tel. (+48) 22 431 81 40

SPIS TREŚCI

ARTYKUŁY

I. Dokonania diagnostyki edukacyjnej

Bolesław Niemierko – Dwie dekady dojrzewanania diagnostyki edukacyjnej w Polsce	5
Sylwia Jaskuła, Leszek Korporowicz – Odrębność i współgranie. Polskie relacje diagnostyki i ewaluacji edukacyjnej	27
Maria Groenwald – Diagnozowanie – między etyką bliskiego a dalekiego zasięgu	39
Piotr M. Skorupiński – Recepcja koncepcji trafności w diagnostyce edukacyjnej	55
Grażyna Szyling – Teoria i praktyka oceniania szkolnego. W poszukiwaniu pedagogicznych przesłanek hybrydyczności oceny szkolnej	67
Zofia Lisiecka, Wioletta Kozak – O potrzebie badań nad efektem zwrotnym egzaminu maturalnego z języka polskiego	83
Jerzy Chodnicki – Dylematy dydaktyki matematyki czy dylematy diagnostyki?	93
Teresa Sadoń-Osowiecka – Dydaktyczne dyskursy w diagnostycznych rozważaniach o przedmiotach przyrodniczych na Konferencjach Diagnostyki Edukacyjnej	101

II. Nowe obszary zainteresowań

Małgorzata Żytko – Gotowość dzieci i gotowość szkoły do uczenia się. Rozważania wokół problematyki diagnozowania jakości wczesnej edukacji	111
Barbara Ciżkowicz – O potrzebie ewaluacji efektów kształcenia w szkołach wyższych: stan, problemy, perspektywy	123

RELACJE Z BADAŃ

Beata Udzik – Dyskurs edukacyjny wpisany w komunikowanie ewaluacji (wokół zagadnień pragmatyczno-lingwistycznych)	135
Alina Kalinowska – Wyniki badań trzecioklasistów jako diagnoza kontekstów nauczania matematyki w klasach najmłodszych	147

RECENZJE

Jolanta Kruk – „Teraźniejszość – Człowiek – Edukacja”, nr 4/2013	157
Michał Daszkiewicz – Potencjał poznawczy wybranych monografii z kręgu diagnostyki edukacyjnej. Perspektywa językoznawcza	162

KRONIKA

Anna Kot – Profesor Bolesław Niemierko i diagnostyka edukacyjna w Polsce	167
Barbara Przychodzeń – Czynniki podnoszące efektywność kształcenia ogólnego. Kronikarska relacja z badań Okręgowej Komisji Egzaminacyjnej w Gdańsku	173

CONTENTS

ARTICLES

I. The achievements of educational diagnostics

Bolesław Niemierko – Educational Diagnostics: Two Decades of Maturing in Poland	5
Sylwia Jaskuła, Leszek Korporowicz – Distinctiveness and synergy. Polish relations between educational diagnostics and evaluation	27
Maria Groenwald – Diagnosing – between close-range and remote-range ethics	39
Piotr M. Skorupiński – Reception of the validity concept in educational diagnostics	55
Grażyna Szyling – Theory and practice of classroom assessment. In search for premises of hybridity of school grade	67
Zofia Lisiecka, Wioletta Kozak – On the need of studies on feedback from matriculation exam in the Polish language	83
Jerzy Chodnicki – Teaching-of-mathematics dilemmas or diagnostics dilemmas?	93
Teresa Sadoń-Osowiecka – Didactic discourse in diagnostic rumination on natural sciences at the Conference of Educational Diagnostics	101

II. New areas of interest

Małgorzata Żytka – Children’s school readiness and school readiness to learn. Discussion on the quality of early childhood education	111
Barbara Ciżkowicz – On the need for evaluation of learning outcomes in institutions of higher education: state, problems and prospects	123

RESEARCH REPORTS

Beata Udzik – Educational discourse enshrined in the communication of evaluation (around pragmatic-linguistic issues)	135
Alina Kalinowska – Results of research concerning third-formers as a diagnosis of maths-teaching contexts in youngest classes	147

REVIEWS

Jolanta Kruk – The review of monographic issue of the quarterly „Present-Man-Education” 2013/4, devoted to educational diagnosis	157
Michał Daszkiewicz – Cognitive potential of selected monographs from the circle of educational diagnostics. The prospect of linguistic	162

CHRONICLE

Anna Kot – Professor Boleslaw Niemierko and educational diagnostics in Poland	167
Barbara Przychodzeń – Factors increasing the effectiveness of general education. Chronicle report of the research of the Regional Examination Committee in Gdańsk	173

ARTYKUŁY

I. Dokonania diagnostyki edukacyjnej

RUCH PEDAGOGICZNY 2/2014

Bolesław Niemierko
Szkoła Wyższa Psychologii Społecznej
Wydział Zamiejscowy w Sopocie

DWIE DEKADY DOJRZEWANIA DIAGNOSTYKI EDUKACYJNEJ W POLSCE

Jedną z cech nauki, odróżniająca ją od innych form przekazywania doświadczeń, jest **kumulatywność**, rozumiana jako systematyczne zbieranie dorobku poznawczego. By dorobku nie stracić, a zarazem nie utonąć w jego nadmiarze, dokonywane są przeglądy: w osobnych dziełach, na wstępie każdej liczącej się publikacji badawczej, w wyspecjalizowanych czasopismach (np. „Review Educational Research”) i w monograficznych numerach czasopism o szerszym profilu. W tym ostatnim przypadku przegląd dziedziny służy dodatkowo jej rozpowszechnianiu wśród czytelników. Te dwa aspekty monografii naukowej, porządkujący i wyjaśniający, pojawią się w moim opracowaniu.

Swoją **syntezę rozwoju diagnostyki edukacyjnej w Polsce** oprę na dwiętnastu tomach sprawozdań z ogólnopolskich konferencji odbytych się w latach 1993 – 2013 w piętnastu miastach: Krakowie (3), Gdańsku (3), Wrocławiu (2), Łodzi, Lublinie, Szczecinie, Katowicach, Kielcach, Toruniu, Opolu, Łomży, Wałbrzychu, Elblągu, Legnicy i Gnieźnie. Tomy liczą łącznie 9324 strony druku i zawierają 902 autorskie teksty. Duża liczba tekstów stanowi wartość, ale i ryzyko dla syntezy, gdyż zwiększa liczbę zasad i porządków, według których mogłaby być dokonywana. Moim założeniem jest potrzeba refleksji na temat diagnoz uprawianych w edukacji, a co za tym idzie – rozwoju dyscypliny, którą – na wzór diagnostyki medycznej, psychologicznej i technicznej – można nazwać diagnostyką edukacyjną. W chronologicznym porządku prześledzę:

- 1) formułowanie paradygmatu dyscypliny i modeli diagnoz edukacyjnych,
- 2) ogarnianie obszarów edukacji jako dziedziny diagnostyki,
- 3) kształtowanie się stylu publikacji naukowych z tej dziedziny.

Start diagnostyki edukacyjnej, 1993–1994

Odróżniam **diagnostykę**, jako teorię diagnozy, od **diagnozy** jako czynności rozpoznawania rzeczywistości, a także nominalny start dziedziny, czyli pojawienie się jej nazwy, od wcześniejszych poglądów pedagogów. Praktyka diagnostyczna

towarzyszyła edukacji od zarania dziejów i wydała takich mistrzów, jak Janusz Korczak. Czym innym jest jednak tworzenie dyscypliny naukowej.

Pierwsze seminarium o nazwie „Diagnostyka edukacyjna”, przeprowadzone przez Zakład Metodologii Badań Pedagogicznych w Instytucie Pedagogiki Uniwersytetu Gdańskiego oraz Centrum Edukacji Nauczycieli w Gdańsku, mimo skromnej liczby uczestników (około 50), przyniosło kilka do dziś aktualnych tez teoretycznych i sporą liczbę doniesień z badań empirycznych. Zaproponowano (Niemierko, 1994), by nowa dziedzina:

- 1) koncentrowała się na rozpoznawaniu przebiegu uczenia się,
- 2) równolegle obejmowała aspekt poznawczy i motywacyjny kształcenia,
- 3) była uprawiana równolegle na czterech szczeblach: (a) krajowym i międzynarodowym, (b) regionalnym i lokalnym, (c) szkolnym i nauczycielskim, (d) uczniowskim.

Uczenie się zostało potraktowane szeroko, jako dowolnie uregulowane nabywanie gotowości do wykonywania nowych czynności. Można dyskutować, czy kolejność aspektów i szczebli uczenia się była właściwa. Obecnie wydaje się, że jej odwrócenie, to jest wysunięcie na plan pierwszy motywacji poszczególnych uczniów, odświeża nie tylko diagnostykę, lecz całą pedagogikę. Nadal jednak wiązanie diagnoz poznawczych z motywacyjnymi oraz przenikanie modeli i metod z wielkich badań do małych pozostaje postulatem dyscypliny.

W jednym z referatów, obok zabawnych analogii z *Podróżami Guliwera*, znalazło się celne objaśnienie różnic między **diagnozą jakościową** a **diagnozą ilościową**, stanowiących stały przedmiot sporów w edukacji. Oto ono (Krope, 1994, s. 200):

„Pięć zasad posługiwania się słowami i liczbami

- I. Słowa uwalniają nas od sytuacji, o której mówmy. Zastępują pokazywanie rzeczy.
- II. Liczby zastępują wielokrotność jednakowych okoliczności.
- III. To, co określamy jako „jednakowe” i „niejednakowe”, jest pochodną celu, do którego potrzebujemy liczenia.
- IV. Ktokolwiek posługuje się liczbami bez wcześniejszego posługiwania się słowami, nie może wyjaśnić, o czym mówi. Najpierw słowa, potem liczby.
- V. Terminy naukowe różnią się od słów języka potocznego tym, że są używane zawsze w tym samym znaczeniu”.

Byłoby znakomicie, gdyby te zasady uwolniły nas od pokusy nadużywania słów i liczb w raportach diagnostycznych! Takich nadużyć staraliśmy się ustrzec w sprawozdaniach zamieszczonych w pionierskim wydawnictwie konferencyjnym. Badania zostały dokonane w szkołach podstawowych, średnich i wyższych w przedmiotach: język polski, matematyka, fizyka, statystyka, informatyka i plastyka. Te ostatnie (Hamryszak, 1994) miały bodaj największą doniosłość metodologiczną, gdyż ujawniały **strategie oszacowań holistycznych** stosowane przez diagnostów.

Ocenianie prac twórczych przez profesjonalistów i przez nieprofesjonalistów

Czworo magistrów sztuki, absolwentów Wyższej Szkoły Sztuk Plastycznych, i czworo absolwentów kierunku technika z plastyką w studium nauczycielskim oraz jedna

osoba po studiach polonistyki z historią sztuki, wszyscy z długim stażem nauczania przedmiotu, oceniało prace malarskie martwej natury wykonane przez 31 uczniów klasy I liceum ogólnokształcącego.

Pierwsza grupa nauczycieli wystawiła większość ocen bardzo dobrych (55%), a średnia wyniosła 4,5, a druga grupa wystawiła najwięcej (41%) ocen dostatecznych, przy średniej 3,6. Na podstawie szczegółowych uzasadnień tych ocen, autor badań sformułował następujące wnioski (tamże, s. 211): „Nauczyciele profesjonalści, rozumiejąc twórcze zmagania się w akcie tworzenia dzieła plastycznego, oceniają przede wszystkim wrażliwość ucznia na kolor i światło, ład kompozycyjny i autentyczność tworzenia. Nieprofesjonaliści zaś kładą większy nacisk na realistyczne odzwierciedlanie rzeczywistości plastycznej, doszukują się treści, pomijają deformacje twórcze młodocianych artystów”. Należy sądzić, że zauważona prawidłowość występuje w wielu przedmiotach kształcenia szkolnego: wyższe kwalifikacje pedagoga pozwalają mu dostrzec więcej twórczości ucznia, a niższe kwalifikacje prowadzą do sztywnej formalizacji wymagań. Diagnoza musi być oparta na solidnej znajomości dziedziny, której dotyczy, a diagnoza twórczości – na twórczych (naukowych, artystycznych) kompetencjach diagnosty.

Nowa dyscyplina zyskała wkrótce przychylne opinie ekspertów. Heliodor Muszyński (1998, s. 24) uznał dokonywanie diagnoz za niezwykle silnie związane z wszelkim celowym, a więc praktycznym działaniem. Rozróżnił diagnozy (a) zorientowane praktycznie (optymalizacyjne) i (b) zorientowane teoretycznie, czyli wyjaśniające, dlaczego zachodzi jakieś zjawisko. W podsumowaniu poskarżył się na niedocenywanie diagnoz edukacyjnych.

Diagnoza a zmiana edukacyjna

„Diagnozowanie w edukacji nie jest z pewnością mniej potrzebne niż w innych dziedzinach życia, a tym bardziej w tych dziedzinach, w których mamy do czynienia z działalnością człowieka – pisał H. Muszyński (1998, s. 41). A jednak nie ma z pewnością innej – poza edukacją – dziedziny życia, w której diagnozowanie byłoby tak mocno kwestionowane, niedoceniane i zaniedbywane. (...) Dlaczego tak się dzieje? (Ze względu na) odległą i trudną drogę, jaka prowadzi od diagnozy rzeczywistości edukacyjnej do celowej, skutecznej zmiany tej rzeczywistości. Można więc powiedzieć, że to nasza bezradność na tym polu sprawia, że lepiej czujemy się w rzeczywistości wygodnie zmitologizowanej niż surowo obnażonej”.

Także Tadeusz Lewowicki (1998, s. 47) wyznaczył diagnostyce edukacyjnej doniosłe zadania. Stwierdził, że „diagnostyka ta obejmować powinna w coraz większym stopniu wielodyscyplinarne obszary badań – integrować badania pedagogiczne, psychologiczne, socjologiczne i inne, a w dobie przemian oświatowych poszerzenie obszaru zainteresowań i zastosowań diagnostyki jest szansą zarówno rozwoju tej subdyscypliny, jak i ulepszenia oświaty”.

Wśród szkolno-przedmiotowych doniesień badawczych pojawiło się świetne studium Krzysztofa Konarzewskiego (1998), kwestionujące konstrukcję **testu sprawdzającego wielostopniowego** (TSW), to jest testu opartego na warstwowaniu treści kształcenia według skali stopni szkolnych. Założenia i wynik jego badań wskazywały na ograniczenia pomiaru.

Ograniczona przydatność pomiaru sprawdzającego w diagnostyce

Badanie przeprowadzone przez K. Konarzewskiego (1998) objęło po 20 nauczycielek języka polskiego i matematyki w warszawskich szkołach podstawowych. Polegało na próbach normowania gotowych testów sprawdzających. Wykazało nasilone myślenie życzeniowe polonistek (72% niedostatecznych, u matematyczek tylko 2%) i ich podatność na ideologie oświatowe (biurokratyczną, romantyczną, liberalną), a ogólnie – kapryśne i pełne wyjątków wyniki. Na tej podstawie autor badań wnioskuje, by diagnozować indywidualne systemy dydaktyczne nauczycieli i odsłonić mechanizmy, które rządzą ich praktyką, a nie próbować projektowania zmian praktyki, której narzędziem miałby być pomiar sprawdzający. „Rola eksperta jest pomóc praktykom w ściślejszym sformułowaniu ich intuicji, a rolą uczonego jest pomóc praktykom w ściślejszym zrozumieniu tych intuicji” (tamże, s. 73). Ostrzeżenie sformułowane przez Konarzewskiego wskazuje na różnicę między diagnozą edukacyjną, dostarczającą informacji o wybranych procesach uczenia się, a **ewaluacją w edukacji**, będącą wartościowaniem wyników diagnozy uzasadniającym decyzje o zmianie oświatowej.

Pomiar sprawdzający, zwłaszcza w wersji wielostopniowej, może pomóc nauczycielom w ściślejszym sformułowaniu i zrozumieniu ich własnych intuicji, ale gdyby miał oznaczać nakazowe normowanie ocen, przyniósłby więcej szkody niż pożytku. Po ekspansji psychologii poznawczej i pedagogicznego **konstruktywizmu**, jaki nastąpił w kolejnych latach, stało się jasne, że zarówno nauczyciel, jak i uczeń, mają prawo nadawać oryginalne znaczenie faktom i budować własne konstrukcje pojęciowe z dostępnej informacji. Wstępne warstwowanie przedmiotowej treści kształcenia może pomóc w rozpoznawaniu tych konstrukcji. Najlepiej zaś, gdy nauczyciel sam jest konstruktorem i analitykiem testu, bo wtedy sprawdza nie tylko osiągnięcia uczniów, ale i własne koncepcje struktury treści kształcenia.

Przed powstaniem Polskiego Towarzystwa Diagnostyki Edukacyjnej, 1999–2003

Gdy po kilkuletniej przerwie między rokiem 1994 a 1999 konferencje diagnostyki edukacyjnej zostały przywrócone, w oświacie polskiej trwał okres „burzy i naporu”, poprzedzający reformę systemu, w tym wprowadzenie egzaminów zewnętrznych.

Próba konsolidacji nowej dyscypliny naukowej było sformułowanie jej czterech celów (Niemierko, 1998):

- 1) dostarczanie uczniom i nauczycielom informacji wyjaśniającej przebieg i wyniki uczenia się,
- 2) podnoszenie trafności oceniania szkolnego,
- 3) poznawanie zależności między kontekstem, wejściami i działaniami a wynikami kształcenia w różnych systemach dydaktycznych,
- 4) poznawanie wpływu postaw uczestników procesów pedagogicznych wobec tych procesów na przebieg procesów.

Hasłami dyscypliny miały być więc: (1) rozwój ucznia, (2) trafność oceniania, (3) podejście systemowe i (4) znaczenie motywacji. To zapowiadało paradygmat **diagnostyki rozwojowej**, skoncentrowanej na rozpoznawaniu postępów w ucze-

niu się, odmienny od **diagnostyki interwencyjnej**, polegającej na rozpoznawaniu zaburzeń w uczeniu się. O niebezpieczeństwie dominacji tej drugiej przekonywał nas Krzysztof Konarzewski w swym niezwykle sugestywnym referacie (1999). Jego tezy były następujące:

Przeostrog przed edukacyjną diagnostyką zaburzeń

Gdy diagnozę edukacyjną potraktujemy jako „rozpoznawanie choroby na podstawie jej objawów” (słownikowe objaśnienie terminu), stanie się ona narzędziem przemocy symbolicznej wobec uczniów. Będzie „przypisywaniem ich niepożądanych zachowań trwałym cechom ich psychiki na podstawie głośnych orzeczeń, bo przypadki badania trafności i rzetelności oświatowych narzędzi pomiarowych są mi w Polsce nieznanne” (tamże, s. 34). Diagnostowanie edukacyjne nie jest przyjazne uczniom, bo niszczy postawę bezstronnej nadziei na to, że uczenia stać na wszystko w procesie rozwoju. Domniemania nauczycieli co do przyczyn niepowodzeń w uczeniu się „wyrastają bowiem z konfliktów interpersonalnych, są zawsze niepewne i uruchamiają mechanizm samospełniającego się proroctwa. Diagnostyka edukacyjna jako subdyscyplina pedagogiczna byłaby więc szkodliwa, potrzebna tylko pomysłodawcom” (tamże, s. 57).

Równie krytyczne były oceny ewaluacji w edukacji sformułowane przez Waldemara Kozłowskiego (1999). Jego zdaniem, „kłopoty z ewaluacją” są potrójne: kontekstowe (upolitycznienie oświaty, zadufanie decydentów), metodologiczne (nadmernie uproszczone procedury) oraz personalne (konflikty wartości, uzależnienie służbowe ewaluatorów). Na pytanie: „Lepsza zła ewaluacja czy żadna?” – autor referatu odpowiada jednak wiarą, że może być „niezła”, pod warunkiem właściwej promocji jej zasad, gruntownego wykształcenia ewaluatorów i nadania im statusu niezależnych kontrahentów ośrodków władzy.

Krytycyzm nie oszczędził metod diagnozy edukacyjnej. **Nadużywanie ankiet** było plagą „mierzenia jakości pracy szkoły” (Bednarek, 2001). Zbyt wiele także spodziewano się po pisemnych **zadaniach wyboru wielokrotnego** (WW), w których budowanie odpowiedzi przez ucznia jest zastępowane wyborem gotowej odpowiedzi. Wprawdzie krótki czas pracy nad takim zadaniem (średnio około 1,5 minuty) nakazuje porównywać wartość zadania otwartego krótkiej odpowiedzi z co najmniej trzema zadaniami WW, a zadania rozszerzonej odpowiedzi z dziesięcioma takimi zadaniami, ale wykazuje ono wady nie do usunięcia (Niemierko, 2000, s 9).

Wady zadań wyboru wielokrotnego

1. Zniewolenie myślenia uczniów, brak pola dla ich własnych modeli teoretycznych.
2. Fałszywy obraz świata jako spójnego systemu pojęciowego.
3. Dominacja jednolitej formy zadań nad różnorodną treścią kształcenia.
4. Podatność na zgadywanie i posługiwanie się przez ucznia niepełną wiedzą.
5. Trudność i pracochłonność konstruowania oraz dokonywania niezbędnych analiz.
6. Zadanie WW z pewnością nie nadaje się do testów nauczycielskich, ale jego potencjał może pójść dalej, jak w referacie Krzysztofa Konarzewskiego (2000). Erudycyjny wywód poparty odpowiednimi przykładami doprowadził go do wniosku, że test wyboru należy do kultury taśmy produkcyjnej, w której robotnik

ma opanować procedury, a koordynatorów produkcji kształci się poza szkolnictwem państwowym. Tak więc kraje anglosaskie, w których zadania zamknięte dominują w badaniach osiągnięć uczniów, straciły w naszych oczach, a do metodologii diagnostyki edukacyjnej wkroczył kontekst ekonomiczno-polityczny.

Pojawiły się problemy **etyki diagnozowania**, rozumianej jako ogół norm moralnych regulujących ten proces. Zainaugurował ją referat Marii Groenwald (2000), w którym autorka rozważyła odpowiedzialność diagnosty wobec siebie samego i przed światem zewnętrznym, w tym przede wszystkim – przed uczniami. „Niestety, nauczycielskie dylematy nie są zadaniami wyboru wielokrotnego, w których tylko jedna odpowiedź jest prawdziwa. Najczęściej są zadaniami otwartymi, złożonymi i bardzo trudnymi” (tamże, s 75).

Począwszy od roku 2001 konferencje diagnostyczne zdominowała problematyka **egzaminów zewnętrznych**, prowadzonych w całym kraju przez Centralną Komisję Egzaminacyjną i komisje okręgowe. Takim egzaminom można postawić cztery wymagania (Niemierko, 2001):

1. Egzamin jest maksymalnie sprawny i ekonomiczny.
2. Zadania egzaminacyjne są naturalnym sprawdzianem umiejętności.
3. Wynik egzaminu obrazuje zarówno odtwórcze, jak i twórcze osiągnięcia uczniów.
4. Egzamin bezwarunkowo wspiera indywidualny rozwój każdego ucznia.

O tych postulatach można powiedzieć, że każdy kolejny na liście jest trudniejszy od poprzedniego. Dwa początkowe, o charakterze technologicznym, (1) ekonomiczny i (2) „ekologiczny”, z roku na rok były bliższe spełnienia, a liczba niedociągnięć i braków malała, ale dwa końcowe, o charakterze humanistycznym, (3) kreatywny i (4) prorozwojowy, do dziś pozostały polem żmudnych poszukiwań. Swych obaw co do nich nie krył H. Muszyński (2001), którego stanowisko jest przytoczone poniżej.

Niedostatek humanizmu w egzaminowaniu zewnętrznym

„Truizmem byłoby powiedzieć, że diagnoza edukacyjna swoimi następstwami sięga zawsze – w sposób bardziej lub mniej zasadniczy – losów osób poddanych procesom edukacji. (...) Nadanie nauczycielskiej indywidualnej diagnozie edukacyjnej charakteru zewnętrznego niesie z sobą cały szereg niebezpieczeństw: 1) koncentrację wyłącznie na formalnej procedurze, 2) ograniczenie diagnozy nauczycielskiej do treści diagnozy zewnętrznej, 3) standaryzację kryteriów oceny rozwoju ucznia, 4) sumowanie wyników według algorytmu i sprowadzanie ich do średniej arytmetycznej. (Niewątpliwie) wszystkie te tendencje stwarzają istotne zagrożenie dla tego, co zdaje się jednak być głównym hasłem współczesnej szkoły: jej humanizacja” (tamże, s. 34).

Podobnie zabrzmiał wynik etycznej analizy zmiany, jaką do edukacji wniosły egzaminy zewnętrzne (Groenwald, 2001). Zdaniem M. Groenwald, „odkąd elementy diagnostyki zostały wprowadzone do systemu egzaminowania niezależnego od szkoły, zmiany w edukacji nabrały tempa, a egzaminy zewnętrzne zyskały na znaczeniu. Wobec doniosłego charakteru tych egzaminów powszednie problemy pracy pedagogicznej zeszły na margines działalności edukacyjnej” (tamże, s. 343). (Natomiast), nasilają się prace nad poszukiwaniem i konstruowaniem nowych form zadań oraz

wdrażaniem ich do praktyki szkolnej i egzaminacyjnej. Co więcej, szermując pojęciem «obiektywizmu egzaminu» usiłuje się do minimum ograniczyć kontakt nauczyciela czy egzaminatora z uczniem, a nawet całkowicie go wyeliminować, stosując egzamin na odległość” (tamże, s. 346).

Już wtedy pojawił się zwiastun psychometrycznego postępu egzaminów zewnętrznych: pierwsza próba wdrożenia **teorii wyniku zadania**, czyli probabilistycznej teorii związku wyników zadania testowego ze **zmienną ukrytą** – właściwością mierzoną przez wszystkie zadania danego testu. Henryk Szaleniec (2001) przedstawił założenia tej teorii, komputerowy program analiz, przykład jego wykorzystania do analizy wyników złożonych zadań egzaminacyjnych oraz pożytki z analizy w postaci nowej informacji o właściwościach zadania.

Zmiana systemowa w oświacie, jaką było wprowadzenie egzaminów zewnętrznych, spowodowała **efekt zwrotny** w kształceniu i wewnątrzszkolnym ocenianiu osiągnięć uczniów. Pojawiły się analizy tego efektu oparte na Samuela Messicka teorii **trafności konsekwencyjnej** pomiaru dydaktycznego, nadającej specjalne znaczenie skutkom społecznym pomiaru, nie w pełni przewidywalnym przed jego dokonaniem. Ta nowa teoria, sformułowana (w Stanach Zjednoczonych) dopiero w 1989 roku, szybko została włączona do repertuaru diagnostyki edukacyjnej (Niemierko, 2003, s. 19–21).

Główną konsekwencją egzaminu zewnętrznego dla ucznia są określone szanse jego przyjęcia do wybranej szkoły wyższego szczebla. Prześledził je Piotr Skorupiński (2002) na przykładzie „renomowanego warszawskiego liceum”, w którym ponad 500 uczniów kandydowało do dwóch oddziałów o profilu humanistycznym.

Niska trafność procedur selekcji uczniów

Do liceum przyjęto 64 uczniów na podstawie sumy punktów uzyskanych (1) za oceny na świadectwie gimnazjalnym i (2) w egzaminie zewnętrznym, obu kryteriom dając równą wagę. Gdyby zastosowano wyłącznie pierwsze kryterium, do tego liceum dostałoby się tylko 15 uczniów z liczby 64 uszeregowanych według najlepszych średnich stopni na świadectwie, w tym z pierwszej dziesiątki tylko uczniowie o numerach 1 i 9, a pozostali musieliby szukać innej szkoły. Gdyby zastosowano wyłącznie drugie kryterium, do liceum dostałoby się tylko 31 uczniów z liczby 64 uszeregowanych według wyników egzaminu. „Jakich celów poznawczych, motywacyjnych i praktycznych nie zdołali osiągnąć ci 33 uczniowie (...), którzy zostaliby przyjęci do liceum jedynie na podstawie wyników z egzaminu i którzy znaleźli się w sytuacji porażki na skutek wprowadzenia (dodatkowego – B.N.) kryterium stopni ze świadectwa gimnazjalnego?” – zastanawiał się autor badań (Skorupiński, 2002, s. 48). Podobnie można by spytać o 49 uczniów o wysokiej średniej stopni na świadectwie i niezadowolającym wyniku egzaminu zewnętrznego. Za miarami, punktami i wskaźnikami kryją się nadzieje, dramaty i żale uczniów, które diagnostyka edukacyjna powinna ujawnić i zracjonalizować.

Krokiem ku rozwikłaniu tajemnic oceniania wewnątrzszkolnego były badania Grażyny Szyling (2002; 2003) nad **drugim układem wymagań**, obejmującym wymagania programowe ograniczone ze względu na specjalne potrzeby edukacyjne uczniów. Poniżej główne wyniki jej badań.

Obniżanie wymagań programowych jako zjawisko pedagogiczne

1. Ocena osiągnięć uczniów zależy od systemu wartości nauczyciela, w tym zwłaszcza od sposobu przezwycięzania konfliktu ról „sędziego” i „advokata”.
2. Około połowy trudnych sytuacji decyzyjnych oceniania szkolnego nauczyciele rozstrzygają przez obniżanie wymagań programowych.
3. Wrozumiałość nauczycieli rośnie wraz ze znaczeniem wyniku dla dalszych losów ucznia.
4. Zastosowanie drugiego układu wymagań programowych dotyczy nie tylko słabych, lecz także i wybitnych uczniów, w tym starających się o przyjęcie do ekskluzywnych szkół i uczelni.
5. Obniżanie wymagań jest zjawiskiem powszechnym i nie zależy od poziomu pracy szkoły mierzonego średnimi wynikami egzaminu zewnętrznego absolwentów.

„Wielkość pola obniżonych wymagań podstawowych coraz wyraźniej będzie ujawniana przez wyniki egzaminów zewnętrznych. Problemu nie rozwiąże obarczanie odpowiedzialnością nauczycieli i rozliczanie szkół z efektów kształcenia mierzonych testami standaryzowanymi” (Szyling, 2002, s. 86). „Nieuchronny wydaje się konflikt między oczekiwaniami odwołującymi się do uzasadnień trafności oceny na podstawie dowodów a postawami nauczycieli, którzy przy podejmowaniu decyzji doniosłych dla uczniów biorą pod uwagę przede wszystkim trafność konsekwencyjną” (Szyling, 2003, s. 130).

Wszyscy udziałowcy procesów edukacyjnych wydają się doceniać potrzebę równoległego stosowania dwu rodzajów wymagań programowych: obiektywnych i obniżonych. Ankietowe badania Elżbiety Kowalik (2003) pokazały, że wprawdzie świadomość ich wartości dydaktycznej rośnie od uczniów gimnazjum (28% potwierdzeń), przez maturzystów (48%) i nauczycieli (73%) do dyrektorów (93%), ale ani uczniowie, ani nauczyciele, ani dyrektorzy szkół nie kwestionują potrzeby istnienia obu systemów oceniania, dostrzegając ich komplementarność.

Nowym wątkiem analiz, jaki nabrał znaczenia przy ustalaniu trafności konsekwencyjnej egzaminów, stało się **komunikowanie oceny osiągnięć**, rozumiane jako porozumiewanie się egzaminatora z uczniem na temat jego wyniku uczenia się. Klemens Stróżyński (2003), za Stephenem Coveyem, powiązał skuteczność komunikowania oceny z zaufaniem, jakie egzaminator uzyskał, i współdziałaniem, jakie egzamin poprzedzało. By przełamać **postawę obronną** wobec diagnozy, wyrażającą się utrudnianiem jej przebiegu i podważaniem wyniku, należy zapewnić uczniom wiarę we własne siły i oswoić ich z procedurami. To pozwoli, zdaniem autora referatu, zapobiec „toksycznym”, zdroworozsądkowym interpretacjom liczb.

W ramach Polskiego Towarzystwa Diagnostyki Edukacyjnej, 2004–2013

Utworzenie Polskiego Towarzystwa Diagnostyki Edukacyjnej (PTDE) zapewniło inicjatywom konferencyjnym ciągłość i korzystnie wpłynęło na rozwój dziedziny. Liczba członków Towarzystwa przekroczyła stu, liczba uczestników konferencji – dwustu, a liczba wystąpień na konferencjach – pięćdziesiąt referatów i doniesień badawczych.

Egzaminy zewnętrzne nadal stanowiły główną tematykę dorocznych konferencji krajowych. Ponieważ utrzymała się zasada budowania testów, nazywanych „arkusza-

mi egzaminacyjnymi”, w dwu częściach: (a) złożonej z zadań wyboru wielokrotnego i (b) złożonej z zadań otwartych, istotne stało się pytanie o funkcje tych dwu części w egzaminie. Analiza ich funkcji (Niemierko, 2004) przyniosła następującą odpowiedź na to pytanie:

Trafność przedmiotowa arkusza egzaminacyjnego

Założeniem sprawdzianu osiągnięć po szkole podstawowej był pomiar umiejętności uczniów w następujących proporcjach: czytanie 25% (zadania zamknięte), pisanie 30% (zadania otwarte), rozumowanie, posługiwanie się informacją i stosowanie wiedzy 45% (zadania zamknięte i otwarte). Na średni wynik krajowy w 2004 roku (25,6 pkt) złożyły się: czytanie 32%, pisanie 32%, rozumowanie i pochodne 36%, a zadania zamknięte (w liczbie 20), które miały dostarczyć 50% punktów, dostarczyły ich 55%. Zatem test **sprawdzał** w przewadze czytanie i pisanie (64% uzyskiwanych punktów), a zadania zamknięte okazały się łatwiejsze.

Na rozrzut wyników krajowych, określony odchyleniem standardowym wyników testowania i wskaźnikami rzetelności zadań, złożyły się: czytanie 18%, pisanie 31%, rozumowanie i pochodne 51%, a zadania otwarte (w liczbie 6) wniosły 56% odchylenia standardowego. Zatem test **różnicował** osiągnięcia uczniów głównie przez zadania otwarte, trudniejsze i bardziej odporne na podawanie odpowiedzi przy niepełnej wiedzy.

Rozwinięcie tej problematyki przedstawiła Maria Krystyna Szmigel (2004). Pokazała, że w grupie stanina 1 (najsłabsi) uczniowie zdobywają punkty niemal wyłącznie za czytanie (rozumienie czytanego tekstu), w grupie stanina 5 (przeciętni) – zdobywają ich dwukrotnie więcej za czytanie i pisanie niż za zadania z matematyzowanymi, a dopiero w grupie staninów 8 i 9 (najlepsi) typ treści oraz forma zadań przestają odgrywać istotną rolę.

Zadania o największej mocy różnicującej Ewa Stożek nazwała **zadaniami rozrywającymi**, gdyż powodują spłaszczenie rozkładu wyników pomiaru, a nawet ich dwumodalność, czyli „rozerwanie” zbioru wyników na dwie części (Stożek, 2006; Stożek, Dąbrowski, 2006; Dubiecka, 2006). Gdyby celem diagnozy były tylko różnicowanie i selekcja uczniów, wystarczyłoby kilka takich zadań, by silnie posegregować uczniów, pozbawiając słabszych spośród nich możliwości pokazania, jakie prostsze umiejętności dobrze opanowali.

Strategie egzaminacyjne uczniów mogą obejmować **oszustwo egzaminacyjne**, będące ukrytym posługiwaniem się informacją uzyskaną z innych źródeł niż własna wiedza i praca, dokonywanym dla zdobycia pożądanego wyniku egzaminu. Dla diagnostyki edukacyjnej są istotne jego następujące właściwości (Niemierko, 2006):

Zmaganie się z nieuczciwością egzaminacyjną

1. Oszustwo egzaminacyjne ucznia ma wiele form: wykradanie zadań przed egzaminem, podglądanie i korzystanie z podpowiedzi, przepisywanie gotowych tekstów i inne.
2. Ma podobne uwarunkowania społeczne jak niepowodzenia szkolne i brak motywacji do uczenia się, jest wspieranie przez nieformalne grupy oporu wobec szkoły.
3. W żadnym kraju nie udało się go całkowicie wyeliminować, nawet wtedy, gdy są stosowane, owiane legendą skuteczności, „kodeksy honorowe”.

4. Bywa wspierane oszustwem egzaminacyjnym nauczyciela: tolerowaniem niesamodzielności ucznia, dostarczaniem niedozwolonych materiałów, tendencyjnym punktowaniem odpowiedzi.
5. Środki zaradcze obejmują zróżnicowany poziom trudności zadań, podnoszenie motywacji uczniów do wysiłku, wersje równoległe arkuszy egzaminacyjnych, nadzór nad przebiegiem egzaminu.

Skalę oszustwa egzaminacyjnego ujawnił egzamin maturalny z języka angielskiego, w którym, nie informując o tym uczniów, zastosowano trzy warianty układu odpowiedzi na zadania wyboru wielokrotnego, co okazało się przyczyną większej niż w poprzednich latach liczby niepowodzeń (21%). Henryk Szaleniec (2006) wykazał, że większość (62%) tych uczniów, którzy ponieśli porażkę, uzyskałaby lepszy wynik, gdyby dodczytywania wyników wykorzystano szablon punktowania innego wariantu. Zatem „istnieje konieczność stosowania nie tylko wariantowości, ale pełnych wersji równoległych arkuszy egzaminacyjnych – przede wszystkim tam, gdzie zawierają one zadania zamknięte, które stwarzają szczególną okazję do oszustwa egzaminacyjnego” (tamże, s. 203).

Punktowanie zadań otwartych zostało także poddane analizom statystycznym. Tropiono **efekt egzaminatora**, rozumiany jako wpływ doboru punktującego wyniki takich zadań na wynik egzaminu. Ustalono, że egzaminatorzy zbyt łagodni pojawiają się wśród kandydatów stosunkowo częściej niż zbyt surowi, ćwiczenia w punktowaniu przykładowych zadań zmniejszają rozrzut ocen, a styl pracy egzaminatora zależy od składu zespołu oceniającego (Dubiecka i in., 2006). Efekt egzaminatora nasila się w egzaminach ustnych, gdy doświadczenie metodyczne i ćwiczenia przygotowawcze nauczycieli punktujących wypowiedzi uczniów okazują się nieskuteczne (Smolik, 2006).

A jak uczniowie traktują egzaminy zewnętrzne? M.K. Szmigiel (2005) starannie przeanalizowała wypowiedzi niemal czterech tysięcy wylosowanych uczniów o sprawdzianie po szkole podstawowej. A oto wyniki analizy:

Szóstoklasiści oceniają egzamin zewnętrzny

1. Większość uczniów klasy VI, niezależnie od wykształcenia rodziców i ocen wewnętrzzszkolnych, spodziewała się wyniku punktowego zbliżonego do średniej krajowej z zeszłego roku, ale uzyskane wyniki okazały się bardzo zróżnicowane i w wysokim stopniu zbieżne z tymi dwiema zmiennymi kontekstowymi.
2. „Sprawdzian był wielkim wydarzeniem w szkole i w moim w życiu” – napisali uczniowie. Prawie wszyscy (98%) brali udział w sprawdzianach próbnych, a znaczna większość (73%) rozwiązywała nadto zadania publikowane w prasie i w Internecie.
3. Zadania były oryginalne, a matematyczne – trudne, przydałoby się więcej czasu na ich rozwiązywanie. Rady dla przyszłych egzaminowanych: „Korzystajcie z zajęć szkolnych! Powtarzajcie materiał! Nie denerwujcie się na egzaminie! Czytajcie kilka razy polecenia! Myślcie samodzielnie!”
4. Zdaniem uczniów (97%), „warto przykładać się do nauki”. Wyższe studia są pragnieniem prawie wszystkich ankietowanych uczniów i większości (77%) rodziców.
5. Egzaminy kształtują **wspólnotę dydaktyczną** uczniów, opartą na uspołecznieniu motywacji do uczenia się. Oto uderzająco szczerą wypowiedź ankieto-

wa (s. 181): „Pani Minister niech wie, że ja nie mam motywacji, żeby uzyskać dużo punktów. One dla mnie nie są aż tak ważne. Gimnazja i tak czekają na nas z otwartymi drzwiami. No, ale Pani Dyrektor prosiła, żeby nie było wstydu, że jesteśmy zdolni, a wynik będziemy mieli najgorszy. A na to nie zasługujemy! Chyba ma rację”.

Z tych badań wyłania się sympatyczny obraz ucznia walczącego ze stresem i niedostatkami motywacji, nie kryjącego nadziei na szczęśliwy wynik, dającego dojrzałe rady swoim następcom. Trzynastolatki potrafią też pięknie pisać o swoich kolegach i koleżankach z klasy (Groenwald, 2007). Badania gimnazjalistów (Rappe, 2007) nie brzmią już tak pomyślnie:

Gimnazjaliści oceniają egzamin zewnętrzny

Pogląd typu: trudny, beznadziejny, bez sensu, nie podobał mi się; dobrze, że się skończył – wyraziło aż 40% ankietowanych. Autorka raportu tak wyraża swoje rozczarowanie (s. 462): „Po przeczytaniu wszystkich (ponad 2 tysięcy) opinii o egzaminie nie znalazłam ani jednej, w której zdający byłby ciekaw, jaka jest jego wiedza i umiejętność rozwiązywania testów na tle innych uczniów. Wśród przeczytanych ankiet nie było ani jednej, w której zdający chciał sprawdzić, czy jego wiele lat spędzonych w szkole pozwoli na osiągnięcie sukcesu egzaminacyjnego. Jeśli uczeń ma pozytywny stosunek do egzaminu, to dlatego, że «egzamin był łatwy», «rozwiązałam/em zadania», a nie ma wypowiedzi o chęci sprawdzenia swojej wiedzy, chęci sprawdzenia umiejętności egzaminacyjnych, czy porównania swojej wiedzy z wiedzą uczniów w szkole, regionie czy kraju”.

Komunikowanie wyników **egzaminu doniosłego**, dokonywanego przez instytucje zewnętrzne wobec szkoły, staje się upartą walką o świadomość rozwojowego znaczenia diagnozowania osiągnięć uczniów i powstrzymanie dziennikarskich zapędów do traktowania egzaminów jako publicznej rozgrywki, „wyścigów konnych” drużyn szkolnych (Szmigel, 2005). Ta walka toczy się do dzisiaj, a jej wynik jest wciąż niepewny.

Postęp cywilizacyjny szkolnictwa światowego przedstawił Maciej Sysło (2005). Objaśnił trzy scenariusze dla przyszłości oświaty, opracowane w Ośrodku Badań Edukacyjnych i Innowacji, afiliowanym przy OECD: (1) ekstrapolacja obecnego systemu: umocnienie hierarchicznej struktury administracyjnej i orientacja na zaspokojenie potrzeb rynku pracy, (2) uspołecznienie edukacji: silniejsze powiązanie szkół z potrzebami społeczności lokalnej w zakresie kształcenia nieformalnego i ustawicznego, (3) internetowe sieci kształcenia w miejsce szkół: rolę nauczycieli przejmują konsultanci i tutorzy, kształcenie różnicuje się merytorycznie i kulturowo. Diagnostyka edukacyjna jest niezbędna w każdym z tych scenariuszy, ale jej model ewoluuje od scentralizowanego (scenariusz 1) do wieloparadygmatycznego, realizowanego przez wykorzystanie technologii informacyjno-komunikacyjnej (scenariusz 3).

Akcentowanie **rozwaju ucznia**, jako ciągu zmian długotrwałych i względnie nieodwracalnych, uporządkowanych w sekwencję rosnącą, a spowodowanych jego wewnętrznymi siłami, prowadzi do pojęcia **wartości dodanej** jego osiągnięć (edukacyjnej wartości dodanej, EWD) jako przyrostu osiągnięć w wybranym zakresie programowym w określonym czasie. Rysuje się perspektywa **dydaktyki postępu**, opartej na planowaniu, kontraktowaniu i sprawdzaniu korzystnej zmiany osiągnięć

poszczególnych uczniów i **diagnostyki postępu**, prorozwojowej orientacji diagnostyki edukacyjnej (Niemierko, 2005, s. 24–29; Janowski, 2007). Oto jej założenia:

Diagnostyka edukacyjnej wartości dodanej

1. Edukacyjna diagnostyka postępu podnosi znaczenie umiejętności poznawczych ucznia (czytanie, pisanie, obserwowanie zjawisk, interpretacja zależności, posługiwanie się komputerem), a pomniejsza znaczenie jego encyklopedycznych wiadomości.
2. Ze względu na zwiększoną niepewność pomiarową podwójnego szacowania stanu osiągnięć diagnozowanie postępu musi obejmować okresy uczenia się nie krótsze niż kilkumiesięczne.
3. Przy szacowaniu **wartości dodanej względnej**, jako różnicy między pozycją ucznia w zbiorze wyników końcowych a pozycją ucznia oczekiwaną w tym zbiorze na podstawie jego wyniku początkowego, konieczne jest zastosowanie statystycznej **analizy regresji**, uwzględniającej błąd standardowy pomiaru wyniku początkowego.
4. Bardziej rzetelne niż wskaźniki EWD pojedynczych uczniów są wskaźniki **EWD szkół**, jako średnie arytmetyczne EWD diagnozowanych uczniów.
5. Rozwój ucznia podlega prawidłowości **efektu wachlarzowego**, wzrostu rozrzutu wyników uczenia się w toku kształcenia, a więc poprawnie oszacowana EWD uczniów wykaże dodatnią korelację z ich poziomem osiągnięć na wejściu.

Diagnostyka postępu ucznia i szkoły wymaga **skal pionowych**, rejestrujących rozwój osiągnięć w kolejnych etapach (Tomkowicz, 2008). Dwa główne typy takich skal to: (a) **równoważniki klas**, czyli lata i kolejne miesiące nauki szkolnej, w których średni (krajowy) wynik punktowy ma daną wartość i (b) **skale standardowe**, oparte na odchyleniu standardowym wyniku od (krajowej) średniej.

Równoważniki klas zastosowali Radosław Sterczyński, Teresa Kutajczyk i Barbara Przychodzeń (2008) do analizy zależności EWD w gimnazjach wiejskich od właściwości uczniów, nauczycieli i szkół. Wyniki tej analizy przyniosły wiele tzw. „efektów odwróconych”, przeczących teorii pedagogicznej (np. wyższa motywacja – niższe osiągnięcia), wywołanych zapewne deklaratywnymi wskaźnikami zmiennych niezależnych, obciążonymi kompensacją braku osiągnięć szkolnych i swoistych dla środowiska wiejskiego. Dalsze badania wymienionej trójki autorów (Kutajczyk i in., 2009) sugerują, że EWD odgrywa większą rolę w przedmiotach ścisłych niż humanistycznych, wewnątrzszkolna segregacja uczniów na oddziały zwiększa efekt wachlarzowy, na wysokie EWD szkoły składa się głównie postęp słabszych uczniów (staniny 1 – 3), a duże znaczenie dla nich mają aspiracje edukacyjne.

Również badania wykorzystujące skale standardowe EWD skłaniają do ostrożności w interpretowaniu wyników. Roman Dolata (2008) uznał, że jego badanie nie przyniosło jednoznacznego potwierdzenia wpływu **wewnątrzszkolnej segregacji** gimnazjalistów na oddziały A, B, C i dalsze, zróżnicowane co do wyników sprawdzianu po szkole podstawowej, na postępy gimnazjalistów. Znakomicie natomiast wykazał, stosując analizę wariancji, negatywny *efekt zwrotny* systemu egzaminów zewnętrznych w postaci wzrostu różnic między placówkami wywołanych publikowaniem ich rankingów (Dolata, 2009).

Spróbowano też zastosować standardową EWD do zbadania wpływu wielkości oddziałów szkolnych oraz finansowania gimnazjów przez jednostki samorządu te-

rytorialnego na postępy uczniów (Jakubowski, 2008). I tutaj, mimo odpowiedniego zróżnicowania modeli statystycznych, nie udało się ustalić istotnych zależności, ale autor pozostał zwolennikiem metody, stwierdzając, że „EWD to potężne narzędzie ewaluacji polityki edukacyjnej państwa” (tamże, s. 224).

Elegancki model statystyczny EWD o dużej użyteczności praktycznej przedstawił Artur Pokropek (2009). Jego założenia są następujące:

Wskaźnikowanie edukacyjnej wartości dodanej szkół

1. Średnie wyniki egzaminu zewnętrznego polskich szkół w wybranych latach są poddane *normalizacji*, to jest upodobnieniu do rozkładu gaussowskiego.
2. Wyniki są przedstawione w skali standardowej o średniej 100 punktów i odchyleniu standardowym 15 punktów (skala IQ).
3. Wartość dodana (EWD) szkoły jest przedstawiona graficznie w przestrzeni dwuwymiarowej, skrzyżowana z wynikami końcowymi, wraz z błędem obu oszacowań (5%) tworząc dużą elipsę.
4. Położenie małej elipsy reprezentującej wybraną szkołę na dużej elipsie populacji jest interpretowane w czterech kategoriach: (a) „szkoła sukcesu”, o wysokim EWD i wysokiej średniej, (b) „szkoła wspierająca”, o wysokim EWD i niskiej średniej, (c) „szkoła niewykorzystanych możliwości”, o niskim EWD i wysokiej średniej, (d) „szkoła wymagająca pomocy”, o niskiej EWD i niskiej średniej.
5. Warto też zauważyć, że probabilistyczny model skalowania Rascha (jednoparametryczne IRT) okazał się w pełni przydatny do analiz jakości punktowania zadań otwartych, a więc do znacznie skromniejszego celu (Węziak, 2005; Smolik, 2005). Model Rascha został także zastosowany do wieloletniego **zrównywania wyników** egzaminu zewnętrznego, czyli przedstawiania ich w skali wybranego (kotwiczącego) roku po skorygowaniu ze względu na wahania trudności zadań. Taką próbę, ukierunkowaną na ewaluację systemu egzaminacyjnego, podjął Henryk Szaleniec (2010). Wkrótce do celu zrównywania zastosowano mocniejszy, dwuparametryczny model logistyczny (Kondratek, 2011). Te kalkulacje pozwoliły zniwelować różnice punktowe krajowych średnich egzaminu gimnazjalnego w latach 2002–2010 w części humanistycznej z 7 punktów w skali IQ (o średniej 100 i odchyleniu standardowym 15) do 2 punktów, ale w części matematyczno-przyrodniczej wzrosły z 5 do 6 punktów na skutek silnego spadku w latach 2002–2005 (Szaleniec, 2012). Przyszłość to niewątpliwie wyprzedzające zrównywanie arkuszy egzaminacyjnych, bo bez niego daleko naszemu rodzimemu systemowi egzaminacyjnemu do międzynarodowych standardów (Pokropek, 2011).

Mocnym metodologicznie wejściem w problematykę międzynarodowych badań osiągnięć szkolnych był artykuł Macieja Jakubowskiego (2009). Autor porównał w nim badania PISA (Organizacji Współpracy Ekonomicznej i Rozwoju, OECD) z badaniami TIMSS i PIRLS (Międzynarodowego Stowarzyszenia Badań Osiągnięć Pedagogicznych, IEA) ze względu na odtworzenie „prawdziwych zależności między cechami ucznia, jego rodziny i szkoły a osiągnięciami ucznia”. Znakomitym uzupełnieniem tego studium była diagnoza Pasi Reinikainena (2009), dokonana przez reprezentanta kraju niedoścignionego jeszcze wówczas dla Polski pod względem wyników kształcenia. Jego wnioski były następujące:

Czynniki sukcesu edukacyjnego kraju

1. Wysoko wykwalifikowani nauczyciele (z tytułem magistra) na całym obszarze.
2. Szkoły heterogeniczne, wyrównane co do poziomu.
3. Niewielkie oddziały szkolne (w Finlandii średnio 18 uczniów).
4. Elastyczne programy kształcenia i swobodny wybór podręczników przez nauczycieli.
5. Wysoka ranga kształcenia językowego i czytelniczego.
6. Skuteczna eliminacja drugoroczności w drodze zajęć wyrównawczych.
7. Jedność kulturowa kraju przy poszanowaniu mniejszości etnicznych.
8. Zaangażowanie uczniów w pracę nad testami (mała liczba opuszczanych zadań).

Wnikliwe analizy wyników badań PISA (2011) pokazały, że Polska należy do krajów, w których w latach 2000–2009 nie tylko wzrosły osiągnięcia, zwłaszcza dziewcząt, lecz także mocno zmalała wariancja średnich między szkołami. Badania międzynarodowe, które mogą być potraktowane przez uczniów jako **egzamin powszedni** („niskich stawek”), gdyż ich wyniki ujawniają osiągnięcia i braki do wyrównania, ale nie jako **egzamin doniosły** („wysokich stawek”), wpływający na karierę szkolną i zawodową ucznia, wykazują międzykulturowe różnice roli motywacji do uczenia się. W krajach trzeciego świata związek między motywacją do uczenia się a wynikiem testu jest bliski zeru, w krajach Europy Środkowej jest niski, a w obydwu tych grupach krajów występuje umiarkowana korelacja (około 0,4) między pozycją ekonomiczno-społeczną rodziny a wynikiem. W krajach zachodnich przeciwnie: motywacja wykazuje umiarkowany, a pozycja ekonomiczno-społeczna słaby związek z wynikami (Dolata, Pokropek, 2010).

Oryginalnym wątkiem diagnostyki edukacyjnej było zaprezentowanie przez Stanisława Walukiewicza (2009) ekonomiczno-socjologicznej teorii **kapitału ludzkiego**, obejmującego wszystkie zasoby związane z rozważaną osobą traktowaną jako pojedyncza osoba ludzka, a więc zdrowie, energia, zdolności, wiedza i doświadczenie, oraz **kapitału społecznego**, obejmującego zasoby związane z relacjami w grupie. Operacjonalizacja tych pojęć powinna otworzyć drogę pomiarowi dydaktycznemu i opartej na nim diagnostyce (Walukiewicz, 2010, s. 79–80):

Diagnostyka edukacyjna w ujęciu systemowo-ekonomicznym

1. Kapitał finansowy i kapitał materialny naszego świata są i muszą być ograniczone, natomiast kapitał ludzki, rozumiany jako możliwości uczenia się, oraz kapitał społeczny, rozumiany jako możliwości coraz to lepiej zorganizowanej współpracy grup, społeczności oraz narodów, są nieograniczone i rosną szybko, szczególnie w erze Internetu.
2. W logicznie spójnym systemie edukacji od przedszkola do doktoratu, „od domu do dyplomu” pojęcia kapitału ludzkiego i kapitału społecznego powinny odgrywać kluczową rolę.
3. W systemie edukacji niezbędne jest przejście od nauczania poszczególnych przedmiotów do nauczania umiejętności, takich jak umiejętności poznawcze, umiejętności analityczne, umiejętności pracy w zespole.
4. W tym systemie wszystkie egzaminy powinny tworzyć jedną spójną całość, w której wszystkie pomiary są dokonywane w tych samych jednostkach i z tą samą dokładnością.

5. Egzaminacje te powinny zachowywać równowagę między centralizmem i regionalizmem, zarówno w konstrukcji, jak i w organizacji.

Interesujące studium na temat związku biologicznego zasobu rozwojowego z wynikami egzaminów zewnętrznych przedstawili R. Dolata i A. Pokropek (2012). Dowiedli, że różnice wieku mają niewielki (miesiąc życia przekłada się na $\frac{1}{4}$ punktu w skali IQ), ale istotny wpływ na te wyniki. Ten wpływ maleje w toku nauki szkolnej, jednak po wprowadzeniu obowiązku szkolnego sześciolatków powinien być brany pod uwagę. Istotnym diagnostycznie pytaniem o wpływ wieku ucznia określonej klasy na jego samoocenę zajął się K. Konarzewski (2013). Stwierdził, że w klasie III szkoły podstawowej samoocena jest wtórna w stosunku do osiągnięć szkolnych, wyjąwszy dwie grupy skrajne: „przyśpieszeni” (wyprzedzający obowiązek szkolnych) mają systematycznie wyższą samoocenę, a „opóźnieni” (starsi niż przewiduje obowiązek) – systematycznie niższą.

Z zakresu diagnostyki pracy nauczyciela warto odnotować badania Kornelii Rybickiej (2009), która, stosując test perspektyw czasowych Zimbardo, wykazała różnice postaw nauczycieli i uczniów gimnazjum. Tylko ci pierwsi są głównie nastawieni na przyszłość, podczas gdy dla drugich „istotna jest chwila, która niesie z sobą przyjemności, ekscytujące i nowe wyzwania, co pozwala im żyć intensywnie, unikając przy tym rutyny i nudy” (tamże, s. 287). Ta różnica wyjaśnia wiele trudności wychowawczych (2010).

K. Rybicka (2013) zaapelowała o pełniejsze wykorzystanie teorii psychologicznych w diagnostyce edukacyjnej. Jej zdaniem, „konieczność korzystania z teorii psychologicznych w diagnozach pedagogicznych jest jak matematyka w modelach astronomicznych: bez niej nie ma teorii naukowych, jest opowiadanie popularnonaukowe” (tamże, s. 319). Obok Zimbardo teorii spostrzegania czasu, odwołuje się do Kahnemana teorii dwóch systemów myślenia: intuicyjnego i refleksyjnego. Ten pierwszy przeważa u uczniów nastawionych fatalistycznie, niechętnych czytaniu i analizom lektur szkolnych, a ten drugi – w (mniejszej) grupie uczniów nastawionych przyszłościowo, czerpiących z lektur wzory życia.

Dla typowego ucznia liczą się w edukacji głównie teraźniejszość i przeszłość, nie zawsze idące w parze. Luka między **ocenianiem kształtującym**, dostarczającym uczniowi informacji potrzebnej do podniesienia umiejętności, a **ocenianiem sumującym**, zamykającym wybrany etap kształcenia, została już w ewaluacji i diagnostyce dostrzeżona. Ocenianie pośrednie, sumująco-kształtujące, w którym uczeń i nauczyciel interpretują osiągnięcia ze względu na cele uczenia się, zostało nazwane **ocenianiem orientującym**, trudnym dydaktycznie, ale cennym diagnostycznie (Niemierko, 2010).

Krytyczne studia rozwoju ewaluacji w edukacji, a w szczególności jej zastosowania do sprawowania nadzoru pedagogicznego w Polsce, przedstawili Leszek Korporowicz (2011) oraz Henryk Mizerek (2011). W pierwszym czytamy, że wprawdzie ewaluacja jest dużym postępem w stosunku do „mierzenia jakości pracy szkoły”, systemu o metodologicznie bardzo eklektycznym charakterze, ale ewaluacja w Polsce wpadła we wszystkie pułapki myślenia technokratycznego i jej funkcja kulturotwórcza wymaga odbudowy. Drugi z autorów dostrzega „kłatwę ewaluacji biurokratycznej” i skłonność administracji oświatowej do koniunkturalnego naginania obowiązujących procedur, bez niezbędnego łączenia wątków etycznych z technicznymi.

Zjawisko nadmiernego upraszczania procedur diagnostyczno-ewaluacyjnych w praktyce edukacyjnej L. Korporowicz (2012) nazwał **redukcjonizmem w badaniach ewaluacyjnych**. Powierzchnowe rozumienie pojęć prowadzi do bezdusznych działań, a nawet do pseudoewaluacji. Także nieświadomość ograniczeń rzetelności i trafności pomiaru przynosi duże szkody diagnostyce edukacyjnej, więc „ujawnianie błędów pomiaru jego użytkownikom stanowi powinność moralną nauk społecznych” (tamże, s. 49).

O wiązanie w diagnostyce badań o charakterze socjologiczno-statystycznym z badaniami w skali pojedynczego nauczyciela-wychowawcy upomniał się Bolesław Niemierko (2013). Analizom kapitału ludzkiego i społecznego oraz edukacyjnej wartości dodanej brak przełożenia na codzienną praktykę szkolną, potrzeba natomiast szerszych badań nad twórczością szkolną ucznia i Okoniowskiimi drogami uczenia się. Badacze podzielili się na „ilościowych” i „jakościowych”, stwarzając sztuczne bariery między tymi dwoma środowiskami. Pomocy w ich przewyżczeniu warto szukać w teoriach psychologicznych, takich jak wspomniana już Kahnemana teoria równoległych systemów myślenia.

Znaczącą dla diagnostyki edukacyjnej psychoanalityczną teorię **analizy transakcyjnej** (AT), wprowadzającej kategorie komunikacyjne: (a) Ja-Rodzic, osoby posiadającej władzę, (b) Ja-Dorosły, osoby dojrzałej (c) Ja-Dziecko, osoby niedojrzałej, przedstawił Jarosław Jagieła (2012). Jego zdaniem, ta metoda stanowi w jakimś sensie *terra incognita* współczesnej edukacji. Jej znaczenie dla diagnostyki edukacyjnej jest następujące:

Analiza transakcyjna sytuacji diagnostycznej

1. W każdej interakcji z uczniami nauczyciel przyjmuje jedną z trzech postaw (stanów Ja), na ogół zbyt często postawę Rodzica (autorytaryzm), a zbyt rzadko postawę Dorosłego (humanizm).
2. Uczeń przyjmuje jedną z trzech postaw wobec treści kształcenia (autorytetów naukowych), zbyt często – Dziecka (przyswajanie), zbyt rzadko Dorosłego (działanie) i Rodzica (odkrywanie).
3. Bywa, że uczeń podejmuje grę transakcyjną z nauczycielem, na przykład symulując rolę skrzywdzonego Dziecka lub brutalnego Rodzica.
4. Postawy nauczycieli-diagnostów także mogą być zróżnicowane i symulowane, co jest szybko spostrzegane przez uczniów.

Jeszcze głębiej sięga teoria neurobiologiczna, zaproponowana przez Marka Kaczmarzyka (2013) na podstawie **dydaktyki ewolucyjnej**, wykorzystującej genetycznie zaprogramowaną zdolność człowieka do zorganizowanego uczenia się. Antropologicznie uzasadniona ocena osiągnięć miałaby wykorzystywać pojęcie **memotypu**, zgeneralizowanego schematu poznawczego, który byłby rozpoznawany jako indywidualna konstrukcja **memów** – jednostek zaakceptowanej informacji. Autor referatu zakończył go pytaniem: „Czy biorąc pod uwagę zmienność memotypów, mamy jakąkolwiek szansę na ich świadome kształtowanie?” Jeśli nawet odpowiemy mocnym „tak” na to pytanie, pozostaje szansa na diagnozę memotypów ku indywidualnemu i społecznemu pożytkowi. Gdy rozpoznajemy już ludzki genotyp, to nowe zadanie wydaje się na tyle realne, że wspierająca je struktura pojęciowa może pomóc w przewyżczeniu ograniczeń tradycyjnych modeli uczenia się.

Podsumowanie

Dwadzieścia lat rozwoju diagnostyki edukacyjnej w Polsce przyniosło adaptację do jej potrzeb zbioru podstawowych pojęć, co pozwoliło na stopniowe wyłanianie się teorii diagnozy edukacyjnej. Pole zastosowań tej teorii jest na razie wąskie, gdyż zbyt wiele badań skupia się na gimnazjum, a zbyt mało na przebiegu edukacji domowej i wczesnoszkolnej, na kształceniu zawodowym i wyższym.

Przyczyną dominacji problematyki gimnazjalnej jest względna łatwość uzyskania danych statystycznych z egzaminów zewnętrznych i międzynarodowych badań osiągnięć piętnastolatków, a trudność dotarcia do zjawisk pedagogicznych w codziennych, nieunormowanych sytuacjach. Analizy ilościowe (z daleka) przeważają nad jakościowymi (z bliska), co ogranicza możliwość wykorzystania ich dorobku w pracy nauczyciela-wychowawcy.

Szansą diagnostyki edukacyjnej jest jej chłonność teoretyczna. Po przyswojeniu elementów psychologii poznawczej, teorii programów szkolnych, klasycznej i probabilistycznej teorii testów, oceniania szkolnego i komunikacji międzyludzkiej oraz po rozwinięciu teorii pomiaru sprawdzającego i edukacyjnej wartości dodanej, czekają w kolejce ekonomiczne teorie kapitału niematerialnego, psychologiczne – analizy transakcyjnej i być może nawet biologiczne – memotypów.

Diagnostyce edukacyjnej – od jej obwieszczenia w Polsce – niezczędzono zachęt (Muszyński, Lewowicki, Janowski) i konstruktywnej krytyki (Konarzewski, Muszyński). Zasada ograniczonego zaufania do jej metod i wyników obowiązuje nadal. Jej podstawą teoretyczną jest Messicka *teoria trafności* pomiaru, w tym zwłaszcza jej centralna część, oparta na badaniu konsekwencji (skutków społecznych) dokonywanych diagnoz. Bardzo ważne jest rejestrowanie i analizowanie postaw uczniów-wychowanków, ich rodziców, nauczycieli-wychowawców i innych udziałowców (administracji, opinii publicznej) wobec warunków, przebiegu i wyników edukacji.

Dla dalszego rozwoju diagnostyki edukacyjnej optymalny byłby powolny wzrost jej znaczenia, w tym stopniowe wprowadzanie do programów kształcenia pedagogicznego jako osobnego przedmiotu. Zredukowanie jej natomiast do wymiaru jednego z obszarów diagnostyki pedagogicznej, obok diagnostyki potrzeb, środowiska, dysfunkcji i zaburzeń (Wysocka, 2013), grozi rozkawałkowaniem. Z drugiej zaś strony, oficjalne (rozporządzeniem ministerialnym) wprowadzenie diagnostyki do obowiązków nauczycieli-wychowawców mogłoby skutkować sformalizowaniem na wzór praktyki ewaluacji pracy szkół.

Najlepiej byłoby pozwolić rozwijać się diagnostyce edukacyjnej samodzielnie, przy nieustannej rewizji założeń, dorobku i zadań do wykonania.

Literatura

Bednarek K. (2001): *Diagnoza szkoły bez ankiety?* W: K. Wenta (red.): *Pomiar edukacyjny jako kompetencje pedagogiczne*. Wyd. Naukowe Uniwersytetu Szczecińskiego, Szczecin.

Dolata R. (2008): *Czy segregacja uczniów ze względu na uprzednie osiągnięcia szkolne zwiększa efektywność nauczania mierzoną metodą EWD?* W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Dolata R. (2009): *Cicha rewolucja w polskiej oświacie: proces różnicowania się gimnazjów w dużych miastach*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Dolata R., Pokropek A. (2010): *Motywacja a wynik testu z nauk przyrodniczych. Studium na przykładzie PISA 2006*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Dolata R., Jakubowski M., Pokropek A. (2011): *Nowe spojrzenie na pozycję Polski w międzynarodowych badaniach umiejętności uczniów*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Dolata R., Pokropek A. (2012): *Czy warto urodzić się w styczniu? Wiek biologiczny a wyniki egzaminacyjne*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Kraków: Polskie Towarzystwo Diagnostyki Edukacyjnej.

Dubiecka A. (2006): *Skąd te garby? O tym, co może być powodem „nienormalności” rozkładu wyników sprawdzianu dla szóstoklasistów z kwietnia 2006 roku*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Dubiecka A., Szalenc H., Węziak D. (2006): *Efekt egzaminatora w egzaminach zewnętrznych*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Groenwald M. (2000): *Odpowiedzialność nauczyciela jako konstruktora i użytkownika zadań testowych*. W: B. Niemierko, J. Mulawa J. (red.): *Diagnoza edukacyjna. Zadania wyboru wielokrotnego*. Instytut Badań Kompetencji, Wałbrzych.

Groenwald M. (2001): *Czy godność nauczyciela jest zagrożona egzaminowaniem zewnętrznym?* W: B. Niemierko, W. Malecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław.

Groenwald M. (2007): *Nasza przyjaźń nigdy się nie rozstanie – inspirowana Arystotelesem analiza rozwiązań jednego zadania*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Hamryszak J. (1994): *O sprawdzaniu i ocenianiu osiągnięć uczniów w zakresie plastyki*. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Jagiela J. (2012): *Analiza transakcyjna jako metoda diagnozy pedagogicznej*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Jakubowski M. (2008): *Zastosowanie modeli EWD do ewaluacji polityki edukacyjnej*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Jakubowski M. (2009): *Międzynarodowe badania umiejętności uczniów a badania w Polsce*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Janowski A. (2007): *Poznanie uczniów jako podstawa indywidualizacji kształcenia*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Kaczmarzyk M. (2013): *Dyskretny urok oceny*. W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Konarzewski K. (1998): *Jak nauczyciele normują test osiągnięć szkolnych?* W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Konarzewski K. (1999): *Komu jest potrzebna diagnostyka oświatowa?* W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Ośrodek Diagnozy, Egzaminów Szkolnych i Informacji Pedagogicznej, Legnica.

Konarzewski K. (2000): *Miejsce testów wyboru w kulturze oświatowej*. W: B. Niemierko, J. Mulawa (red.): *Diagnoza edukacyjna. Zadania wyboru wielokrotnego*. Instytut Badań Kompetencji, Wałbrzych.

Konarzewski K. (2013): *Względny wiek uczniów w oddziale a ich samoocena*. W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Kondratek B. (2011): *Zrównywanie wyników egzaminów zewnętrznych z wykorzystaniem modeli IRT*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Korporowicz L. (2011): *Zmienne losy polskiej ewaluacji. Pomiędzy nadzieją, animacją i konfuzją*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Korporowicz L. (2012): *Refleksja jako działanie. O stałej potrzebie przelamywania redukcji w pojmowaniu badań ewaluacyjnych w edukacji*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Kowalik E. (2003): *Dwa rodzaje egzaminów w opiniach uczniów, nauczycieli oraz dyrektorów szkół*. W: B. Niemierko i in. (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej, Łódź.

Kozłowski W. (1999): *Ewaluacja pedagogiczna: niezbędna i niemożliwa*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Ośrodek Diagnozy, Egzaminów Szkolnych i Informacji Pedagogicznej, Legnica.

Krope P. (1994): *Ocena opisowa w pedagogice źródłem nieporozumień*. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Kutajczyk T., Przychodzeń B., Sterczyński R. (2009): *Od jakich czynników zależy skuteczność kształcenia w gimnazjach wiejskich i wielkomiejskich? Czynniki skuteczności kształcenia ogólnego w gimnazjach wielkomiejskich*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Lewowicki T. (1998): *Współczesne wyzwania wobec diagnostyki edukacyjnej – przemiany teleologii oświatowej i pojmowania pedagogiki a szanse rozwoju diagnostyki*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Mizerek H. (2011): *Tam gdzie kończą się zaklęcia, a zaczyna ewaluacja*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Muszyński H. (1998): *Diagnoza edukacyjna i jej miejsce w praktyce i teorii pedagogicznej*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Muszyński H. (2001): *Diagnostyka edukacyjna a humanizacja szkoły*. W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław.

Niemierko B. (1994): *Diagnostyka edukacyjna*. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Niemierko B. (1998): *Diagnostyka edukacyjna – cele, metody i szanse rozwoju*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Niemierko B. (2000): *Czy zadanie wyboru wielokrotnego nadaje się do diagnozowania procesów edukacyjnych?* W: B. Niemierko, J. Mulawa (red.): *Diagnoza edukacyjna. Zadania wyboru wielokrotnego*. Instytut Badań Kompetencji, Wałbrzych.

Niemierko B. (2001): *Ku czemu zmierzają egzaminy szkolne?* W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław.

Niemierko B. (2003): *Jaki pomiar dydaktyczny jest nam potrzebny?* W: B. Niemierko i in. (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej, Łódź.

Niemierko B. (2004): *Założone i uzyskane, sprawdzające i różnicujące znaczenie treściowe egzaminacyjnej skali pomiarowej*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Niemierko B. (2005): *Horyzonty diagnostyki edukacyjnej*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Niemierko B. (2006): *Oszustwo egzaminacyjne*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Niemierko B. (2010): *Między prawdą a skutecznością – perspektywy oceniania szkolnego*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Niemierko B. (2012): *Jawne i ukryte błędy pomiaru dydaktycznego*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Niemierko B. (2013): *Diagnostyka edukacyjna duża i mała*. W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

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*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Pokropek A. (2011): *Zrównywanie wyników egzaminów zewnętrznych w kontekście międzynarodowym*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Rappe A. (2007): *Opinie uczniów o egzaminie gimnazjalnym w 2006 roku*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Reinikainen P. (2009): *Sukces fińskich uczniów w badaniach PISA*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Rybicka K. (2009): *Dydaktyczne tendencje polskich nauczycieli mierzone w psychologicznej perspektywie postrzegania czasu*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Rybicka K. (2010): *Uczniowskie planowanie czasu w luce między ocenianiem wewnątrzszkolnym a ocenianiem zewnętrznym*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Rybicka K. (2013): *Szkolna diagnoza z psychologią myślenia i postrzegania czasu w tle*. W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Skorupiński P. M. (2002): *Trafność interpretacji wyników egzaminu gimnazjalnego w rekrutacji kandydatów do publicznej szkoły ponadgimnazjalnej*. W: B. Niemierko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnątrzszkolne i zewnętrzne a jakość pracy szkoły*. Regionalny Ośrodek Metodyczno-Edukacyjny „Metis”, Katowice.

Smolik M. (2005): *Ku porównywalności oceniania. Wykorzystanie wieloaspektowego modelu Rascha (model Facets) w procesie szkolenia egzaminatorów. Już rzeczywistość czy jeszcze science fiction?* W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Smolik M. (2006): *Każdy każdego własną miarą. Kilka uwag o kryteriach oceniania egzaminu ustnego z języka angielskiego na poziomie podstawowym*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Sterczyński R., Kutajczyk T., Przychodzeń B. (2008): *Od jakich czynników indywidualnych zależy edukacyjna wartość dodana w szkołach wiejskich?* W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Stożek E. (2006): *Zadania „rozrywające” w testach na przykładzie sprawdzianu ze szkoły podstawowej*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Stożek E., Dąbrowski H. (2006): *Zadania „rozrywające” w testach na przykładzie zadań maturalnych z matematyki*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Stróżyński K. (2003): *Pedagogiczny sens oceniania – komunikowanie wyników*. W: B. Niemierko i in. (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej, Łódź.

Sysło M. (2005): *Rozwój technologii informacyjnej a edukacja – stan, kierunki, wyzwania*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Szaleniec H. (2001): *Zastosowanie teorii analizy zadania testowego (IRT) w procesie egzaminowania zewnętrznego*. W: B. Niemierko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. PAN-DIT, Kraków.

Szaleniec H. (2006): *Oszukiwanie na egzaminie istotnym źródłem majowej porażki*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Szaleniec H. (2010): *Czy egzaminy zewnętrzne wpływają na wewnątrzszkolne ocenianie i politykę edukacyjną?* W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Szaleniec H. (2012): *Możliwości wykorzystania zrównanych wyników egzaminu gimnazjalnego w szkolnej diagnozie*. W: B. Niemięro, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Polskie Towarzystwo Diagnostyki Edukacyjnej – Kraków.

Szmigel M. K. (2004): *Treściowe znaczenie wyniku egzaminacyjnego ucznia i szkoły na przykładzie sprawdzianu w 2004 roku*. W: B. Niemięro, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Szmigel M. K. (2005): *Uczniowie o sprawdzianie 2004 i jego wyniku*. W: B. Niemięro, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Szmigel M. K. (2005): *O trudnościach i odpowiedzialności za komunikowanie wyników egzaminu maturalnego*. W: B. Niemięro, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Fundacja Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Szyling G. (2002): *„Idź w świat i radź sobie jak potrafisz”*. *Dругi układ wymagań programowych w oczach uczniów i nauczycieli*. W: B. Niemięro, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnątrzszkolne i zewnętrzne a jakość pracy szkoły*. Regionalny Ośrodek Metodyczno-Edukacyjny „Metis”, Katowice.

Szyling G. (2003): *„Być może będzie w przyszłości ozdobą polskiej piłki nożnej?”*. *Wybrane aspekty trafności oceny końcowej osiągnięć ucznia*. W: B. Niemięro i in. (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej, Łódź.

Tomkowicz J. (2008): *Konstrukcja i zastosowanie skal pionowych w pomiarze osiągnięć uczniów*. W: B. Niemięro, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Walukiewicz S. (2009): *Kapitał ludzki i społeczny jako przedmiot badań pedagogicznych*. W: B. Niemięro, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Walukiewicz S. (2010): *Kapitał ludzki w pomiarze dydaktycznym*. W: B. Niemięro, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Węziak D. (2005): *Zastosowanie wieloczynnikowego skalowania Rascha do porównania sposobu oceniania egzaminatorów*. W: B. Niemięro, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Wysocka W. (2013): *Diagnostyka pedagogiczna. Nowe obszary i rozwiązania*. OW „Impuls”, Kraków.

Dwie dekady dojrzewania diagnostyki edukacyjnej w Polsce

Diagnostyka edukacyjna jest teorią i praktyką rozpoznawania warunków, przebiegu i wyników uczenia się dzieci, młodzieży i dorosłych. Jako subdyscyplina pedagogiczna została zainicjowana w Polsce w 1993 roku. Od tego czasu odbyło się dziewiętnaście krajowych konferencji naukowych w tej sprawie, a od 2004 roku istnieje Polskie Towarzystwo Diagnostyki Edukacyjnej, koordynujące działalność w tej dziedzinie.

Artykuł zawiera przegląd głównych wątków konferencyjnych wraz z ich charakterystyką metodologiczną. Obok ogólnej teorii i etyki diagnozowania, w ich programie dominowała tematyka oceniania szkolnego, egzaminowania zewnętrznego, międzynarodowych badań osiągnięć szkolnych, zastosowań klasycznej i probabilistycznej teorii pomiaru, edukacyjnej wartości dodanej. Większość zaawansowanych metodologicznie badań obejmowała gimnazjum ze względu na dostępność wyników pomiaru osiągnięć ogółu kandydatów i absolwentów tego szczebla kształcenia.

Autor artykułu postuluje (1) śmiałe sięganie do dorobku pokrewnych pedagogice dyscyplin naukowych, takiego jak ekonomiczna teoria kapitału ludzkiego i społecznego, psychologiczna teoria analizy transakcyjnej, neurobiologiczna teoria memotypów, (2) stopniowe rozszerzanie diagnoz na cały zakres kształcenia ustawicznego i na wszystkie środowiska: rodzinne, szkolne, pozaszkolne, zawodowe, (3) wprowadzenie diagnostyki edukacyjnej do szkół wyższych jako przedmiotu kształcenia zawodowego pedagogów.

Słowa kluczowe: diagnostyka edukacyjna, pomiar dydaktyczny, egzamin zewnętrzny, ocenianie szkolne, Polskie Towarzystwo Diagnostyki Edukacyjnej.

Educational Diagnostics: Two Decades of Maturing in Poland

Educational diagnostics constitutes the theory and practice of recognizing contexts, processes, and outcomes of children's, youth's and adults' learning. As a pedagogical subdomain it was launched in Poland in 1993. Since then nineteen nation-wide conferences have been held in this field and in 2004 the Polish Association of Educational Diagnostics was established to coordinate the activities.

The paper carries a review of the main conferential threads as well as their methodological characteristics. Alongside the general theory and the ethics of diagnosis their programs were predominated by achievement grading, external examinations, international surveys, applications of classical and probabilistic measurement theory, educational value-added studies. Most of methodologically advanced research concerned junior high-schools (gymnasia) on the ground of availability of standardized measurement outcomes for unbiased populations of applicants and graduates.

The author of the paper claims (1) fearless reaching for achievements of the disciplines kindred to education, such as the economical theory of human and social capitals, psychological theory of transactional analysis, neurobiological memotype theory, (2) gradual introduction the diagnoses to the whole area of life-long education and to all the family, school, extracurricular, and vocational environments, (3) initiating educational diagnostics into regular teacher-training courses.

Keywords: educational diagnostics, educational measurement, mandatory examination, achievement grading, Polish Association for Educational Diagnostics.

Sylvia Jaskuła

Państwowa Wyższa Szkoła Informatyki i Przedsiębiorczości w Łomży

Leszek Korporowicz

Uniwersytet Jagielloński

ODRĘBNOŚĆ I WSPÓŁGRANIE

Polskie relacje diagnostyki i ewaluacji edukacyjnej

W warunkach polskiego systemu edukacyjnego diagnostyka edukacyjna znacząco wyprzedziła badania zwane ewaluacyjnymi, które w Polsce pod tą nazwą pojawiły się dopiero w drugiej połowie lat dziewięćdziesiątych. Warto jednak pamiętać, że w innej formie polskie fundamenty tego typu badań – np. socjotechnika, pojawiły się już w końcu lat siedemdziesiątych. W porównaniu z dojrzałym stanem rozwoju, świadomością metodologiczną i metodyczną diagnostyki edukacyjnej, badania ewaluacyjne w edukacji, niezależnie od swoich wcześniejszych analogii w naukach społecznych, definiowały nie tylko swój przedmiot, ale i metodologię korzystając z kilkudziesięcioletniej tradycji wypracowanej w świecie zachodnim, szczególnie kulturze anglosaskiej. Od początku jasne było, iż pomiędzy tymi dziedzinami istnieje wiele podobieństw, a nawet synergii, ale także daleko posuniętych odmienności, zważywszy, iż badania ewaluacyjne były i są aplikowane daleko poza sferą edukacji.

Sytuacja taka utrzymuje się w istocie do dnia dzisiejszego i wynika z obiektywnych różnic na wielu polach, choć po niespełna dwudziestu latach współistnienia są one lepiej rozumiane, traktowane z respektem i świadomym wkomponowaniem w całość celów teoretycznych, badawczych i społecznych obu tych form diagnozowania i refleksyjności w obrębie systemu polskiej edukacji.

Diagnostyka edukacyjna – trzymając się bardziej celów poznawczych, choć respektując całość procesów rozwojowych ucznia – preferowała zoperacjonalizowaną metodykę pomiaru, czyniąc z niej konkretny instrument wpływu i kształtowania rzeczywistości szkolnej. Ewaluacja zakładała włączenie różnych metod diagnostycznych, analitycznych i operacyjnych do rozpoznawania, refleksyjności i realizacji celów społecznych, ujmując wszystkie elementy procedury badawczej jako komponenty szerzej pojętego procesu o własnych wartościach i strategii rozwojowej, realizowanej w całej społeczności szkolnej oraz jej środowisku. Dlatego też nie tylko wynik, ale cały proces badawczy, jako uwikłana społecznie interakcja wszystkich jego uczestników, stał się obszarem postulowanej refleksji. Z perspektywy ewaluacyjnej żadna czynność poznawcza nie jest wolna od wartościowania, które powinno stanowić przedmiot społecznego dyskursu, a nawet dialogu jego społecznych aktorów. Same procedury badawcze, ale także czynności ich planowania i realizacji, stanowią więc tylko komponent złożonej sieci interakcji i wyborów. Niewątpliwie są one koniecznym elementem nie tylko taktyki, ale i strategii procesu ewaluacyjnego jako

profesjonalnej umiejętności widzenia „rzeczy” i ludzi w ich relacjach oraz horyzontach przestrzeni wartości.

Współcześnie na tyle lepiej rozumiemy specyfikę, jak i odmienność obu form „badania w działaniu”, aby mogły się wspierać, a nie wykluczać, kumulować, a nie redukować. Przy współczesnym obniżeniu profesjonalnych standardów działania w obszarze ewaluacji edukacyjnej wkomponowanej w procedury nadzoru, a więc daleko od idei „polskiej drogi do ewaluacji”, jakie towarzyszyły demokratycznym przeobrażeniom rzeczywistości polskiej połowy lat dziewięćdziesiątych, diagnostyka edukacyjna staje się sprzymierzeńcem w staraniach nie tylko o merytoryczną, ale i społeczną odpowiedzialność badań edukacyjnych, społecznych, jak i całego systemu edukacji.

Trudne początki. Rewolucja „od góry”

Gdy w drugiej połowie lat dziewięćdziesiątych na zlecenie ówczesnego ministerstwa edukacji rozpoczęły się starania o utworzenie zespołu ewaluacyjnego do programu edukacyjno-projektującego TERM, niewiele osób, łącznie z ministerialnymi animatorami tego przedsięwzięcia, zdawało sobie sprawę z rzeczywistej roli, jaką spełnić miały planowane zaledwie badania o dziwnej wówczas nazwie „ewaluacja”. Co istotne, od początku wskazywano, iż mają one wkomponować się w szerszy kontekst zmian społecznych zmierzających do demokratyzacji, uspołecznienia, wspierania oraz rozbudzania wyobraźni kadry kierowniczej i nauczycieli całego systemu edukacji. Dodawano także, że badania te wykorzystywać powinny istniejące już formy badań edukacyjnych, które mają podobne założenia i cele. Wskazania te były więcej niż zbieżne z diagnozą potrzeb społecznych, jaką realizowano w badaniach socjologicznych, ale także pedagogicznych i politologicznych. Przynosiło to znaczącą konsolidację „obozu reform”, co nazwano „polską drogą do ewaluacji”.

W klimacie poszukiwania i budzenia się inicjatywności wielu środowisk zainteresowanie ewaluacją stanowiło o emanacji takich wartości, jak:

- demokratyzacja kultury organizacyjnej polskich instytucji, szkół i życia publicznego,
- otwarcie dyskursu społecznego na procesy uspołecznienia i partycypacji,
- animacja potencjałów społecznych środowisk lokalnych, wspólnot pracowniczych, edukacyjnych i badawczych,
- kooperacja międzynarodowa organizacji uczących się i zorientowanych na rozwój,
- wspieranie myślenia i wzorców „refleksyjności instytucjonalnej” zorientowanych na dialog, środowiskowy i interakcyjny model edukacji,
- wyraźne postawienie kwestii jakości edukacji jako elementu samoświadomości, uspołecznienia, modernizacji i otwarcia na procesy rozwojowe (Korporowicz, 2011).

W okresie tym badania ewaluacyjne podejmowano pod bezpośrednim i szkoleniowym kierunkiem ekspertów brytyjskich, szczególnie Helen Simons, której udział, zaangażowanie i kompetencje trudne były do przecenienia. Zaproszeni przez nią szkoleniowcy, często najwybitniejsze postaci ewaluacji brytyjskiej, stanowili nie-dościgniony wzorzec projektowania, ale także wrażliwości społecznej, umiejętnego mobilizowania potencjałów społecznych oraz profesjonalnej realizacji badań ewaluacyjnych. To z jej inspiracji powstały wówczas książki, długo wykorzystywane nie

tylko środowisku edukacyjnym, złożone z tłumaczeń na język polski najbardziej klasycznych tekstów współczesnej teorii ewaluacji (Korporowicz, 1997; Mizerek, 1997).

Teoria i praktyka diagnostyki edukacyjnej w realiach polskiego systemu edukacji w inicjatywnym okresie rozwoju ewaluacji była nieporównywalnie bardziej rozwinięta, świadoma swych zadań, dojrzała akademicko, a prace i działania jej najwybitniejszego animatora – Bolesława Niemierki – stwarzały grunt właściwego kierunku dojrzewiania, choć także borykały się z systemową aplikacją (Niemierko, 1998, 2009, 2010).

Sięganie po ewaluację jako rodzaj „zachodniej nowinki” mogło w tej sytuacji wydawać się nie do końca zrozumiałym eksperymentem (Simons, 1987). Dopiero z dalszej perspektywy czasowej docenić można otwartość, analityczność, przykładową gotowość rekompozycji, wyspecyfikowania i skoordynowania działań dla dobra całości systemu, czego strategicznie podjął się B. Niemierko, zapoznany już wówczas z zachodnią praktyką badań ewaluacyjnych. Warunki polskie stawiały wszak wyzwania nowe. Diagnostyka edukacyjna, postrzegana z perspektywy ewaluacyjnej, podejmowała epistemologiczne i metodyczne problemy pomiaru doświadczenia rozwojowego ucznia, natomiast ewaluacja, korzystając także z rezultatów tej pierwszej, ogarnąć chciała aksjologiczne, interakcyjne i społeczne wymiary badania jako działania w konkretnej społeczności szkolnej, uwikłanej w kulturę pedagogiczną i charakterystykę jej środowiska. Z tego punktu widzenia ewaluacja stawiała się niemal antropologią diagnostyki wkomponowaną przy tym w idee partycypacji, zmiany społecznej i kulturowej o zmiennych uwarunkowaniach konkretnego miejsca, w jakim funkcjonowała szkoła i jej podmioty.

Studia podyplomowe w dziedzinie diagnostyki edukacyjnej, prowadzone przez B. Niemierkę na Uniwersytecie Gdańskim, były prawdziwą szkołą poszukiwania i dialogu, czasami polemiki, ale zawsze otwartego dyskursu diagnostyki i ewaluacji, wzajemnego uczenia się, wielokrotnego redefiniowania stanowisk, budowania tożsamości, ale i kierunku możliwych przewartościowań. Było to wówczas bardzo twórcze miejsce na tego typu dyskusje, prowadzone z udziałem licznej dzisiaj rzeszy absolwentów, ale nade wszystko osób, które stanowią o potencjale rozwojowym obu tych dyscyplin. Podobne studia, rozpoczęte znacznie później przez Leszka Korporowicza, przy wydatnym wparciu Henryka Mizerka, poświęcone były ewaluacji programów społecznych, ale posiadały i posiadają wiele odniesień do systemu edukacji. Prowadzone są do dzisiaj w Instytucie Socjologii Uniwersytetu Warszawskiego. Ich cele były i są odmienne, dlatego nie powtórzyły tego dialogu i nie wypracowały takiej synergii dyscyplin, jaka może być modelem dla wielu innych studiów, współcześnie coraz bardziej koniecznych.

Analogiczne, choć mniej wyspecyfikowane cele udało się realizować w innych ośrodkach szkoleniowych i akademickich w Polsce, jak np. w Studium Menedżerów Oświaty w Kaliszu, czy wielozadaniowych studiach podyplomowych podejmowanych w wielu ośrodkach szkoleniowych i akademickich w pierwszych latach dwutysięcznych. Niestety sama ewaluacja edukacyjna zmieniła się w jej nieokreślony rodzaj uproszczenia, a mianowicie w „mierzenie jakości pracy szkoły”, które wbrew pozorom szybko utraciło swoje przesłanie, porzuciło doświadczenia diagnostyki edukacyjnej, ale nie zbliżyło się także do ewaluacji. Był to niewątpliwie koniec „rewolucji od góry”, która utraciła wszystko, co było jej siłą, zamieniła się w faktyczną „kontrewolucję”, poprzez instytucjonalizację, biurokratyzację i technokratyzację, autonomię systemów, egzaminów i pseudoewaluacji.

Znаниеcki jako prekursor ewaluacji z współczynnikiem humanistycznym

Poszukiwanie polskiego modelu ewaluacji, jej tożsamości, ale i twórczych form współgrania z dynamicznie rozwijającą się diagnostyką edukacyjną, wyrastało z głębokiego przekonania, że naczelną wartością nie tylko edukacji, ale całego systemu społecznego cywilizacji zachodniej, jaki warto upowszechniać i wspierać, jest budzenie się i konsekwentne stymulowanie kilku podstawowych wartości. Są nimi: podmiotowość, partycypacja, demokratyzacja, samorealizacja, prawo do tworzenia oraz kultywowania własnych systemów wartości i tożsamości, o ile tylko nie degradują i nie są skierowane przeciw innym wspólnotom, a także nie naruszają godności innych ludzi. Aby to osiągnąć, każda ze społeczności, w tym społeczności edukacyjne, muszą mieć prawo do poznawania samych siebie, do własnych kryteriów wartościowania, niezależnie jak bardzo i w jakim zakresie potrzebne są im badania porównawcze i standaryzacje ponadśrodowiskowych osiągnięć. Tradycje takie nie były w polskim dorobku nauk społecznych czymś nowym, chociaż coraz bardziej technokratyczne fale rozmaitych strategii zarządzania wydawały się skutecznie produkować ich swoistą „niepamięć”, wbrew głoszonym często założeniom. Tradycją i realnym dorobkiem uwzględniania postulatu podmiotowości w badaniach społecznych, w tym ewaluacji, była spuścizna Floriana Znаниеckiego. Nawet dzisiaj odkrywamy w niej ciągle mało jeszcze znane elementy. Wyprzedziły one swoją epokę, współcześnie są czynnikiem humanizującym badania społeczne, edukacyjne i ewaluacyjne.

F. Znаниеcki zwracał uwagę na potrzebę badania wartości systemu edukacyjnego w perspektywie zmian społecznych, jakich wymagało modernizujące się społeczeństwo amerykańskie lat trzydziestych ubiegłego wieku. Pod wieloma względami były to studia pionierskie, zwłaszcza zupełnie nieznanymi raporty ewaluacyjne, których znaczenie ujawnia się nam dopiero współcześnie, w sytuacji większego skomasowania, ale i wiedzy o znaczeniu badań ewaluacyjnych (Znаниеcki, 1998). Ich konceptualizację należało odnieść do podstawowych pytań i potrzeb ówczesnego społeczeństwa, rozpoznając rzeczywistych i potencjalnych liderów zmiany, a tym samym rozwojowo zorientowanej ewaluacji. Idea główna tak zorientowanej ewaluacji wynikała z przekonania, iż o podstawie czynności wartościowania stanowić mają nie badacz i zewnętrzne w stosunku do badanej społeczności kryteria. Podstawy tej należy szukać w realnych doświadczeniach, potrzebach, potencjałach i intencjach przedstawicieli konkretnej społeczności, która poszukuje swojej tożsamości lub chce ją jedynie zachować. Oznacza to również, że wartościowane działania i inicjujące je systemy wartości traktujemy w sposób podmiotowy, jako wartości o własnym prawie do istnienia, w ramach których dokonać się musi właściwe projektowanie równie podmiotowo zorientowanej ewaluacji. Tylko w ten sposób całościowo pojęty proces ewaluacji, który powinien rozpoczynać się od rozpoznania systemu wartości danej grupy, może mieć autentycznie rozwojowy, wspierający i nie-przedmiotowy charakter.

Z opisanych powyżej względów, podmiotowo zorientowana postawa społeczna badacza – o jaką apelował F. Znаниеcki w znanym w kręgu humanistycznie zorientowanej socjologii postulatcie „współczynnika humanistycznego” – oraz znakomicie zaprojektowana strategia badań to jedne z kluczowych założeń ewaluacji humani-

stycznej w dobie strategicznych wyzwań zmiany społecznej. Jej sens coraz bardziej ujawnia się w sytuacji, gdy systemy wyznaczające kryteria oceny mają charakter anonimowy, wynikają z zewnętrznych wobec ewaluowanych społeczności odniesień, zwłaszcza gdy uwikłane są w obce im interesy i nie są przedmiotem żadnej negocjacji. Innymi słowy, projekt F. Znanieckiego wynika z głębokiego respektu dla podmiotowego znaczenia wartości poszczególnych grup, niezależnie od możliwości i potrzeby ich zmiany w wyniku świadomego i refleksyjnego zamysłu.

Wrażliwość badacza ewaluatora na system wartości analizowanych działań i realizujących je ludzi jest więc fundamentalnym przesłaniem społecznej racjonalności i odpowiedzialności badań ewaluacyjnych. Wrażliwość ta powinna otwierać się na sytuacje, w których rozpoznać należy realną odmiennność systemów wartości konkretnych społeczności, jak również zrekonstruować ów system poprzez realny udział osób, które go praktykują, doświadczają i respektują. Tak zorientowana ewaluacja staje się *de facto* nie tylko kulturowa, a więc wrażliwa na kontekstualny i lokalny charakter wyznaczników wartościowania, ale w równej mierze partycypacyjna, czasami nawet międzykulturowa, poprzez próby negocjowania akceptowalnych kryteriów ewaluacji w programach realizowanych w różnych środowiskach społecznych, ale nie „przygniecionych” ciężarem standaryzacji. Odwołanie się do koncepcji współczynnika humanistycznego może uratować nie tylko kulturową, a tym samym środowiskową adekwatność ewaluacji, ale nade wszystko odpowiedzialność i społeczną racjonalność badań ewaluacyjnych. Jest to jedyna droga do skutecznej profilaktyki przed jej biurokratyczną rytualizacją, a czasami wręcz alienacją (Korporowicz, 2012). Co jednak istotne, współczynnik humanistyczny inspiruje nas do czegoś więcej niż tylko „rozpoznania” lokalnych kontekstów wartości oraz ich kulturowych odmienności, inspiruje do partycypacyjnej formuły w całym procesie ich „eksplorowania”, a więc nie tylko rozpoznania, ale i analizy, wnioskowania i aplikacji. Polskie Towarzystwo Ewaluacyjne, ale także wiele innych środowisk badawczych widziało w promowaniu tak pojętej ewaluacji swoją misję, społeczne i edukacyjne powołanie.

Dialogiczny i partycypacyjny sens ewaluacji inspirowanej społeczną aksjologią F. Znanieckiego jest więc bardzo świadomym celem humanizacji stosowanych badań społecznych, o którym współcześnie zapomniano. Ewaluacja „czwartej generacji”, zwana ewaluacją demokratyczną lub dialogiczną, o którą zaczęto się upominać od końca lat siedemdziesiątych ubiegłego wieku, ma pod tym względem bardzo ważkiego poprzednika, który praktykował ją blisko pięćdziesiąt lat wcześniej. Współcześnie postulat ewaluacji ze współczynnikiem humanistycznym rozbija się o dyrektywne instrukcje ewaluacji biurokratycznych, które przyjmują z wielką atencją zestandaryzowane kryteria wartościowania, podobnie jak kontrolnie i nadzorczo zorientowane pomysły użycia ewaluacji w zadaniach na skalę masową, ponadnarodową, ale także sektorową. Wieloetapowy i rozłożony na wiele podmiotów proces ewaluacji zostaje tam bardzo okrojony i po scentralizowanej fazie projektowania skoncentrowany na procedurach generowania, analizy i prezentowania danych, a zespoły ewaluacyjne sprowadzone do roli ekspertów „realizacji badań”. W tak pomyślanym procesie nie mieści się także diagnostyka edukacyjna, która w jednakowym stopniu, jak ewaluacja, pozostaje w cieniu procesu nadzorczego i w równym stopniu zredukować musi uspołecznienie czynności projektowania i realizowania badań.

Ewaluacja jako inter-akcja

Przyjmując zainicjowane przez F. Znanieckiego widzenie edukacji i ewaluacji jako proces intencjonalnej zmiany, widzimy, jak wiele zależy w nim od relacji, jakie zachodzą pomiędzy jego uczestnikami. Ewaluacja jawi się wówczas nie jako jednorazowa procedura zrobienia fotografii, ale jako bardzo wielowątkowy film, w którym są aktorzy, jest akcja, jest teatralna prawie scena, a więc odwołać musimy się do metod badających dyskurs, wymianę doświadczeń, konstrukcję i destrukcję wspólnych działań. Polska szkoła ewaluacji świadomie podkreśla wagę jej drogi uspołecznienia, pobudzenia potencjału rozwojowego ewaluowanego programu, ale także interakcji pomiędzy ewaluatorami i wszystkimi innymi podmiotami procesu. Ewaluacja nabrać może w ten sposób cech animacji, jest wezwaniem do działania poprzez umiejętne pytanie, analizę, interpretację i pobudzanie zasobów, jakimi dysponują ludzie, proces i program kształcenia. Nie bez powodu wskazuje się na innowacyjne aspekty procesu ewaluacyjnego, ale także jego wartość pedagogiczną, co klasyfikuje ten rodzaj działania w obszarze zainteresowania pedagogiki społecznej. Taka funkcja ewaluacji wychodzi poza cechy trzeciej generacji, powstałej w latach siedemdziesiątych, wkomponowanej bardzo utylitarnie w procesy podejmowania decyzji. Wzmacnia ona cechy czwartej generacji powstałej w latach osiemdziesiątych, nazwanej często dialogiczną, uspołecznioną czy demokratyczną i wytycza drogi piątej generacji funkcjonującej już w tworzących się dopiero warunkach społeczeństwa wiedzy.

W relacji do diagnostyki edukacyjnej – tak pomyślana i praktykowana ewaluacja wykazuje się „słabszymi” walorami twardej diagnostyki. Z tych to powodów oskarżana wręcz była o to, że staje się formą „interwencji społecznej”, że zamienia funkcje badawcze na funkcje konsultacyjne, że przejmując zadania zbyt dosłownie pojętej pedagogiki środowisk społecznych. Zarzuty te mogą, aczkolwiek nie muszą, być słuszne. W obecnej fazie wkomponowania ewaluacji w system oświatowego nadzoru brzmią one niemal jako komplementy z racji zupełnie innych oskarżeń praktyki ewaluacyjnej o zdecydowanie nadmierny poziom przedmiotowości i technokratycznego stosunku do szkoły. Zagrożenia osłabienia badawczej wartości ewaluacji są niestety realnym zagrożeniem, szczególnie w pośpiesznie i rzeczywiście amatorsko organizowanym procesie badawczym, który nie tylko nie jest badaniem, ale także właściwie uprawianą ewaluacją. Brak profesjonalnego przygotowania osób, które sztucznie i doraźnie postawione zostały w roli ewaluatora, skutkuje serią błędów, ewidentnych wypaczeń, brakiem jakichkolwiek zasad nie tylko projektowania, ale i doboru metod badawczych, nieumiejętnością doboru prób badawczych i nieporadnością, a czasami wręcz bardzo niskim poziomem analizy i interpretacji danych. Tak realizowana ewaluacja jest prawdziwym nieszczęściem jakkolwiek pojętego działania i stosunkowo często karykaturą ewaluacji. Z tego też punktu widzenia realizacja badań ewaluacyjnych bez należytego przygotowania zarówno samych ewaluatorów, jak i społeczności szkolnych, przynosi konsekwencje odwrotne od zamierzonych – uczy tolerowania działań pozornych. Ewaluacja w takiej formie nie realizuje większości przypisywanych jej celów i podważa wiarygodność przywoływanych przy tej okazji motywów jej podejmowania, jak: innowacyjność, refleksyjność, doskonalenie jakości, dialogiczność i nastawienie na rozwój.

Ujawnienie interakcyjnych cech, ale i możliwości procesu ewaluacyjnego, na który składają się: a) ustalenie społecznych podmiotów ewaluacji, b) uspołeczniona czynność planowania i projektowania, c) zbierania danych, d) analizy danych, e) pi-

sania raportu, f) komunikowania i g) aplikacji wyników, jest szczególnie istotne w dobie licznych biurokratycznych zagrożeń, przed jakimi stoi ewaluacja jako strategia wspomagająca procesy zarządzania. Na takie niebezpieczeństwa narażona jest ewaluacja nie tylko systemu oświaty, ale także działań finansowanych z funduszy strukturalnych Unii Europejskiej. Ta ostatnia musi stać się umiejętnym poszukiwaniem relacji między często sprzecznymi jej cechami, jak biurokratyzacja i uspołecznienie, ale także centralizacja i regionalizacja, instrumentalizacja i upodmiotowienie, profesjonalizacja i społeczna partycypacja, specjalizacja i całościowość, depersonalizacja i personalizacja wartości.

Proces ewaluacyjny jest formą interakcji nie tylko ludzi, ale także instytucjonalnych jego podmiotów, systemów myślenia i postaw, jest także możliwością otwierania się na inspiracje, doświadczenia i konkretny dorobek innych form badawczego zaangażowania w rzeczywistość edukacyjną (Groenwald, 2011). Od wielu już lat coraz bardziej potrzebne jest nie tylko poszerzenie, ale i uwzględnienie konkretnych doświadczeń zdobywanych w trakcie diagnostyki edukacyjnej, przede wszystkim pod względem profesjonalnego przygotowania kadry, poziomu zaangażowania środowiska naukowego, zarówno społeczności akademickich, jak też zapotrzebowania na taki udział centralnych instytucji edukacyjnych. Współczesny stan rzeczy jest niemal zaprzeczeniem modelu współdziałania. Badania, ale także prace koncepcyjne i szkoleniowe w dziedzinie ewaluacji skutecznie omijają uniwersytety i wydziały pedagogiczne. Ujawnianie pedagogicznego wymiaru badań ewaluacyjnych i swoistej „pedagogiki ewaluacji” jest nawet w środowiskach edukacyjnych postrzegane jako niezbyt uprawnione „naciąganie” czysto badawczej, nadzorczej i zarządczej funkcji tego typu działań (Jaskuła, 2013). Wiąże się to z brakiem właściwego miejsca problematyki ewaluacyjnej w uczelniach pedagogicznych, co system edukacji pozbawia akademicko wykształconych ewaluatorów. Najczęściej skazani są oni na doraźne kursy i szkolenia prowadzone często przez słabo przygotowane firmy szkoleniowe. Sytuacja ta jest nie do utrzymania w dalszej perspektywie rozwoju i świadczy o daleko idącym braku komunikacji, a więc interakcji wymienionych podmiotów zarówno na płaszczyźnie strategicznej, określanej przez instytucje centralne i regionalne, jak i w praktyce rzeczywistości szkolnej.

W obecnej fazie zaawansowania polskich badań ewaluacyjnych równie silny jest brak interakcji między różnymi jej formami, ustawiczna tendencja mylenia ich funkcji, a w konsekwencji nieumiejętność wypracowania obszarów komplementarnych. Nie dostrzega się dużej różnicy w funkcjach i profilu, jaki charakteryzuje ewaluację zewnętrzną i wewnętrzną. Ta pierwsza zbliża się ku stronie technokratycznej, porównawczej, z koniecznością standaryzacji kryteriów. Ewaluacja wewnętrzna skłania ku procedurom zindywidualizowanego i podmiotowego traktowania wymogów konkretnej szkoły i konkretnej społeczności, procesu planowania i wykorzystania ewaluacji. Wymogi ewaluacji zewnętrznej nakierowane są na kumulatywność wyników w obrębie całego systemu oświaty. Nawet jednak ewaluacja zewnętrzna posiada duże możliwości dialogiczne i animacyjne, realizowana jest bowiem w konkretnej sytuacji i zawsze służyć ma konkretnej szkole lub organizacji. Takie zakorzenienie procesu ewaluacyjnego stanowi o tożsamości ewaluacji, jeśli nie ma się ona przeistoczyć w typowe badania naukowe.

Ewaluacja nie jest badaniem naukowym, choć posługuje się metodami naukowymi w takim stopniu, w jakim jest to tylko możliwe. Jest to jedno z najczęściej

spotykanych nieporozumień, które odbiera ewaluacji jej tożsamość, zawsze nakierowaną na konkretny obiekt, a nie zmierzającą do odkrywania ogólnych prawidłowości, praw o zasięgu ponadkontekstualnym. Taki jest także zakres animacyjny ewaluacji, skoncentrowanej na służebności w stosunku do zdefiniowanego przedmiotu w konkretnym miejscu i czasie. Dlatego im szerszy zakres przedmiotowy badania ewaluacyjnego, tym bardziej narażony jest on na skłanianie się ku dekontekstualizacji, centralizacji i biurokratyzacji. Sprawą bardzo dojrzałego i świadomego zamyśłu, umiejętnej interakcji ludzi, potrzeb oraz wartości jest równoważenie i konieczne uspołecznienie cech systemu ewaluacji oświaty, aby nie stał się on rytualnym zabiegiem dla anonimowych potrzeb biurokratycznie zorientowanej administracji. Ewaluacja taka staje się formą audytu i akredytacji, które też mają swoje ważne funkcje, ale traci się to, co w ewaluacji może być najistotniejsze: odkrywanie i animowanie endogennych, a więc wewnętrznych potencjałów rozwoju.

Ewaluacja w relacji do diagnostyki edukacyjnej – posiada stosunkowo najwięcej celów wspólnych, nawet wtedy gdy dotyczą one obszarów sprawozdawczych lub korekcyjnych. W obu przypadkach konieczny jest wymóg dużej świadomości i refleksyjnej metodologii, niwelowania udziału czynników administracyjnych, zewnętrznych w stosunku do badanych procesów, a więc duży udział wiedzy i kompetencji o charakterze merytorycznym (Szyling, 2013). Na przestrzeni ostatnich kilkunastu lat relacje te przybrały pewne powtarzalne i warte usystematyzowania formy. Jak każda klasyfikacja i typologia, wykazuje się ona uproszczeniami. W wielu sytuacjach relacje te są znacznie bardziej dynamiczne i wielowymiarowe, a na podkreślenie zasługuje duża liczba obszarów współdziałania. Propozycja syntetycznego ujęcia relacji wydaje się jednak potrzebna. Zawarta została w tabeli, w środku której znajdują się pola charakteryzujące wspólne cechy diagnostyki i ewaluacji.

Tabela 1. Diagnostyka edukacyjna i ewaluacja – odrębność i współgranie

DIAGNOSTYKA EDUKACYJNA		EWALUACJA W EDUKACJI
ETYMOLOGIA OKREŚLEŃ		
gr. <i>διγνωσις</i> – rozpoznanie		ang. <i>valuation</i> – wartościowanie
ISTOTA		
uściślone rozpoznawanie warunków, przebiegu i wyników uczenia się (Niemierko, 2009, s.30)		uspołeczniony proces zmierzający do określenia wartości danego obiektu lub działania według przyjętych kryteriów w celu jego udoskonalenia, lepszego rozumienia i rozwoju
GŁÓWNE CELE		
pomiar edukacyjny	rozpoznanie, wspomaganie, wyjaśnianie	uczenie się, zmiana, interakcja społeczna
GŁÓWNE FUNKCJE		
sprawozdawczo-kontrolna	formatywna, informacyjna, poznawcza, wyjaśniająca, pragmatyczna, rozwojowa	rozumiejąca, animacyjna
ZAKRES		
wyniki uczenia się	przebieg, uwarunkowania i rezultaty procesu uczenia się	celowe działania, ich dynamika, konsekwencje i realizowanie przez nie wartości
GŁÓWNE CECHY		
rzetelność, trafność, obiektywizm, dokładność,	wiarygodność, adekwatność, komplementarność, bezstronność, nastawienie na rozwój	uspołecznienie, dialogiczność, demokratyczność, nastawienie na proces

DIAGNOSTYKA EDUKACYJNA		EWALUACJA W EDUKACJI	
CHARAKTER			
klasyfikacyjny	utylnitarny, który ma przede wszystkim wspomagać konkretnych odbiorców w realnej sytuacji, a także wywoływać pozytywne zmiany i poprawiać jakość zarówno bieżących, jak i przyszłych działań	elastyczny, kontekstualny	
PODEJŚCIE DO PROCESU BADANIA			
weryfikacja stopnia realizacji celu, dyscyplinowanie i kontrola	świadome konstruowanie procedury diagnostycznej i badawczej	Interaktywność i negocjacyjne podejście do procesu badania, wzajemne oddziaływanie badacza i badanego, będącego jednocześnie podmiotem uczenia się	
PLANOWANIE			
na podstawie kompetencji eksperta	uspołecznione pośrednio poprzez reprezentantów i odpowiednio doboranych ekspertów	w środowisku ewaluowanej instytucji z udziałem badanych	
PORÓWNYWALNOŚĆ WYNIKÓW			
pożądana	porównywalność w wybranych zakresach	często brak możliwości dokonywania porównań, ze względu na indywidualny charakter ewaluacji, który uwzględnia czynniki zewnętrzne i wewnętrzne badanego programu lub instytucji	
RAPORT KOŃCOWY			
ściśle określona struktura i procedura	dopasowanie do potrzeb odbiorcy	możliwość interaktywnej korekty raportu	
EFEKT KOŃCOWY			
przedstawienie informacji, oszacowanie wyników badanego działania	zgrupowanie wiedzy i rozwój refleksyjności działań	określenie wartości badanego działania, doświadczenie i kompetencje powstałe w wyniku interaktywnego procesu uczenia się oraz aplikacja wyników	

Z powyższego zestawienia (Jaskuła, 2012) wynika, iż istnieje wiele możliwości porównań zakresu i charakteru zastosowań diagnostyki i ewaluacji edukacyjnej, które jak w przypadku każdej rozwijającej się dyscypliny ulegają modyfikacjom, aktywnie reagują na potrzeby społeczne, ale także modelowane są przez regulacje, a niekiedy i wymogi administracyjne. Dlatego regularnie powraca pytanie o cele systemowe, tożsamość oraz stopień wspierania nie tyle ich funkcji kontrolnej, nadzorczej i regulacyjnej, ile o wspieranie podmiotowości, refleksyjności oraz umiejętności rozwojowych konkretnych osób i społeczności, w realiach organizacyjnych oraz edukacyjnych współczesnego świata.

Zakończenie

Pomimo wielorakich uwikłań i odmiennych warunków powstania diagnostyka i ewaluacja edukacyjna w latach dziewięćdziesiątych próbowały wkomponować się w system reform edukacyjnych, ściśle powiązanych z całokształtem reform społecznych i politycznych. Każda z tych dziedzin widziała swoją rolę w nieco inny,

choć dopełniający się sposób. W przypadku ewaluacji początkowy okres „rewolucji od góry” ujawnił wszystkie swoje mankamenty, brak zakorzenienia w świadomości rozpoznawanych potrzebach środowiska edukacyjnego na poziomie szkół oraz fluktuacje ekip rządowych, które utraciły przekonanie i zainteresowanie konsekwentnym rozwojem badań ewaluacyjnych, pograżając się w biurokratycznych wyzwaniach nowych systemów zarządzania. Uspołeczniona wizja ewaluacji jako dialogicznej formy refleksyjności została niemal całkowicie zapomniana, a polska droga do ewaluacji to jedynie przypomniany z rzadka postulat – ocalały w nielicznych ośrodkach akademickich oraz w środowisku Polskiego Towarzystwa Ewaluacyjnego.

Ponowne zainteresowanie ewaluacją z powodu jej obligatoryjnego charakteru w programach finansowanych z funduszy strukturalnych Unii Europejskiej, nie od razu dotyczyło sfery edukacji, powodując faktyczne przeniesienie ośrodków podejmujących badania ewaluacyjne na teren Ministerstwa Rozwoju Regionalnego oraz administracji publicznej i samorządowej. Zmiana ta nie wiązała się jednak z utrzymaniem tej samej koncepcji ewaluacji. W nowych warunkach niemal całkowicie odeszła ona od funkcji demokratyzujących na rzecz sprawozdawczych, monitorujących i kontrolnych, wkomponowując się w techniki, nowomowę oraz kompletne wymieszanie konwencji zarządczych, dalekich od misji i standardów ewaluacji. W zestandaryzowanych kanonach biurokratycznych redukcji, cenne w ewaluacyjnym procesie dynamizmy wspierania refleksyjności, dialogiczności i partycypacji zamienione zostały na aplikację gotowych i przysyłanych w załącznikach „warunków specyfikacji umowy”, a procedury i proces ewaluacyjny w zwyczajne techniki monitorowania i sterowania działaniami badanych społeczności.

Współcześnie jesteśmy niestety zaledwie w fazie rodzenia się krytycznej świadomości mankamentów procedur uprzedmiotowienia procesów ewaluacyjnych i zamiany ich w reprodukcję stosunków władzy. Dzieje się tak poprzez standaryzację kryteriów wartościowania ustalanych przez ośrodki decyzyjne w zgodzie z preferowaną przez nie wizją i interesem zarządzanych przez nie organizacji. W ślad za takim podporządkowaniem badań diagnostycznych i ewaluacyjnych postępuje preferowany profil celów ewaluacji, nastawiony na dość wąskie spektrum wartości. Omijają one demokratyzację, partycypację czy dialogiczność ewaluowanych procesów. Większość komercyjnych ośrodków usług badawczych, w tym ewaluacyjnych, nie jest zainteresowana krytyczną analizą tak powstałej sytuacji, stanowiąc swoiste alibi dla instytucji decydenckich, które chętnie korzystają z tak jednowymiarowo pojętego, profesjonalnego sposobu działania. Powstałe w ten sposób techniki ewaluacyjnego wzmacniania efektywności programów, jako koronnego kryterium ich wartościowania, nie mogą degradować ich sensu społecznego i podporządkowywać zewnętrznie ustalonym wyznacznikom oceny. Społeczny etos badań ewaluacyjnych spotyka się z etosem obrony profesjonalnych kompetencji i tożsamości niezależnych badań, jaki reprezentuje Polskie Towarzystwo Diagnostyki Edukacyjnej, w którym otwarcie mówi się nie tylko o ściśle profesjonalnych, ale także etycznych standardach badań stosowanych, szczególnie w dziedzinie edukacji. Podobne standardy wypracowane zostały przez Polskie Towarzystwo Ewaluacyjne.

Dalszy rozwój badań ewaluacyjnych w sytuacji dominacji technokratycznych metod zarządzania z pewnością nie będzie zadaniem ani łatwym, ani szybkim i z naturalnych powodów narażony będzie na funkcjonalne i koncepcyjne podporządkowanie realnym ośrodkom zarządzania. One to wytwarzają własny język, sposób

formułowania i wyboru celów, w tym także styl i kulturę ewaluacji nowych, instytucjonalnych „mandarynów władzy”. Celem strategicznym byłoby upublicznienie polityki diagnostycznej i ewaluacyjnej zarówno na szczeblu centralnym, jak i lokalnym, w postaci powołania publicznych ośrodków jej wdrażania. Oczywiście nie jest to możliwe poprzez wkomponowanie badań ewaluacyjnych w systemy kontroli, nadzoru oraz stworzenie czegoś w rodzaju policji ewaluacyjnej i promowanie nie dialogicznych, ale autokratycznych procedur badawczych. Stworzenie programów i instytucji edukacyjnych w dziedzinie ewaluacji, następnie ośrodków nastawionych na rozwój, a nie kontrolę działań ewaluacyjnych, traktowanych jako element wsparcia placówek edukacyjnych w ich świadomej i możliwie profesjonalnej zmianie społecznej, to pałaca potrzeba polskiego systemu społecznego.

O podobnym programie pisał Florian Znaniecki, tworząc tradycję polskiej szkoły ewaluacji. Wspólna inicjatywa ośrodków akademickich oraz stowarzyszeń badawczych, twórczych i środowiskowych rozpocząć się musi od rekonstrukcji tej właśnie tradycji, wkomponowania jej we współczesne realia organizacyjne i społeczne, przypomnienia i analizy procesu ewaluacyjnego, właściwego projektowania ewaluacji, szczególnie w warunkach społeczeństwa wiedzy, a więc konstytuując założenia i metody ewaluacji „piątej generacji”. To ona wykreować powinna nie tylko nową umiejętność obecności w społeczeństwie sieci, ale także hamowania wielkich zapędów na ukryte formy sterowania, reprodukcji i formowania nieznanych jeszcze w historii technik oddziaływania, wpływu, a nawet dominacji.

Literatura

Groenwald M. (2011): *Etyczne aspekty egzaminów szkolnych*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Guba E. G., Lincoln Y. S (1989): *Fourth generation evaluation*. Sage, London.

Jaskuła S. (2010): *O potrzebie re-ewaluacji w społeczeństwie nadmiaru informacji*. W: K. Kasprzyk (red.): *Ewaluacja programów operacyjnych – konteksty, dylematy, praktyki*. Urząd Marszałkowski Województwa Kujawsko-Pomorskiego w Toruniu, Toruń.

Jaskuła S. (2011): *Ewaluacja versus akredytacja – dwa światy w jednej przestrzeni*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Wyd. Grupa Tomami, Kraków.

Jaskuła S. (2012a): *Ewaluacja rozwojowa jako wyzwanie współczesnego systemu edukacji*. „Zarządzenie Publiczne. Zeszyty Naukowe Instytutu Spraw Publicznych UJ”, nr 3 (19).

Jaskuła S. (2012b): *Mutualizm ewaluacji i diagnozy edukacyjne*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. Grupa Tomami, Kraków.

Jaskuła S. (2013): *Pedagogika ewaluacji*. „Teraźniejszość–Człowiek–Edukacja”, nr 4 (64).

Korporowicz L. (red.) (1997): *Ewaluacja w edukacji*. Oficyna Naukowa, Warszawa.

Korporowicz L. (2001a): *Ewaluacja i społeczeństwo*. W: E. Hałas (red.): *Rozumienie zmian społecznych*. TN KUL, Lublin.

Korporowicz L. (2001b): *Rozumienie ewaluacji. Historia, która ma przyszłość*. W: J. Ł. Grzelak, M. J. Sochoczek (red.): *Ewaluacja profilaktyki problemów dzieci i młodzieży*. Pracownia Profilaktyki Problemowej, Warszawa.

Korporowicz L. (2010a): *Interakcyjna misja ewaluacji*. W: G. Mazurkiewicz (red.): *Ewaluacja w nadzorze pedagogicznym. Konteksty*. Wyd. UJ, Kraków.

Korporowicz L. (2010b): *Współczesne koncepcje ewaluacji w kontekście wyzwań rozwoju regionalnego*. W: K. Lotko-Czech (red.): *Ewaluacja programów operacyjnych na poziomie regionalnym – teoria i praktyka*. Urząd Marszałkowski Województwa Opolskiego, Opole.

Korporowicz L. (2011): *Zmienne losy polskiej ewaluacji. Pomiędzy nadzieją, animacją i konfuzją*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Wyd. Grupa Tomami, Kraków.

Korporowicz L. (2012): *Refleksja jako działanie. O stałej potrzebie przełamywania redukcji w pojmowaniu badań ewaluacyjnych w edukacji*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Wrocław.

Korporowicz L. (2013): *Spoleczna etyka ewaluacji*. „Teraźniejszość–Człowiek–Edukacja”, nr 4 (64).

Mizerek H. (1997): *Ewaluacja w szkole*. Wyd. MG, Olsztyn.

Mizerek H. (2009): *Po co szkole ewaluacja*. „Problemy Wczesnej Edukacji”, nr 1.

Mazurkiewicz G. (red.) (2010): *Ewaluacja w nadzorze pedagogicznym. Konteksty*. Wyd. UJ, Kraków.

Niemierko B. (2009): *Diagnostyka edukacyjna. Podręcznik akademicki*. Wyd. Naukowe PWN, Warszawa.

Niemierko B. (2010): *Między prawdą a skutecznością – perspektywy oceniania szkolnego*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. PTDE, Kraków.

Simons H (1987): *Getting to Know School in a Democracy. The Politics and Process of Evaluation*. The Falmer Press, London.

Szyling G. (2013): *Diagnoza i ewaluacja a teoretyczne horyzonty refleksji nad działaniem*. „Teraźniejszość–Człowiek–Edukacja”, nr 4 (64).

Znanięcki F. (1998): *Education and Social Change*. ed. E. Hałas. Peret Lang, Frankfurt am Main.

Odrębność i współgranie. Polskie relacje diagnostyki edukacyjnej i ewaluacji

Historia polskich badań ewaluacyjnych w dziedzinie edukacji rozpoczęła się w połowie lat dziewięćdziesiątych jako bardzo niedojrzały rodzaj badań, ale podejmujący wielką rolę animatora zmian strukturalnych i demokratycznych transformacji. W przeciwieństwie do diagnostyki edukacyjnej polskie badania ewaluacyjne nie posiadały wówczas ani wystarczająco określonych celów, ani instrumentów z powodu wielorakiej otwartości na różnorodność możliwych zastosowań. Ewaluacja zorientowana jest w kierunku społecznej refleksyjności i traktuje proces badawczy jako dynamiczną interakcję swoich instytucjonalnych i indywidualnych „aktorów”. Z tego właśnie powodu ewaluacja nie może być traktowana jako rodzaj pomiaru, akredytacji czy audytu, ale jako wyzwanie dla konkretnej społeczności oraz programu jej rozwoju poprzez respektowanie własnego systemu wartości. Współczesne rozporządzenia ministerialne niestety nie wspierają tego typu ewaluacji i redukują dialogiczność i odpowiedzialność studiów ewaluacyjnych.

Słowa kluczowe: badania ewaluacyjne, refleksyjność społeczna, podmioty procesu ewaluacji, ewaluacja demokratyczna i dialogiczna, modele ewaluacji, ewaluacja międzykulturowa

Distinctiveness and synergy. Polish relations between educational diagnostics and evaluation

The history of Polish evaluation studies started in the field of education in mid nineties as very immature kind of research but with a great role in processes of structural and democratic transformations. In contrast to educational diagnostics, Polish evaluation studies did not have operational, concrete instruments and specific objectives because of great openness for potential applications. Nevertheless, there are a lot of fields for constructive encounters, common values and mutual learning. Evaluation is oriented to social reflectiveness, and treats the research process as a dynamic interaction of its institutional and individual actors. From this point of view evaluation should not be treated as a kind of assessment, accreditation or audit, but as a challenge for a specific community and program on the way of their development according to one's own systems of values. Contemporary administrative regulations do not support this kind of evaluation and reduce responsiveness and responsibility of the real educational and evaluation studies.

Key words: evaluation studies, social reflexivity, subjects of evaluation process, democratic and dialogic evaluation, evaluation models, intercultural evaluation

Maria Groenwald
Uniwersytet Gdański

DIAGNOZOWANIE – MIĘDZY ETYKĄ BLISKIEGO A DALEKIEGO ZASIĘGU

Zagadnienia etycznych aspektów diagnozowania oraz społecznych konsekwencji sformułowanych diagnoz są jednym z wielu tematów podejmowanych w konferencyjnych opracowaniach. Zwraca uwagę to, że ich autorzy, obierając różne perspektywy etyczne, przedstawiają zróżnicowane interpretacje rzeczywistości. Tym samym podnoszonym problemom nadają nie neutralny, a dyskursywny charakter (Klus-Stańska, 2009, s. 35; Rapley, 2013, s. 23 i 200). W niniejszym opracowaniu zamierzam prześledzić analizę tego dyskursu w poszukiwaniu odpowiedzi na pytanie: Które aspekty moralne i etyczne zainteresowały autorów tekstów oraz jak je prezentowali? Będzie mnie interesowało nie tylko to, co powiedziano i napisano, ale również, co zostało przemilczane (Peräkylä, 2009, s. 328–326). Ponadto planuję wiodące problemy usystematyzować (pogrupować) i zinterpretować w świetle koncepcji etycznej dobranej adekwatnie do poruszanych zagadnień.

Założenia metodologiczne

O źródłach danych

Źródłem informacji na temat interesującego mnie zagadnienia są te opracowania opublikowane w wydawnictwach konferencyjnych w latach 1994–2013, w których podejmowane są kwestie etycznych aspektów diagnozowania realizowanego w jednej z dwu perspektyw: „z bliska” bądź „z daleka” (Groenwald, 2007, s. 31–37; Niemierko, 2013, s. 26–28). Oznacza to, że wykorzystam dane zastane, obejmujące zarówno opisy zdarzeń „z pierwszej ręki”, jak i tzw. zapisy wtórne, czyli dane „nie wywołane”, ale bazujące na opisach przypadków zaistniałych w innym czasie i miejscu (Peräkylä, 2009, s. 325).

Po dokonaniu przeglądu tomów konferencyjnych, do dalszych analiz zakwalifikowałam osiemdziesiąt tekstów, w zaledwie niewielkiej części przywołanych w tym artykule. Dodam jedynie, że wybrane opracowania cechuje zróżnicowane ugruntowanie w realiach lokalnych i niejednakowe zainteresowanie kontekstem etycznym (Miles, Huberman, 2000, s. 10), prezentowanym i analizowanym albo ilościowo – z dominacją liczb, albo jakościowo – poprzez ukazywanie znaczeń moralnych nadawanych diagnozowaniu i jego konsekwencjom społecznym.

O analizie danych

Prowadzonej w tym opracowaniu analizie dyskursu (jak wszelkich tego typu analiz) nie sposób podporządkować „żelaznym zasadom” i prostym technikom, by przebiegła ona według pewnych i jasnych wskazówek, które wystarczy bezrefleksyjnie zastosować (Rapley, 2013, s. 28). Istotne natomiast okazują się w niej:

- zachowanie otwartości w analizowaniu tekstów, przyjmującej formę swoistego z nimi dialogu (Urbaniak-Zajac, 2013, s. 256);
- ujawnianie kwestii etycznych przemilczanych lub opuszczanych w niektórych tekstach;
- dostrzeganie związku między rolami społecznymi pełnionymi przez autorów prac a przedstawianymi przez nich analizami i interpretacjami zdarzeń (Rapley, 2013, s. 191–202).

Z powodu bogactwa i obszerności opublikowanych materiałów poświęconych etycznemu aspektowi diagnozowania, w analizie konferencyjnych opracowań konieczne było zastosowanie redukcji i kondensacji (Miles, Huberman, 2000, s. 11). W ramach redukcji dokonałam selekcji publikowanych tekstów, przeprowadzonej z zachowaniem chronologii opisywanych zdarzeń. Natomiast w zakresie kondensacji pomocne okazało się wykreślenie mapy będącej formą reprezentacji częściowo uporządkowanej. W myśl wymogów metodologii jakościowej, tej formy graficznej użyłam w celu ukazania wzajemnych relacji pomiędzy wybranymi aspektami analizowanych zagadnień lub grup (tamże, s. 107). W rozpatrywanym przypadku „współrzędne” mapy wyznaczyły: (1) czas obejmujący dwudziestolecie diagnostyki edukacyjnej, w którym opublikowano analizowane przeze mnie teksty na tematy etyczne; (2) zasięg diagnozowania – z bliska lub z daleka. Treścią wypełniającą tę mapę były etyczne aspekty diagnozowania, tworzące w kilku przypadkach obszary zagęszczeń, czyli tzw. „sytuacje graniczne”, na których skoncentruję się w dalszej analizie.

O teoretycznej perspektywie interpretacji danych

Już na etapie selekcji materiałów można było dostrzec wyłaniającą się z tekstów polaryzację treści etycznych, które najogólniej ujawniały się jako:

- dylematy moralne doświadczane jednostkowo przez poszczególnych nauczycieli oraz egzaminatorów, które dotyczyły refleksji nad sobą samym, pracą z uczniami, zdarzeniami ze szkolnej codzienności;
- problemy etyczne ujawniające się w relacjach między uczniami i nauczycielami a władzami oświatowymi, politycznymi, czy nawet grupami słabo powiązаныmi z uczestnikami edukacji.

Równocześnie dało się zauważyć pewną prawidłowość, polegającą na tym, że w artykułach opisujących diagnozowanie „z bliska” zazwyczaj przeważały kwestie natury moralnej, natomiast w diagnozowaniu „z daleka” dominowały problemy kodeksów etycznych i wykładni prawa stanowionego. Tak zarysowująca się polaryzacja danych była dobrą podstawą przyjęcia koncepcji etyki bliskiego i dalekiego zasięgu (Abriszewski, 2005, 2007, 2010; Jonas, 1996), jako podstawy interpretacji adekwatnej do rozpatrywanego zagadnienia¹.

Etyka bliskiego zasięgu (Abriszewski, 2007, s. 82) obejmuje obszar indywidualnych doświadczeń moralnych jednostek pozostających w bezpośrednich relacjach z innymi osobami z tej samej gromady czy społeczności. W jego centrum znajduje się człowiek doświadczający świata, codzienności i samego siebie, oraz rozwijający własną wrażliwość moralną. Takie ujęcie stosunków międzyludzkich świadczy o tym, że mamy do czynienia z etyką antropocentryczną, z właściwą jej czasową oraz przestrzenną bliskością relacji i zdarzeń, małym zasięgiem działań oraz stosunkowo niewielką za nie odpowiedzialnością, pozostawianą przypadkowi, losowi bądź opatrności (Jonas, 1996, s. 27). Współcześnie określa się tę etykę jako tradycyjną i odchodzącą w przeszłość.

W szkolnej codzienności etyka bliskiego zasięgu znajduje egzemplifikację w indywidualizacji nauczania–uczenia się możliwej dzięki diagnozowaniu bezpośrednio, realizowanemu „z bliska”. Ujawnia się ona w dokonywanych przez nauczyciela osądach własnych działań lub relacji z uczniami danej klasy. To etyka więzi nacechowanych bliskością, ale i też stosunkowo krótkim czasem ich trwania, gdyż obejmującym tylko okres współbycia ze sobą wychowanków i wychowawców we wspólnej im teraźniejszości, formalnie wyznaczonej przez cele oraz przedmiot kształcenia. Równocześnie właśnie z racji owej bliskości i krótkotrwałości, moralna ocena relacji nauczyciela i uczniów zazwyczaj jest wyrazista i daje się określać w kategoriach dobra lub zła, aczkolwiek przypisywanych określonym sytuacjom z dużą dozą niejednomyślności wartościujących je osób.

Etyka dalekiego zasięgu, której istotą są narracje o moralności, skupia się na odległych czasowo i przestrzennie skutkach działań (Abriszewski, 2005, s. 411). Zwraca się w niej uwagę na to, że tradycyjne małe wspólnoty współcześnie ulegają przekształcaniu, a nawet niszczeniu na masową skalę przez procesy modernizacyjne będące splotem współczesnej technonauki, rynku i polityki (Abriszewski, 2010 s. 175). Pod wpływem tych procesów bezpośrednie relacje łączące ludzi zmieniają się lub zanikają, gdyż jednostki rozplývają się w szerszych i często anonimowych układach zależności. W efekcie wspólnoty i wspólnotowość zostają wyparte przez sferę stowarzyszeniową i rozrastający się świat instytucji pozostających poza horyzontem rzeczywistości życia codziennego (Abriszewski 2007, s. 83). W związku z tym, że działania realizowane przez jednostki zostają wpisane w szersze (nawet „globalne”) układy zależności, ich rezultaty zazwyczaj pojawiają się w miejscu i czasie odległym od tego, w którym sytuuje się ich sprawca. Jeszcze bardziej odległe okazują się długofalowe konsekwencje społeczne tych działań, a moralna refleksja nad nimi jest podejmowana niechętnie bądź wcale. Na gruncie edukacji tego typu relacje widoczne są, na przykład, w kształceniu *online* lub w diagnozowaniu z daleka, w którym anonimowość jest postrzegana jako jego cenna własność.

Wrażenie rozrastającej się w sferze etyki dalekiego zasięgu **etycznej pustki** zaburzą tzw. „sytuacje graniczne”, czyli sytuacje niezwykle trudne (bądź nieznośne) dla człowieka, a radzenie sobie z nimi okazuje się kluczowe dla jego egzystencji i tożsamości. Jako ich przykład wskazywane jest doświadczanie: porażki, krzywdy bądź winy, skłaniające do znajdowania możliwych sposobów reagowania na nie, zwłaszcza poszukiwanie wsparcia (Abriszewski, 2010, s. 168–169). Etyczny wymiar tych sytuacji jest więc wyraźny. Z sytuacjami granicznymi mamy do czynienia także w edukacji i jej diagnozowaniu, zwłaszcza realizowanym „z daleka”, co zostanie pokazane w dalszej części opracowania.

Etyczne aspekty diagnozowania w świetle wystąpień konferencyjnych

Intuicje moralne w diagnozowaniu z bliska. Lata 1994–1999

Z analizy tekstów konferencyjnych opublikowanych w latach 1994–1999 wynika, że wspólną płaszczyzną tematyczną, w obrębie której spotykają się diagnozowanie i etyka, jest problematyka wartości, rozpatrywanych w kontekście diagnozowania z bliska oraz etyki bliskiego zasięgu. W tych pracach jako szczególnie cenna i ważna dla diagnostyki jest wskazywana odpowiedzialność badacza-diagnosty (i ewaluatora) za diagnozowanie (i ewaluację), ponoszona wobec uczniów, nauczycieli i innych osób zainteresowanych diagnozą (i ewaluacją), na podstawie takich kryteriów normatywnych, jak: sumienie, prawo etyczne lub prawo stanowione (Kozłowski, 1999, s. 194). Pojawiają się też ostrzeżenia, by nie traktować pomiaru jako nowej religii pedagogicznej, w której „bóstwo objawia się pod postacią narzędzia pomiaru, liturgia ma postać spójnego zestawu reguł postępowania z upostaciowionym bóstwem, wierni mają tak świat porządkować, by bóstwo mogło zapanować nad nim” (Kruszewski, 1998, s. 77). Przestroga to o tyle poważna, że ów pomiar jest mocno uwikłany w ideologię oświatową (Konarzewski, 1998, s. 58), która – równie znacząco jak religia – kształtuje świadomość społeczną, oddziałuje na wychowanie i oświatę, zarazem nie ujawniając tego, że w istocie odzwierciedla dążenia dominującej klasy, grupy bądź partii politycznej (Gutek, 2003, s. 256–158). Niewykluczone zatem, że jeśli pomiar będzie realizowany w upolitycznionej oświacie, stanie się dostarczycielem wyników, wykorzystywanych jako „narzędzia walki, sposób wywierania wpływu, a nawet obiekt manipulacji” (Kozłowski, 1999, s. 189).

Kwestie etyczne ujawniane w pracach „młodej” jeszcze diagnostyki edukacyjnej są raczej intuicyjnym przewidywaniem obecności aspektów moralnych przyszłych diagnoz i rozterek doświadczanych przez diagnostów kierujących się nie tylko wymogami poprawności metodologicznej, lecz również osobowym systemem wartości moralnych i zgromadzonym indywidualnie doświadczeniem (Aleksander, 1994, s. 39; Lewowicki, 1998, s. 47). Ponadto diagnozy te (co się okazało po latach) trafnie zapowiadają trudności w odizolowaniu diagnozowania z bliska (powiązanego z etyką bliskiego zasięgu) od kontekstu społecznego, polityczno-ekonomicznego i kulturalnego. Ów wpływ miał przybrać na sile i znaczeniu wraz z upowszechnieniem przez egzaminy diagnozowania z daleka, z wpisanymi wewnątrz procedurami i formalnoprawnymi zasadami. Stąd obecność w diagnostycznym dyskursie głosów wyrażających zaniepokojenie negatywnymi konsekwencjami egzaminowania (Hall, 1999), szczególnie zamykaniem szans uczniom słabszym (Wieczorkowska, Radkiewicz, 2000, s. 145).

W kolejnych latach zagadnienia dotyczące dylematów moralnych doświadczanych przez nauczycieli na płaszczyźnie diagnozowania szkolnego pojawiają się coraz rzadziej, a opracowania konferencyjne zostają zdominowane przez problematykę egzaminacyjną i etykę dalekiego zasięgu.

Etyczne sytuacje graniczne w diagnozowaniu prowadzonym w latach 2000–2013

Wprowadzenie egzaminów i realna perspektywa rozwoju systemu oceniania zewnętrznego ożywiły dyskurs na temat etycznych aspektów diagnozowania. Zainicjowany w roku 2000, w kolejnych latach (do chwili obecnej) jest toczony podczas

konferencyjnych spotkań z różnym natężeniem i zaangażowaniem uczestniczących w nich osób, co sprawia, że tę problematykę obieram jako wiodącą w dalszej części opracowania. W zakresie zagadnień dotyczących wpływu egzaminów na diagnostykę edukacyjną następuje przyspieszenie rozwoju diagnozowania – zwłaszcza „ilościowego”, realizowanego w instytucjach profesjonalnie zajmujących się pomiarem. Podkreśla się, że dzięki egzaminom zostaje pokazany uczniom i nauczycielom wyraźny cel kształcenia – osiągnięcie wysokiego wyniku i korzyści z tego płynących, jak również: podnoszenie jakości uczenia się, gromadzenie informacji o uczniach i szkołach, dostrzeganie uczniów o specjalnych potrzebach edukacyjnych, rozwój zawodowy nauczycieli itd. (Niemierko, 2003, s. 28).

Wraz z wprowadzeniem egzaminów uczestnicy konferencji sygnalizują także pojawianie się problemów będących konsekwencją długofalowego oddziaływania egzaminacyjnych wyników na losy uwikłanych w nie osób, szkół, ośrodków egzaminacyjnych i innych instytucji oświatowych (Szaleniec, 2001, s. 275). Na płaszczyźnie etyki niektóre z nich okazują się problemami złożonymi i tak trudnymi, że można je określić mianem sytuacji granicznych. W analizie konferencyjnego dyskursu ograniczę się do przywołania tych sytuacji z nich, które są dyskutowane i opisywane najczęściej, a dotyczą kwestii: **nieufności, oszustwa, niekorzystnych konsekwencji społecznych** diagnozowania z daleka oraz jego **uwikłania w politykę**.

Nieufność wobec egzaminatorów i próby jej rozwiewania

Jednym z częściej podnoszonych problemów natury moralnej, na który diagności zaczynają zwracać baczniejszą uwagę dopiero w powiązaniu z egzaminami, jest **zaufanie** do wartości wspólnie tworzonych i praktykowanych na co dzień (Szyling, 2005, s. 377). Przejawianie go wobec egzaminatorów wymaga, jak pisano, spełniania przez nich wymagań etycznych wyższych niż przez innych ludzi, zwłaszcza wykonywania „dobrej roboty” i przestrzegania etosu zawodowego (Bednarek, Lempa, 2006, s. 176). W tym kontekście jako istotny pojawia się postulat odwzajemniania zaufania egzaminatorów do siebie, niemożliwego, gdy w relacjach między nimi wolę współpracy zastępuje wola rywalizacji. Ponadto zwraca się uwagę, że budowanie zaufania utrudnia centralizacja zarządzania pewnymi wycinkami pracy merytorycznej i organizacyjnej oraz najczęściej informacyjny sposób kontaktowania się okręgowych komisji egzaminacyjnych z komisją centralną (Szmigel, 2001, s. 369), postrzegany jako pochodna silnej władzy CKE nad komisjami okręgowymi (Niemierko, 2000, s. 17; Niemierko 2009, s. 25). Równolegle w konferencyjnym dyskursie są przedstawiane przez pracowników okręgowych komisji egzaminacyjnych działania, które podejmują oni na rzecz podnoszenia zaufania do egzaminów i do egzaminatorów, jak prowadzenie szkoleń oraz selekcji kandydatów na egzaminatorów, mające na celu podniesienie rzetelności oceniania arkuszy egzaminacyjnych. Ponadto, jako nie mniej ważne dla ugruntowania zaufania, wskazują udostępnianie uczniom ich prac do wglądu (Mazur, Tyralska-Wojtyca, 2008, s. 610).

Mimo tych przedsięwzięć, wciąż powraca problem niedostatku zaufania do egzaminów, dlatego prawidłowości ich przebiegu przyglądają się egzaminacyjni obserwatorzy. Mają oni, między innymi, „uwiarygodnić (...) nauczycielską dojrzałość emocjonalną i posiadane wartości etyczne” (Szmigel, Szaleniec, 2001, s. 227). Tymczasem wydaje się, że każdą zewnętrzną kontrolę można odbierać jako przejaw

nieufności, a obecnego na egzaminie obserwatora – jako przedstawiciela uogólnionej społecznej nieufności wobec uczniów i nauczycieli. Jak zatem ma się ugruntować oczekiwane zaufanie do egzaminów wobec jawnie prezentowanej nieufności wobec egzaminowanych (Bednarek, Lempa, 2005, s. 204; Groenwald, 2008)? Nieufności skądinąd częściowo uzasadnionej problemem egzaminacyjnego oszustwa.

Egzaminacyjne oszustwo

W konferencyjnych publikacjach problemy oszustwa egzaminacyjnego przez kilka lat dominowały wśród etycznych kwestii diagnozowania. Pierwsze sprawdziany po szóstej klasie i egzaminy gimnazjalne, a potem matura uświadamiają wielorakość oszustw (Furgoł, Szafraniec, 2002, s. 156; Niemierko, 2006, s. 39–54). O ich skali natomiast niech świadczą przytoczone przez Henryka Szaleńca dane, z których wynika, że na obszarze działania OKE w Krakowie podczas egzaminu gimnazjalnego w 2002 roku 44% uczniów miało możliwość odpisania odpowiedzi od kolegi, 60% miało przy sobie ściągawki, 24% twierdziło, że nauczyciele pozwalali uczniom porozumiewać się między sobą, 11% – że widziało jak nauczyciele udzielają uczniom pomocy (Szaleniec, 2006, s. 196). Oszustwo okazuje się bardzo trudnym doświadczeniem dla młodego jeszcze systemu egzaminów zewnętrznych i równie problematycznym dla samej diagnostyki, gdyż wpływa ono na obniżenie własności psychometrycznych testu i pomiaru, zwłaszcza na jego bezstronność oraz rzetelność. W związku z tym w kolejnych latach podejmowane są działania mające na celu wyeliminowanie go za pomocą: precyzyjnych regulacji proceduralnych (Czarnota-Mączyńska, 2006), prowadzących nawet do unieważnienia egzaminów (Tyralska-Wojtyca, 2006). Pojawia się także próba opracowania etycznego kodeksu egzaminatora, na wzór kodeksów obowiązujących w Wielkiej Brytanii i Irlandii (Bartmańska, 2003, s. 331–332).

Byłoby jednak nieuprawnionym uproszczeniem wskazywanie wyłącznie egzaminów jako „sprawcy” nakłaniającego do oszustw, gdyż – co podnoszono na konferencjach – jest ono elementem kultury szkolnej codzienności, w której „zdecydowana większość uczniów przyznaje się do zachowań nieuczciwych podczas sprawdzianów, oszukuje podczas odrabiania prac domowych, przedstawia opracowania pozyskane z Internetu jako własne, wykorzystuje «bryki» w przygotowaniu do lekcji” (Lipska, 2006, s. 242). W tej codzienności oszustwo uzyskuje status uprawomocnionego, uczciwość zaś jest postrzegana jako szkodliwa, o czym świadczy wypowiedź nauczyciela stwierdzającego, że: „około 200 uczniów naszego Gimnazjum czuje się przez nas pokrzywdzonych, ponieważ nie pozwoliliśmy (podczas egzaminu) odpisywać. W innych gimnazjach nauczyciele «pomagali» w różnorodny sposób” (Boba i in., 2002, s. 161).

W diagnostycznym dyskursie oszustwo staje się płaszczyzną, na której etyka bliskiego i dalekiego zasięgu spotykają się we wspólnym przeciwdziałaniu mu. W ramach tegoż pojawiają się propozycje: przeprowadzenia kampanii na rzecz uczciwości edukacyjnej, której celem byłoby wychowanie ku uczciwości (Lisiecka, 2006), co nie wyklucza podejścia bardziej pragmatycznego, polegającego na proceduralnym wprowadzeniu precyzyjnych zasad przeprowadzania egzaminów. Z czasem to działanie okaże się najskuteczniejsze w walce z egzaminacyjnym oszustwem. Trzeba jednak podkreślić, że wraz ze skupieniem się diagnostów na wyeliminowaniu go, zostaje z pola widzenia usunięte zainteresowanie innymi etycznymi aspektami diagnozowania, ujawniającymi się w międzyludzkich relacjach.

Konsekwencje społeczne egzaminowania

W ogólnie ujętej kwestii społecznych skutków egzaminowania zawiera się kilka wywołanych nimi problemów, ale przytoczę jedynie niektóre spośród nich, poruszone podczas konferencji diagnostycznych. I tak, Tomasz Szkudlarek (2005, s. 100) zwraca uwagę na to, że poprzez egzaminacyjną selekcję ulega pogłębieniu zróżnicowanie społeczne, utrwalane na co dzień przez szkołę. Zbigniew Kwieciński, analizując zmiany w edukacji na przestrzeni ostatniego dwudziestolecia, dostrzega szereg działań pozorowanych. Wśród nich, na przykład, wskazuje: (a) wzmacnianie mechanizmów społecznej segregacji realizowane pod hasłem wyrównywania szans edukacyjnych; (b) zamykanie kształcenia w powszechnej, liberalnej szkole egzaminami, które okazują się przesiewowe; (c) upowszechnienie matury, która ulega dewaluacji w związku z wprowadzeniem 30% progu zaliczeń; (d) świadome niedostrzeganie tego, że skok skolaryzacji akademickiej dokonuje się kosztem jakości i poziomu kształcenia wyższego itd. (Kwieciński, 2010, s. 26). Do założeń szkolnictwa powszechnego i jednolitego nawiązują także badania Romana Dolaty (2009, s. 112–122), zwracającego uwagę na proces różnicowania się gimnazjów w dużych miastach. Zastrzega przy tym, że choć niezwykle złożoność mechanizmu tego procesu nakłada konieczność podejmowania kolejnych badań diagnostycznych, z już uzyskanych danych można wnioskować, iż za różnicowanie tych szkół odpowiadają między innymi: nieformalne i ukryte procesy autoselekcyjne i selekcyjne, uruchamiane jako segregacja uczniów już na progu gimnazjum i znajdujące potem odzwierciedlenie w efektywności nauczania.

Niemożliwość przerwania przez uczniów kręgu dziedziczenia biedy, sprzyja rozbudowaniu nadziei na sprawiedliwe, bo – jak uzasadniają autorzy referatów i dyskusji konferencyjnych – zobiektywizowane ocenianie zewnętrzne (np. Janicka-Panek, 2007, s. 429). W przeciwieństwie do uczniów, nauczyciele do egzaminacyjnej sprawiedliwości podchodzą z większym dystansem. Dostrzegają bowiem, że nie odzwierciedlają one nakładów ich pracy, a mimo to są traktowane jako argument w przydzielaniu nagród dyrektorom szkół przodujących w egzaminacyjnej punktacji (Suckiel, 2008, s. 567). Wydaje się, że owa rozbieżność opinii uczniów i nauczycieli ma swoje źródło w stosowaniu przez nich różnych kryteriów sprawiedliwości, nie zawsze uświadamianych i przejrzystych (Groenwald, 2010). W efekcie argumenty uznawane przez jednych za sprawiedliwe, dla innych są całkowitym zaprzeczeniem tej wartości.

O tym, jak ważne dla uczniów i ich rodziców są wyniki egzaminowania, świadczy zainteresowanie korepetycjami. Autorzy konferencyjnych opracowań przyznają, że wcześniej korzystanie z nich tłumaczono autentycznymi brakami w uczniowskiej wiedzy przewidzianej programem kształcenia, co skądinąd świadczyło o niskiej jakości pracy nauczycieli, nieumiejących pomóc swoim wychowankom (Maciąg, 2005, s. 271–272). Obecnie natomiast korzystanie z korepetycji jest uzasadnione przede wszystkim potrzebą uzyskania wyższego wyniku na egzaminie (Szmigiel, 2007, s. 164–165). Równocześnie skala pozaszkolnego uczenia się jest trudna do oszacowania, gdyż uczniowie niechętnie o tym rozmawiają nawet między sobą, przywoływane zaś na konferencji dane liczbowe różnią się wartościami zależnie od przyjętych założeń metodologicznych prowadzonych badań (Putkiewicz, 2005; Hawrot, Jasińska, 2013).

Uwikłanie edukacji i diagnozowania w politykę

Już w latach budowania systemu oceniania zewnętrznego pojawiały się obawy o nadmierny wpływ jego centralizacji na swobodę mierzenia efektów kształcenia. Przypuszczano, że „okręgowe komisje egzaminacyjne, działające w ministerialnym, silnie scentralizowanym systemie nie (będą) miały takiej swobody mierzenia i komentowania wyników pracy szkoły, jak niepaństwowe ośrodki diagnozy typu wałbrzyskiego” (Niemierko, 2000, s. 15). Sformułowane wówczas obawy znajdują potwierdzenie w praktyce i po latach Bolesław Niemierko uznaje, że „Centralna Komisja Egzaminacyjna w Warszawie uzyskała silną władzę nad komisjami okręgowymi, a postulowany przez (J. Marquand) «margines swobody na eksperymenty» okazał się bardzo wąski i miał rozwijać się, system egzaminacyjny uległ petryfikacji” (Niemierko, 2009, s. 25). Na polityczne i kulturowe uwikłanie diagnozowania (wraz z edukacją) zwraca także uwagę Z. Kwieciński (2010, s. 26), pisząc, że hasła, pod którymi wprowadzane są wielkie reformy, zawierają też „mętne wizje” i obietnice bez pokrycia, wspieranie przez naukowców, polityków, dziennikarzy. Nierówności społeczne generowane lub utrwalane przez działania mające na celu podnoszenie skuteczności pracy szkoły – zdaniem Doroty Klus-Stańskiej (2010, s. 31) – przestają być ujmowane w kategoriach merytokracji, a nabierają tła politycznego.

Konferencyjny dyskurs diagnostyczny dostarcza przykładów pokazujących, że decyzje polityczne okazują się ważkim inicjatorem nie tylko diagnozowania z daleka i opracowanych dzięki nim diagnoz, ale również czynnikiem oddziałującym na etyczny wymiar tego diagnozowania. Choć zazwyczaj jego konsekwencje społeczne i aksjologiczne są długofalowe, rzadko stanowią przedmiot refleksji instytucji oświatowych (i innych), z którymi pozostają w relacji współzależności (i politycznego uwikłania). Czy słusznie? Zdaniem Jacka Jaśtala, „wprowadzenie etyki do polityki nie daje żadnej gwarancji poprawy jakości życia politycznego, (...) [a próba] podniesienia indywidualnych standardów moralnych [sprzyja] jeszcze ostrzejszemu zarysowaniu konfliktu pomiędzy tym, co etyczne, a tym, co polityczne. (...) Konflikt ten nie wydaje się możliwy do wyeliminowania (...) i z jego istnieniem musimy się pogodzić” (Jaśtala, 2005, s. 108). Równocześnie Simona Critchley (2006, s. 33) stwierdza, że „jeśli etyka bez polityki jest pusta, to polityka bez etyki jest ślepa”.

Czy oznacza to, że przestrzeń pomiędzy etykami bliskiego i dalekiego zasięgu jest pusta etycznie? I czy właśnie dlatego prowadzone w niej diagnozowanie, opisywane i prezentowane na konferencjach, zostaje wyswobodzone z wszelkich norm moralnych i etycznych?

Między obecnością etyki a jej brakiem w diagnozowaniu

Jeśli za Haliną Promieńską (2005, s. 47) przyjąć, że „dominującą funkcją etyki jest wprowadzanie ładu aksjologicznego w wewnętrzny (osobowy) i zewnętrzny (kulturowy, społeczny i polityczny) system motywacji i zachowań ludzkich”, to w przypadku diagnostyki wyraźnie daje się wyodrębnić dwa rozłączne obszary realizowania tego ładu. W diagnostycznym dyskursie konferencyjnym jeden z nich stanowi sfera diagnozowania „z bliska”, zazwyczaj prowadzonego w bezpośredniości (Wysocka, 2013, s. 44). W jego obrębie pojawiają się kwestie wzajemnego szacunku nauczycieli

uczniów, zaufania, odpowiedzialności diagnozujących wychowawców. Pokazują one relacje człowieka do siebie samego i do drugiego, zdolność jednostki do pogłębionej autorefleksji i rozwijania wrażliwości moralnej (Promieńska, 2005, s. 47). Są to kwestie właściwe **etyce bliskiego zasięgu** i dlatego mogą być ujmowane przez pryzmat formalizmu etycznego (Kanta lub etyki kantowskiej) albo utylitaryzmu bądź koncepcji rozwijanych na gruncie etyki cnót (Brandt, 1996; Szutta, 2007; Domańska, 2012).

Jako druga wyróżniona została sfera diagnozowania z daleka, przybierająca kształt zobiektywizowanych badań ilościowych i prowadzona przez instytucje (edukacyjne, egzaminacyjne i badawcze) i stowarzyszenia, oraz zlecane przez władze polityczne. Sprawne funkcjonowanie tych instytucji osiąga się dzięki tworzonej przez nie strukturze biurokratycznej i relacjom ustalonym proceduralnie. Obowiązująca etyka jest **etyką dalekiego zasięgu**, która „nie odwołuje się do uniwersalnych prawd moralnych, ale do określonych praktyką zasad współdziałania i współpracy” (Podrez, 2005, s. 78). Przestrzeń objętych przez nią relacji, porównana ze sferą etyki bliskiego zasięgu, sprawia wrażenie etycznie pustej oraz sankcjonującej własną autonomizację i oddzielenie od etyki.

Tymczasem analiza diagnostycznego dyskursu konferencyjnego wydaje się przemawiać za przyjęciem nieco innego określenia etyczności diagnozowania z daleka i nazywanie go **etycznie ubogim**, gdyż etycznych aspektów niezupełnie pozbawionym. Świadczyć o tym mogą prace, w których mówi się nie tyle o zaufaniu, co raczej o jego niedostatku, albo opisuje oszustwo pojawiające się obok zachowań uczciwych, bądź ukazuje społeczne skutki diagnozowania – korzystne i szkodliwe dla diagnozowanych.

Jakie są zatem etyczne aspekty diagnozowania usytuowanego pomiędzy wyodrębnionymi powyżej sferami, spośród których jedna jest etyką bliskiego zasięgu powiązaną z diagnozowaniem „z bliska”, a druga – diagnozowaniem „z daleka” i właściwą mu etyką dalekiego zasięgu, obejmującą działania na rzecz społeczeństwa, dobra wspólnego, w tym – zapewnienia dzieciom i młodzieży skutecznej edukacji (Promieńska, 2005, s. 47–48)?

W świetle przeprowadzonych analiz, etyczne aspekty diagnozowania prezentowane podczas corocznych spotkań diagnostów, okazują się złożone ontologicznie i poznawczo oraz niejednoznaczne etycznie. Jednak zinterpretowane przez pryzmat etyk bliskiego i dalekiego zasięgu, odkrywają sieć wzajemnych zapośredniczeń etycznych i diagnostycznych, ukrytych przed bezpośrednim poznaniem. Poddanie ich dalszej interpretacji może pokazać, na czym owe zapośredniczenia polegają i jakie są ich konsekwencje, dlatego warto podjąć się jej w kolejnej części pracy.

Etyka w przestrzeni diagnostycznych zapośredniczeń

Choć według Krzysztofa Abriszewskiego przestrzeń etyczna usytuowana „pomiędzy” etyką bliskiego i dalekiego zasięgu sprawia wrażenie pustej, to dla Joanny Byrskiej (2005, s. 232) jest to pustka szczególna, gdyż problemem jest w niej nie tyle brak pewnych wartości, co raczej nadmiar wszystkiego: możliwości, celów, sposobów i stylów życia, wielości idei, ideałów i religii. To także sfera przemieszanych ze sobą norm: moralnych, etycznych i stanowionych przez prawo. Okazuje się jednak, że owa etyczna wielorakość i złożoność odpowiada diagnozom ustalonym w przestrzeni „pomiędzy”, czyli w sferze znajdującej się między opozycyjnymi paradygmatycznie

diagnozowaniami z bliska i z daleka. Przykładów diagnozowań „pośrednich” dostarczają te spośród prezentowanych podczas konferencji opracowań, w których przytaczane są opisy nauczycieli przenoszących do własnego warsztatu pracy metody oraz normy stosowane w diagnostyce z daleka. Taki scenariusz ma miejsce wówczas, gdy za pomocą wybranych zadań z arkuszy egzaminacyjnych mierzą oni dokonany przez uczniów postęp w zakresie przygotowania się do egzaminu właściwego. Mniemają bowiem, że zadania z arkuszy mierzą sprawiedliwiej, a skonstruowane przez specjalistów – wzbudzają większe zaufanie co do ich psychometrycznych własności niż testy budowane przez nauczyciela (Niemierko, 2002, s. 267–268). Ponadto zapewniają nauczycielowi poczucie bezpieczeństwa w tym sensie, że skoro nie jest autorem zadań, nie może ponosić odpowiedzialności ani za ich jakość, ani za jakość pomiaru, ani tym bardziej za stopnie szkolne, które wystawi uczniom w wyniku mierzenia.

W świetle powyższych ustaleń, w konceptualizacji przestrzeni „pomiędzy” wiodąca rola przypada **zapośredniczeniom** – wieloaspektowym oraz wielostronnym, nie tylko diagnostycznym, ale i etycznym. Przebiegają one w dwu kierunkach. Jeden z nich odnosi się do (przywołanych powyżej) implementacji diagnozowań z daleka do nauczycielskiego badania w działaniu. Drugi dotyczy przenoszenia (bez przekształceń) znaczeń moralnych przypisywanych głównie diagnozowaniu z bliska (a więc ze sfery etyki bliskiego zasięgu) na diagnozowanie z daleka. Przejawia się to w oczekiwaniu, że do sfery etyki dalekiego zasięgu zostaną włączone wartości wywiedzione ze sfery etyki bliskiego zasięgu. W ramach tegoż, na przykład, pomiar ilościowy przedstawia się uczniom i nauczycielom nie tylko jako niezbędny dla funkcjonowania systemu oświaty, ale również jako przyjazny jednostkom i grupom. W rezultacie takiego wielokierunkowego zapośredniczania powstaje sieć powiązań łączących etykę i diagnostykę bliskiego oraz dalekiego zasięgu w sposób przewidywalny wówczas, gdy wiązania te odnoszą się do sytuacji ustabilizowanych (Latour, 2010, s. 56, 83).

Do zapośredniczania nie dochodzi samoistnie, lecz jest ono inicjowane przez tzw. centra decyzyjne, określane również mianem „centrów kalkulacyjnych” (Latour, 2010, s. 257, 260; Abriszewski, 2012, s. 21–22). Są to miejsca podejmowania decyzji i sprawowania kontroli nad ich realizacją. W tym celu centrum wykorzystuje nie tylko urządzenia techniczne, lecz także „wszystkie negocjacje, intrygi, kalkulacje, akty perswazji i przemocy, dzięki którym aktor czy też siła przyznają sobie lub powodują, że zostaje im nadana władza, by mówić lub działać w imieniu innego aktora lub siły” (Abriszewski, 2010, s. 24). Etyczny wymiar tych działań jest więc wyraźny. Autorzy przywołanych w tej pracy wypowiedzi konferencyjnych lokują władzę podejmowania decyzji w dwu centrach:

- w instytucjach zajmujących się profesjonalnym diagnozowaniem edukacyjnym,
- w obrębie nauczycielskich badań własnych, w których jednak wykorzystywane są elementy diagnozowania z daleka.

Pokrótkie przedstawię działania diagnostyczne podejmowane przez te centra, koncentrując się na ich etycznych aspektach.

Za punkt wyjścia obiorę **diagnozowanie inicjowane przez nauczycieli**. Mocą nadanej im władzy używanej wyłącznie w obrębie „izby lekcyjnej”, zarządzają oni m.in. przebiegiem lekcji oraz doбором treści nauczania, do którego włączają przygotowania do egzaminu. Nawet powszednie sprawdziany wzorują na egzaminacyjnych testach. Wytlumaczeniem ich postępowania może być, przywołana przez Henryka

Szaleńca, zasada George'a F. Madausa (*The influence of testing on the curriculum*, Chicago 1988), który stwierdza, iż jeśli „uczniowie, nauczyciele i administracja oświatowa wierzą i uważają, że wyniki testowania są istotne, to nieważne, czy to rzeczywiście prawda, czy fałsz. Wpływ [egzaminów na proces kształcenia] powodowany jest przez to, jak ludzie postrzegają doniosłość egzaminu” (Szaleniec, 2011, s. 42). Dlatego też nauczanie-uczenie się w szkole polega przede wszystkim na przygotowaniu do niego.

Z myślą o dobru ucznia nauczyciele również wzorują ocenianie „powszednie” na egzaminacyjnych kryteriach oceniania. Wykorzystując arkusze i normy egzaminacyjne oczekują, że – z jednej strony – zastąpią one diagnozowanie nauczycielskie i zarazem – z drugiej – przyjmą wszystkie jego własności. Potwierdzają to wypowiedzi badanych uczniów i nauczycieli, którzy w konferencyjnym dyskursie diagnostycznym apelują o uwzględnianie w diagnozowaniu z daleka takich wartości moralnych właściwych etyce bliskiego zasięgu, jak: poszanowanie godności ucznia i nauczyciela, kierowanie się sprawiedliwością w ocenianiu i zachowanie odpowiedzialności za decyzje i działania. Opowiadając się za nimi prezentują swoje przywiązanie do uniwersalnych wartości moralnych, strzeżonych przez tradycyjne systemy normatywne (obyczaje i zwyczaje, moralność), których nie znajdują w diagnozach z daleka. I znaleźć nie mogą, gdyż (o czym już wspomniałam) diagnozowanie z daleka współwystępuje z etyką dalekiego zasięgu, która nie kieruje się normami moralnymi.

W konferencyjnym dyskursie diagnostycznym podnoszone są także kwestie selekcyjnego charakteru egzaminów, odbieranego jako przejaw niesprawiedliwości. Tymczasem trudno w diagnozowaniu z daleka o taką sprawiedliwość, jakiej oczekują ich uczestnicy, a więc uwzględniającą w ocenianiu wkład pracy i możliwości każdego. Wydają się oni nie pamiętać, pisze Roland Meighan (1993, s. 385), że „system egzaminów oparty na normach ma w zamierzeniu dokonywać selekcji uczniów odnoszących sukcesy, a innych klasyfikować jako porażki. To uprawomocnia nierówny system rozdziału szans życiowych”. Poczucie krzywdy i niesprawiedliwości wydadają się więc nieuchronne.

Ów niedostatek „diagnostycznej etyczności” usiłują wypełnić **kodeksy etyczne**. Z konferencyjnych wystąpień wynika, że najpilniejsze byłoby opracowanie kodeksu etyki egzaminatora. Apelują o niego zwłaszcza osoby postrzegające diagnozowanie z daleka jako realizowane z bliska i doświadczające w związku z tym poczucia erozji wartości czy sytuacji „zapaści aksjologicznej” i „pustki etycznej” w diagnozowaniu z daleka. Postulują więc, by kodeks etyczny, poprzez poczynione w nim zapisy, przypominał o zagubionych normach moralnych odniesionych do konkretnych sytuacji egzaminacyjnych. Jednak wówczas stałby się on rodzajem instrumentu bądź narzędzia pragmatycznego (Podrez, 2005, s. 78). Trudności w opracowaniu takiego kodeksu spowodowały, że uczestnicy konferencji diagnostycznych po kilku nieudanych próbach, ostatecznie z inicjatywy się wycofali. Natomiast do rozwiązywania problemów natury moralnej włączyły się: teoria psychometryczna i władze oświatowe, opracowując dokumenty formalnie rozstrzygające niejasności (Niemierko, 2012, s. 47). Dlatego miejsce kodeksów etycznych zajęły: „Standardy dla testów stosowanych w psychologii i pedagogice” (2007) oraz corocznie uaktualniane wykazy egzaminacyjnych procedur. Choć w stosunku do kodeksów etycznych okazują się one bardziej precyzyjne, rozwiązują wyłącznie problemy techniczne, pomijając kwestie etyczne i moralny kontekst sytuacji. Zarazem w konsekwencji utożsamienia sankcji

moralnych z sankcjami prawnymi dochodzi do zawłaszczenia tych pierwszych przez państwo i instytucje publiczne oraz ich zinstrumentalizowanie (Jaśtał, 2005, s. 101).

Z konferencyjnego dyskursu diagnostycznego wynika, że wobec doświadczanego przez badanych niedostatku etyczności egzaminów i innych diagnozowań z daleka, podejmowane są działania mające na celu zaprezentowanie diagnostyki jako zdolnej do uwzględniania wartości i potrzeb diagnozowanych. W staraniach na rzecz „uetycznienia” diagnozowania z daleka daje się wyróżnić kilka nurtów, w praktyce nie zawsze rozłącznych. Należą do nich:

1. Działania na rzecz zapewnienia bezstronności diagnozowaniu z daleka, zwłaszcza gdy jest ono egzaminem zewnętrznym. Jak wynika z konferencyjnych doniesień, ku temu zmierza wyznaczanie zewnętrznych obserwatorów egzaminów, od których oczekuje się nie tylko postępowania zgodnego z egzaminacyjnymi procedurami, lecz również kierowania się normami moralnymi.

2. Ukazywanie diagnozowania z daleka jako etycznego poprzez prezentowanie w związku z nim:

- „oburzenia moralnego”, czyli spektakularnego narzędzia stosowanego zwłaszcza podczas egzaminów zewnętrznych w sytuacjach, w których dochodzi do ujawniania przypadków nieuczciwych zachowań dyrektorów szkół, bądź gdy w artykułach prasowych opisywane są zdarzenia, które według instytucji egzaminacyjnych nie miały miejsca, a zatem godzą w dobre imię egzaminatorów;
- poprawności politycznej diagnoz, odzwierciedlającej w gruncie rzeczy masowe i stereotypowe sądy przewidujące, które spośród zawartych w nich informacji okażą się słuszne lub nieetyczne, dla kogo będą one korzystne, a dla kogo szkodliwe;
- „moralności” diagnozowania, w którym wartości uniwersalne są eksponowane i ukazywane jako leżące u podstaw pomiaru dokonywanego w imię dobra, prawdy i innych wartości (Czakon, 2005, s. 94–96).

Trzeba jednak pamiętać, że wartość tego typu działań jest zazwyczaj ograniczona i zamyka się w sferze moralnościowej retoryki, zapośredniczanej ze sfery etyki bliskiego zasięgu.

Zakończenie

W prowadzonych od lat konferencyjnych dyskursach diagnostycznych zawsze były obecne etyczne aspekty diagnozowania. Odczytane przez pryzmat dwudziestoletniego dorobku diagnostyki, umożliwiają dostrzeżenie ich narastającej złożoności, będącej konsekwencją przemieszania diagnozowań „z bliska” i „z daleka”, oraz pomijania odmiennych założeń leżących u podstaw etyk bliskiego i dalekiego zasięgu. Tymczasem ich włączenie do diagnostycznej refleksji uświadamia, że:

- jednym z dążeń diagnozowania „z daleka” nie jest sprostanie zasadom etyki, lecz zapewnienie psychometrycznej poprawności, zwłaszcza rzetelności;
- oprócz podporządkowania diagnozowania „z daleka” wymogom poprawności metodologicznej, oczekuje się, że zachowa ono poprawność polityczną nakładaną na diagnostów przez instytucje zlecające realizację badań.

Tradycyjne zalecenia „etyki bliźniego”, głoszące sprawiedliwość, miłosierdzie, uczciwość – nadal, oczywiście, obowiązują, ale bywają przestrzegane tylko w najbliższej sferze codziennych wzajemnych stosunków międzyludzkich i diagnozowania

„z bliska”. Jednak i w tej sytuacji sukcesywnie tracą one na swoim znaczeniu, gdyż zostają zaanektowane przez diagnozowaną „z daleka”, dziedzinę działań kolektywnych i bezimiennych, w których sprawca, czyn i skutek już nie są takie same, jakimi były w sferze etyki bliskiego zasięgu (Jonas, 1996, s. 30). Można odnieść wrażenie, że impulsy moralne przekazywane przez pozostające w bliskim kontakcie osoby okazują się w obecnej rzeczywistości za słabe, by mogły obowiązywać w liczniejszych grupach, za wątle w pokonywaniu odległości, zbyt mało ważne i cenne, by je przekazywać przyszłym pokoleniom oraz uwzględniać w diagnozowaniu „z daleka”.

Dziś już nie sposób wyodrębnić „czystych” postaci: diagnozowania i etyki „z bliska”, oraz diagnozowania i etyki „z daleka”. Wzajemnie uwikłane tworzą **sferę zapośredniczeń diagnostycznych i etycznych**. Opracowane w niej diagnozy sprawiają wrażenie patchworku, collage’u, bricolage’u bądź hybryd. Jednak przyjęcie owych konstruktów jako perspektywy myślenia o diagnostycznych zapośredniczeniach uświadamia, że nie stanowią przypadków czystych ontologicznie (Abriszewski, 2005, s. 415), gdyż wpisana jest w nie wielorakość. Łączy je także pewna wspólna im właściwość, jaką jest „interakcja między odmiennymi aspektami rzeczywistości, która pozwala między innymi na budowanie własnej tożsamości i własnego osądu rzeczywistości bez lekceważenia bądź odrzucania tego, co inne” (Szyling, 2013). W takim ujęciu właśnie sfera zapośredniczeń ujawnia się jako dominujący obszar eksploracji diagnostyki edukacyjnej. Właściwe jej: różnorodność i wieloraka złożoność, otwierają o wiele więcej możliwości badawczych niż jednowymiarowość i wierność modelom. Dlatego w przestrzeni zapośredniczeń widzę rozpościerające się przed diagnostyką edukacyjną interesujące perspektywy rozwoju w kolejnych latach.

Przypisy

¹ Przyjęłam, za Danutą Urbaniak-Zajac (2013, s. 255), że gdyby wybrana teoria w trakcie analizy i interpretacji okazała się nieprzystająca do rozpatrywanego zagadnienia, wówczas dokonam jej zmiany lub uzupełnię o inną koncepcję.

Literatura

Abriszewski K. (2005): *Etyka, polityka i rzeczy*. W: D. Probuca (red.): *Etyka i polityka*. OW „Impuls”, Kraków.

Abriszewski K. (2007): *Sytuacje graniczne jako sytuacje logicznie pierwotne*. W: D. Probuca (red.): *Etyka wobec sytuacji granicznych*. OW „Impuls”, Kraków.

Abriszewski K. (2010): *Wszystko otwarte na nowo. Teoria Aktora-Sieci i filozofia kultury*. Wyd. Naukowe Uniwersytetu Mikołaja Kopernika, Toruń.

Abriszewski K. (2012): *Poznanie, zbiorowość, polityka*. Wyd. Uniwersytetu, Kraków.

Aleksander Z. (1994): *Interakcja młodzież – dorośli: analiza komunikacyjna*. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Bartmańska A. (2003): *Kodeks etyki egzaminatora*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Bednarek K., Lempa Cz. (2005): *Wykorzystanie modelu dialogicznego oceniania w komunikowaniu wyników egzaminu zewnętrznego*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

- Bednarek K., Lempa Cz. (2006): *Egzaminator – osoba zaufania publicznego*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. gRUPA TOMAMI, Kraków.
- Boba M., i in. (2002): *W trosce o etykę egzaminu gimnazjalnego*. W: B. Niemierko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnętrzne i zewnętrzne a jakość pracy szkoły*. Wyd. BiP, Katowice.
- Brandt R. B. (1996): *Etyka. Zagadnienia etyki normatywnej i metaetyki*, tłum. B. Stanosz. Wyd. Naukowe PWN, Warszawa.
- Byrska J. (2005): *Pochodzenie treści etycznych w życiu publicznym*. W: D. Probuca (red.): *Etyka i polityka*. OW „Impuls”, Kraków.
- Critchley S. (2006): *Nieustające żądanie*, tłum. R. Dobrowolski, M. Gusin. Wyd. Naukowe DSW, Wrocław.
- Czakon T. (2005): *Etyka i polityka. Między moralizatorstwem a nihilizmem*. W: D. Probuca (red.): *Etyka i polityka*. OW „Impuls”, Kraków.
- Czarnota-Maczyńska J. (2006): *Niektóre prawne, organizacyjne i etyczne aspekty zapewniania obiektywizmu egzaminowania zewnętrznego*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. gRUPA TOMAMI, Kraków.
- Dolata R. (2009): *Cicha rewolucja w polskiej oświacie – proces różnicowania się gimnazjów w dużych miastach*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Wyd. gRUPA TOMAMI, Kraków.
- Domańska E. (2012): *Historia egzystencjalna*. Wyd. Naukowe PWN, Warszawa.
- Ficner C. (2004): *Zmierzyć wychowanie*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Wyd. gRUPA TOMAMI, Kraków.
- Furgoł S., Szafraniec M. (2002): *Rozterki egzaminatora*. W: B. Niemierko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnętrzne i zewnętrzne a jakość pracy szkoły*. Wyd. BiP, Katowice.
- Groenwald M. (2007): *Przestrzeń jako kategoria strukturotwórcza w badaniu osiągnięć uczniów*. W: M. Strzyż, A. Zieliński (red.): *Region w edukacji przyrodniczo-geograficznej*, Wyd. Instytutu Geografii Akademii Świętokrzyskiej, Kielce.
- Groenwald M. (2008): *Egzaminy a zaufanie*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczycieli*. Wyd. gRUPA TOMAMI, Kraków.
- Groenwald M. (2010): *Patchworkowa sprawiedliwość oceniania*. W: B. Niemierko, M. K. Szmigel (red.): *Terazniejszość i przyszłość oceniania szkolnego*. Wyd. gRUPA TOMAMI, Kraków.
- Gutek G.L. (2003): *Filozoficzne i ideologiczne podstawy edukacji*, tłum. A. Kacmąjor i A. Sulak. GWP, Gdańsk.
- Hall K. (1999): *Wpływ egzaminów zewnętrznych na pracę szkoły i nauczyciela na podstawie prac nad programem nauczania Klocki Autonomiczne*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna, Oczekiwania, problemy, przykłady*. Wyd. ODESiP, Legnica.
- Hawrot A., Jasińska A. (2013): *Kto się uczy poza szkołą? Predyktory korzystania z płatnej i bezpłatnej pomocy naucz.* W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Wyd. gRUPA TOMAMI, Kraków.
- Janicka-Panek T. (2007): *Samoocena osiągnięć edukacyjnych szóstoklasistów a wyniki sprawdzianu*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Wyd. gRUPA TOMAMI, Kraków.
- Jaśtał J. (2005): *W kierunku polityki cnót? Partykularyzm etyczny a decyzje polityczne*. W: D. Probuca (red.): *Etyka i polityka*. OW „Impuls”, Kraków.
- Jonas H. (1996): *Zasada odpowiedzialności*, tłum. M. Klimowicz. Wyd. Platan, Kraków.
- Klus-Stańska D. (2009): *Dyskursy pedagogiki wczesnoszkolnej*. W: D. Klus-Stańska, M. Szczepka-Pustkowska (red.): *Pedagogika wczesnoszkolna – dyskursy, problemy, rozwiązania*. Wyd. Akademickie i Profesjonalne, Warszawa.
- Klus-Stańska D. (2010): *Ukryta polityczność oceniania szkolnego w świetle radykalnych nurtów socjologicznych – tezy do wystąpienia*. W: B. Niemierko, M. K. Szmigel (red.): *Terazniejszość i przyszłość oceniania szkolnego*. Wyd. gRUPA TOMAMI, Kraków.
- Konarzewski K. (1998): *Jak nauczyciele normują test osiągnięć szkolnych*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Kozłowski W. (1999): *Ewaluacja pedagogiczna: niezbędna i niemożliwa*. W: B. Niemierko, B. Machowska (red.): *Diagnostyka edukacyjna, oczekiwania, problemy, przykłady*. Wyd. ODESiP, Legnica.

- Kruszewski K. (1998): *Test wiadomość złego i dobrego* W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Kwieciński Z. (2010): *Zmiany w edukacji – próba bilansu dwóch dekad*. W: B. Niemierko, M. K. Szmiigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.
- Latour B. (2010): *Splatając na nowo to, co społeczne*, tłum. A. Dera, K. Abriszewski. Universitas, Kraków.
- Lipska M. (2006): *Uczciwość edukacyjna w opinii przyszłych maturzystów*. W: B. Niemierko, M. K. Szmiigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.
- Lisiecka Z. (2006): *O polską kampanię na rzecz uczciwości edukacyjnej*. W: B. Niemierko, M. K. Szmiigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.
- Lewowicki T. (1998): *Współczesne wyzwania wobec diagnostyki edukacyjnej – przemiany teleologii oświatowej i pojmowania pedagogiki a szanse rozwoju diagnostyki*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Maciąg A. (2005): *Trafność oceny postępów z języka angielskiego w zreformowanej szkole ponadgimnazjalnej*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej*. Wyd. Fundacji Uniwersytetu Gdańskiego, Gdańsk.
- Mazur U., Tyralska-Wojtyca E. (2008): *Budowanie zaufania społecznego w kontekście współpracy z egzaminatorami i komunikowania wyników – refleksje z doświadczeń Pracowni Egzaminu Gimnazjalnego OKE w Krakowie*. W: B. Niemierko, M. K. Szmiigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Wyd. GRUPA TOMAMI, Kraków.
- Meighan R. (1993): *Socjologia edukacji*, tłum. Z. Knutsen i in. Wyd. UMK, Toruń.
- Miles M.B., Huberman A. M. (2000): *Analiza danych jakościowych*, tłum. S. Zabielski. Wyd. Uniwersyteckie Trans Humana, Białystok.
- Niemierko B. (2000): *Czy zadanie wielokrotnego wyboru nadaje się do diagnozowania procesów edukacyjnych*. W: B. Niemierko, J. Mulawa (red.): *Zadania wielokrotnego wyboru*. Wyd. Instytut Badań Kompetencji, Wałbrzych.
- Niemierko B. (2002): *Ocenianie szkolne bez tajemnic*. WSiP, Warszawa.
- Niemierko B. (2003): *Jaki pomiar dydaktyczny jest nam potrzebny?* W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.
- Niemierko B. (2006): *Oszustwo egzaminacyjne*. W: B. Niemierko, M. K. Szmiigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.
- Niemierko B. (2009): *Zachodni wiatr w edukacji. Od unijnego prelegenta do wspólnych badań*. W: B. Niemierko, M. K. Szmiigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Wyd. GRUPA TOMAMI, Kraków.
- Niemierko B. (2012): *Jawne i ukryte błędy pomiaru dydaktycznego*. W: B. Niemierko, M. K. Szmiigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.
- Niemierko B. (2013): *Diagnostyka edukacyjna duża i mała*. W: B. Niemierko, M. K. Szmiigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Wyd. GRUPA TOMAMI, Kraków.
- Peräkylä A. (2009): *Analiza rozmów i tekstów*, tłum. A. Figiel. W: N. K. Denzin, Y. S. Lincoln, *Metody badań jakościowych*, t.2. Wyd. Naukowe PWN, Warszawa.
- Podrez E. (2005): *Etyka i polityka a sens ludzkiej egzystencji (pytania otwarte)*. W: D. Probuca (red.): *Etyka i polityka*. OW „Impuls”, Kraków.
- Probuca D. (2005): *Hierarchizacja postaw polityków w świetle etapów życia moralnego*. W: D. Probuca (red.): *Etyka i polityka*. OW „Impuls”, Kraków.
- Promieńska H. (2005): *O trzech funkcjach etyki i ewolucji jej zadań*. W: D. Probuca (red.): *Etyka i polityka*. OW „Impuls”, Kraków.
- Putkiewicz E. (2005): *Korepetycje – szara strefa edukacji*. Wyd. Fundacja Instytut Spraw Publicznych, Warszawa.
- Rapley T. (2013): *Analiza konwersacji, dyskursu i dokumentów*, tłum. A. Gąsior-Niemiec. Wyd. Naukowe PWN, Warszawa.
- Standardy dla testów stosowanych w psychologii i pedagogice*, tłum. E. Hornowska (2007). GWP, Gdańsk.
- Suckiel I. (2008): *Opinie nauczycieli na temat egzaminów zewnętrznych*. W: B. Niemierko, M. K. Szmiigel (red.): *Egzamin i uczenie się w oczach nauczyciela*. Wyd. GRUPA TOMAMI, Kraków.

- Szkudlarek T. (2005), *Ekonomia i społeczne skutki reform*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Szaleniec H. (2001): *Próbny egzamin zewnętrznego systemu oceniania*. W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej, Wrocław.
- Szaleniec H. (2006): *Oszukiwanie na egzaminie istotnym źródłem majowej porażki*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.
- Szaleniec H. (2011): *Wyniki egzaminu a ewaluacja zewnętrzna szkoły*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Wyd. GRUPA TOMAMI, Kraków.
- Szmigel M. K. (2001): *Uzyskiwanie zaufania do egzaminów zewnętrznych*. W: B. Niemierko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. Pandit, Kraków.
- Szmigel M. K. (2007): *Uczenie się w opiniach uczniów szkół ponadgimnazjalnych*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Wyd. GRUPA TOMAMI, Kraków.
- Szmigel M..K., Szalenciec H. (2001): *Obserwator osobą społecznego zaufania w systemie oceniania zewnętrznego*. W: B. Niemierko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. Pandit, Kraków.
- Szutta N. (2007): *Współczesna etyka cnót*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Szyling G. (2005): *„Między młotem a kowadłem”, czyli rzecz o kulturze nauczycielskiego oceniania*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Szyling G. (2013): *Hybrydowy charakter oceny szkolnej we wczesnej edukacji: stan i wyzwania*. „Studia Edukacyjne”, nr 28.
- Tyralska-Wojtycza E. (2006): *Unieważnienie egzaminu próbą wyjścia na prostą*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.
- Urbaniak-Zajac D., Kos E. (2013): *Badania jakościowe w pedagogice*. Wyd. Naukowe PWN, Warszawa.
- Wieczorkowska G., Radkiewicz P. (2000): *Magia liczb o ocenianiu w edukacji*. W: B. Niemierko, J. Mulawa (red.): *Zadania wielokrotnego wyboru*. Wyd. Instytutu Badań Kompetencji, Wałbrzych.
- Wysocka E. (2013): *Diagnostyka pedagogiczna. Nowe obszary i rozwiązania*. OW „Impuls”, Kraków.

Diagnostowanie – między etyką bliskiego a dalekiego zasięgu

W artykule poddano analizie opracowania opublikowane w wydawnictwach konferencyjnych w latach 1994–2013, podejmujące kwestie etycznych aspektów diagnostowania realizowanego: „z bliska” w wymiarze etyki bliskiego zasięgu bądź „z daleka” w kontekście etyki dalekiego zasięgu. W przyjętym konstrukcie przestrzeń „pomiędzy” zajmują wielostronne oraz wielokierunkowe zapośredniczenia etyczne i diagnostyczne. Ich ontologiczna niejednorodność implikuje poznawczą złożoność, czyniąc strefę zapośredniczeń interesującym obszarem eksploracji dla diagnostyki edukacyjnej.

Słowa kluczowe: etyka bliskiego i dalekiego zasięgu, diagnostowanie: „z bliska” i „z daleka”, strefa zapośredniczeń etycznych i diagnostycznych

Diagnosing – between close-range and remote-range ethics

The article constitutes an analysis of papers presented in the conference publications in the period 1994–2013 concerning ethical issues of diagnosis conducted as “close-ups” under close-range ethics or “from afar” in the context of remote-range ethics. In the adopted construct the space “between” is taken by multi-faceted and multi-directional ethical and diagnostic borrowings. Their ontological non-uniformity implies cognitive complexity, making the sphere of borrowings an interesting exploration field for educational diagnostics.

Keywords: close-range and remote-range ethics, diagnosing as “close-ups” and “from afar”, sphere of ethical and diagnostic borrowings

Piotr Maciej Skorupiński
 III Liceum Ogólnokształcące
 im. gen. Józefa Sowińskiego w Warszawie

RECEPCJA KONCEPCJI TRAFNOŚCI W DIAGNOSTYCE EDUKACYJNEJ

„Nie należy się spodziewać, że przewidzimy (...) zachowanie oficera na polu walki na podstawie tego, jak sobie radził na placu treningowym (...). Nie neguję, że testy mają swoje miejsce – jeśli test pozwala przewidzieć ważny rezultat z trafnością 0,20 czy 0,30, należy z niego skorzystać. Jednak nie oczekujemy niczego więcej”.

Daniel Kahneman, *Pułapki myślenia*

Modele trafności

Trafność stanowi fundamentalne pojęcie teorii pomiaru edukacyjnego i – szerzej – metodologii diagnostyki edukacyjnej. Poszukiwania klarownej koncepcji trafności zaowocowały różnorodnymi jej modelami (por. Kane, 2006, s. 18–23; Niemierko, 2009, s. 145–149), między innymi kryterialnym, trynitarnym, unitarnym i argumentacyjnym (Skorupiński, 2013).

W ramach kryterialnego modelu trafności, dominującego w pierwszej połowie ubiegłego stulecia, badanie trafności testu sprowadzało się do określenia jego trafności kryterialnej (za pomocą korelacyjnego współczynnika trafności) lub – w przypadku testów osiągnięć – trafności treściowej. Model trynitarny, kluczowy w II połowie wieku poprzedniego, obejmował już trzy zasadnicze typy trafności: trafność kryterialną, trafność treściową oraz trafność teoretyczną, definiowaną na przykład jako stopień teoretycznego uzasadnienia interpretacji wyników pomiaru. W modelu unitarnym, zaproponowanym przez Samuela Messicka, dotychczas funkcjonujące typologie trafności zastąpiono zunifikowanym i jednocześnie poszerzonym pojęciem trafności, rozumianym jako sąd oceniający, w jakim stopniu dane empiryczne i konstrukcje teoretyczne upoważniają do inferencji i działań opartych na wynikach pomiaru edukacyjnego (por. Messick, 1989, s. 13). Swoistą strategię ustalania tak pojętej trafności – inaczej: argumentacyjny model trafności – popularyzuje obecnie Michael T. Kane (2006).

W polskojęzycznym piśmiennictwie poświęconym pomiarowi edukacyjnemu i psychologicznemu – od czasów Pracowni Psychometrycznej PAN, kierowanej przez

Mieczysława Choynowskiego, do lat dziewięćdziesiątych ubiegłego stulecia, kiedy to pojawiają się pierwsze próby unitarnego ujęcia problematyki trafności – dominuje trynarny model trafności. W ostatnim piętnastolecu trwają prace nad włączeniem unitarnego modelu trafności do klasycznej już teorii pomiaru dydaktycznego, zainicjowane przez Bolesława Niemierkę. W tym ostatnim okresie zainteresowanie koncepcjami unitarnymi, a także wcześniejszymi, wielokrotnie artykułowane było również podczas kolejnych konferencji z cyklu „Diagnostyka edukacyjna”. Głosy te zostaną przypomniane i w finalnych partiach powiązane z pytaniem o przyszłość idei trafności w diagnostyce edukacyjnej czy – wężej: w środowiskach teoretyków i praktyków pomiaru edukacyjnego.

Początki recepcji modelu unitarnego

Podczas konferencji „Problemy diagnozy edukacyjnej” B. Niemierko, wskazując cele diagnostyki edukacyjnej, a wśród nich podniesienie trafności oceniania, odwołał się do S. Messicka (1989) i jego „zhumanizowanego, rozszerzonego pojęcia «trafności», obejmującego zarówno zgodność procedury z założeniami, jak i długofalowe skutki jej zastosowania” (Niemierko, 1999, s. 17). Wówczas także Maria Sobczak (1999, s. 322), powołując się na publikacje B. Niemierki, zgłosiła postulat „nieustannego badania trafności, a zwłaszcza tych jej aspektów, które dotyczą diagnozy skutków podejmowanych działań” oraz wskazała przykładowe działania zmierzające do uzasadnienia trafności pomiaru, a w tym przeprowadzenie dowodu zgodności zamierzonego i rzeczywistego przedmiotu pomiaru, zgodności egzaminu z teorią naukową oraz korzystnych jego konsekwencji.

Podczas konferencji „Zadania wyboru wielokrotnego” B. Niemierko (2000, s. 9–10) określił warunki konieczne diagnozy osiągnięć: dysponowanie „strukturalnym modelem teoretycznym oraz zadowalająco trafnymi miarami (...) osiągnięć, jakie ten model obejmuje”, i stwierdził, że aby zapewnić trafność, należy prowadzić badania trafności. Podczas konferencji „Diagnostyka edukacyjna jako dziedzina kompetencji pedagogicznych” wyraził przekonanie, że „wszyscy [kuratoria, szkoły, ośrodki egzaminacyjne itp. – przyp. M.S.] będą wymagać wsparcia w pomiarze, jako teorii naukowej i zbiorze procedur uściślonego poznawania rzeczywistości” (Niemierko, 2001b, s. 73), i zalecił autorom diagnozy edukacyjnej, w której główną rolę powinny odegrać wówczas nowo powstałe komisje egzaminacyjne, szacowanie jej obiektywizmu, trafności i rzetelności. Podczas konferencji „Teoria i praktyka oceniania zewnętrznego” B. Niemierko (2001a, s. 16), porównując egzaminy doniosłe i powszednie, jeszcze raz przywołał Messickowski – jak to zostało ujęte – „model rozszerzonej trafności teoretycznej, obejmującej nie tylko uzasadnienia pomiaru na podstawie zebranych dowodów, lecz także uzasadnienia na podstawie przewidywanych konsekwencji powziętych decyzji”.

W trakcie tej samej konferencji o „rozszerzeniu (...) znaczenia pojęcia trafności o funkcje aksjologiczne i społeczne” i o „typologii trafności w postaci macierzy postępującej” – w kontekście kilkudzaniowego omówienia ewolucji pojęcia trafności – wspomniła też Arleta Poręba-Konopczyńska (2001, s. 37) we wstępie do rozważań nad relacją między rzetelnością i trafnością. Podobnie Bożena Żuk (2001) podkreśliła wagę zintegrowanego i poszerzonego pojęcia trafności w trakcie ewaluacji egzami-

nu zewnętrznego oraz przedstawiła Messickowską koncepcję sześciu aspektów trafności teoretycznej (treściowego, rzeczowego, strukturalnego, uniwersalizacyjnego, zewnętrznego i konsekwencyjnego). Ponadto o konieczności badania trafności konsekwencyjnej wspomniała Ewa Ludwikowska (2001b), a Roman Sokulski (2001), przywołując B. Niemierkę i S. Messicka, zarzucił Centralnej Komisji Egzaminacyjnej, iż przyczynia się do obniżenia trafności projektowanego pomiaru, rezygnując z przekonywania środowisk nauczycielskich do swoich racji.

Na konferencji „Dawne i nowe formy egzaminowania” B. Niemierko stwierdził między innymi, że idealny test sprawdzający – choć niemożliwy do skonstruowania, lecz ważny, bo wyznaczający kierunek prac – charakteryzować się powinien zupełną rzetelnością i całkowitą trafnością programową. Zidentyfikował także ograniczenia, które powodują, że „trafność programowa testu nie może być zupełna” i że „na ogół nie jesteśmy w stanie oszacować jej poziomu” (Niemierko, 2001c, s. 20). Ewa Ludwikowska (2001a, s. 307) zwróciła uwagę na społeczne skutki egzaminu doniesłego oraz skonstatowała, że „trafność teoretyczna opiera się na połączeniu wielu dowodów, które odnoszą się do znaczenia lub interpretacji wyników” i „obejmuje wszystkie formy dowodów trafności, w tym zarówno stosowność i reprezentatywność treści zadań oraz porównanie z kryteriami” (tamże). Wskazała także na związek badania treści kształcenia z badaniem trafności pomiaru. Elżbieta Kowalik (2001) podjęła próbę określenia trafności treściowej testu laboratoryjnego z chemii.

Podczas konferencji „Dwa rodzaje oceniania szkolnego” B. Niemierko (2002, s. 20) porównał egzaminy zewnętrzne i wewnętrzne pod względem ich właściwości pomiarowych, w tym trafności, którą określił jako „użyteczność wniosków opartych na jego [pomiaru – przyp. M.S.] wynikach do określonego celu”. Piotr M. Skorupiński (2002, s. 44) podjął problem konsekwencji stosowania testów, nawiązując do Messickowskiego pojęcia trafności i podkreślając, że „trafność jest raczej cechą interpretacji wyników sprawdzania osiągnięć uczniów i działań z nimi związanych niż właściwością samego testu”. Z kolei Władysław Mąsior (2002) przywołał – za B. Niemierką – koncepcję aspektów trafności, koncentrując się na aspektach treściowym i strukturalnym w trakcie analizy zadań matematycznych z egzaminu gimnazjalnego. Na S. Messicka – również za pośrednictwem B. Niemierki – powołał się Wojciech Walczak (2002), cytując definicję rozszerzonej trafności i podejmując zagadnienie skutków społecznych pomiaru edukacyjnego.

Podczas konferencji „Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych” B. Niemierko (2003, s. 18) określił koncepcje S. Messicka jako „motor postępu”, a jej autora nazwał inicjatorem rewolucji w wartościowaniu pomiaru edukacyjnego, rozpoczętej w Stanach Zjednoczonych na początku lat dziewięćdziesiątych ubiegłego stulecia i „stamtąd promieniującej na resztę świata”. Charakteryzując te koncepcje, zaprezentował Messickowską macierz postępującą (*progressive matrix*), syntetyzującą zintegrowane pojęcie trafności teoretycznej, obejmujące wyodrębnione wcześniej typy trafności i poszerzone o kontrowersyjną trafność konsekwencyjną, opartą – jak stwierdził – „na niepewnych przewidywaniach przyszłości, nie zaś na jakimkolwiek zbiorze danych z badań (*evidence*)” (tamże). W związku z powyższym, pokrótce relacjonując debatę środowisk pomiarowych nad koncepcjami Messicka, przywołał polemiczne głosy W. Jamesa Pophama i Williama Mehrensa, domagające się uściślenia i zawężenia pojęcia trafności. W tym kontekście kilkakrotnie wyraził wysoką ocenę dorobku naukowego Lee Cronbacha, wskazując między innymi

na ukształtowanie przez niego pojęcia trafności teoretycznej i zmianę przedmiotu dowodu trafności: z testu na decyzje podejmowane na podstawie wyników pomiaru (por. tamże, s. 20–22).

W trakcie konferencji – zogniskowanej na problematyce trafności – Krzysztof Konarzewski (2003, s. 42) podkreślił, że „angielskiemu słowu (*validity* – przyp. aut.) znacznie bliżej do naszej «prawomocności» niż do skromnej, technicznej «trafności»”. W związku z powyższym stwierdził, że odpowiedź na pytanie o trafność testu „uciera się w dyskursie polemicznym i nigdy nie jest ostateczna” (tamże) ze względu na niejasne warunki uznania pomiaru za prawomocny. Wskazał też jedyny – jego zdaniem – sposób wykazania trafności testu osiągnięć szkolnych, jakim jest określenie trafności treściowej, czyli dowiedzenie, że test „rzetelnie mierzy dyspozycje (...) zawarte w realizowanym programie kształcenia” (tamże), zaznaczając, iż uprawomocnienie pojęcia obejmującego te dyspozycje, a więc oszacowanie trafności teoretycznej, będzie „niełatwe”. Z kolei Peter Kroppe i Wilhelm Wolze (2003, s. 59) krótko scharakteryzowali trafność treściową, trafność kryterialną i trafność teoretyczną. Michał Daszkiewicz (2003, s. 70), opierając się na artykule M. Kane’a (2001), opisał procedurę „argumentacyjnego ustalania trafności (*argument-based approach*)”, prowadzącą od/do „dyskursu interpretacyjnego” (*interpretative argument*) do/od „dyskursu nad trafnością” (*validity argument*). P. Skorupiński (2003a) przedstawił genezę pojęcia i terminu *construct validity* oraz odpowiadającej im na gruncie polskim trafności teoretycznej, koncentrując się na publikacjach z lat pięćdziesiątych ubiegłego stulecia. Podkreślił też znaczenie studiów korelacyjnych w badaniu zarówno tradycyjnie pojętej trafności kryterialnej, jak i Messickowskiej trafności teoretycznej (Skorupiński, 2003b). Grażyna Szyling (2003, s. 117) zaprezentowała zmodyfikowaną Messickowską macierz postępującą (*progressive matrix*) i zdefiniowała trafność jako „jednolite pojęcie, którego podstawę stanowi trafność teoretyczna potwierdzona empirycznymi dowodami”, walidację zaś jako „proces polegający na gromadzeniu i porządkowaniu danych empirycznych oraz argumentów (racjonalnych uzasadnień)” (tamże). Bożena Żuk (2003), zwracając uwagę na przemiany zachodzące za sprawą S. Messicka, zaproponowała hermeneutyczne i jakościowe badanie trafności testu. Danuta Król (2003) przedstawiła wyniki korelacyjnego badania trafności prognozy stycznej egzaminów i ocen szkolnych. Krzysztof Bednarek i Cezary Lempa (2003, s. 297) definiowali pojęcie trafności na podstawie jednego ze słowników języka polskiego, stwierdzając, że „najbardziej adekwatne dla naszych rozważań wydają się być określenia «trafiający w cel» oraz «odpowiadający rzeczywistości»”.

Marginalizacja trafności

W trakcie kolejnych dziesięciu konferencji z cyklu „Diagnostyka edukacyjna”, które odbyły się w latach 2004–2013 i zaowocowały obejmującymi około pół tysiąca artykułów publikacjami konferencyjnymi, problematyka trafności stopniowo ulegała marginalizacji. Z jednej strony, B. Niemierko konsekwentnie, powracał do rozbudowanego przez S. Messicka pojęcia trafności teoretycznej. Między innymi twierdził, iż „ma [trafność teoretyczna – przyp. M. S.] u podstaw trafność wewnętrzną, polegającą na porównaniu treści pomiaru z programem kształcenia lub z inną dokumentacją wymagań wobec ucznia”. Uznał, że „kluczowe dla pomiaru dydaktycznego jest

pytanie o jego trafność” (Niemierko, 2004, s. 22–23), po S. Messicku badaną także z zastosowaniem metod właściwych hermeneutyce czy konstruktywizmowi (tamże). Rozwazał także problem oszustwa egzaminacyjnego, eksplorowany za Gregorym Cizkiem – definiującym trafność jako stopień, w jakim dane wspierają wnioski dotyczące egzaminowanych, oraz jako celność konkluzji opartych na wynikach testu (por. Cizek, 1999, s. 5–6). Zwracał uwagę na „ubytek trafności [programowej – przyp. M.S.] egzaminów zewnętrznych (...) w stosunku do (...) metod diagnostycznych stosowanych w sposób ciągły przez nauczycieli” (Niemierko, 2008, s. 24) i relację pomiędzy wysoką ich rzetelnością i trafnością. W 2011 roku stwierdził, iż „spośród wielu odmian trafności i stanowisk metodologicznych (...) wybieramy koncepcję Samuela Messicka” (Niemierko, Walukiewicz, 2011, s. 77) i przywołał po raz kolejny Messickowską macierz, akcentując konieczność uwzględnienia trafności konsekwencyjnej. Rok później przypominał, iż trafność stanowi jedną z pięciu niezbędnych właściwości pomiaru dydaktycznego, i przywołał Messickowską definicję trafności (Niemierko, 2012, s. 45; Messick, 1989, s. 13). Wreszcie podczas XIX Krajowej Konferencji Diagnostyki Edukacyjnej kilkakrotnie powrócił do problemu trafności, zwracając uwagę na kwestie trafności diagnoz, trafności osądów, trafności oceny i trafności pomiaru sprawdzającego (por. Niemierko, 2013, s. 27–31).

O ile w latach 2004–2005 w pozostałych referatach konferencyjnych terminologia związana z problematyką trafności była przywoływana – głównie przez uczestników seminarium prowadzonego przez Niemierkę przy Międzywydziałowym Studium Pedagogicznym Uniwersytetu Gdańskiego – w miarę regularnie, o tyle w latach następnych ustąpiła zupełnie innym dyskursom, a w tym szczególnie tzw. edukacyjnej wartości dodanej (EWD), koncepcji rozwijanej od 2005 roku przez zespół koordynowany przez Romana Dolatę (2006), regularnie prezentujący wyniki realizowanych przedsięwzięć. W ostatnich dwóch latach pojawiły się ponadto liczne doniesienia z badań, prowadzonych przez Instytut Badań Edukacyjnych, ukierunkowanych na przykład na zastosowanie tzw. teorii wyniku zadania (Verhelst, 2009) w polskim systemie egzaminów zewnętrznych.

W tym pierwszym okresie jeszcze Michał Daszkiewicz (2004, s. 125), odwołując się do artykułów z pisma *Educational Measurment. Issues and Practice*, wśród cech oceniania wymienił trafność, rozumianą jako adekwatność bądź do treści (w ocenianiu zewnątrzszkolnym), bądź do procesu (w ocenianiu wewnątrzszkolnym), i powiązał z problematyką trafności pytanie o skutki oceniania. Arleta Poręba-Konopczyńska (2004, s. 145–146; 2005, s. 94) pisała o „rewolucji zapoczątkowanej (...) przez S. Messicka” i przedstawiła swoją autorską modyfikację Messickowskiej macierzy postępującej (*progressive matrix*), zwracając uwagę na metodologiczne skutki „rozszerzenia pojęcia trafności o problematykę nie dającą się mierzyć ilościowo” (tamże). Również Małgorzata Lipska (2005, s. 209–211) sięgnęła – za pośrednictwem *Pomiaru wyników kształcenia* B. Niemierki – po tę macierz i omówiła Messickowską koncepcję trafności teoretycznej. Do tej koncepcji nawiązała też Alina Maciąg (2005, s. 269), poszukując skutków oceniania. O ograniczeniach rozumianej na sposób Messickowski trafności pomiaru dokonywanego za pomocą egzaminu doniosłego rozprawiała G. Szyling (2005a, s. 147–148; 2005b, s. 101–102)¹. P. M. Skorupiński (2004, s. 89; 2005b; 2005a, s. 159) przedstawił Cronbachowską koncepcję walidacji konstruktów i wspominał o wcześniejszym kryterialnym podejściu do problemu trafności. Maria K. Szmigel (2004, s. 351) stwierdziła, iż „trafność pola programowego” dla standaryzowanych arkuszy egzaminacyjnych

„gwarantowana” jest przez przygotowujące je komisje egzaminacyjne. Spoza kręgu seminaryjnego jedynie R. Dolata (2004, s. 86) odniósł się do problematyki trafności, zadając pytanie o trafność oceniania wypracowań z języka polskiego.

W latach 2006–2013 liczba zastosowań aparatu terminologicznego i pojęciowego trafności zdecydowanie spadła, do tego stopnia, iż w materiałach z trzech konferencji z lat 2009–2011 pojawił się on właściwie tylko raz: we wspólnym referacie B. Niemierki i S. Walukiewicza (2011). Jednocześnie autorzy wielu teoretycznie zaawansowanych analiz własności psychometrycznych różnorodnych narzędzi pomiarowych, na przykład Artur Pokropek (2011), kwestię trafności pozostawili do rozwinięcia przy innej okazji.

Jedynie Maria Groenwald (2007, s. 393), analizując zadanie testowe, nawiązała do artykułu S. Messicka *Trafność testu a etyka oceny (diagnozy)*². Także Marcin Smolik (2006a, s. 579–581; 2006b, s. 120–121; 2008, s. 185, 189) w krytycznych wywodach dotyczących egzaminu maturalnego i gimnazjalnego kilkakrotnie powołał się na *Validity* S. Messicka (1989) i artykuły M. Kane’a, przedstawiające argumentacyjne podejście do ustalania trafności (*argument-based approach*). Swoją koncepcję badania trafności konsekwencyjnej oraz trafności opartej na teorii i kilku innych typów trafności egzaminu gimnazjalnego przedstawiła Sabina Nowak (2008). Krzysztof Krużewski (2012, s. 51) zwrócił uwagę na „trafność mierzoną indywidualnymi losami”, a G. Szyling (2012) wykazała korzyści płynące z triangulacji w procesie podnoszenia trafności badania diagnostycznego.

Aleksandra Jasińska i Michał Modzelewski (2012, s. 159, 167) pisali o „wydłużaniu” testu osiągnięć szkolnych w celu „zwiększenia” trafności pomiaru, definiując „trafne zadanie” jako zadanie, które „jest w ogóle w stanie zmierzyć to, co chcemy zmierzyć”. Filip Kulon, Agata Gajewska-Dyszkiewicz i Katarzyna Paczuska (2012, s. 442–443) zauważyli, że rzetelność i trafność są „bardzo istotnymi parametrami wszelkich testów”, i przywołali trynitarny model trafności, wskazując na stosowane metody weryfikowania trafności treściowej, trafności kryterialnej i trafności teoretycznej egzaminu. Do unitarnego rozumienia trafności odwołał się A. Pokropek (2013, s. 113), określając trafność jako „charakterystykę, która odnosi się do relacji między narzędziem pomiarowym a celami, jakie są przed nim stawiane”, i stwierdzając, iż w koncepcji S. Messicka trafność ujmowana jest w „kategoriach zdolności formułowania, na podstawie wyników pomiaru, wniosków dotyczących konkretnie sformułowanych funkcji pomiaru” (tamże). Anna Wiśniewska (2013) zaapelowała o refleksję nad problematyką trafności i rzetelności egzaminu gimnazjalnego, a Jerzy Chodnicki (2013, s. 404), rozprawiając o egzaminie maturalnym z matematyki, sformułował tezę, iż „stosowane na maturze schematy [oceniań – przyp. M.S.] są nie-trafne, niepoprawne pod względem merytorycznym i źle stosowane”. O trafności czynnikowej, trafności programowej, trafności wewnętrznej i zewnętrznej konkretnych narzędzi wspomnieli autorzy trzech innych referatów.

Pytania o przyszłość nauki o trafności

W ogromnym i tematycznie wielostronnym dorobku piśmienniczym środowiska skupionego wokół konferencji z cyklu „Diagnostyka edukacyjna” znalazły się teksty świadczące zróżnicowanym stopniu zaawansowania recepcji teorii trafności. W ob-

szernych materiałach konferencyjnych znaleźć można zarówno pionierskie próby popularyzowania unitarnego modelu trafności S. Messicka oraz modelu argumentacyjnego M. T. Kane'a, jak i nawiązania do starszych koncepcji, takich jak model trynitarny, obejmujący różne typy trafności. Jednocześnie w licznych pracach, poświęconych metodom czy narzędziom diagnozy edukacyjnej, żaden z modeli trafności nie został zastosowany.

Wśród tekstów propagujących nowoczesne koncepcje trafności, ówczesnie i współcześnie dyskutowane w światowej społeczności teoretyków pomiaru edukacyjnego, zdecydowanie dominują referaty B. Niemierki i – do roku 2005 – uczestników prowadzonego przez niego seminarium doktoranckiego na Uniwersytecie Gdańskim. Po tym okresie zintensyfikowanego zainteresowania teorią trafności nie tylko dramatycznie spada liczba wystąpień zawierających aparat pojęciowo-terminologiczny modelu unitarnego, lecz także powracają historyczne koncepcje trafności, mimo niezmiennego przywiązania B. Niemierki do idei zunifikowania i poszerzenia tego pojęcia. Jednocześnie rolę pierwszoplanową zaczynają odgrywać dyskursy zaawansowane pod zdecydowanie innym względem.

Czy zasygnalizowane powyżej – powiedzmy – przesunięcie problemowe, które dokonało się w granicach diagnostyki edukacyjnej jako subdyscypliny pedagogicznej *in statu nascendi*, nie spowoduje jej rychłego zastąpienia przez inny program badawczy? Albo zdecydowanie wężziej: czy możemy się spodziewać, iż stagnacja w zakresie teorii walidacji nie ogranicza w istotny sposób rozwoju pomiaru edukacyjnego? Trwają właśnie przygotowania do XX Krajowej Konferencji Diagnostyki Edukacyjnej, organizowanej przez Zakład Badań nad Dzieciństwem i Szkołą Instytutu Pedagogiki Uniwersytetu Gdańskiego oraz Polskie Towarzystwo Diagnostyki Edukacyjnej. Oczekując na ich efekt, życzyłbym całemu Towarzystwu, aby pesymistyczna diagnoza Zbigniewa Kwiecińskiego (2013) dotycząca opóźnienia rozwojowego, kulturowego, cywilizacyjnego i edukacyjnego naszego kraju – w odniesieniu do diagnostyki edukacyjnej okazała się w najgorszym przypadku trafna w wymiarze określonym przez D. Kahnemana.

Przypisy

¹ Szerzej problematykę trafności G. Szyling omówiła w późniejszych *Nauczycielskich praktykach oceniania poza standardami*, w których nie tylko skomentowała Messickowską macierz postępującą (*progressive matrix*), lecz także przedstawiła ewolucję pojęcia trafności (por. Szyling, 2011, s. 66–70).

² W późniejszych *Etycznych aspektach egzaminów szkolnych* M. Groenwald podejmie próbę powiązania trafnościowych idei S. Messicka z wielostronną refleksją etyczną nad egzaminem szkolnym (por. Groenwald, 2011, s. 20–21).

Literatura

Bednarek K., Lempa C. (2003): *Znaczenie kontekstu dla trafnego komentarza do wyników oceniania zewnętrznego*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Chodnicki J. (2013): *Co kryje się za tymi liczbami?* W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Wyd. GRUPA TOMAMI, Gniezno.

Cizek G. J. (1999): *Cheating on tests. How to do it, detect it, and prevent it*. Lawrence Erlbaum Associates, Mahwah.

Daszkiewicz M. (2003): *Znaczenie i zawartość „argumentu” oraz jego rola w ustalaniu trafności i komunikowaniu wyników egzaminu*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Daszkiewicz M. (2004): *Pierwsze kroki „klasometrii”*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Dolata R. (2004): *Wybrane problemy rzetelności i trafności nowej matury*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Dolata R. (2006): *Informacja o pracy zespołu metodologicznego ds. prognozowania i komunikowania wyników egzaminów, działającego w ramach projektu CKE Badania dotyczące wyników egzaminów zewnętrznych*. „Egzamin. Biuletyn Badawczy Centralnej Komisji Egzaminacyjnej” nr 8, dostępny na: http://www.cke.edu.pl/images/stories/badania/biul_8.pdf (otwarty 23 marca 2014 r.).

Groenwald M. (2007): *Nasza przyjaźń nigdy się nie rozstanie – inspirowana Arystotelesem analiza rozwiązań jednego zadania*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Wyd. GRUPA TOMAMI, Łomża 2007.

Groenwald M. (2011): *Etyczne aspekty egzaminów szkolnych*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Jasińska A., Modzelewski M. (2013): *Można inaczej. Wykorzystanie IRT do konstrukcji testów osiągnięć szkolnych*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.

Kane M. T. (2001): *Current concerns in validity theory*. „Journal of Educational Measurement” t. 38, nr 4.

Kane M. T. (2006): *Validation*. W: R. L. Brennan (red.): *Educational measurement*. American Council on Education i Praeger, Westport.

Konarzewski K. (2003): *Egzamin 2002: rzetelność i trafność testu matematyczno-przyrodniczego*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Kowalik E. (2001): *Trafność testu laboratoryjnego wobec podstawy programowej i standardów wymagań egzaminacyjnych z chemii w liceum ogólnokształcącym*. W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław 2001.

Kropie P., Wolze W. (2003): *Ordinary language and scientific language. Test validity from a constructivist point of view*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Król D. (2003): *O trafności prognostycznej egzaminu gimnazjalnego*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Kruszewski K. (2013): *Poszerzajcie temat błędów pomiaru dydaktycznego (fragment osobistego listu, przedrukowany za zgodą autora)*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.

Kulon F. i in. (2013): *Porównanie wyników Europejskiego Badania Kompetencji Językowych i egzaminu gimnazjalnego z języków obcych*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.

Kwieciński Z. (2013): *Edukacja wobec opóźnienia kulturowego*. „Nauka” nr 1, s. 19–30, dostępny na: http://www.pan.poznan.pl/nauki/N_113_02_Kwiecinski.pdf (otwarty 23 marca 2014 r.).

Lipska M. (2005): *Oceniający ocenianymi...? Nadzór pedagogiczny a wyniki egzaminów zewnętrznych – wybrane aspekty*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Ludwikowska E. (2001a): *„Czego Jaś się nauczył, tego Jan nie będzie umiał”, czyli o okazjach do uczenia się*. W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław.

Ludwikowska E. (2001b): *O potrzebie, komunikowaniu i rozumieniu informacji zwrotnej w edukacji*. W: B. Niemierko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. PANDIT, Kraków.

- Maciąg A. (2005): *Trafność oceny postępów z języka angielskiego w zreformowanej szkole ponadgimnazjalnej*. W: B. Niemięko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.
- Mąsior W. (2002): *Aberacje egzaminacyjne, czyli o wychodzeniu z jaskini*. W: B. Niemięko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnątrzszkolne i zewnętrzne a jakość pracy szkoły*. Katowice.
- Messick S. (1989): *Validity*. W: R. L. Linn (red.): *Educational measurement*. American Council on Education i Macmillan, New York.
- Niemięko B. (1999): *Diagnostyka edukacyjna. Cele, metody i szanse rozwoju*. W: B. Niemięko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. ODESiP, Legnica.
- Niemięko B. (2000): *Czy zadanie wyboru wielokrotnego nadaje się do diagnozowania procesów edukacyjnych?* W: B. Niemięko, J. Mulawa (red.): *Diagnoza edukacyjna. Zadania wyboru wielokrotnego*. Instytut Badań Kompetencji w Wałbrzychu, Wałbrzych.
- Niemięko B. (2001a): *Chłodne oblicze egzaminu zewnętrznego*. W: B. Niemięko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. PANDIT, Kraków.
- Niemięko B. (2001b): *Jak utworzyć sieć ośrodków diagnozy osiągnięć szkół podstawowych i średnich*. W: K. Wenta (red.): *Pomiar edukacyjny jako kompetencje pedagogiczne*. Wyd. Naukowe Uniwersytetu Szczecińskiego, Szczecin.
- Niemięko B. (2001c): *Ku czemu zmierzają egzaminy szkolne?* W: B. Niemięko, W. Malecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław.
- Niemięko B. (2002): *Egzaminy pożądane i niechciane*. W: B. Niemięko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnątrzszkolne i zewnętrzne a jakość pracy szkoły*. Katowice.
- Niemięko B. (2003): *Jaki pomiar dydaktyczny jest nam potrzebny?* W: B. Niemięko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.
- Niemięko B. (2004): *Założone i uzyskane, sprawdzające i różnicujące znaczenie treściowe egzaminacyjnej skali pomiarowej*. W: B. Niemięko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Niemięko B. (2006): *Oszustwo egzaminacyjne*. W: B. Niemięko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.
- Niemięko B. (2008): *Czy egzaminy zakłócają pracę nauczyciela?* W: B. Niemięko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Wyd. GRUPA TOMAMI, Kraków.
- Niemięko B. (2009): *Diagnostyka edukacyjna. Podręcznik akademicki*. Wyd. Naukowe PWN, Warszawa.
- Niemięko B. (2012): *Jawne i ukryte błędy pomiaru dydaktycznego*. W: B. Niemięko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.
- Niemięko B. (2013): *Diagnostyka edukacyjna duża i mała*. W: B. Niemięko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Wyd. GRUPA TOMAMI, Kraków.
- Niemięko B., Walukiewicz S. (2011): *Jak mierzyć kapitał ludzki? Ewaluacyjne perspektywy operacjonalizacji pojęć ekonomicznych*. W: B. Niemięko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Wyd. GRUPA TOMAMI, Kraków.
- Nowak S. (2008): *Trafność egzaminu z języków obcych w gimnazjum*. W: B. Niemięko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Wyd. GRUPA TOMAMI, Kraków.
- Pokropek A. (2011): *Matura z języka polskiego. Wybrane problemy psychometryczne*. W: B. Niemięko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Wyd. GRUPA TOMAMI, Kraków.
- Pokropek A. (2013): *Porównywalność w badaniach międzynarodowych. Przykład wskaźnika motywacji do nauki w badaniu PISA 2006*. W: B. Niemięko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Wyd. GRUPA TOMAMI, Kraków.
- Poręba-Konopczyńska A. (2001): *Moje refleksje nad ocenianiem na podstawie układu rzetelności i trafności wyników sprawdzania zaproponowanego przez Petera T. Knighta*. W: B. Niemięko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. PANDIT, Kraków.
- Poręba-Konopczyńska A. (2004): *Miejsce teorii pomiaru dydaktycznego w naukach społecznych na podstawie przemysłu Kena Wilbera*. W: B. Niemięko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Poręba-Konopczyńska A. (2005): *Testy zespołowe z wyposażeniem jako propozycja humanizacji oceniania*. W: M. Groenwald i in. (red.): *Diagnostyka edukacyjna. Niemiękowskie spotkania i inspiracje*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Skorupiński P. M. (2002): *Trafność interpretacji wyników egzaminu gimnazjalnego w rekrutacji kandydatów do publicznej szkoły ponadgimnazjalnej*. W: B. Niemierko, J. Brzdań (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnątrzszkolne i zewnętrzne a jakość pracy szkoły*. Regionalny Ośrodek Metodyczno-Edukacyjny „Metis”, Katowice.

Skorupiński P. M. (2003a): *Geneza pojęcia trafności teoretycznej*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Skorupiński P. M. (2003b): *Kryterialny wymiar trafności interpretacji wyników egzaminu gimnazjalnego*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Skorupiński P. M. (2004): *Krytyka standardów wymagań egzaminacyjnych ze stanowiska radykalnego konstrukttywizmu*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Skorupiński P. M. (2005a): *Korelacja wyników pomiarów*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej. Perspektywy informatyczne egzaminów szkolnych*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Skorupiński P. M. (2005b): *Walidacja konstruktów*. W: M. Groenwald i in. (red.): *Diagnostyka edukacyjna. Niemierkowskie spotkania i inspiracje*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Skorupiński P. M. (2013): *Modele trafności pomiaru*. W: M. Karwowski (red.): *Ścieżki rozwoju edukacyjnego młodzieży – szkoły pogimnazjalne. Trafność wskaźników edukacyjnej wartości dodanej dla szkół maturalnych*. Wydawnictwo IFiS PAN, Warszawa.

Smolik M. (2006a): *Czy maturę naprawdę trzeba „zdać”? Głos w kwestii amnestii*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.

Smolik M. (2006b): *Każdy każdego własną miarą. Kilka uwag o kryteriach oceniania ustnego z języka angielskiego na poziomie podstawowym. Raport wstępny z badań*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.

Smolik M. (2008): *Edukacyjny i pomiarowy „Miś”, czyli o nowym egzaminie gimnazjalnym z języka obcego*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela*. Wyd. GRUPA TOMAMI, Kraków.

Sobczak M. (1999): *W poszukiwaniu trafności standardów egzaminacyjnych*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. ODESiIP, Legnica.

Sokulski R. (2001): *O potrzebie obiegu informacji w procesie tworzenia egzaminu dorosłego na przykładzie egzaminu maturalnego z języka polskiego*. W: B. Niemierko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. PANDIT, Kraków.

Szmigel M.K. (2004): *Treściowe znaczenie wyniku egzaminacyjnego ucznia i szkoły na przykładzie sprawdzianu w 2004 roku*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Szyling G. (2003): *„Być może będzie w przyszłości ozdobą polskiej piłki nożnej?”. Wybrane aspekty trafności oceny końcowej osiągnięć ucznia*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Szyling G. (2005a): *Między ocenianiem osiągnięć uczniów a ekonomią*. W: M. Groenwald i in. (red.): *Diagnostyka edukacyjna. Niemierkowskie spotkania i inspiracje*. Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Szyling G. (2005b): *O prawie do odrębności, czyli o niektórych aspektach nauczycielskiego oceniania osiągnięć uczniów*. W: B. Niemierko, M.K. Szmigel (red.): *Jak praktycznie wykorzystać pomiar dydaktyczny w oświacie?* Wyd. Rożak, Warszawa.

Szyling G. (2011): *Nauczycielskie praktyki oceniania poza standardami*. OW „Impuls”, Kraków.

Szyling G. (2013): *O „nadwyżce wiedzy” w diagnozowaniu klasy szkolnej*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.

Verhelst N. (2009): *Probabilistyczna teoria wyniku zadania*, tłum. M. Daszkiewicz. W: H. Szaleniec (red.): *Teoria wyniku zadania IRT. Zastosowania w polskim systemie egzaminów zewnętrznych*. CKE, Warszawa.

Walczak W. (2002): *Relacje między ocenianiem wewnątrzszkolnym a zewnętrznym. Doniesienie z pierwszego etapu badań dotyczących trafności pomiaru dydaktycznego w VI kl. szkoły podstawowej*. W: B. Niemierko, J. Brzdań (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnątrzszkolne i zewnętrzne a jakość pracy szkoły*. Regionalny Ośrodek Metodyczno-Edukacyjny „Metis”, Katowice.

- Wiśniewska A. (2013): *Trafności i rzetelności egzaminów zewnętrznych w grupie uczniów z uszkodzeniem narządu słuchu na przykładzie egzaminu gimnazjalnego*. W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. Wyd. GRUPA TOMAMI, Kraków
- Żuk B. (2001): *O potrzebie i sposobach badania trafności testów stosowanych w egzaminach zewnętrznych*. W: B. Niemierko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. PANDIT, Kraków.
- Żuk B. (2003): *Miejsce hermeneutyki w procesie budowania trafności egzaminów powszednich i doniosłych*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Recepcja koncepcji trafności w diagnostyce edukacyjnej

Trafność jest podstawowym konstruktem w teorii pomiaru edukacyjnego. Recepcji tej teorii towarzyszyło zainteresowanie problematyką trafności, wielokrotnie artykułowane podczas dziewiętnastu kolejnych konferencji z cyklu Diagnostyka Edukacyjna, które odbyły się w latach 1997–2013. Głosy te w niniejszym tekście zostaną przypomniane i powiązane z pytaniem o stopień zaawansowania tytułowej recepcji. W kontekście przywołane będą historyczne i aktualnie rozwijane koncepcje trafności pomiaru.

Słowa kluczowe: trafność, pomiar edukacyjny, diagnostyka edukacyjna

Reception of the validity concept in educational diagnostics

Validity is a fundamental construct in educational measurement theory. The reception of the theory has been accompanied by concerns about validity issues, articulated repeatedly during the series of nineteen Educational Diagnostics conferences, held between 1997 and 2013. In this paper those voices are echoed and linked to a question of quality of the reception. In the context, historical and modern concepts of measurement validity will be highlighted.

Keywords: validity, educational measurement, educational diagnostics

Grażyna Szyling
Uniwersytet Gdański

TEORIA I PRAKTYKA OCENIANIA SZKOLNEGO W poszukiwaniu pedagogicznych przesłanek hybrydyczności oceny szkolnej

Ocenianie szkolne jest tworem dalekim od jednoznaczności i w gruncie rzeczy niezwykłym, choć najpowszechniej traktuje się je jak kategorię dydaktyczną i wiąże z celami kształcenia, podkreślając przy tym jego rolę w procesie dynamizowania uczniowskiego uczenia się. Tę „oczywistą” definicję oceniania problematyzuje przede wszystkim rozpatrywanie jej z perspektywy współwystępujących współcześnie odmiennych koncepcji nauczania-uczenia się. Odwołują się one do często opozycyjnych założeń teoretycznych, prowadząc do różnego odczytania pojęć kluczowych dla edukacyjnego myślenia i działania (Klus-Stańska, 2010). Nie ma też wśród pedagogów zgodności w sprawie bliskiego pokrewieństwa oceniania z ewaluacją (Niemierko, 2013, s. 10–11) bądź diagnozowaniem (Kasáčová, Cabanová, 2011, s. 12), a przy tym obie subdyscypliny zdają się od niego w jakimś zakresie odcinać (Jaskuła, 2013). Jest ponadto ocenianie szkolne również sądem wartościującym i praktyką społeczną, co nie tylko wpisuje je w niestabilne relacje interpersonalne i rodzi wiele problemów etycznych, ale także uzależnia od dynamicznych zmian zachodzących w społeczeństwie i jego oczekiwaniach wobec edukacji (Szyling, 2013).

Zarysowane przesłanki uzasadniają przyjęcie założenia, istotnego dla podjętych przeze mnie rozważań, że **na teorię i praktykę oceniania szkolnego składa się wiele heterogenicznych elementów, czyniąc ją tworem hybrydalnym** w rozumieniu tego pojęcia zaproponowanym przez Bruno Latoura (2011). W najprostszym ujęciu hybrydy są efektem powiązania ze sobą różnorodnych elementów, często sprzecznych, ale wzajemnie od siebie zależnych. Mogą łączyć w jedną sieć ludzi, twierdzenia naukowe, dyskursy i instytucje społeczne, lecz także przestrzenie, dokumenty, procedury i rozwiązania technologiczne, przyznając im takie same prawa w oddziaływaniu na społeczną praktykę i na siebie wzajemnie. Tak silnie eksponowana heterogeniczność nie wyklucza jednak możliwości całościowego ujęcia zagadnienia, jeśli obejmie ono również związane z nim kontrowersje. Te bowiem – jak sądzę – decydują o specyfice analizowanego obszaru i dynamice jego rozwoju. Weryfikacja postawionej tezy wymaga jednak odszukania i zrekonstruowania źródeł, konstytuujących obecny stan teorii i praktyki oceniania szkolnego. Zagadnienie to stanowić będzie oś niniejszych rozważań.

Założenia metodologiczne

Przyjęta perspektywa oglądu badawczego sprawia, że najbardziej stosowną do kompleksowej analizy zjawiska oceniania szkolnego jest stworzona przez B. Latoura (2010) **teoria aktora-sieci**, zwana w skrócie **ANT**, która dostarcza „języka do opisu zmiennej, heterogenicznej rzeczywistości” (Bińczyk, 2007, s. 38). Bada ona ustanawianie związków i relacji, ale także mechanizmy stabilizowania powiązań między elementami, które działają jak pewna całość, a przy tym pyta o ich genezę.

Analizą ANT w obszarze diagnozowania edukacyjnego posłużyła się już wcześniej Maria Groenwald (2012). Usytuowała ona cykliczne konferencje diagnostyki edukacyjnej „w aurze dzielenia się wiedzą i diagnostycznym doświadczeniem wynoszonym z filtrowania w górę i w dół” (tamże, s. 250), czyli określiła ich rolę w budowaniu sieci translacji. Konstatacja ta stwarza i zarazem uzasadnia możliwość potraktowania tekstów o szkolnym ocenianiu, które pojawiały się podczas tych wieloletnich ogólnopolskich spotkań diagnostów, jako obszernych zasobów powiązanych ze sobą z różną siłą i kontrolowanych przez różnych aktorów.

Aktor w rozumieniu ANT jest **bytem działającym**: człowiekiem, przedmiotem, urządzeniem, pomieszczeniem, dokumentem, który – wyjaśnia Krzysztof Abriszewski (2010a, s. 21) – „posiada swoją tożsamość jako rezultat relacji, w jakich pozostaje z innymi aktorami. (...). Tożsamość aktora zmienia się wraz z tymi aktorami, z którymi zostaje związany. Kluczowe stają się różnice między bytami (aktorami). Zmiana w owych różnicach powoduje zmianę w kształcie świata”.

Przytoczona definicja tyleż wyjaśnia znaczenie jednego z ważniejszych pojęć słownika ANT, co określa zasadę doboru publikacji konferencyjnych przywoływanych w niniejszym tekście. Sięgnę przede wszystkim do tych spośród nich, w których translacje „nie wyblakły”, „nie stały się przemieszczeniami bez przekształceń” (Latour, 2010, s. 187). W podjętych analizach istotne staje się bowiem **śledzenie różnic i kontrowersji**, dynamizujących postrzeganie oceniania szkolnego i działanie w jego obszarze.

Metodologicznym zaleceniem dla tak określonego procesu badawczego, które wprost sformułował B. Latour (2010, s. 330), jest wielokrotnie powtarzany w jego tekstach slogan: „»podążajcie za aktorami«; tak jest, należy podążać za nimi wtedy, kiedy *mnożą* byty, oraz wtedy, gdy *zmniejszają* ich *ilość*”. Zakres tego **podążania za aktorami** w niniejszych analizach jest jednak ograniczony podwójnie:

- a) głównym problemem rozważań, którego zbadanie wymaga wybrania kilku tylko aktorów działających w rozległym obszarze oceniania szkolnego, w określonym czasie, wyznaczonym przede wszystkim – choć nie jedynie – pierwszymi kilkoma Konferencjami Diagnostyki Edukacyjnej, które odbyły się przed ekspansją problematyki egzaminów zewnętrznych (lata 1994–2000);
- b) szczególnym charakterem próby badawczej, na którą składają się wyłącznie teksty (w znacznej części związane z konferencjami diagnostycznymi).

Uwarunkowania te pozwalają jedynie na podjęcie próby ustalenia pierwszych wiązań i otwarcie kolejnych obszarów sieci odzwierciedlającej kompleksowość analizowanego zjawiska. Istotne przy tym jest przyjęcie założenia, że w ANT **teksty** – nazywane **reprezentacjami** – **traktowane są jak pewien rodzaj translacji**. Gromadzą one bowiem inne reprezentacje (na przykład teksty i wyniki badań z różnych dziedzin), wiążą je w sieć stabilnych relacji i silnych wiązań, by nie dopuścić do swobodnej interpretacji przedstawionych w nich tez lub tylko maksymalnie ją

utrudnić (Abriszewski, 2010a, s. 25). Teksty (reprezentacje) zyskują w ten sposób moc definiowania w określony sposób zjawisk społecznych, w tym również oceniania szkolnego, i oddziałują na praktykę edukacyjną.

Zarysowana perspektywa pozwala poddać analizie nie tyle same teksty (por. Abriszewski, 2010a, s. 34–50), ile zawarty w nich obraz zjawiska edukacyjnego. W tym celu wykorzystam język i narzędzia analityczne ANT, zwłaszcza przedstawione przez B. Latoura (2011, s. 186) dwa odmienne zespoły praktyk: „oficjalną pracę puryfikacji (...) pozostającą w owocnej sprzeczności z nieoficjalną pracą mediacji”. Umożliwiają one zbadanie, na ile szkolne ocenianie jest tworem heterogenicznym, który wymyka się się trwałym ujęciom systemowym, ale działa i tworzy **translacyjną sieć – hybrydę**, definiującą określone praktyki społeczne (Latour, 2011, s. 105–109).

Puryfikacja zmierza do podzielenia hybrydy na jednorodne elementy, które mogą zostać przypisane czystym formom zjawisk, stabilnym i biegunowo odrębnym (Latour, 2011, s. 112), jak na przykład paradygmaty lub dyskursy naukowe, albo modele dydaktyczne wyprowadzone z odmiennych przesłanek teoretycznych. Ten proces „oczyszczania” może następować poprzez wykluczanie wybranych kontrowersji związanych ze szkolnym ocenianiem, przemilczanie lub eliminowanie wyjaśnień określonego typu. Kiedy już – jak pisze B. Latour (2011, s. 105) – zostaną usunięte „miejsca wejścia i wyjścia wszystkich hybryd”, wówczas można „interpretować heterogeniczne układy jako systematyczne całości, gdzie wszystko dobrze do siebie pasuje. (...) Dzieje się tak, gdy elementy stworzą zamkniętą i rozpoznawalną grupę”. Osiągnięcie tej jedności staje się możliwe dzięki **pośrednikom**, którzy postrzegają świat społeczny (także edukacyjny) jak gotowy wyrób, dlatego przenoszą występujące w nim powiązania bezrefleksyjnie, często ich nawet nie dostrzegając i nie naruszając *status quo* (tamże, s. 187).

Translacja z kolei jest takim rodzajem relacji między dwoma elementami, który je przekształca, redefiniuje, tłumaczy, zakłóca, a tym samym tworzy nową całość wraz z nie zawsze rozpoznawymi, bo trudnymi do przewidzenia konsekwencjami (tamże, s. 64). Proces ten dokonuje się dzięki **mediatorowi**: człowiekowi, ale także przedmiotowi czy sytuacji, „wpływającemu zwrotnie na tych, z którymi wchodzi w relację” (Abriszewski, 2010, s. XXX). Mediatorzy – jak twierdzi B. Latour (2010, s. 315–316) – nie są przyczyną działania i nie powodują go, ale dają okazję, by inny działający byt zachował się jak mediator, skłaniają do współlistnienia, „gromadzą i wiążą zbiorowość” (tamże, s. 350).

Efektom zabiegów translacyjnych jest **hybrydalna sieć**, w której każdy element (ludzki i pozaludzki) „ma głos rozstrzygający, czyni różnicę i stawia opór. Niezależność czynników pozaludzkich polega każdorazowo na konkretnym ustabilizowaniu się sieci powiązań. (...) Żaden z elementów sieci (...) nie jest bezpośrednio przyczynowy, ale każdy posiada rzeczywisty głos rozstrzygający i czyni różnicę” (Bińczyk, 2007, s. 37).

Początki diagnostycznego myślenia o ocenianiu i działaniach w tym obszarze: rozpoznanie głównych aktorów i kontrowersji

Ocenianie szkolne jest wątkiem, który stale, acz z różnym natężeniem i w odmiennych ujęciach pojawiał się w rozważaniach podejmowanych podczas Konferencji Diagnostyki Edukacyjnej, a dwie z nich – w 2002 i 2010 roku – uczyniły go

bezpośrednio tematem obrad. Jednak ograniczenie się w niniejszych rozważaniach wyłącznie do tropów odnalezionych w konferencyjnych tomach czyniłoby dokonany ogląd niepełnym, a przede wszystkim prowadziłyby – o czym mówi B. Latour (2011, s. 105) – do „usunięcia wejść” heterogenicznej sieci, czyli rezygnacji z głębszego poznania fragmentu edukacyjnej rzeczywistości, który – zdaniem Krzysztofa Kona-rzewskiego (1998, s. 56) – „ma kilkusetletnią tradycję i jest głęboko zakorzeniony w społecznym świecie”. Nie podejmę się sięgania w odległą o kilkaset lat przeszłość choćby tylko ze względu na objętość niniejszego tekstu, ograniczając się do odszuka-nia głównych kontrowersji dostrzeganych w ocenianiu szkolnym w latach poprze-dzających cykliczne spotkania polskich diagnostów.

Teorie pedagogiczne

Uporządkowania teoretycznych przesłanek oceniania szkolnego podjęła się u progu lat sześćdziesiątych minionego wieku Ida Altszuler (1960). Przyjmując zatem, że do-konane przez nią analizy, będące rodzajem naukowej translacji, reprezentują określo-ny stan wiedzy pedagogicznej na temat oceniania edukacyjnego.

Definiując pojęcie oceny szkolnej, autorka uwzględnia związane z nim spory to-czone latami między zwolennikami teorii pedagogicznych. Ogniskują się one wokół wymienionych poniżej zagadnień (Altszuler, 1960, s. 98–99), które dzięki wzajem-nym powiązaniom między poszczególnymi elementami mogą aspirować do roli ak-torów, stymulujących zmianę w kształcie edukacyjnej praktyki:

- Przedmiot oceny może być rozumiany wąsko i obejmować wyłącznie osiągnięcia poznawcze ucznia, a w wersji poszerzonej – także stałe cechy danej jednostki, ta-kie jak inteligencja i zdolności, bądź w kolejnej, bardziej umiarkowanej: logiczne myślenie, samodzielność i krytycyzm.
- Subiektywizm oceny jest uznawany za niemożliwą do usunięcia właściwość nauczycielskiego oceniania albo postrzegany jako możliwy do zlikwidowania lub tylko zdecydowanego ograniczenia poprzez zastosowanie jednolitych kryteriów i norm ocen.
- Jawność oceniania bywa traktowana jak warunek spełniania przez ocenę funkcji informacyjnej lub zagrożenie dla funkcji wychowawczej i społecznej.
- Forma oceny prowadzi do sporu o stopień szkolny, traktowany przez jego zwolen-ników jako gwarant ścisłości i warunek porównywania wyników, a przez przeciwn-ików odrzucany ze względu na obciążenie subiektywizmem i przeciwstawiany najczęściej komentarzowi pisemnemu lub ustnemu.

Źródło tych nierozstrzygniętych wątpliwości I. Altszuler lokuje nie tyle w histo-rycznych zmianach w myśleniu o edukacji, ile w **opozycyjnym charakterze teorii pedagogicznych**, stanowiących podstawę bądź systemu klasowo-lekcyjnego, bądź edukacji spod znaku Nowego Wychowania. Tym samym – mówiąc językiem ANT – dokonuje wstępnej puryfikacji hybrydalnej definicji oceny szkolnej, czyli dzieli ją na jednorodne elementy, zawieszając tworzącą ją sieć między przeciwstawnymi bie-gunami dyskursów dydaktycznych. Za najbardziej stabilny punkt, spajający często sprzeczne teorie, uznaje natomiast hybrydyczne w gruncie rzeczy założenie, że **ocena szkolna jest niezbędną i nieodłączną częścią procesu dydaktyczno-wy-chowawczego w każdym systemie szkolnym, choć jej przydatność bywa różnie rozumiana i uzasadniana** (Altszuler, 1960, s. 97).

Sprawdzenie trwałości ustanowionych kontrowersji i powiązań możliwe będzie dopiero w następnych częściach analizy, śledzącej kolejnych aktorów. Niemniej zaproponowane przez I. Altszuler dydaktyczne osadzenie oceny szkolnej z jednej strony stabilizuje rozważania na jej temat, a z drugiej – otwiera drogę do uczynienia danej teorii pedagogicznej mediatorem w procesie translacji, stwarzającym warunki określonego wyjaśniania i przekształcania rzeczywistości. Postrzegana z tej perspektywy nauka traktowana jest „nie jako dostarczycielka bezdyskusyjnych danych o faktach, lecz jako obszar niegasnących sporów, kontrowersji i negocjacji. (...) Oznacza to, że dostarczona przez nauki układanka nie jest zamknięta na wejściu i podlega wielorakim dyskusjom różnych zaangażowanych stron” (Abriszewski, 2010, s. XXXIII).

Praktyka edukacyjna

Otwarta przestrzeń możliwości, rozpinająca się między edukacyjnymi koncepcjami zorientowanymi behawioralnie i konstruktywistycznie, zderza się jednak nieuchronnie z przestrzenią regulacji i kolejnych sprzeczności, w której wytwarzane są praktyki szkolnego oceniania (Szyling, 2011). Każde zaś z możliwych do ujawnienia wiązań tej heterogenicznej sieci pozostawia „w przeważającej mierze puste te obszary, które *nie* zostały połączone” (Latour, 2010, s. 189), czyli pozwala odczytywać się również przez pryzmat przemilczeń i opuszczeń.

Zastany przez uczestników Konferencji Diagnostyki Edukacyjnej stan szkolnego oceniania, reprezentowany w ich wystąpieniach (Hamryszak, 1994; Poręba-Konopczyńska, 1998; Wszolek, 1999), nie pozostawia wątpliwości co do swojej heterogeniczności, której przejawem staje się **zjawisko rozbieżności lub zmienności ocen**, wyrażanych stopniem szkolnym. Jednak w omawianych ujęciach aktorem staje się nie forma oceniania, traktowana jak rodzaj przekąznika, ale nauczyciel i stanowione przez niego kryteria oceny. Postrzegając **obiekttywizację oceny** szkolnej jako jej walor, autorzy tekstów zwracają uwagę na te właściwości nauczyciela, które stoją na przeszkodzie do osiągnięcia porównywalności uczniowskich osiągnięć wyrażonych stopniem. Zaliczają do nich: (a) brak merytorycznego przygotowania nauczyciela do prowadzenia określonych zajęć, co współwystępuje z nietrafną oceną nawet wtedy, gdy ta odwołuje się do ustalonych kryteriów (Hamryszak, 1994, s. 221); (b) zróżnicowany poziom surowości osób oceniających, uwikłanych ponadto w poczucie osobistej odpowiedzialności za wyniki osiągnięte przez uczniów (Poręba-Konopczyńska, 1998, s. 269); (c) indywidualne preferencje nauczycieli w doborze kryteriów oceny, warunkowane również psychologicznymi prawidłowościami oceniania (Wszolek, 1999, s. 170).

Główne kontrowersje zdają się zatem lokować między biegunami, z których jeden jest wyznaczany przekonaniem o możliwości ograniczenia zakresu przedmiotu oceny szkolnej poprzez wykluczenie z niego aspektów pozadydaktycznych, a drugi – trudnymi do wyeliminowania w nauczycielskim ocenianiu uwarunkowaniami psychologicznymi, zwanymi z czasem coraz częściej „pułapkami szkolnego oceniania” (Niemierko, 2007, s. 312).

Teoretyczną reprezentacją tych kontrowersji jest dokonany przez Bolesława Niemierkę (1990) podział nauczycielskiego oceniania na dydaktyczne i społeczno-wychowawcze. Zaproponowana klasyfikacja może być odczytana jako przejaw

puryfikacji heterogenicznego zjawiska szkolnego oceniania. Arbitralnie odrywa ona od wiązań łączących sieć tę grupę uczniowskich osiągnięć poznawczych i praktycznych, które łatwo poddają się pomiarowi, co pozwala bezstronnie ustalić ich poziom w odniesieniu do jednego kryterium, jakim są zewnętrzne wobec szkoły wymagania programowe (tamże, s. 25). Z drugiej jednak strony B. Niemierko uwzględnia w swej klasyfikacji impulsy docierające z nauczycielskiej praktyki. Uznaje, że w edukacji występuje także inny rodzaj oceniania, który łączy w trudnych do jednoznacznego ustalenia okolicznościach i proporcjach: wymagania programowe, elementy kontekstu kształcenia oraz aspekty emocjonalno-motywacyjne.

Przedstawiony podział utrwała kontrowersje charakterystyczne dla oceniania szkolnego, na które wskazywała już I. Altszuler (1960). Inicjuje też jednak nową sieć lub rozwija dynamicznie załączek już istniejącej, której wiązania stają się bardziej widoczne dzięki tekstom konferencyjnym. Ich autorzy dokonują empirycznego oglądu szkolnej rzeczywistości z perspektywy, która zasadza się na przekonaniu, że tendencja do „oczyszczania” oceniania z czynników pozadydaktycznych jest jedyną możliwą do zaakceptowania drogą zmiany edukacyjnej. Kierunek tego wiązania zdaje się bardziej wyrazisty, gdy prześledzi się jego relacje z książką B. Niemierki: *Między oceną szkolną a dydaktyką. Bliżej dydaktyki* (1997). Pojawiła się w niej diagnoza ówczesnego stanu oceniania osiągnięć szkolnych: najbardziej zaniedbanego aspektu edukacji, wręcz obszaru porażki pedagogicznej. Jego źródłem B. Niemierko (1997, s. 11) upatrywał w postsocjalistycznym autorytaryzmie szkoły oraz silnym podporządkowaniu oceniania wąsko pojmowanym celom wychowawczym, a zwłaszcza dyscyplinowaniu uczniów. Swoistym antidotum na te problemy miało stać się upowszechnienie wśród nauczycieli teorii pomiaru dydaktycznego, prowadzące w założeniu do **unormowania i obiektywizacji** oceniania szkolnego, pozostającego w bezpośrednich relacjach z działaniami dydaktycznymi. Przykładem siły i kierunku wiązania, biegnącego od pomiaru do organizowania sytuacji dydaktycznych, są sformułowane przez B. Niemierkę (1997) zasady analizy treści programu nauczania danego przedmiotu. Odwoływały się one wprost do pomiarowych procedur budowania koncepcji wielopoziomowego testu sprawdzającego osiągnięcia poznawcze ucznia (Niemierko, 1990, s. 174–271). W ten sposób **dokonała się transformacja**, skłaniająca do współistnienia dwa odrębne byty: **nauczycielską praktykę oceniania i pomiar dydaktyczny**.

Pomiar dydaktyczny

Przeprowadzona analiza ujawniła nowego aktora, oddziałującego na myślenie o ocenianiu szkolnym i jego praktycznym wymiarze, czyli pomiar dydaktyczny, który do polskiej pedagogiki i edukacji został wprowadzony przez B. Niemierkę (1990; 1999). We wczesnych tekstach konferencyjnych pojawia się on przede wszystkim poprzez zapośredniczenia, czyli jako twór gotowy, zamknięty, pozbawiony powiązań, które go konstytuowały (np.: Smętek, Matrenicka, 1998; Zimny, 1999; Hłuszyk, 2001). Silnie utożsamiony z testem, staje się w powszechnym odbiorze uczestników konferencji diagnostycznych narzędziem pozbawionym siły sprawczej i wbudowanej weń intencjonalności. Za pożądane i w pewnym sensie oczywiste uznaje się blokowanie przez procedury pomiarowe tych działań nauczycieli, które mogłyby ograniczyć porównywalność oceny, pomijając milczeniem koszty takich działań.

Z założeń pomiaru dydaktycznego – jak ze swoistego „centrum kalkulacji” lub „centrum decyzyjnego” kontrolującego obszar oceniania edukacyjnego (Latour, 2010, s. 260) – zdaje się zatem wyłaniać początek sieci, będącej efektem mobilizowania sojuszników, czyli działających i przekształcających realia szkolne aktorów-diagnostów: nauczycieli, metodyków, naukowców. Nabywają oni nowych kompetencji w obszarze oceniania szkolnego, w wyniku czego „zostaje im nadana władza, by mówić lub działać w imieniu innego aktora lub siły” (Abriszewski, 2010, s. 24), czyli pomiaru dydaktycznego. Przejawem tego mówienia i działania stają się między innymi teksty konferencyjne zawierające doniesienia z testowych badań wyników nauczania, które dostarczają przesłanek do doskonalenia praktyki nie tylko szkolnego oceniania, ale także nauczania.

Przykładem zaś szczególnym, który ilustruje siłę wiązania biegnącego od pomiaru dydaktycznego ku edukacyjnej rzeczywistości, jest poradnikowa książka, napisana przez Elżbietę Kowalik i Krystynę Wojciechowską: *Szkolny system oceniania oparty na pomiarze dydaktycznym* (2000). Sygnalizuje ona powstanie kolejnego aktora-sieci, czyli **szkolnego systemu oceniania**, posiadającego cechy mediatora. Łączy on w sobie co najmniej trzy odmienne byty: obowiązujące powszechnie wymagania odnoszące się do oceniania określone w Rozporządzeniu MEN (DzU 1999, poz. 131), zakorzenione w społecznym odbiorze właściwości pomiaru dydaktycznego oraz lokalnie uwarunkowane cechy danej placówki edukacyjnej. Ich wzajemne przekształcenia (zwane w ANT translacjami) – zainicjowane przez kolejne „centrum decyzyjne”, czyli Ministerstwa Edukacji Narodowej – często prowadzą do redefinicji kluczowych pojęć tworzących się na styku teorii i potoczności pedagogicznej (Putkiewicz, 2000). W efekcie powstaje nowa hybrydalna całość, która stabilizuje sieć oceniania szkolnego oraz skłania do współlistnienia lokalną zbiorowość: nauczycieli, uczniów i ich rodziców oraz dyrekcję danej szkoły.

Śledzenie tych translacji, które zachodzą także między tym, co globalne i lokalne, nie tylko przekracza ramy niniejszego tekstu, ale też jest nieobecne we wczesnych rozważaniach diagnostycznych. Zaledwie kilka z tekstów prezentuje stosunkowo nieodległe w czasie aspekty praktycznego funkcjonowania szkolnych systemów oceniania, ujawniając między innymi: specyficzny charakter wykorzystania pojęć pomiarowych w szkolnych dokumentach regulujących ocenianie (Szafraniec i in., 2003, s. 243–246), a także opór, jaki wdrażaniu nowych procedur stawia zmienność ocen wyrażanych stopniem szkolnym (Jagodzińska, 2004, s. 452–456).

Ideologie i standardy edukacyjne

Z prowadzonych rozważań wyłania się także, istotna dla początków diagnostyki edukacyjnej, **kontrowersja między rozumieniem pomiaru dydaktycznego jako działającego aktora albo jako formy**. W ujęciu ANT (Latour, 2010, s. 325) cechą tego drugiego z bytów społecznych jest przenoszenie bez deformacji jednego obszaru – czyli sprawdzanych umiejętności uczniów, do drugiego – czyli oceny szkolnej. Za jego teoretyczną reprezentację można uznać pojęcie znaczenia treściowego oceny szkolnej, zdefiniowane przez B. Niemierkę (2002, s. 223) jako „takie powiązanie stopni szkolnych z określoną treścią kształcenia, że o uczniu uzyskującym każdy stopień można zasadnie wnioskować, jakie elementy tej treści opanował, a jakich nie opanował”. Zapewnienie wysokiej jakości tej właściwości oceny jest możliwe dzięki

wykorzystaniu procedur pomiarowych i ma być dla nauczyciela najważniejszym zadaniem. Natomiast komentarz do informacji o wyniku – twierdzi autor w jednym z konferencyjnych tekstów (Niemierko, 1999a, s. 16–17) – jest w szkolnym ocenianiu także potrzebny, ale wpisuje się w obszar diagnozowania i może posłużyć „zarówno do wzbogacenia oceny szkolnej, jak i do osłabienia jej znaczenia treściowego i podważenia sensu”.

Dokonane przez B. Niemierkę rozróżnienie jest przejawem puryfikacji, czyli dążenia do nadania ocenianiu szkolnemu przejrzystości i jednoznaczności na przekór jego praktycznemu wymiarowi, z drugiej jednak eksponuje obiektywizującą rolę wymagań programowych jako stałego i bezdyskusyjnego punktu odniesienia wyniku osiągniętego przez ucznia. Dzięki temu ocena ustalona na drodze pomiaru dydaktycznego staje się tylko formą, która niesie ze sobą zobiektywizowaną informację o poziomie spełnienia przez ucznia standardów edukacyjnych.

Z kolei Krzysztof Kruszewski (1998) zdecydowanie przychylił się do nadania pomiarowi statusu aktora-sieci, uznając „pedagogikę pomiarową [za] jedną z wielu religii pedagogicznych” (tamże, s. 75). Istotę translacji między pomiarem dydaktycznym a rzeczywistością edukacyjną określa jako dążenie tego pierwszego do „takiego uporządkowania świata, by bóstwo mogło nad nim zapanować” (tamże, s. 77). Autor nie ma też wątpliwości, że osiągnięcie tak sformułowanego celu wymaga upraszczania i redukcji złożoności realiów społecznych, czego jednym z przejawów jest „odbieranie prawomocności temu, co nie istnieje, bo jest niemierzalne” (tamże, s. 78).

Z odczytaniem pomiaru dydaktycznego jako aktora-sieci, oddziałującego na rzeczywistość edukacyjną, zgadza się też K. Konarzewski (1998). Twierdzi on, że zorientowana technologicznie koncepcja pedagogiczna, wykorzystująca założenia pomiaru, rości sobie prawo do „sterowania życiem społecznym”, czyli codzienną pracą szkoły (tamże, s. 56). Natomiast w realiach edukacyjnych, argumentuje autor (tamże, s. 72–73), ogromną rolę odgrywają „nauczycielskie intuicje”, które stoją za wartościowaniem osiągnięć uczniów, a są osadzone w odmiennych, często sprzecznych systemach przekonań. Ponieważ to określona ideologia implikuje rozumienie takich między innymi pojęć, jak osobista odpowiedzialność nauczyciela oraz rozwój ucznia, nie można oczekiwać, by stopnie szkolne były ze sobą porównywalne i niesły wyłącznie informację o poziomie realizacji wymagań programowych.

Konstatacja ta wprowadza w obszar teorii szkolnego oceniania kolejne kontrowersje, **wikłając go w ideologie oświatowe**, rozumiane jako „zespół ogólnych przekonań o misji szkoły, roli nauczyciela i naturze ucznia, do których odwołują się nauczyciele (lub inne oświatowe »grupy interesu«) w legitymizującym dyskursie” (tamże, s. 58).

Podobnie jak K. Konarzewski (1998), źródła zmienności szkolnego oceniania i interpretuje Aleksander Nalaskowski (2004, s. 264)¹, który nazywa nauczycielskie wymagania edukacyjne „standardem na miękko”, biorąc pod uwagę ich niedookreśloność i nierzadko intuicyjność oraz uzależnienie od indywidualnych ideologii. W jego tekście pojawiają się jednak także inne argumenty znaczące dla rozpoznania przesłanek diagnostycznego myślenia o ocenianiu jak o obszarze hybrydalnym. Autor wskazuje wprost na **translacyjny charakter standardów edukacyjnych**, określając dwa źródła zniekształceń zachodzących w ich społecznym funkcjonowaniu (tamże, s. 265–267). Pierwszym z nich jest rodzaj ideologii, zdefiniowany w omawianym tek-

ście zgodnie z klasyfikacją Laurence'a Kohlberga i Rochelle Mayer (1993). Spojrzenie na standard edukacyjny przez pryzmat określonej ideologii oświatowej wpływa na radykalnie czasem odmienne definiowanie wykształcenia czy ostatecznego celu edukacji, jakim – w mniemaniu A. Nalaskowskiego (2004, s. 266) – jest „człowiek wyedukowany”. W konsekwencji innego znaczenia nabierają podstawowe pojęcia związane z ocenianiem oraz jego umiejscowienie w procesie kształcenia, a zwłaszcza kategorie takie, jak: wiedza, umiejętności, osiągnięcia, wymagania, rozwój czy samoocena. Przemilczanie konotacji ideologicznych prowadzi do odmiennych odczytań tych samych sytuacji edukacyjnych, stając się źródłem nieporozumień czy wręcz chaosu poznawczego.

Drugim źródłem redefiniowania standardów, wskazanym przez A. Nalaskowskiego (2004, s. 263), jest hierarchiczna („pionowa”) relacja między standardem jednostkowym, społecznym i instytucjonalnym. Autor wskazuje przy tym, że jednostkowy standard w przypadku oceniania szkolnego jest definiowany dla każdego ucznia przez każdego uczącego go nauczyciela, którego decyzje (uwikłane ideologicznie i osobowościowo) są nie tylko rozstrzygające, ale i brzemiennie w skutki, niekoniernie korzystne dla ocenianego (tamże, s. 264).

Przedstawione rozważania, dotyczące standardów edukacyjnych (głównie nauczycielskich) i ich relacji z ideologiami, w pewnym zakresie wpisują się w koncepcję ANT, w której standard jest rodzajem hybrydalnej sieci, silnie wiążącej „w dół” globalne regulacje z ich rozlicznymi lokalnymi zastosowaniami. Dzieje się jednak tak tylko wtedy, gdy „nigdzie w przekazie nie pojawi się żadne przerwianie, uszkodzenie, luka ani jakakolwiek niepewność” (Latour, 2010, s. 332), co w sytuacji szkolnego oceniania – jak wskazują dotychczasowe analizy – jest warunkiem niemożliwym do spełnienia. W rzeczywistości społecznej częściej spotkać można „quasi-standardy”, które też łączą zbiorowości, takie jak na przykład klasa szkolna czy szkoła, koordynując działanie tworzących je podmiotów. W takich sytuacjach trzeba jednak zawsze spytać – twierdzi B. Latour (2010, s. 335): „Czy dana kategoria nas uprzedmiotawia czy *upodmiotawia*?”

Pytanie to można uznać za kluczowe dla obszaru szkolnego oceniania, ale jego personalny wymiar realizuje się w działaniu kolejnych aktorów, czyli **nauczyciela i ucznia**. Pierwszy z nich jest jednak we wczesnych wystąpieniach konferencyjnych obecny w szczególny sposób: jako ten element sieci translacji między pomiarem edukacyjnym a praktyką szkolną, który z różnych powodów **stawia opór zmianie prowadzącej do obiektywizacji** oceniania. Uczeń zaś występuje tylko jako odległy punkt odniesienia, choć troska o jakość jego kształcenia jest deklarowana mniej lub bardziej bezpośrednio we wszystkich omalże tekstach.

Pierwsze wiązania sieci i kolejne otwarcia

Procesem podjętej w tym artykule analizy kierowało „podążanie za aktorami”, czyli oddziałującymi na praktykę edukacyjną bytami, które reprezentowane są w tekstach związanych z początkami diagnostyki edukacyjnej. Wędrowka ta okazała się inspirującym doświadczeniem naukowym, ponieważ – mówiąc językiem B. Latoura (2010, s. 364) – wymagała ode mnie, jako badaczki, „otwarcia i zdefragmentowania” zasobów dobrze znanych i oswojonych, by odnalezione tropy starać się spleść na nowo.

To splatanie jest kolejną translacją, swoistym przekładem, który pomoże znaleźć uzasadnienia dla postawionej na wstępie tezy, że szkolne ocenianie jest tworem heterogenicznym i dynamicznym, który wymyka się się trwałym regulacjom systemowym, ale działa i tworzy translacyjną sieć – hybrydę.

Zawiązywanie sieci nie wymaga przestrzegania kolejności, w jakiej prezentowani byli aktorzy, dlatego zacznę od **praktyki nauczycielskiego oceniania**, nazwanej przez B. Niemierkę (2001, s. 16) „szkolnym matecznikiem”. Widziana przez pryzmat konferencyjnych tekstów jawi się ona jako obszar niekonsekwencji, subiektywizmu, intuicyjnych decyzji oraz rozlicznych uwikłań: od prawidłowości psychologicznych począwszy, a skończywszy na stabilności uwarunkowanych ideologicznie przekonań nauczycieli. Szkolna codzienność oceniania, **sprowadzona głównie do działań nauczyciela** – które zdaje się, że są niezależne od jakichkolwiek zewnętrznych regulacji – okazuje się nie tylko hybrydyczna, ale też trudna do odczytania przez osoby znajdujące się na zewnątrz niej. „Szkoła skrywa swoje tajemnice przed społeczeństwem, władzami i instytucjami wyższego szczebla, a postęp uzyskiwany przez kodyfikację wewnątrzszkolnych systemów oceniania nie może być ani raptowny, ani równomierny” – pisał B. Niemierko (2001, s. 18), wskazując tym samym, że zakłócenia występujące w wiązaniach sieci wiodą nie tylko od praktyki do teorii edukacyjnej (i z powrotem), ale także od szkoły do otaczającego ją świata społecznego.

To drugie z połączeń miały wzmacniać i stabilizować wewnątrzszkolne systemy oceniania. Okazały się one jednak słabo rozpoznany przez diagnostykę obszarem nie tylko w zakresie efektów scalania heterogenicznych założeń, leżących u ich podstaw, ale także konsekwencji ich normalizującego wpływu na codzienne nauczycielskie i uczniowskie działania. Można z pewnym prawdopodobieństwem lokować **źródła tych opuszczeń** w pojawieniu się kolejnego aktora-sieci, a właściwie – stwierdza M. Groenwald (2012) – ważnego „centrum decyzyjnego”, czyli systemu egzaminów zewnętrznych, który okazał się czynnikiem dużo bardziej stabilizującym niż lokalne regulacje. Postawiona teza wymagałaby empirycznej weryfikacji, ale pewnego jej uzasadnienia dostarczają wyniki rekonosansu badawczego przeprowadzonego przez Urszulę Opłocką (2008). Jednak zarówno jej badania, jak i nieco późniejsze eksploatacje Piotra W. Parczewskiego (2012) pojawiają się na obrzeżach głównych nurtów zainteresowań diagnostyki edukacyjnej i dotyczą nauczycielskich (a nie wewnątrzszkolnych) systemów oceniania. Z perspektywy niniejszych rozważań ważne natomiast wydaje się, że z badań, w których przyjęto odmienne koncepcje teoretyczne, wyłania się **obraz nauczycielskiego oceniania jako tworzywa niejednorodnego**, powiązanego rozlicznymi interakcjami, czyli w gruncie rzeczy tworzącego hybrydyczną sieć o lokalnym zasięgu, zależną w różnym zakresie od centralnych regulacji.

Zanim jednak translacje egzaminacyjne „przebiegające jako filtrowanie w dół” (Groenwald, 2012, s. 250) zaczęły redefiniować szkolne praktyki oceniania, rolę tę odgrywał lub chciał odgrywać pomiar dydaktyczny, wspierany i popularyzowany przez licznych sojuszników. Jego teoretyczne założenia działać miały „oczyszczająco” na zastaną nieprzejrzystość i stanowić wsparcie dla zmiany oblicza nauczycielskiego oceniania poprzez ograniczenie jego subiektywizmu. Pociągało to za sobą zawężenie przedmiotu oceny do osiągnięć poznawczych ucznia, sprawdzanych zgodnie z procedurami pomiarowymi i wartościowanymi według jednego kryterium, jakim był poziom zrealizowania wymagań programowych. Zarazem jednak wyniki pomiaru traktowano jak ważny element doskonalenia procesu nauczania, a ocenianie sys-

tematycznie i coraz wyraźniej zaczęło rościć sobie prawo do sterowania dydaktyką. Zarówno wówczas, jak i w kolejnych latach, ten szczególnie translacyjny splot interpretowany był jako zagrożenie dla praktyki edukacyjnej tylko przez nielicznych uczestników konferencji diagnostycznych (Kruszewski, 1998, 2005; Szkudlarek, 2005; Klus-Stańska, 2010a; Kwieciński, 2010), raczej ich gości niż stałych uczestników. Ci drudzy przez pewien czas zdawali się nawet nie dostrzegać **kontrowersji związanych z pozorną neutralnością aksjologiczną pomiaru**, przyjmując jako pewnik jego niezależność ideologiczną, przeciwstawianą różnym uwikłaniom nauczycielskich ocen.

Izolowanie praktycznych zastosowań pomiaru dydaktycznego od świata wartości okazało się jednak niemożliwe. Szansa ponownego odczytania zachodzących między nimi translacji, przebiegających w obu kierunkach, pojawiła się wraz z unitarną koncepcją trafności Samuela Messicka (2005). W zakres tematyczny konferencji diagnostycznych wprowadził ją i konsekwentnie upowszechniał B. Niemierko (np. 2002a, 2003). Należy przy tym zaznaczyć, że na ówczesnym etapie rozwoju diagnostyki edukacyjnej pomiar dydaktyczny nabierał nowych znaczeń pod wpływem egzaminów zewnętrznych, traktowanych jako „inny rodzaj oceniania” (Niemierko, 2001, 2002), co z kolei wymagało przedefiniowania wielu kategorii związanych z nauczycielskim ocenianiem. Teoretyczną reprezentacją tych ostatnich jest na przykład relatywizacja jasnego, psychometrycznego podziału występującego między rzetelnością i trafnością oceniania szkolnego (Poręba-Konopczyńska, 2001).

Teoria S. Messicka otworzyła też możliwości ujawnienia się nowego oblicza kontrowersji powiązanych z nauczycielskim ocenianiem, które dotychczas lokowano między biegunami: subiektywizm – obiektywizm. Odmienność tych kontrowersji wynikała z innego niż dotychczas kierunku wiązań translacyjnej sieci, której końce zaczepiane były na przykład w świecie wartości etycznych (Groenwald, 2010) lub obszarze przewidywanych przez nauczycieli konsekwencji, jakie ocena może nieść dla konkretnego ucznia (Szyling, 2003). Zmiana perspektywy oglądu szkolnych realiów sprawiła, że w wielu konferencyjnych tekstach pojawiły się przesłanki empiryczne **czyniące nauczyciela rzeczywistym aktorem**, a nie elementem zakłócającym translację między pomiarem dydaktycznym a praktyką edukacyjną. Prowadziło to do odnajdowania licznych powiązań występujących wewnątrz szkolnego oceniania, których nie sposób przywołać.

Jednym z najlepiej rozpoznanych zjawisk okazał się drugi układ wymagań, czyli sformułowany przez B. Niemierkę (2002a) konstrukt teoretyczny, porządkujący nieprzejrzystość nauczycielskich praktyk oceniania, który poddany został empirycznej weryfikacji przez Grażynę Szyling (2002, 2003, 2011). Przeprowadzone przez nią badania ujawniły znaczną niezależność szkolnych ocen zarówno od określonych zewnętrznie wymagań, jak i od wyników egzaminów zewnętrznych, jakie uzyskują uczniowie szkoły. Dane te nie tyle redefiniują, ile rozwijają tezę A. Nalaskowskiego (2004) o translacyjnym charakterze instytucjonalnych standardów edukacyjnych i jego ideologicznych źródłach. W niniejszych rozważaniach istotne jest empiryczne ustalenie (Szyling, 2011, s. 369), że ocenianie szkolne jest obszarem splecionym ze sprzecznych oczekiwań podmiotów zaangażowanych w proces kształcenia i pozostających ze sobą w bezpośrednich relacjach. Nauczyciele próbują te oczekiwania zaspokoić wyłącznie lub głównie za pomocą oceny szkolnej, której funkcja informacyjna została poważnie ograniczona przez podporządkowanie jej doraźnym celom, zwłaszcza behawioralnie rozumianej motywacji do uczenia się.

Wyniki powyższych analiz pozwalają na wyciągnięcie wniosku, że nauczycielskie ocenianie – mówiąc językiem ANT – staje się lokalną hybrydalną siecią, rozpiętą między dążeniem nauczyciela do zachowania znaczenia treściowego oceny, odniesionego do „globalnych” wymagań programowych, a jego otwarciem na „lokalnie” zróżnicowane potrzeby i możliwości ucznia. Równocześnie jednak ta sieć jest silnie stabilizowana przez **stopień szkolny** – byt prawie nieobecny w rozważaniach konferencyjnych lub – co najwyżej – zmarginalizowany i sprowadzony do roli empirycznego wskaźnika poziomu uczniowskich osiągnięć albo nauczycielskiej bezstronności. W wypowiedziach diagnostów pojawia się najczęściej w formie ukrytej, jako informacja o wyniku nauczania uzupełniana komentarzem dydaktycznym (co można uznać za peryfrazę) lub po prostu jako ocena (co jest metonimią). Jego mocy sprawczej diagnostyka edukacyjna nie doceniała i nadal zdaje się nie doceniać, co skłania do postawienia hipotezy, że jest on postrzegany w kategoriach „naturalnego”, bo zsocjalizowanego elementu szkolnego oceniania, choć w rzeczywistości jest hybrydą, pozbawioną punktów zaczepienia, czyli miejsc, z których wyszedł i do których może dotrzeć w procesie translacji (Szyling, 2013a).

Można zatem przypuszczać, że uczynienie ze stopnia szkolnego **aktora-sieci** pozwoliłoby dostrzec jeszcze inny obraz oceniania, nie mniej heterogeniczny niż ten, który spleciony został w tych rozważaniach, ale mobilizujący innych aktorów i inne między nimi relacje. Przyjęcie takiej perspektywy analitycznej jest kolejnym w tym tekście otwarciem na nowe translacje, a jego zasadność może być wzmacniana przez mniej lub bardziej jawne kontrowersje, które można odnaleźć w konferencyjnych tekstach poświęconych – zyskującemu na popularności – ocenianiu kształtującemu i jego relacjom z ocenianiem sumującym (np. Jakubowicz i in., 2008; Niemierko, 2010; Szyling, 2010). Odkrywanie występujących między nimi powiązań, napięć i redukcji jest otwartym polem badań, jednym z wielu dotyczących teorii i praktyki szkolnego oceniania. O należne w nich miejsce upomina się także uczeń, traktowany jak aktor-mediator, a nawet lokalne „centrum decyzyjne”, stabilizujące zewnętrzne wobec niego oczekiwania i próbujący na nie oddziaływać na swój sposób, mimo licznych przeszkód.

Zakończenie

„Czyjeś działanie spowoduje różnicę jedynie w świecie zbudowanym z różnic”. Te słowa B. Latoura (2010, s. 367) wskazują, dlaczego identyfikowanie hybrydalnych sieci – a taką jest szkolne ocenianie – nie jest sztuką dla sztuki, ale kompleksowym rozpoznaniem sytuacji, bez czego niemożliwy staje się namysł nad zmianami wpisanymi immanentnie we współczesną rzeczywistość społeczną, czyli także edukację. W tym kontekście odrzucanie lub ograniczanie kontrowersji, które mogą być odczytywane, jako zakłócenie ustalonego porządku, oznacza zgodę na zawężenie zagadnienia do „rutynowych, zimnych i spokojnych sytuacji” (tamże, s. 37), czyli stagnację.

Hybrydalne sieci bywają jednak nieczytelne, a ich nieoficjalne czy ukryte rozprzestrzenianie się niesie trudne do przewidzenia konsekwencje społeczne, co dostarcza argumentów za ich porządkowaniem, czyli puryfikacją: praktyką charakterystyczną dla naukowych reprezentacji. Jednak w naukach społecznych, a do nich zaliczana

jest pedagogika, mogą nakładać się na siebie klasyfikacje i interpretacje prowadzone z różnych perspektyw teoretycznych, co prowadzi do kolejnych translacji.

W takim środowisku jedynym rozwiązaniem staje się – uważa B. Latour (tamże, s. 370) – „stopniowe układanie jednego wspólnego świata”, czyli demokratyczne kontrolowanie hybryd. W procesie tym potrzebne jest jednak porozumienie i zgoda na to, by o jednej rzeczy mówić z różnych, pozornie wykluczających się perspektyw. Stanowisko to podkreśla rolę translacji międzyparadygmatycznych, nad czym już kilka lat temu podjęła namysł Joanna Rutkowiak (2009) i co w nowszym ujęciu proponuje Dorota Klus-Stańska (2013), mówiąc o ekumenicznym spotkaniu paradygmatów, „jako interdyscyplinarnych (czy może inter-paradygmatycznych) głosów na określony temat, na przykład diagnozy i ewaluacji. Zwielokrotnienie możliwych perspektyw poznania otwiera nowe horyzonty pytań, uwrażliwia na nieoczywistość naszych założeń, uczy dystansu do dokonywanych rozstrzygnięć”.

Heterogeniczność jest zatem warunkiem rozwoju zarówno subdyscypliny naukowej, jaką jest diagnostyka edukacyjna, jak i szkolnego oceniania, rozpatrywanego w kategoriach dydaktycznych, społecznych i aksjologicznych, przed czym współcześnie nie da się już po prostu uciec.

Przypisy

¹ Tekst referatu wygłoszony został na KDE w Legnicy, w 1999 r., a opublikowany kilka lat później w jubileuszowym wyborze tekstów pod redakcją B. Niemierki (2004).

Literatura

Abriszewski W. (2010): *Splatając na nowo ANT. Wstęp*. W: B. Latour: *Splatając na nowo to, co społeczne. Wprowadzenie do teorii aktora-sieci*. Tłum. A. Derra i K. Abriszewski. Wyd. Universitas, Kraków.

Abriszewski W. (2010 a): *Wszystko otwarte na nowo. Teoria Aktora-Sieci i filozofia kultury*. Wyd. Naukowe UMK, Toruń.

Altszuler I. (1960): *Badania nad funkcją oceny szkolnej*. PZWS, Warszawa.

Bińczyk E. (2007): *Program badawczy Bruno Latoura i jego zalety w kontekście badań nad światem współczesnym*. „Przegląd Kulturoznawczy”, nr 3.

Groenwald M. (2010): *Patchworkowa sprawiedliwość oceniania*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.

Groenwald M. (2012): *Co na płaszczyźnie diagnostyki edukacyjnej łączy Henryka Szaleńca z Bruno Latourem i co na to Bolesław Niemierko?* W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. PTDE. Kraków.

Hamryszak J. (1994): *O sprawdzaniu i ocenianiu osiągnięć uczniów w zakresie plastyki*. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Hłuszyk H. (2001): *Wykorzystanie pomiaru dydaktycznego w ocenie i sprawdzaniu osiągnięć studentów biologii z dydaktyki biologii*. W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej Edukacji, Wrocław.

Jagodzińska M. (2004): *Wewnątrzszkolne systemy oceniania – założenia i rzeczywistość szkolna*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Wyd. GRUPA TOMAMI, Kraków.

Jakubowicz S., Plebański S., Rybicka K., Udzik B. (2008): *Uczyć i oceniać, czy oceniając, uczyć?* W: B. Niemierko, M.K. Szmigel (red.): *Egzamin i uczenie się w oczach nauczyciela*. Wyd. GRUPA TOMAMI, Kraków.

Jaskuła S. (2012): *Mutualizm ewaluacji i diagnozy edukacyjne*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.

- Kasáčová B., Cabanová M. (2011): *Pedagogická diagnostika. Teória a metódy diagnostikovania v elementárnej edukácii*. Pedagogická fakulta, Univerzita Mateja Bela, Banská Bystrica.
- Klus-Stańska D. (2013): *Myslenie poza diagnozą i ewaluacją. Pytanie o możliwość zmiany logiki tworzącej edukację*. „Teraźniejszość–Człowiek–Edukacja”, nr 4 (64).
- Klus-Stańska D. (2010): *Dydaktyka wobec chaosu pojęć i zdarzeń*. Wyd. Akademickie „Żak”, Warszawa.
- Klus-Stańska D. (2010a): *Ukryta polityczność oceniania szkolnego w świetle radykalnych nurtów socjologicznych – tezy do wystąpienia*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.
- Kohlberg L., Mayer R. (1993): *Rozwój jako cel wychowania*. Tłum. P. Kwieciński i A. Nalaskowski. W: Z. Kwieciński, L. Witkowski (red.): *Dylematy i kontrowersje we współczesnych pedagogiach*, IBE, Warszawa
- Konarzewski K. (1998): *Jak nauczyciele normują test osiągnięć szkolnych?* W: B. Niemierko, E. Kowalik (red.), *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Kowalik E., Wojciechowska K. (2000): *Szkolny system oceniania oparty na pomiarze dydaktycznym*. Wyd. Podkowa BIS, Gdańsk.
- Kruszewski K. (1998): *Test wiadomości dobrego i złego*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Kruszewski K. (2005): *Test wiadomości dobrego i złego dziesięć lat później*. W: M. Groenwald, G. Szyling, M. Daszkiewicz (red.): *Diagnostyka edukacyjna. Niemierkowskie spotkania i inspiracje*. Wyd. Fundacji Rozwoju UG, Gdańsk.
- Kwieciński Z. (2010): *Zmiany w edukacji – próba bilansu dwóch dekad*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.
- Latour B. (2010): *Splatając na nowo to, co społeczne. Wprowadzenie do teorii aktora-sieci*. Tłum. A. Derra i K. Abriszewski. Wyd. Universitas, Kraków.
- Latour B. (2011): *Nigdy nie byliśmy nowoczesni. Studium z antropologii symetrycznej*. Tłum. M. Gdula. Oficyna Naukowa, Warszawa.
- Messick S. (2005): *Trafność oceny psychologicznej*. W: J. Brzeziński (red.): *Trafność i rzetelność testów psychologicznych. Wybór tekstów*. Tłum. M. Zakrzewska. GWP, Sopot.
- Nalaskowski A. (2004): *Ideologiczny gen oceny*. W: B. Niemierko (red.): *Diagnostyka edukacyjna. Teoria i praktyka*. PTDE, Kraków.
- Niemierko B. (1990): *Pomiar sprawdzający w dydaktyce*, PWN, Warszawa.
- Niemierko B. (1997): *Między oceną szkolną a dydaktyką. Bliżej dydaktyki*. Wyd. 2 zm. WSiP, Warszawa.
- Niemierko B. (1999): *Pomiar wyników kształcenia*, WSiP, Warszawa.
- Niemierko B. (1999a): *Diagnostyka edukacyjna. Cele, metody i szanse rozwoju*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Wyd. Ośrodek Diagnozy, Egzaminów Szkolnych i Informacji Pedagogicznej, Legnica.
- Niemierko B. (2001): *Chłodne oblicze egzaminu zewnętrznego*. W: B. Niemierko, M.K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. Pandit, Kraków.
- Niemierko B. (2002): *Ocenianie szkolne bez tajemnic*. WSiP, Warszawa.
- Niemierko B. (2002a): *Egzaminy pożądane i niechciane*. W: B. Niemierko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnętrzne i zewnętrzne a jakość pracy szkoły*. Regionalny Ośrodek Metodyczno-Edukacyjny „Metis”, Katowice.
- Niemierko B. (2003): *Jaki pomiar dydaktyczny jest nam potrzebny?* W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.
- Niemierko B. (2007): *Kształcenie szkolne. Podręcznik skutecznej dydaktyki*. Wyd. Akademickie i Profesjonalne, Warszawa.
- Niemierko B. (2010): *Między prawdą a skutecznością – perspektywy oceniania szkolnego*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Niemierko B. (2013): *Czy diagnostyka edukacyjna i ewaluacja w edukacji zachowają odrębność?* „Teraźniejszość–Człowiek–Edukacja”, nr 4 (64).
- Opłocka U. (2008): *Emergencja nauczycielskiej pedagogii oceniania – doniesienie badawcze*. W: B. Niemierko, M.K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczycieli*. Wyd. GRUPA TOMAMI, Kraków.
- Parczewski P. W. (2012): *Nauczycielskie systemy oceniania jako systemy autopojetyczne. Teoria systemów*. W: B. Niemierko, M.K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. PTDE. Kraków.

Poręba-Konopczyńska A. (1998): *Analiza rozbieżności ocen pracy maturalnej z biologii*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Poręba-Konopczyńska A. (2001): *Moje refleksje nad ocenianiem na podstawie układu rzetelności i trafności wyników sprawdzania zaproponowanego przez Petera T. Knighta*. W: B. Niemierko, M. K. Szmigel (red.): *Teoria i praktyka oceniania zewnętrznego*. Wyd. PANDIT, Kraków.

Putkiewicz E. (2000): *O ocenianiu szkolnym i ocenianiu nauczania*. W: K. Kruszewski (red.): *Pedagogika w pokoju nauczycielskim*. WSiP, Warszawa..

Rozporządzenie Ministra Edukacji Narodowej z dn. 19 kwietnia 1999 r., w sprawie warunków i sposobu oceniania i promowania uczniów i słuchaczy oraz przeprowadzania egzaminów i sprawdzianów w szkołach publicznych: „Dziennik Urzędowy” 1999, poz. 131 ze zmianami.

Rutkowiak J. (2009): *Wielość paradygmatów dydaktyki, a wspólny mianownik rzeczywistości życia. Ku pytaniom o przekładalność paradygmatów*. W: D. Klus-Stańska, L. Hurlo, M. Łojko (red.): *Paradygmaty współczesnej dydaktyki*. Oficyna Wydawnicza „Impuls”, Kraków.

Smętek L., Maternicka K. (1998): *Test integracyjny sprawdzający umiejętności myślenia przyczynowo-skutkowego uczniów klasy III szkoły podstawowej*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Szafraniec M., Szafraniec M., Szlecht P. (2003): *Szkolny system oceniania – mit i rzeczywistość*. W: Niemierko B. (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Szudlarek T. (2005), *Ekonomia i społeczne skutki reform*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Szyling G. (2002): „*Idź w świat i radź sobie jak potrafisz*”. *Drugi układ wymagań programowych w oczach uczniów i nauczycieli*. W: B. Niemierko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnątrzszkolne i zewnętrzne a jakość pracy szkoły*. Regionalny Ośrodek Metodyczno-Edukacyjny „Metis”, Katowice.

Szyling G. (2003): „*Być może będzie w przyszłości ozdobą polskiej piłki nożnej?*”. *Wybrane aspekty trafności oceny końcowej osiągnięć ucznia*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.

Szyling G. (2010): *Ocenianie kształtujące, czyli o niejednoznaczności*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Szyling G. (2011): *Nauczycielskie praktyki oceniania poza standardami*. Oficyna Wydawnicza Impuls, Kraków.

Szyling G. (2013): *Obiektywizm i sprawiedliwość – (nie)wygodne kategorie szkolnego oceniania*. W: M. Dudzikowa, K. Knasiecka-Falberska (red. nauk.): *Sprawcy i/lub ofiary działań pozornych w edukacji szkolnej*. Oficyna Wydawnicza Impuls, Kraków.

Szyling G. (2013a), *Hybrydowy charakter oceny szkolnej we wczesnej edukacji: stan i wyzwania*. „*Studia Edukacyjne*”, nr 28.

Wszolek K. (1999): *Nauczycielskie kryteria oceny osiągnięć dydaktycznych uczniów*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Wyd. Ośrodek Diagnostyki, Egzaminów Szkolnych i Informacji Pedagogicznej, Legnica.

Zimny T. (1999): *Test dydaktyczny jako narzędzie ewaluacji i diagnozy edukacyjnej*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Wyd. Ośrodek Diagnostyki, Egzaminów Szkolnych i Informacji Pedagogicznej, Legnica.

Teoria i praktyka oceniania szkolnego. W poszukiwaniu pedagogicznych przesłanek hybrydyczności oceny szkolnej

W artykule zanalizowano teksty poświęcone szkolnemu ocenianiu, pomieszczone w publikacjach związanych z Konferencjami Diagnostyki Edukacyjnej (1984–2013). Punktem wyjścia podjętych rozważań uczyniono założenie, że na teorię i praktykę oceniania szkolnego składa się wiele heterogenicznych elementów, co nie wyklucza możliwości ich całościowego ujęcia. W analizie posłużono się językiem i metodologią zaczerpniętą z teorii aktora-sieci (Latour, 2010), ujawniając elementy stabilizujące badany obszar, a także kontrowersje, które dynamizują relacje między teorią i praktyką

ocenywania oraz wewnątrz nich. Uzyskany obraz potwierdził, że szkolne ocenianie można rozpatrywać jako twór hybrydalny, w rozumieniu tego pojęcia zaproponowanym przez Bruno Latoura (2011). Warunkiem zidentyfikowania i zrozumienia występujących w nim relacji jest jednak zgoda na to, by o jednej rzeczy mówić z różnych, pozornie wykluczających się perspektyw, co podkreśla rolę translacji międzyparadygmatycznych, a zarazem stanowi interesujący obszar poszukiwań dla diagnostyki edukacyjnej.

Słowa kluczowe: teoria i praktyka oceniania szkolnego, teoria-aktora sieci, heterogeniczność i hybrydyczność oceniania szkolnego

Theory and practice of classroom assessment. In search for premises of hybridity of school grade

The article analyses texts devoted to classroom assessment which have been included in publications related to the Conferences of Educational Diagnostics (1984–2013). The starting point of the undertaken reflections was an assumption that the theory and practice of classroom assessment is comprised of many heterogeneous components, which does not exclude the possibility of their joint treatment. In the analysis language and methodology derived from Bruno Latour's actor-network theory (2010) are applied, with it revealing the elements which stabilise the area studied and the controversies that dynamise relationships between the theory and practice of classroom assessment and those within the two. The picture obtained confirmed that classroom assessment can be viewed as a hybrid entity, the concept being construed in accordance with Bruno Latour's proposal (2011). The prerequisite for identification and comprehension of the relationships existing therein is, however, that it is agreed that one issue can be spoken about from various perspectives which supposedly exclude one another, which emphasises the role of inter-paradigmatic translations and at the same time constitutes an interesting area of exploration for educational diagnostics.

Key words: theory and practice of classroom assessment, actor-network theory, heterogeneity and hybridity of classroom assessment

Zofia Lisiecka

Centralna Komisja Egzaminacyjna w Warszawie

Wioletta Kozak

Centralna Komisja Egzaminacyjna w Warszawie

Instytut Badań Edukacyjnych w Warszawie

O POTRZEBIE BADAŃ NAD EFEKTEM ZWROTNYM EGZAMINU MATURALNEGO Z JĘZYKA POLSKIEGO

Przed dziesięcioma laty, wyznaczając perspektywy badań wspomagających system egzaminowania zewnętrznego, prof. Bolesław Niemierko wskazał obszar badań dotyczących wpływu egzaminów na treść i formę kształcenia poprzedzającego egzamin, jako jeden z najpilniejszych do eksploracji przez naukowców. Postulat dotyczył zbadania tzw. efektu zwrotnego¹ egzaminów, w tym ustalenia zależności „naprzód” i „wstecz” pomiędzy założeniami (programem) a egzaminem (Niemierko, 2004, s. 7–8). Przegląd literatury przedmiotu pozwala stwierdzić, iż postulat ów nie trafił na podatny grunt, gdyż w Polsce nie prowadzono dotychczas szeroko zakrojonych badań nad efektem zwrotnym egzaminów. Tym bardziej nieznaną jest literatura opisująca kompleksowo, z wykorzystaniem strategii ewaluacyjnych, oddziaływanie egzaminów doniosłych z zakresu przedmiotów humanistycznych (sprawdzian, egzamin gimnazjalny do roku 2012) i języka polskiego (nowy egzamin gimnazjalny, matura) na system edukacji, szkołę, nauczyciela czy ucznia. Równocześnie w powszechnej opinii zdecydowanie przeważa pogląd o negatywnym, a nawet niszczącym wpływie egzaminowania zewnętrznego. Dodajmy, pogląd w zdecydowanej większości nieoparty empirią, gdyż „powszechne mniemania”, „intuicja pedagogiczna” oraz „doświadczenie dydaktyczne”, zamiast rzetelnych badań naukowych, przeważają nad rzadkimi przyczynkarskimi doniesieniami z mikrobadań szkolnych bądź regionalnych. Przegląd tekstów towarzyszących dziewiętnastu krajowym konferencjom diagnostyki edukacyjnej potwierdza ten stan rzeczy.

Nie bez wpływu na tak wąskie i jednostronne postrzeżenie zjawiska efektu zwrotnego miały zapewne pejoratywne konotacje towarzyszące implementacji pojęcia w czynnym słowniku diagnostów² oraz charakterystyczne dla każdej znaczącej zmiany obawy przed jej skutkami³. Tymczasem przywołana przez Elżbietę Zawadowską-Kittel bogata literatura przedmiotu pozwala określić efekt zwrotny egzaminów doniosłych⁴ jako wpływ tychże egzaminów na szeroko rozumiany proces uczenia się i nauczania oraz na wszystkich jego uczestników, a nawet na cały system edukacji, przy czym wpływ ów może być zarówno negatywny, jak i pozytywny, zamierzony bądź nieplanowany, uświadomiony lub nie, długo- bądź krótkotrwały, o silnym lub słabym charakterze (Zawadowska-Kittel, 2012, s. 73, 81–84). Tak szeroko rozumiane pojęcie *efektu zwrotnego* obejmuje swym zakresem nie tylko znaczenia słownikowe

przypisane takim pojęciom, jak *washback (effect) / backwash effect*, ale też *impact* (oddziaływanie), a nawet *consequences* (konsekwencje, trwałe rezultaty) egzaminu (zob. Kussmaul, 1995; Shohamy, 2001).

Niniejsza publikacja nie aspiruje do miana kompendium wiedzy o przedmiocie badań, nie daje też pełnego obrazu wszelakich wpływów, jakie egzaminowanie zewnętrzne wywarło na edukację oraz jak zmiany treściowe i formalne w nauczaniu języka polskiego wpłynęły na proces egzaminacyjny, ale ma ambicje stać się przyczynkowym głosem w dyskusji o konieczności pogłębionych prac pedagogicznych, psychologicznych, społecznych (socjologia, politologia), badających efekt zwrotny egzaminów na zaangażowane w ów proces podmioty oraz programy i strategie edukacyjne.

Koncepcja matury z języka polskiego a rezultaty dydaktyczne

Rozpatrując wpływ egzaminów doniosłych na proces uczenia się i nauczania języka polskiego w ostatnich kilkunastu latach, przede wszystkim zwrócić należy uwagę na to, co przyświecało wprowadzanym w roku 2002 zmianom w koncepcji tzw. nowej matury. Odejście od sprawdzania wiedzy na temat literatury i pójście w stronę tekstocentryzmu miało doprowadzić do pogłębienia rozumienia kultury, umiejętności odczytywania tekstów jako formy komunikacji symboliczno-aksjologicznej. Badacze i dydaktycy pozostawali zgodni, iż tekst kultury, z którym pracuje się zarówno na lekcji, jak i na egzaminie, powinien być traktowany jako „wypowiedź antropologiczna, określająca miejsce człowieka w przestrzeni kulturowej, dająca w języku relacji przestrzennych i czasowych odpowiedź na podstawowe pytania o sens ludzkiego istnienia” (Ożdżyński, 1999, s. 341). Konsekwencją tego myślenia stała się konieczność odczytania przez maturzystę znaków, symboli, przywoływania kontekstów interpretacyjnych, tym samym ujawniania zachowań intelektualnych, emocjonalnych i aksjologicznych, świadczących o rozumieniu tekstu. Rozumienie to w metodologii hermeneutycznej polega na interpretacji i dialogu z tekstem (Myrdzik, 1999, 2000).

W wymiarze edukacyjnym działania te powinny prowadzić do zrozumienia wpływu przeszłości na kształtowanie takich umiejętności młodego człowieka, które pozwolą mu na samodzielne zaistnienie w kulturze, budowanie postaw wobec wartości poprzez kształcenie potrzeb i umiejętności ich odkrywania, poznawania, określania, aktywnego ustosunkowania się do nich, czyli ich uznawania, odczuwania, wybierania, realizowania, tworzenia, hierarchizowania, formułowania sądów o wartościach i sądów wartościujących. Umiejętności te można, jak twierdzi ks. Marek Sołtysiak, traktować jako wyraz poczucia tożsamości kulturowej, zakorzeniania w tradycji (będącej podstawą rozumienia) oraz ciągłości pamięci (Sołtysiak, 2004) albo – w ujęciu Ewy Bieńkowskiej – jako przekonanie, że osoba wykształcona to ta, która czuje się prawowitym spadkobiercą, chociaż swą rolę widzi skromnie: jako kogoś obdarowanego i zobowiązanego do dalszego przekazywania dziedzictwa (Bieńkowska, 1995, s. 20–27). Nowa formuła egzaminu maturalnego miała więc wyposażyć młodego człowieka w narzędzia intelektualne i wywoływać w nim refleksję nad rolą tekstów kultury, wprowadzać go w świat wartości. U podstaw przemian stały prowadzone w połowie lat dziewięćdziesiątych obserwacje zachowań uczniów, wynikające z badań o negatywnym wpływie rozwoju cywilizacyjnego, zwłaszcza mass mediów. „Szo-

kiem dla wielu z nas jest odkrycie, że uczniowie przestają rozumieć język (języki) tekstów, które są przedmiotem lektury. Utożsamiony on bywa najczęściej z głosem spikera czytającego dialogi w filmach lub ze strzępami informacji dialogowej zawartej w komiksowych «dymkach». Na plan kontekstu, z dominującym życiem rodzinnym ekranem telewizyjnym, wchodzi sugestywny styl reklamy, seks i brutalny dreszczowiec” (Kryda, 1995, s. 57). Na ile więc formuła egzaminu maturalnego, wprowadzona powszechnie w roku 2005, przyczyniła się do zatrzymania procesu „nierozumienia” języka tekstów kultury? Czy spowodowała, że szkoła zaczęła wyposażać uczniów w odpowiednie narzędzia służące świadomej, samodzielnej lekturze? Jak nauczyciele budowali proces prowadzący do pogłębionego rozumienia, wywołania refleksji w młodym czytelniku?

Badania pokazały (m.in. Kozak i Mlonka, 2008; Udzik, 2010), że wpisane w filozofię **nowej matury** założenia nie do końca zostały w praktyce szkolnej zrealizowane. Uczniowie na poziomie podstawowym odczytują tekst powierzchownie – na poziomie faktycznym (Chrzastowska, 2000, s. 265–272), a często nie rozumieją wszystkich treści. Ponadto piszą bez pogłębionej refleksji nad zawartym w temacie problemem, wreszcie nie umiejscawiają dokonywanych rozważań w szerokim kontekście kulturowym. Niewielu maturzystów cechuje dociekliwość poznawcza i traktowanie tekstu literackiego jako partnera dialogu, a tym samym nie podejmują oni próby prowadzenia dyskursu nad sensem życia, oceną postaw i wartości. W większości prac przeważa podejście odtwórcze, polegające na streszczeniu opisywanych wydarzeń, tekst nie stanowi projekcji, czy też inspiracji do głębszej refleksji kulturowej, historycznej czy socjologicznej. Można więc – z dozą ostrożności – przyjąć, że niekoniecznie maturzysta rozumie czytany tekst i zjawiska kultury, o których się wypowiada. Jeśli przyjmiemy, iż rozumienie czytanego tekstu może przebiegać na poziomie percepcyjnym, semantycznym, receptywnym, krytycznym i twórczym, to znaczna część prac słabych nie kwalifikuje uczniów do poziomu semantycznego, obejmującego rozumienie i interpretowanie jednostek w ciągu rosnącym: od znaczenia wyrazów, zdań do fraz, fragmentów, wreszcie całego tekstu. Pozostają oni na poziomie percepcyjnym, odnoszącym się do wyodrębniania wyrazów i zdań tekstu. Ponadto w niewielkim stopniu wyeksponowana jest w pracach maturalnych świadomość kulturowa odbiorcy. By ten proces mógł zaistnieć, odbiorca musi zauważyć zjawiska kultury, rozumieć je, przeżyć, wreszcie akceptować lub nie (Jędrychowska, 1994). Aby czytać ze zrozumieniem teksty kultury, należy odnotować ich znakowy i komunikacyjny charakter (tworzywo, relacje nadawczo-odbiorcze, funkcje, związki), ale też wznieść się ponad wiedzę teoretyczną i historycznoliteracką, skupiając na przesłaniu, jakie te teksty niosą, przesłaniu, które pomaga w rozumieniu sensów, wartości, zachowań, wkracza w symboliczny wymiar kultury.

Wydaje się, że szkolna polonistyka nie przygotowała też należycie uczniów do stosowania różnorodnych strategii kształcenia literaturoznawczego i językoznawczego, a tym samym nie wyposażyła ich w różnorodne narzędzia interpretacyjne. Badania wykazały, iż w przeanalizowanych pracach maturalnych zdecydowanie dominuje strukturalizm i poststrukturalizm, zwłaszcza na poziomie rozszerzonym matury przy analizie i interpretacji poezji, natomiast z rzadka daje się dostrzec hermeneutyczny dialog między tekstem a odbiorcą (Kozak i Mlonka, 2008). Tymczasem postulaty stosowania w edukacji polonistycznej różnorodnych podejść i metodologii literaturoznawczych pojawiały się już od dawna (Jaskółowa, 2008; Bortnowski, 2008;

Koziółek, 2008). „Koezystencja metodologii, niewielkie wahania w tym zakresie, umiarkowany postęp i umiarkowany konserwatyzm niechaj wyznaczają drogi polonistyki szkolnej (...). Więc nie zastój, ale też daleko od popłochu, że jesteśmy gdzieś na rubieżach nauki. Pluralizm i związany z tym eklektyzm, współbieżność szkół badawczych z niewielką tendencją do spokojnych przekształceń – to przewiduję i to postuluję” (Bortnowski, 2008, s. 25). Ten postulat Stanisława Bortnowskiego zakreśla przestrzeń dla badań związanych efektem zwrotnym matury stojącej po raz kolejny u progu zmiany. Jednocześnie pozwala określić efekt zwrotny egzaminu jako w zamierzeniach w pełni świadomy, głęboki, niszczący nauczycielskie schematy i przyzwyczajenia, ożywiający szkolną polonistykę. Czy tak się stało w praktyce? Co sprawia, iż w powszechnym odczuciu wyżej opisane zmiany nie są przez środowisko szkolne odbierane jako uświadomione, trwałe, pozytywne? Jakie działania podjęto, by owe idee i koncepcje przybliżyć uczniom? Czy nie wypaczono założeń *nowej matury*? Dlaczego wciąż silnie brzmią głosy tęsknoty za starymi koncepcjami egzaminu? Te i wiele innych pytań badawczych należy w końcu postawić.

Casus ustnej matury

Oddziaływania egzaminów doniosłych dotyczą nie tylko procesów poznawczych, ale obejmują też cały kompleks zjawisk związanych ze zmianami zachowań, jakie zachodzą w związku z egzaminem (Shohamy, 2001). Tak pojęty wpływ egzaminów na proces nauczania-uczenia się i związane z nim postawy ludzkie może mieć charakter nie tylko edukacyjny, ale też społeczny. Przykładem takiego, niezamierzonego, wpływu na działania uczniów, nauczycieli i inne osoby zaangażowane w proces związany z przygotowaniem się i realizacją egzaminu, jest ustna matura z języka polskiego. *Casus* tego egzaminu mógłby stać się wyzwaniem dla wnikliwego socjologa czy etyka edukacji, gdyż to właśnie względy społeczne związane z nagłośnionym przez media procederem oszustwa egzaminacyjnego, a także inercja środowiska szkolnego, które w założeniu miało wspomagać i nadzorować proces uczenia się ucznia, wydaje się, spowodowały, że po 10 latach stosowania tej formy egzaminowania, odchodzi ona do edukacyjnego lamusa.

Rzecz jednak w tym, że nie w lamusie powinno znaleźć się miejsce egzaminu wciąż nowoczesnego, znakomicie wpisującego się w strategię uczenia się wytyczone przez Unię Europejską⁵, pozwalającego – jak objaśniają przyszlých maturzystów Artur Lewczuk i Zbigniew Kosiński – przyjąć rolę Badacza, wcielić się w rolę Odkrywcy, stać się Konstrukтором, a podczas egzaminu – Mówcą i Ekspertem, jak również Twórcą Nowej Rzeczywistości (Lewczuk, Kosiński, 2006, s. 3), łącząc kompetencje polonistyczne z jakże pożądanymi społecznie, w tym na rynku pracy. Ani w publikacjach z konferencji diagnostycznych, ani w polonistycznych pismach branżowych czy opracowaniach naukowych nie znajdziemy głosów kwestionujących cele egzaminu, sprawdzającego m.in. umiejętności wyszukiwania, selekcjonowania, doboru i samodzielnego interpretowania literatury podmiotu i przedmiotu, myślenia naukowego, używania poprawnego języka mówionego, stosowania adekwatnej terminologii, komunikowania się ze słuchaczami, w tym rozumienia intencji rozmówcy, formułowania odpowiedzi i dostosowania ich szczegółowości do pytań, uzasadniania stanowiska, wartościowania, formułowania problemów badawczych.

Niszczący wpływ tego egzaminu, z założenia zewnętrznego, realizowanego jednakże jako część wewnętrzna, związany jest bowiem nie z jego merytoryczną i wciąż wysoce trafną programowo koncepcją, ale z procesem realizacji i postawami zaangażowanych w ten proces ludzi. Egzamin obnażył (by nie rzec – wygenerował) moralnie naganne postawy uczniów „idących na skróty”, których to celem nie jest uczenie się, ale zdanie egzaminu, co przy społecznym przyzwoleniu na oszustwo egzaminacyjne (Lisiecka, 2006, s. 224–232) zaowocowało konsekwencjami szeroko opisywanymi w mediach i przez badaczy (Gromkowska-Melosik, 2007; Groenwald, 2011): kupczeniem gotowymi prezentacjami, w tym przygotowywanymi przez nauczycieli, uczeniem się cudzych wystąpień na pamięć i bezrefleksyjnym wygłaszaniem ich przed komisją egzaminacyjną. Obnażył też postawę zaniechania dostrzeganą u niektórych nauczycieli i dyrektorów szkół, niechętnych kolejnemu nieodpłatnemu zadaniu prowadzenia konsultacji i nadzorowania postępów ucznia, a na koniec – „taśmowego” egzaminowania. Badaczy zastanowił fakt niskiej korelacji pomiędzy wynikami matury ustnej i pisemnej, słaby obiektywizm egzaminu oraz niska rzetelność wyników (zob. Ślusarz, 2006, s. 528, 534; Ślusarz, 2008, s. 401–408; Krasuska i Modrzewska, 2007, s. 520–523; Finkstein i Koziół, 2010, s. 541), czego źródłem jest – ich zdaniem – ukonstytuowanie ustnej matury jako egzaminu wewnętrznego w systemie zewnętrznym.

Z racji realizacji części ustnej egzaminu w szkole, trudno go nazwać egzaminem porównywalnym czy obiektywnym. Czas jego realizacji (na koniec kształcenia) powoduje z kolei, że traci on też walor egzaminu powszedniego, w którym znaczenie komentarza dydaktycznego jest większe od informacji o wyniku uczenia się. Maturzyści i ich nauczyciele czują się też zawiedzeni niskim prestiżem egzaminu, którego wyniku nie uwzględnia się w rekrutacji na żadne studia, nawet polonistyczne (Ślusarz, 2006, s. 530), co deprecjonuje jego doniosłość, czyni przyzwolenie na oszustwo egzaminacyjne łatwiejszym (Groenwald, 2011, s. 111–119), a niektórym pozwala pytać nawet o sens jego kontynuacji. To sprawia, iż ten – z założenia – twórczy i motywujący egzamin (Niemierko, 2001, s. 26), stał się egzaminem „do poprawki”. Tyle, że owa „poprawka”, obecnie wprowadzana na maturę 2015 roku, koncentruje się nade wszystko na zmianach merytorycznych, a nie tych związanych z zarządzaniem organizacją, jaką jest szkoła, ani na ludzkich postawach. Perspektywa badawcza kolejnych lat powinna być więc skierowana także na nową maturę ustną, a pytania o efekty zwrotne tego egzaminu – nie tylko dydaktyczne, ale też społeczne i psychologiczne – powtórzone.

Od wymagań edukacyjnych do oceniania osiągnięć ucznia

Silny wpływ na dydaktykę szkolną, ściśle powiązaną z dominującym wciąż w polskiej edukacji tradycyjnym modelem oceniania osiągnięć ucznia⁶ (Szyling, 2011, s. 33), miały wprowadzone wraz z pierwszą podstawą programową i egzaminami zewnętrznymi postulowane przez B. Niemierkę idee kształcenia według wymagań oraz oceniania kryterialnego (zob. Niemierko, 1997, s.119–222). Reforma edukacyjna z roku 1999 przyniosła szkole podstawę programową – kluczowy dokument wyznaczający państwowe standardy edukacyjne, oraz usankcjonowała konieczność komunikowania przez nauczycieli własnych wymagań programowych uczniom i ich

rodzicom. Jednak w opinii wielu badaczy (zob. m. in. Lisiecka i Stróżyński, 2001; Lisiecka, 2002; Chorąży, 2004; Palkij, 2004; Zapieraczyński, 2004) najistotniejszą rolę w budowaniu świadomości nauczycielskiej, na czym tworzenie, realizacja i ewaluacja wymagań edukacyjnych polega, odegrały standardy wymagań egzaminacyjnych, materiały stanowiące obudowę narzędzi wykorzystywanych w egzaminowaniu zewnętrznym oraz raporty komisji egzaminacyjnych z prowadzonych diagnoz i badań. Publikacja planów testów (kartotek), schematów poprawnych odpowiedzi i oceniania czy szczegółowych sprawozdań o osiągnięciach uczniów pokazały nauczycielom, iż sięganie do wymagań staje się nie tylko koniecznością, ale też sprzyja planowaniu pracy dydaktycznej, projektowaniu diagnoz nauczycielskich i szkolnych, wdrażaniu różnorodnych strategii oceniania (w tym kształtującego), prowadzeniu autoewaluacji. Jednocześnie wnikliwe badania nad wymaganiami w ocenianiu szkolnym przeprowadzone od kilku lat przez Grażynę Szyling (2008, 2010, 2011) implikują potrzebę kolejnych pytań o wpływ wymagań egzaminacyjnych na budowanie wymagań nauczycielskich.

Efektom zainteresowania problematyką kształcenia wg wymagań są też postulaty badaczy upominających się o wymagania podłużne, a co za tym idzie, potraktowanie egzaminów zewnętrznych na zakończenie poszczególnych etapów edukacyjnych jako kolejnych odsłon panelowych badań osiągnięć uczniów. Wydaje się, że oparcie egzaminów zewnętrznych na nowej podstawie programowej i odwrót od koncepcji egzaminu ponadprzedmiotowego na rzecz egzaminu z przedmiotu, może stać się krokiem w tym kierunku. Efektem takich działań byłoby podniesienie wartości diagnostycznej, w tym ewaluacyjnej, wyników badania.

Egzaminy zewnętrzne stały się także forpocztą w procesie doskonalenia się nauczycieli w trudnej sztuce oceniania ucznia. Czy efekt zwrotny owego doskonalenia pozwolił zrodzić zdrowe owoce? Polonistyka szkolna, przedkładająca podejście z „czuciem i wiarą” ponad „szkiełko i oko”, ze swej natury różnoraka metodologicznie, wielowymiarowa w zakresie treści kształcenia, ukierunkowana nie tylko na zdobywanie wiedzy i umiejętności, ale także na kształtowanie postaw, wspierająca indywidualizm i kreatywność uczniów, początkowo zdecydowanie silniej broniła się przed strategiami oceniania analitycznego, niż tzw. przedmioty ścisłe. Tymczasem w opinii B. Niemierki poczucie wyższości oceniania intuicyjnego nad innymi jego „okielznanymi” formami prowadzi na manowce – do intuicyjnej dydaktyki lub w najlepszym razie – do dydaktyki kierowanej zewnętrznymi racjami (politycznymi, ekonomicznymi) i zapożyczonymi racjami naukowymi (logicznymi, psychologicznymi). „Uduchowienie” oceniania, w założeniu podnoszące teorię nauczania na piedestał, w rzeczywistości prowadzi ją do ubezwłasnowolnienia, stając się formą ucieczki od odpowiedzialności czy wręcz nadużyciem, nie jest też skutecznym narzędziem do oceny treści uważanych za „nieuchwytnie”, bo co najwyżej umożliwia ocenianie powierzchowne, niewolne od popadania w tzw. pułapki oceniania (Niemierko, 1997, s. 124).

Tej skrajnej opinii można przeciwstawić liczne argumenty badaczy i publicystów sprzeciwiające się ocenianiu „technologicznemu”, opartemu na osławionym analitycznym kluczu odpowiedzi (zob. np. Żuk, 2001; Bulska, 2006; Opłocka, 2010; Lewczuk, Kosiński 2010; Groenwald, 2011; Rybicka, 2012). Szczególnie mocno, choć dość często populistycznie, w tej debacie o istocie polonistycznego oceniania brzmiały anonse medialne, iż ów klucz uśmierca kreatywność i samodzielność myślenia,

zniewala, propaguje strategie „uśredniania”, przywilejami obdarza najsłabszych uczniów, poszukując w ich wypracowaniach jakichkolwiek śladów myśli, które pozwalają egzaminatorowi „usklądać” sumę punktów potrzebną do zaliczenia egzaminu. Tymczasem wyniki nielicznych badań pokazywały, że tworzone przez maturzystów analizy i rozprawki są w większości na tyle sztapowe i odtwórcze, iż egzaminatorzy z rzadka odczuwali trudności związane z oceną prac „nietypowych”, równie rzadko mogli też przydzielić zdającym ekstra-punkty za szczególne walory pracy, wyróżniające tekst maturzysty ze względów merytorycznych, formalnych czy językowych (Lisiecka, Firsruk, 2007; Kozak, Mlonka, 2008). Dotyczyło to niespełna 4% zdających w 2007 roku egzamin na poziomie podstawowym i 14% na poziomie rozszerzonym. Nie znaczy to jednak, iż badacze zjawiska efektu zwrotnego nie powinni dać posłuchu opiniom uczniów i nauczycieli, którzy czują się stłamszeni przez system egzaminacyjny. „Praca z uczniami zobowiązuje nauczyciela (...) do działania na rzecz ich dobra. Jak ma postąpić, kiedy tym dobrem staje się pomyślny wynik egzaminacyjny osiągnąty jako efekt kształcenia podporządkowanego egzaminowi? – pyta Maria Groenwald. Dodaje też: „Okazuje się, że bycie nauczycielem-egzaminatorem jest zawieszeniem «pomiędzy» dwoma systemami normatywnymi. Prowadzi ono do ukształtowania swoistej «tożsamości rozproszonej», niepotrafiącej określić się ani w byciu nauczycielem, ani w byciu egzaminatorem» (Groenwald, 2011, s. 94, 95).

Pogłębione badania pokazują, iż stosowanie procedur oceniania kryterialnego podnosi trafność oceniania wypracowań pisanych przez zdających (Niemierko, 1994, s. 72–81; Niemierko, Majkowska, 1998, s. 28–41). Jednakże „z psychometrycznego punktu widzenia egzamin ten (matura z języka polskiego) pozostawia wiele do życzenia. Bez głębszych zmian nie będzie można traktować tego egzaminu jako rzetelnego, obiektywnego, sprawiedliwego czy choćby w pełni porównywalnego” (Pokropek, 2011, s. 449). Czy sytuacja ta ulegnie zmianie wraz z wprowadzeniem na maturze 2015 nowej koncepcji oceniania, chyba zbyt na wyrost zwanego holistycznym? Czy nauczyciele egzaminatorzy są przygotowani na tę dość radykalną zmianę? Czy rezygnacja z klucza w ocenianiu wypowiedzi pisemnych to rzeczywiście panaceum, jak twierdzą akademicy⁷, na uzdrowienie matury z polskiego? Jak owa zmiana wpłynie na rozwój ucznia, szkoły, systemu egzaminacyjnego? Wreszcie – czy nowy egzamin spowoduje nawrót zainteresowania nauczycieli różnymi strategiami oceniania szkolnego, stosowanymi świadomie i zgodnie z celem? Czy to obroni uczniów przed dotychczasową praktyką nadużywania oceniania analitycznego „na wzór i podobieństwo” matury? Kiedy nauczyciele dostrzegą niszczący wpływ strategii „kopiuj-wklej”, nagminnie stosowanej w budowaniu wymagań, narzędzi, kryteriów i schematów oceniania, niekiedy zupełnie nieadekwatnych i do wybranej przez nauczyciela koncepcji programowej i do jego stylu nauczania? Postawienie także i tych pytań o wpływ zmian w egzaminowaniu na edukację polonistyczną staje się ewidentną koniecznością.

Przypisy

¹ W literaturze przedmiotu, zdominowanej przez publikacje anglojęzyczne, alternatywnie używa się pojęć *washback (effect)* / *backwash effect*. W artykule konsekwentnie posługujemy się polskim terminem efekt zwrotny.

² Wprowadzając pojęcie **efektu zwrotnego** na grunt polskiej diagnostyki edukacyjnej, B. Niemierko wyeksponował jedynie przykład efektu niszczącego zastany system edukacyjny (Niemierko, 2002, s. 265).

³ Na dowód badacze chętnie cytują nauczycieli i dyrektorów szkół: „System egzaminacyjny zabija wszystko” (Niemierko, 2004, s. 8, przypis 9); „Nowa matura nie jest już przepustką do świata elit” (Kossakowska, 2006, s. 559); „Matura z polskiego nie jest miarodajna; coś z tą maturą trzeba zrobić! Dalej tak nie może być!” (Lewczuk, Kosiński, 2010, s. 545); „EWD w obecnej postaci to, co najwyżej, egzaminacyjna wartość dodana. Nasza szkoła prowadzi znacznie szerszą działalność edukacyjną” (Niemierko, Walukiewicz, 2011, s. 72).

⁴ Rozróżnienie na **egzamin doniosłe** (ang. *high-stake test* – ‘egzamin wysokich stawek’), czyli takie, w których znaczenie informacji o wyniku uczenia się jest większe od komentarza dydaktycznego, i **powседневnie** (ang. *low-stakes test* – ‘egzamin niskich stawek’) czynimy za B. Niemierką (2002, s. 252).

⁵ W wydanym w 2000 r. przez Komisję Europejską dokumencie: *A Memorandum on Lifelong Learning: Making a European Area of Lifelong Learning a Reality*, propaguje się model uczenia się przez całe życie, definiowanego jako „wszelkie formy nauki podejmowane przez całe życie, mające na celu doskonalenie, pogłębianie wiedzy, umiejętności i kompetencji z perspektywy osobistej (indywidualnej), obywatelskiej, społecznej i/lub zawodowej”.

⁶ Autorka opisuje model **tradycyjny i postępowy** oceniania szkolnego. Testowanie i egzaminowanie, stosowane w celu różnicowania psychometrycznego, przypisuje wyłącznie ocenianiu tradycyjnemu.

⁷ Zob.: *Opinia Konferencji Rektorów Akademickich Szkół Polskich (KRASP) o informatorach maturalnych od 2015 roku*. W: *Informator o egzaminie maturalnym z języka polskiego od roku szkolnego 2014/2015* (2013, s. 155). CKE, Warszawa; <http://www.cke.edu.pl/files/file/Matura-2015/Informatory-2015/Jezyk-polski.pdf> (19.02.2014).

Literatura

- Bieńkowska E. (1995): *Mala historia pamięci*. „Znak”, nr 5.
- Bortnowski S. (2008): *Pluralizm metodologiczny jako konieczność*. „Egzamin. Biuletyn Badawczy Centralnej Komisji Egzaminacyjnej”, nr 16.
- Bulska T. (2006): „Wstrzelić się w model” czy rozwiązać zadanie? W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Chorąży E. (2004): *Jaka rolę mogą odegrać standardy wymagań egzaminacyjnych z historii w kształceniu według wymagań?* W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Chrzastowska B. (2000): *Jak sprawdzać czytanie symboliczne?* „Polonistyka”, nr 5/2000.
- Finkstein A., Kozieł J. (2010): *Ustna matura z języka polskiego. Wewnętrzny egzamin w systemie egzaminów zewnętrznych?* W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Groenwald M. (2011): *Etyczne aspekty egzaminów szkolnych*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Gromkowska-Melosik A. (2007): *Ściagi, plagiaty, fałszywe dyplomy*. GWP, Gdańsk.
- Informator o egzaminie maturalnym z języka polskiego od roku szkolnego 2014/2015* (2013). CKE Warszawa, <http://www.cke.edu.pl/files/file/Matura-2015/Informatory-2015/Jezyk-polski.pdf> (19.02.2014).
- Jaskółowa E. (2008): *Różnorodność metodologiczna w interpretacji tekstu literackiego a przygotowanie do egzaminu maturalnego*. „Egzamin. Biuletyn Badawczy Centralnej Komisji Egzaminacyjnej”, nr 16.
- Jędrzychowska M. (1994): *Lektura i kultura*. Wyd. Edukacyjne, Warszawa-Kraków.
- Kossakowska B. (2006): *Matura na falach społecznych nastrojów*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Kozak W., Mlonka B. (2008): *Osiągnięcia uczniów w zakresie rozumienia, analizy i interpretacji tekstów kultury na przykładzie matury z języka polskiego*. „Egzamin. Biuletyn Badawczy Centralnej Komisji Egzaminacyjnej”, nr 16.

- Koziółek K. (2008): *Od obcości do rozumienia. Przekład, objaśnienia, interpretacja w przygotowaniu ucznia do egzaminu maturalnego*. „Egzamin. Biuletyn Badawczy Centralnej Komisji Egzaminacyjnej”, nr 16.
- Krasuska A., Modrzewska E. (2007): *Założenia egzaminu ustnego z języka polskiego a praktyka szkolna*. W: B. Niemierko, M.K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Kryda B. (1995): *Klasyka literacka w odbiorze uczniów*. W: B. Chrzastowska, Z. Uryga (red.): *Zjazd Polonistów. Zagadnienia edukacyjne*. Wyd. Naukowe WSP, Kraków.
- Kussmaul P. (1995): *Training the Translator*. Benjamins, Amsterdam/Philadelphia.
- Lewczuk A., Kosiński Z. (2006): *Jak przygotować się do ustnego egzaminu maturalnego z języka polskiego? Okręgowa Komisja Egzaminacyjna, Łomża*.
- Lewczuk A., Kosiński Z. (2010): *Kryzys polonistyki szkolnej a system oceniania zewnętrznego*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Lisiecka Z., Stróżyński K. (2001): *Szkolne przygotowywanie uczniów do egzaminu*. W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. DSWE, Wrocław.
- Lisiecka Z. (2002): *Ocenianie wewnętrzne i zewnętrzne. Relacje międzysystemowe*. W: B. Niemierko, J. Brzdąk (red.): *Dwa rodzaje oceniania szkolnego. Ocenianie wewnętrzne i zewnętrzne a jakość pracy szkoły*. Regionalny Ośrodek Metodyczno-Edukacyjny „Metis”, Katowice.
- Lisiecka Z. (2006): *O polską kampanię na rzecz uczciwości edukacyjnej*. W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Lisiecka Z., Firsuk M. (2007): *Analiza wyników matury 2006 i 2007 z języka polskiego w zakresie kryterium „Szczególne walory pracy”*. Centralna Komisja Egzaminacyjna, Warszawa.
- Myrdzik B. (1999): *Rola hermeneutyki w edukacji polonistycznej*. Wyd. UMCS, Lublin.
- Myrdzik B. (2000): *O roli tradycji w szkolnym dialogu z tekstami literackimi*. W: B. Myrdzik (red.): *Nowoczesność i tradycja*. Wyd. UMCS, Lublin.
- Niemierko B. (1994): *Diagnoza systemu oceniania na podstawie analizy recenzji prac maturalnych*. „Warsztaty Polonistyczne” nr 4 (11).
- Niemierko B. (1997): *Między oceną szkolną a dydaktyką. Bliżej dydaktyki*. WSiP, Warszawa.
- Niemierko B. (2001): *Ku czemu zmierzają egzaminy szkolne?* W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. DSWE, Wrocław.
- Niemierko B. (2002): *Ocenianie szkolne bez tajemnic*. WSiP, Warszawa.
- Niemierko B. (2004): *Perspektywy badań psychologicznych, dydaktycznych, socjologicznych i politologicznych wspomagających system egzaminowania zewnętrznego*. „Egzamin. Biuletyn Badawczy Centralnej Komisji Egzaminacyjnej” nr 1.
- Niemierko B., Majkowska I. (1998): *Co warta jest wielokryterialna ocena pisemnych prac maturalnych*. Edukacja: Studia, Badania, Innowacje”, nr 3.
- Niemierko B., Walukiewicz S. (2011): *Jak mierzyć kapitał ludzki? Ewaluacyjne perspektywy operacjonalizacji pojęć ekonomicznych*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Opłocka U. (2010): *Belfer zniewolony*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Ożdżyński J. (1999): *Konteksty kulturowe w nauce o literaturze*. W: Z. Staszewska (red.): *Wartościowanie w dyskursie edukacyjnym*. Wyd. UŁ, Łódź.
- Palkij H. (2004): *Od podręczników do praktyki szkolnej – wpływ standardów wymagań z historii i wiedzy o społeczeństwie na kształcenie przedmiotowe*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Pokropek A. (2011): *Matura z języka polskiego. Wybrane problemy psychometryczne*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji: koncepcje, metody, perspektywy*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Rybicka K. (2012): *Rozpoznawanie szkolnego oceniania*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.
- Shohamy E. (2001): *The power of tests: A critical perspective on the use of language tests*. Pearson Education, London.
- Sołtysiak M. (2004): *Rozumienie i tradycja w hermeneutyce filozoficznej H.-G. Gadamera*. Wyd. Naukowe PAT, Kraków.

Szyling G. (2008): *Strategia przejścia czy przetrwania, czyli dokąd zmierza ocenianie szkolne?* W: D. Klus-Stańska (red.): *Dokąd zmierza polska szkoła?* Wyd. Akademickie „Żak”, Warszawa.

Szyling G. (2010): *W pułapce sprzecznych oczekiwań. Wybrane aspekty oceniania osiągnięć szkolnych uczniów w kontekście ekonomicznego myślenia o edukacji.* W: E. Rodziewicz i in. (red.): *Teraźniejszość jako impuls dla pytań o ogólność pedagogiki.* Wyd. Fundacji Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Szyling G. (2011): *Nauczycielskie praktyki oceniania poza standardami.* OW „Impuls”, Kraków.

Ślusarz A. (2006): *O konieczności uzewnętrznienia ustnej matury z języka polskiego.* W: B. Niemierko, M. K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne.* Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Ślusarz A. (2008): *Jak ratować ustną maturę?* W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczyciela.* Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Udzik B. (2010): *Kryteria egzaminu maturalnego z języka polskiego a ocenianie wewnątrzszkolne.* W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego.* Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Zapieraczyński M. (2004): *Standardy wymagań egzaminacyjnych w szkole podstawowej i gimnazjum – rozumienie konstrukcji oraz wpływ tego na komunikowanie wyników i sposób ich odbioru.* W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej.* Polskie Towarzystwo Diagnostyki Edukacyjnej, Kraków.

Zawadowska-Kittel E. (2012): *Efekt zwrotny (washback) – próba opisu zjawiska.* „The Linguistic Academy Journal of Interdisciplinary Language Studies. Roczniki Naukowe Lingwistycznej Szkoły Wyższej w Warszawie” nr 2/2012.

Żuk B. (2001): *Współczesny polonista – bezduszny mierniczy czy twórczy przyjaciel ucznia.* W: B. Niemierko, W. Małecki (red.): *Dawne i nowe formy egzaminowania.* Wyd. DSWE, Wrocław.

O potrzebie badań nad efektem zwrotnym egzaminu maturalnego z języka polskiego

Artykuł jest próbą włączenia się w debatę o konieczności prowadzenia refleksji nad efektywnością kształcenia polonistycznego w szkole na poziomie ponadgimnazjalnym, popartej szerokimi badaniami empirycznymi. Dokonany przez autorki przegląd zarówno publikacji teoretycznych, jak i doniesień z badań pokazuje szerokie spektrum tematyki podejmowanej w różnych środowiskach, jednocześnie odsłania „białe plamy”, miejsca niezagospodarowane w polskiej diagnostyce edukacyjnej. Ukazuje, że przełożenie wyników prowadzonych badań na konkretne rozwiązania praktyczne, dydaktyczne czy metodyczne, określenie wpływu egzaminów doniosłych na proces uczenia się i nauczania nie doczekało się pogłębionych diagnoz i opracowań. Przedstawiony artykuł nie jest opracowaniem wyczerpującym obszerną problematykę kształcenia polonistycznego w szkole, ale miejmy nadzieję, że pozwoli taki dyskurs rozpocząć.

Słowa kluczowe: egzamin maturalny z języka polskiego, efekt zwrotny

On the need of studies on feedback from matriculation exam in the Polish language

The article is a voice in the debate concerning the need for reflection on research-based effectiveness of teaching Polish at upper-secondary (post-*gimnazjum*) level. The authors' overview of both theoretical and research papers reveals a wide spectrum of topics addressed by various stakeholders. At the same time it also reveals a number of under- or yet not researched areas in Polish educational diagnosis. The overview clearly shows that transferring research translates into practice, delineating the washback effect of high-stakes exams on the teaching-learning process, is yet to become the topic of in-depth analyses and papers. The article by no means provides a definitive discussion on the vast issue of teaching Polish at schools, but it is hoped to provide a good starting point for such a debate.

Keywords: matriculation exam in the Polish language, feedback

Jerzy Chodnicki
 III Liceum Ogólnokształcące w Łomży

DYLEMATY DYDAKTYKI MATEMATYKI CZY DYLEMATY DIAGNOSTYKI?

Początki działalności PTDE zbiegły się w czasie z przemianami ustrojowymi oraz głębokimi zmianami organizacyjnymi i programowymi polskiej szkoły. W awangardzie przemian znalazła się matematyka, w czym największe zasługi mieli profesorowie Stefan Turnau oraz Waclaw Zawadowski. Wystarczy tylko przypomnieć, że pierwszym programem spoza oficjalnej listy ministerialnej, dopuszczonym do użytku szkolnego, była *Matematyka 2001*.

O ile w innych przedmiotach po fali reform nastąpiła niezbędna szkole stabilizacja, to matematyka była i jest reformowana permanentnie. W ciągu dwudziestu lat podstawę programową zmieniano aż sześciokrotnie. Łatwo policzyć, że każdy rocznik przychodzący do szkoły uczył się na podstawie innych programów. Nie ułatwiało to pracy szkole, ale i przed diagnostyką edukacyjną stawiało duże, a często bardzo trudne wyzwania. W tym świetle oczywiste jest, iż sumując dorobek diagnostyki w zakresie matematyki nie sposób uciec od uwikłań dydaktycznych oraz zawirowań programowych.

Meandry nauczania matematyki

W latach sześćdziesiątych ubiegłego wieku przez świat przetoczyła się fala reform nazywana później „nową matematyką”. W drugiej połowie lat siedemdziesiątych w wielu krajach pojawiły się głosy kontestujące „nową matematykę” i podważające sens reform z nią związanych. Na ich gruncie powstał głośny *Raport Cockrofta* (Cockroft, 1982), który stał się inspiracją do poszukiwań nowych rozwiązań programowych i dydaktycznych. Zgodnie z duchem raportu, zmiany dotyczyły innego niż dotychczas rozłożenia treści programowych oraz stylu nauczania matematyki. Najpełniej zostały one wyartykułowane w publikacjach NCTM (1992), które W. Zawadowski (1993, s. 2–3) uważa za „białe księgi” dydaktyki matematyki, a komentując ich zawartość, zwraca uwagę na odmienny od tradycyjnego styl nauczania, osiągnięty poprzez:

1) zwiększenie nacisku na:

- a) wciąganie wszystkich uczniów do działania, indywidualnie lub w grupach przy pracy o charakterze badawczym, stawianiu hipotez, analizowaniu zarówno sytuacji realnych, jak i zagadnień matematycznych,
- b) używanie odpowiednich technologii przy obliczeniach i przy badaniach innego typu,

- c) ułatwianie samodzielnego uczenia się,
 - d) dynamiczne ocenianie procesu indywidualnego uczenia się jako nieodłącznej składowej nauczania i edukacji;
- 2) zmniejszenie nacisku na:
- a) nauczanie obliczeń bez podania kontekstu i musztrę rachunkową,
 - b) podawanie izolowanych faktów, wzorów i reguł, nadmierne akcentowanie zapamiętywania,
 - c) nauczyciela w roli jedyne go źródła wiedzy,
 - d) testy i sprawdziany przeprowadzane tylko w celu postawienia stopni.

Idee te przeniknęły do Polski na przełomie lat osiemdziesiątych i dziewięćdziesiątych ubiegłego wieku, między innymi dzięki Stowarzyszeniu Nauczycieli Matematyki i współpracy międzynarodowej, np. w ramach projektu KREATOR. Powstały nowe programy nauczania, podręczniki i materiały dydaktyczne wykorzystujące nowe technologie. Wydawało się wtedy, że zmiana nauczania matematyki jest tylko kwestią czasu – i tak się stało, zwłaszcza w przypadku szkoły podstawowej. Nastąpił także wyraźny wzrost zainteresowania ocenianiem wspierającym, zarówno od strony teoretycznej, jak i rozwiązań praktycznych, a ocena opisowa w klasach I–III znalazła umocowanie prawne.

Jednak po pewnym czasie fala reform zaczęła wyraźnie słabnąć i wygasać, co W. Zawadowski na XV Konferencji PTDE podsumował w sposób następujący: „Mimo wielkiego zrywu do poprawy edukacji w naszym kraju w latach dziewięćdziesiątych ubiegłego stulecia, wpływ na to, co się dzieje w klasie szkolnej na lekcjach matematyki był i jest niewielki, i nie zawsze działał w dobrym kierunku” (Zawadowski, 2009, s. 40–49).

Wyjaśnienie przyczyn tego zjawiska nie jest proste, albowiem nauczanie matematyki, tak jak każdego innego przedmiotu szkolnego, nigdy nie jest tylko i wyłącznie domeną dydaktyki. Jest ono wypadkową systemu edukacji i postrzegania szkoły w społeczeństwie. Wiele trafnych spostrzeżeń w tym zakresie formułuje Trevor J. Fletcher. Analizując różne style nauczania matematyki, szuka przyczyn, które nastawione na behawiorystyczny przekaz metody nauczania są irytująco skuteczne. Odnajduje je, uwzględniając wychodzący poza dydaktykę czynnik, którym okazują się uwarunkowania społeczne: „«Złe» metody są rzeczywiście bardzo skuteczne, jeśli mierzyć skuteczność zgodnie z powszechnymi oczekiwaniami wobec szkoły. To jest właśnie brakujący czynnik. Nauczyciele wcale nie wybierają złych metod. Oni robią to, czego społeczeństwo od nich oczekuje, i robią to ze znacznym powodzeniem” (Fletcher, 1986, s. 93–108).

W naszym kraju największy wpływ na wyhamowanie reform miał splot takich czynników, jak: zmiana ustroju szkolnego, wprowadzenie systemu egzaminów zewnętrznych i obowiązkowej matury z matematyki, brak społecznego lobby na rzecz edukacji matematycznej. Żaden z nich z osobna nie mógł okazać się decydujący, co więcej ich zharmonizowanie doprowadziłoby do pożądanych zmian w edukacji matematycznej. Niestety, niespójność posunięć i nietrafność rozwiązań spowodowały zapaść nauczania matematyki. Wprowadzenie gimnazjów zaowocowało skróceniem cykli pracy szkół i nauczycieli; stały się one wyraźnie krótkoterminowe, bo zaledwie trzyletnie. Egzaminy zewnętrzne wyznaczyły cele pracy. Powszechność liceów i obowiązkowa matura z matematyki doprowadziły do znaczących cięć treści programowych i bardzo wyraźnego zubożenia wymagań, określanych niekiedy w pośpiechu, o czym świadczy częstość zmian podstawy programowej.

Zmiana treści programowych była konieczna i nie ma nic złego w ich ograniczeniu, ale matematyka jednakowa dla wszystkich nie mogła być dobrym rozwiązaniem. Niefortunne okazało się znaczące spłaszczenie umiejętności, sprowadzenie ich do odtwarzania wyuczonych schematów i izolowanych faktów. Norman L. Webb (Webb, 1999) zauważa – rzecz nie w tym, aby stawiać uczniom wygórowane lub zbyt niskie wymagania, bo obie sytuacje są szkodliwe, ale aby raczej ciągle podnosić podłogę niż opuszczać lub, co gorzej, windować sufit. W polskim systemie edukacji w przypadku matematyki wybrana została jeszcze inna droga – ciągle opuszczania podłogi.

Duży wpływ na styl nauczania matematyki mają egzaminy zewnętrzne. Dobrze pomyślane, mogą być stymulatorem pożądanych zmian. W polskich realiach egzaminy ze względu na ich behawioralny charakter stały się czynnikiem hamującym. Najogólniej rzecz biorąc, można wskazać następujące zaniechania w tym obszarze:

- nie została określona rola egzaminów zewnętrznych w systemie edukacji,
- nie zostały, pomijając stronę werbalną, sprecyzowane cele egzaminu na żadnym poziomie kształcenia,
- nie zostały wystarczająco mocno rozróżnione i precyzyjnie opisane ścieżki kształcenia w poszczególnych typach szkół,
- poziom podstawowy i rozszerzony zostały bardzo słabo rozróżnione i to tylko w zakresie treści, dlatego oba egzaminy są do siebie bliźniaczo podobne,
- systemowi egzaminowania zapewniona została całkowita, niczym nieskrępowana swoboda.

Klienci systemu edukacyjnego szybko się zorientowali w wytworzonej sytuacji i nastąpiło samoograniczenie poziomu kształcenia, co zaowocowało kolejnym obniżeniem wymagań. Mamy do czynienia z bardzo silnym dodatnim sprzężeniem zwrotnym, co pozwala wysunąć tezę, że do cofnięcia reform edukacji matematycznej w Polsce w znaczący sposób przyczynił się system egzaminów zewnętrznych.

Diagnostyka edukacyjna wobec wyzwań edukacji matematycznej

Z zaprezentowanych, chociaż pobieżnych rozważań wynika, iż w matematyce diagnostyka stanęła przed wieloma bardzo trudnymi problemami. W ciągu ostatniego dwudziestolecia można wyróżnić trzy, oczywiście nierozłączne, obszary zainteresowań diagnostyki w zakresie matematyki: zadanie i jego rola w procesach kształcenia oraz egzaminowania, ocenianie wspierające, ocenianie zewnętrzne.

Zadanie jest podstawową „cegielką” edukacji matematycznej. Wielu dydaktyków uważa, że edukacja matematyczna sprowadza się do rozwiązywania różnego rodzaju zadań, stąd tak duże znaczenie tego zagadnienia w teorii i praktyce oceniania. Można przy tym rozróżnić dwie grupy zagadnień: analizy jakościowa i ilościowa uczniowskich rozwiązań oraz problemy związane z budową zadań. Pierwsza z nich cieszyła się dużym zainteresowaniem praktyków i teoretyków oceniania. Dość szybko przeniknęły do Polski metody IRT i z dobrym skutkiem wspierały diagnostykę w obszarach analizy ilościowej wyników testowania. Analizy z wykorzystaniem IRT były także używane do oceny długości skal punktowych stosowanych przy ocenianiu zadań maturalnych (Węziak-Białowolska, 2009, s. 345–354). W tej materii należy jednak postępować z dużą ostrożnością, pamiętając o dwóch czynnikach: zada-

nie musi być zbudowane poprawnie, a nadrzędną rolę odgrywa funkcja dydaktyczna i pomiarowa zadania w teście, zwłaszcza jeśli nie ma on wyłącznie charakteru różnicującego.

Ocenianie zadań otwartych było stałym przedmiotem zainteresowania diagnostyki, zwłaszcza w kontekście egzaminów zewnętrznych. Najczęściej badania dotyczyły efektu egzaminatora (Zapieraczyński, 2010, s. 633–639), znacznie rzadziej problemów związanych z konstrukcją schematów oceniania (Kubat, 2010, s. 618–625). Próba zwiększenia porównywalności oceniania zadań otwartych na egzaminach skłoniła do sięgnięcia, zarówno na świecie, jak i w Polsce, po nowe technologie. Od kilku lat podejmowane są, w tej chwili bardzo już zaawansowane, próby wdrożenia e-oceniania (Szaleniec, 2009, s. 124–134), przy czym matematyka była poligonem doświadczalnym. Bez wątplenia e-ocenianie może dość szybko przyczynić się do zwiększenia rzetelności oceniania, jednak już znaczący wzrost trafności jest wątpliwy bez odpowiednich prac teoretycznych wspierających nowe narzędzie. Wykorzystanie elektronicznego medium stwarza na tyle nową sytuację jakościową, że tradycyjne schematy oceniania mogą słabo przystawać do nowego narzędzia. Ponadto zachowanie istniejących schematów zmienia tylko formę pracy egzaminatorów, nie dotykając meritum. Drugą istotną kwestią e-oceniania jest zatomizowanie oceniania arkusza, co w przypadku matematyki ma bardzo istotne znaczenie. Wystarczy, że np. w trzech zadaniach sprawdzana jest ta sama umiejętność, co wcale nie jest rzadkością, wówczas uczeń, który jej nie opanował, jest „karany” trzykrotnie. Problem można rozwiązać na poziomie agregacji danych, ale to wymaga odmiennie pomyślanych schematów oceniania.

Podczas konferencji diagnostycznych pojawiały się także propozycje nowego spojrzenia na rolę zadania egzaminacyjnego. W tym kontekście warto choćby wspomnieć o ciekawym od strony teoretycznej pojęciu zadań rozrywających (Stożek, 2007). Wnioski wypływające z analiz tego typu zadań wpisują się, w smutne skądinąd, wyniki innych badań wskazujących, że w zakresie edukacji matematycznej nasza szkoła nie przyczynia się do podnoszenia kompetencji uczniów – słabi pozostają słabymi.

Zupełnie natomiast umknęły diagnostyce edukacyjnej tematy kluczowe – konstrukcji zadań matematycznych i procesu ich rozwiązywania. Tylko w kontekście zastosowań IRT Bejar (Fredirksen, 1993) pisał, że jak do tej pory nie były konstruowane zadania spełniające jej wymogi, a dotychczasowe zastosowania pokazywały jedynie jej możliwości i dopiero w ostatnich latach zaczęły pojawić się odpowiednie zadania. W Polsce ta kwestia nie została zauważona. Nie doczekał się także głębszych opracowań proces rozwiązywania zadania otwartego, a bez tego trudno o dobrze zbudowane zadania. Być może dlatego na egzaminach zewnętrznych pojawiają się zadania z błędami lub uchybieniami konstrukcyjnymi, a złe zadania oznaczają zły test, błędne wyniki i nietrafne analizy.

Wiele wystąpień konferencyjnych przestrzegało przed nagminnym stosowaniem zadań zamkniętych i wiarą w ich wartość pomiarową. W przypadku matematyki głosy te należy traktować szczególnie poważnie, albowiem znacząca część arkuszy egzaminacyjnych składa się właśnie z takich zadań. Moim zdaniem sedno problemu kryje się jednak gdzie indziej. W Polsce zadania zamknięte z matematyki konstruowane są tak, jak na świecie robiono to trzydzieści lat temu: dotyczą niższych poziomów taksonomicznych, nastawione są na odtwarzanie faktów i algorytmów. Trudno znaleźć także publikacje, których tematem byłyby zadania zamknięte przeznaczone

do wykorzystania w procesie oceniania wspierającego. Zazwyczaj korzysta się z zadań typowo egzaminacyjnych, ale do celów diagnostycznych ich konstrukcja nie jest odpowiednia.

Tymczasem świat nam uciekł zarówno w konstruowaniu samych zadań (Hala-dyna, 1989, s. 37–50), jak i stosowania zadań zamkniętych wykraczających daleko poza egzaminy (Zhimei, 2011), używania ich do śledzenia rozwoju wyższych procesów myślowych, co nawiązuje do oceniania wspierającego. Jego losy oceniania były podobne do losów reform nauczania matematyki. Na przełomie wieków budziło duże zainteresowanie, które z czasem słabło. O ile było zawsze obecne w głównym nurcie konferencji PTDE, to w przypadku matematyki pojawiała się albo w kontekście edukacji wczesnoszkolnej (Nowik, 2010, s. 697–704), albo też nieśmiały prób rozluźnienia oceniania sumującego (Zawadowski, 2010, s. 56–61). W tym obszarze diagnostyka matematyki nie dotrzymała kroku innym przedmiotom i ma wiele do zrobienia.

Perspektywy i wyzwania diagnostyki edukacyjnej w zakresie matematyki

Przed diagnostyką edukacyjną w zakresie matematyki stoją duże wyzwania. W dającej się przewidzieć przyszłości wyraźnie rysują się następujące tendencje, które można uwzględnić w myśleniu o przyszłych spotkaniach diagnostów:

- wzrost znaczenia oceniania wspierającego,
- pojawienie się nowych jakościowo form oceniania wykorzystujących IT,
- malejąca rola egzaminów zewnętrznych w ich obecnej formie.

W dzisiejszym świecie ciężar uczenia się wyraźnie przesuwają się poza szkołę, gdzie, o ironio – ma charakter konstruktywistyczny. W szkole będziemy mieli do czynienia z uczniami w tym samym wieku, ale na bardzo różnych poziomach, gotowych zdać egzaminy nie na komendę, ale wtedy, kiedy będą do nich gotowi. W związku z tym głównym zadaniem szkoły będzie nie przekazywanie wiedzy, ale organizowanie środowiska pracy ucznia, śledzenie jego postępów i wspieranie rozwoju. Naturalną konsekwencją jest zmiana celów diagnostyki i oceniania. Już dzisiaj W. James Popham (2008) używa terminu *transformative assessment* dla podkreślenia roli, jaką ocenianie powinno odgrywać w procesie ciągłego przeorganizowywania pracy dydaktycznej w klasie.

W naszej szkole ocenianie wspierające nadal w dużej mierze pozostaje w sferze deklaratywnej. Wprowadzenie go wiązałoby się z kłopotliwym i trudnym organizacyjnie obowiązkiem gromadzenia dużej ilości informacji. Nadzieję na zmiany przynoszą dynamicznie rozwijane metody Computer Adaptive Testing, które nawet obecnie pozwalają nie tylko zapanować nad wielką ilością danych, ale przede wszystkim potrafią zdiagnozować mocne i słabe strony ucznia, przekazując mu przy tym unikalną informację zwrotną. Do tej pory narzędzia IT były wykorzystywane w ocenianiu w celu usprawnienia pracy, a komputerowe testy najczęściej były powieleniem papierowych. W najbliższych latach można spodziewać się znaczącego skoku jakościowego, który spowoduje zniknięcie egzaminów zewnętrznych takich, jakie znamy. Przez analogię do gier strategicznych, pojawią się systemy komputerowe o nieliniowej, mocno rozgałęzionej strukturze. Będą one wymagały od ucznia podejmowania decyzji o charakterze strategicznym – związanym z wyborem drogi rozwiązania problemu

lub wielu problemów. Pozwoli to na ocenę nie tylko umiejętności o charakterze odtwórczym czy algorytmicznym, ale także, a może przede wszystkim, wyższych procesów myślowych. Egzamin zewnętrzny przestanie być papierowym testem zdawanym na komendę. Uczeń będzie mógł przystąpić do niego w dogodnym dla niego momencie, jak też wielokrotnie powtarzać próby. Aby świat nam nie uciekł, widzę pilną potrzebę podjęcia studiów związanych z ocenianiem wewnątrzszkolnym nie tylko o charakterze teoretycznym, ale także prac nad wdrożeniem systemów CAT.

Stoją przed diagnostyką także wyzwania małe, ale wcale nie mało znaczące, związane z wypełnieniem „białych plam” w praktyce egzaminów zewnętrznych w naszym kraju. Z pewnością należą do nich problemy związane z konstrukcją zadań, budową schematów oceniania czy wreszcie z budową spójnego od strony merytorycznej systemu obejmującego wszystkie egzaminy zewnętrzne z matematyki. Bez takich prac trudno będzie o sensowne egzaminy.

Literatura

- Cockroft W. H. (1982): *Mathematics Counts*. HMSO, Londyn.
- Fletcher T. J. (1986): *Ulepszanie programów nauczania matematyki w zmieniających się społeczeństwach*. „Roczniki Polskiego Towarzystwa Matematycznego. Seria II: Wadomości matematyczne”, XXV II.
- Fredirixsen R. M. (1993): *Test theory for a new generation of tests*. Lawrence Erlbaum Associates, Hillsdale.
- Haladyna T. M. (1989): *A Taxonomy of Multiple-Choice Item – Writing Rules*. Arizona State University, „Applied Measurement in Education” 2(1),.
- Kubat J. (2010): *Czy nowy klucz punktowania ma wpływ na komunikowanie wyników sprawdzianu 2010 roku?* W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.
- NCTM (1992): *Professional Standards for Teaching Mathematics*. National Council of Teachers of Mathematics, Reston.
- Nowik J. (2010): *Ocena opisowa w edukacji wczesnoszkolnej – diagnoza czy fikcja?* W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.
- Popham W. J. (2008): *Transformative assessment*. ASCD, Alexandria.
- Stożek E. (2007): *Prognostyczna funkcja zadania matematycznego?* W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Wyd. GRUPA TOMAMI, Kraków.
- Szaleniec H. (2009): *E-ocenianie to nie tylko zmiana technologii*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Wyd. GRUPA TOMAMI, Kraków.
- Webb N. L. (1999): *Alignment of Science and Mathematics Standards and Assessments in Four States*. The University of Wisconsin, Wisconsin.
- Weźniak-Białowolska D. (2009): *Analiza skal punktowania zadań – matura z matematyki 2007*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Wyd. GRUPA TOMAMI, Kraków.
- Zapieraczyński M. (2010): *Perspektywy oceniania odpowiedzi do zadań otwartych na egzaminach zewnętrznych – sprawdzianie i egzaminie gimnazjalnym*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.
- Zawadowski W. (1993): *Amerykańskie standardy*. „Nauczyciele i Matematyka”, nr 8.
- Zawadowski W. (2009): *Wpływ współpracy międzynarodowej na nauczanie matematyki w Polsce*. W: B. Niemierko, M. K. Szmigel (red.): *Badania międzynarodowe i wzory zagraniczne w diagnostyce edukacyjnej*. Wyd. GRUPA TOMAMI, Kraków.
- Zawadowski W. (2010): *Ocenianie przez zdobywanie sprawności*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków..
- Zhimei G. (2011): *Maximizing the Potential of Multiple-Choice Items for Cognitive Diagnostic Assessment*. University of Toronto, Toronto.

Dylematy dydaktyki matematyki czy dylematy diagnostyki?

Rozwój diagnostyki edukacyjnej w Polsce w zakresie matematyki zbiegł się z burzliwymi czasami reformowania szkolnej matematyki, kiedy to podstawa programowa w ciągu dwudziestu lat była zmieniana aż sześciokrotnie. Był to także okres wprowadzania egzaminów zewnętrznych. Stawiało to niełatwe zadania przed teoretykami i praktykami oceniania. Artykuł jest próbą spojrzenia na rozwój oceniania zewnętrznego i wewnątrzszkolnego, zwłaszcza działań podejmowanych w ramach PTDE oraz próbą zarysowania wyzwań czekających polską diagnostykę edukacyjną w zakresie matematyki.

Słowa kluczowe: dydaktyka matematyki, ocenianie wspierające, egzaminy zewnętrzne

Teaching-of-mathematics dilemmas or diagnostics dilemmas?

The development of educational diagnostics in Poland within field of mathematics coincided with the turbulent times of reforms in school mathematics when the Core Curriculum had been changed six times within twenty years. It was also the time of introduction of external examinations. This put before theorists and practitioners of evaluation and assessment difficult tasks. This article is an attempt to look at the development of the external and formative assessment in-school, including actions taken by PTDE and an attempt to depict challenges standing before Polish educational diagnostics within the field of mathematics.

Keywords: teaching mathematics, external exam, formative assessment

Teresa Sadoń-Osowiecka
Uniwersytet Gdański

DYDAKTYCZNE DYSKURSY W DIAGNOSTYCZNYCH ROZWAŻANIACH O PRZEDMIOTACH PRZYRODNICZYCH NA KONFERENCJACH DIAGNOSTYKI EDUKACYJNEJ

Wielkie rocznice zazwyczaj wiązane są z wielkimi słowami, którymi podsumowuje się osiągnięcia minionych lat. Podniosły ton może utwierdzać w przekonaniu o nieomyślności i słuszności minionych zdarzeń i konserwować to, co kiedyś było odkrywcze, ale – ze współczesnej perspektywy – nie spełniło pokładanych nadziei lub stało się martwe. Można też odkryć znaczenia kiedyś zapomniane, a ważne dziś. Podsumowania mogą być jednak także okazją do analizy i przyjrzenia się sposobom mówienia oraz poruszonym niegdyś tematom, a dzięki temu ożywić je oraz pośrednio wpłynąć na sposób myślenia i mówienia o teraźniejszości.

W artykule analizuję teksty odnoszące się do edukacji w zakresie szkolnych przedmiotów „przyrodniczych”: biologii, chemii, fizyki, geografii, wygłoszone podczas Konferencji Diagnostyki Edukacyjnej w ciągu ostatnich dwudziestu lat. Podejmuję próbę wyłonienia pewnych kategorii, odnoszących się do treści publikacji, podejść badawczych (świadomych lub nie), używanego języka, sposobu prezentacji zagadnień, odniesień do innych tekstów, nawiązując do „archeologii wiedzy” Michaëla Foucaulta (1977). Tem analiz są inne artykuły publikowane w materiałach konferencyjnych i problemy przez nie poruszone, potraktowane jako odpowiadające czasowi, w jakich powstawały.

Podjęta w niniejszym artykule „analiza lokalnej dyskursywności” (Foucault, 1980, s. 85, za: Scheurich, McKenzie, 2009, s. 299) może pozwolić na odkrycie w diagnostyce edukacyjnej przedmiotów przyrodniczych znaczeń nie do końca uświadomionych. Ponieważ diagnoza edukacyjna musi siłą rzeczy obejmować różnorodne czynniki wpływające na uzyskiwane rezultaty, analiza dotyczy także relacji dyskursu diagnostycznego z innymi sposobami mówienia o edukacji. W założeniu mogą to być badania „gier prawdy” w ramach przemieszczenia teoretycznego dotyczącego przede wszystkim „pytania o formy praktyk dyskursywnych, przez które wypowiada się wiedza (*savoir*)” (tamże), będąca czymś więcej niż tylko „znajomością” zagadnienia (*connaissance*).

Tak rozumiana analiza dyskursu zbliża się do podejścia etnometodologicznego, którego jedną z właściwości jest ujawnienie, jak „stanowiący rozpoznawalne środowisko i znany porządek czynności kontekst maskuje i nie pozwala uczestnikom życia społecznego dostrzec generujących ten porządek praktyk innych uczestników, a zatem pozwala uważać właściwości tego kontekstu, w tym zawarte w nim wytłumaczenia, za «skończone i niezależne» obiekty” (Garfinkel, 2007, s. 354–355). Jednak teksty

powstałe na konferencjach diagnostyki edukacyjnej są wyrazem „znajomości” tematu i jego kontekstu przez ich autorów. Harold Garfinkel pisze: „Dociekania uczestników życia społecznego na nieskończenie wiele sposobów współtworzą kontekst, będący przedmiotem tych dociekań” (tamże, s. 19). To stwierdzenie może mieć szczególne zastosowanie do tekstów poświęconych edukacji. Tym bardziej że uczestnikami i odbiorcami wystąpień na konferencjach diagnostyki edukacyjnej są osoby prowadzące na co dzień działalność naukową, a także praktycy – uczestnicy życia społecznego: nauczyciele, pracownicy ośrodków metodycznych, komisji egzaminacyjnych itp. Różnorodność podejmowanych przez nich tematów i sposobów mówienia może usprawiedliwić potraktowanie tych wypowiedzi jako zdarzeń dyskursywnych w rozumieniu M. Foucaulta, który pisał o tym następująco: „Wypowiedź, choćby najbardziej banalna, choćby najmniej istotna w swoich skutkach, choćby natychmiast zapomniana po swoim zjawieniu się, choćby uznana za mało słyszalną lub źle odczytaną, stanowi zawsze zdarzenie, którego ani język, ani sens nie mogą w całości wyczerpać. Zdarzenie z pewnością niezwykle (...) dlatego również, że choć jednostkowe jak każde zdarzenie, podlega wszakże powtórzeniu, transformacji, reaktywizacji; wreszcie dlatego, że związane jest nie tylko z sytuacjami, które je wywołały, oraz z konsekwencjami, które zapowiada, ale też równocześnie – i to na całkiem innej zasadzie – z wypowiedziami, które je poprzedzają i które po nim następują” (Foucault, 1977, s. 52–53).

Przyjmując taką perspektywę, nie mogą rościć prawa do absolutyzowania swoich analiz w duchu obiektywistycznych podejść. Wiąże się to też z odrzuceniem linearności i pojęcia rozwoju dyscypliny, choć dwudziestolecie konferencji diagnostycznych mogłoby sugerować taką historyczną ciągłość. Badanie dyskursu według M. Foucaulta jest śledzeniem nieciągłości (Foucault, 1977, s. 52), jest z samej istoty interpretatywne. Rola jakościowych badań nad dyskursem wynika z przemian współczesnego świata, „wiąże się z (...) przemianami społecznymi i wizją współczesnego społeczeństwa jako społeczeństwa zdominowanego przez praktyki komunikacyjne oraz przekonaniem, że rzeczywistość jest konstruowana w procesie komunikacji” (Grzymała-Kazłowska, 2004, s. 13).

Diagnostyka edukacyjna

Przed analizowaniem tekstów dotyczących diagnozy w przedmiotach przyrodniczych, niezbędne jest odwołanie się do ogólnych założeń diagnostyki, które mogą stanowić dla nich miarę odniesienia. Bolesław Niemierko (1994a, s. 3) na pierwszej konferencji zdefiniował ją jako „diagnozę osiągnięć uczniów we współczesnych systemach pedagogiki technologicznej i humanistycznej”. Na kolejnych konferencjach spotykały się obydwie podejścia badawcze, mimo że w tym samym tekście B. Niemierko jednoznacznie zadeklarował, że: „Diagnostyka edukacyjna, zajmująca się systematyzowaniem informacji o przebiegu i wynikach uczenia się, ze swej natury należy do kierunku technologicznego w pedagogice” (tamże). W definicji zwraca uwagę posłużenie się terminem „uczenie się”, a nie powszechnie występującym w publikacjach pedagogicznych tamtych czasów zwrotem „nauczanie-uczenie się”, które sugeruje, że za wynik kształcenia odpowiada zarówno nauczyciel („nauczanie”), jak i uczeń („uczenie się”). Jakkolwiek diagnostyka w ujęciu B. Niemierki jest technologiczna, to sam proces uczenia się, jaki bada, jest udziałem odpowiedzialnego, uczą-

czego się, a nie nauczanego podmiotu, ucznia, uwikłanego w złożony kontekst dydaktyczny. Sama gramatyczna forma pojęcia „uczenie się” wpływa na postrzeganie tego procesu jako niedokonanego, niezakończzonego, z racji czego efekty kształcenia nie są jednoznaczne i określone czasowo (por. Zamojski, 2010).

B. Niemierko stwierdził także, że: „Pedagoga interesuje głównie uczenie się przebiegające w zorganizowanym systemie edukacyjnym, będącym celowym układem sytuacji dydaktyczno-wychowawczych. Aczkolwiek regulacyjna rola systemu edukacyjnego bywa przeceniana i jest coraz silniej podważana przez skrajnie humanistyczne kierunki pedagogiki, (...) to jednak, krytykując systemy, ich przeciwnicy wydają się *a rebours* potwierdzać konieczność rozpoznawania procesu uczenia się wraz z systemem, w jakim ono przebiega” (Niemierko, 1994b, s.12). Zatem diagnoza edukacyjna może się odnosić do ucznia, ale dotyczy też całego kontekstu, w tym jakości pracy szkoły, która uczenie się ma umożliwić.

Pojęcie uczenia się z powodu swojej procesualności, zawartej w samym wyrażeniu, sugeruje możliwości zastosowania podejść humanistycznych w diagnostyce edukacyjnej. Być może jest to przejaw relacji między sposobem funkcjonowania wiedzy naukowej (zwanej przez M. Foucaulta *connaissance*) a wiedzą „społeczną”, mądrością (*savoir*), niedającą się zamknąć w jednoznacznych formułach technologicznego myślenia. Wyrazem równoległego funkcjonowania tych dwóch rodzajów wiedzy może być zakończenie artykułu Andrzeja Janowskiego: „Gdy razem z prof. Niemierką kończyliśmy w 1971 roku w Szwecji kurs konstrukcji programów nauczania, głowy mieliśmy pełne formuł statystycznych i «krzywych uczenia się» oraz zastanawialiśmy się, czy dobrze rozumiemy różnice między «formative» a «summative evaluation». Wtedy kierownik kursu, prof. Benjamin Bloom powiedział nam na pożegnanie mniej więcej tak: «a teraz odłóżcie na bok te abstrakcyjne mądrości i pomyślcie o zwykłym dziecku»” (Janowski, 2007).

Zapowiedzią charakteru następnych konferencji jest stwierdzenie B. Niemierki (1994a), że treść tej pierwszej, z 1994 roku, była „treścią rozstrzeloną”. Jej praktyczną realizację znaleźć można w fakcie zapraszania referentów z różnych środowisk naukowych, którzy zajmowali się różnymi dziedzinami pedagogiki, prezentowali bardzo różne poglądy i podejścia badawcze. Wśród nich znaleźli się między innymi (w kolejności alfabetycznej): Dorota Klus-Stańska, Krzysztof Konarzewski, Leszek Korporowicz, Peter Kroppe, Zbigniew Kwieciński, Tadeusz Lewowicki, Maria Mendel, Henryk Mizerek, Heliodor Muszyński, Maciej Sysło, Tomasz Szkudlarek. Wszyscy oni otwierali nowe pola refleksji nad diagnostyką edukacyjną, odległe od technologicznego kierunku, choć nie zawsze skutkujące zmianą perspektywy patrzenia na proces uczenia się i efekty kształcenia u innych uczestników konferencji.

Diagnostyka w tekstach dotyczących przedmiotów przyrodniczych

Różnorodność tematów i metod badawczych diagnostyki jest tłem dla analizy tekstów odnoszących się do przedmiotów przyrodniczych. Można rozważyć, jak wkracza w ten obszar diagnostyka, jak traktowany jest w obrębie tej grupy przedmiotów przebieg i wyniki uczenia się, jakie podejścia badawcze i sposoby postrzegania są reprezentowane, jaka jest ich relacja do ogólnych założeń diagnostyki, naszkicowanych w poprzednim rozdziale.

Dyskurs wiedzy

„Pionierską” pracą poświęconą diagnostyce w zakresie przedmiotów przyrodniczych jest tekst Henryka Szaleńca (1994), analizującego umiejętności uczniów szkół podstawowych w zakresie posługiwania się wykresem na lekcjach fizyki. Autor, powołując się w nim na Piagetowską teorię rozwoju poznawczego, określa formalizację matematyczną jako najwyższą formę przedstawiania wiedzy. Według niego, rozwój poznawczy w dziedzinie fizyki przebiega od przedstawień czynnościowych, poprzez graficzne, do badania związków czysto formalnych „pomiędzy symbolami reprezentującymi wielkości fizyczne. Zaczyna dominować funkcyjny opis prawidłowości przyrodniczych oraz jego graficzne przedstawienie za pomocą diagramów i wykresów” (Szaleniec, 1994, s. 155), przy czym „pojęcia i operacje nie pojawiają się bez wcześniejszego przygotowania polegającego na przejściu przez łańcuch zachowań bardziej prymitywnych, w stosunku do których są one tylko nowym zróżnicowaniem i innym układem. (...) Wykres zależności pomiędzy wielkościami, należący już do języka sformalizowanego, stanowi istotny pomost pomiędzy drugą oraz trzecią formą przedstawiania wiedzy i powinien (...) poprzedzać zapis algebraiczny w procesie dochodzenia do prawa czy innego związku” (tamże).

Przeciw wagą dla definitywnego, formalnego i normatywnego ujęcia rozwoju wiedzy uczniów na lekcjach fizyki jako hierarchicznego i kumulatywnego – od „prymitywnego” doświadczania konkretnych zjawisk do matematycznej abstrakcji – jest postawienie na pierwszym miejscu eksperymentowania i doświadczania zjawisk na lekcjach chemii. Prowadzi ono do zrozumienia ich istoty przez uczniów. Jak pisze Elżbieta Kowalik (2001, 313): „Eksperymentowanie to kwintesencja chemii. Poznawanie chemii jako nauki o symbolach i wzorach jest pomysłem absurdalnym i zbliża nas raczej do nauczania języka obcego, niż do dydaktyki przyrodoznawstwa”. Abstrahowanie kojarzone jest raczej z myśleniem, głębszym rozumieniem procesów chemicznych, niż z ich matematyczną formalizacją, choć jej nie odrzuca. Inne też jest znaczenie myślenia abstrakcyjnego. Symbole i wzory traktowane są tylko jako forma języka umożliwiająca opis rzeczywistości, a nie sedno i ukoronowanie uczenia się chemii. Chemia, jako nauka „przyrodnicza”, dotyczy „konkretnej”, istniejącej realnie przyrody, nie abstrakcji.

W omawianym tekście pojawia się też krytyka specyficznego werbalizmu, przejawiającego się w zadaniach opisujących procesy chemiczne. Oczekuje się w nich od ucznia, by ten wyobraził sobie przedstawione sytuacje i przewidział rezultat bez naocznego sprawdzenia przebiegu reakcji: „Wątpliwa jest trafność takich pseudolaboratoryjnych zadań dla ucznia gimnazjum, który nigdy na lekcjach chemii nie trzymał w ręku próbowki ani zlewki. Można uznać, że przytoczone wyżej przykłady zadań w pierwszej kolejności sprawdzają umiejętność rozumienia czytanego tekstu” (tamże, s. 314).

Charakter wiedzy szkolnej w liceach poddawany jest krytyce także przez E. Kowalik. Twierdzi ona, że: „Kształcenie erudycyjne w liceach ogólnokształcących zbyt często bliższe jest kultowi rozumu, a więc spuściznie modernizmu, w której szkoła tworzy kanon wiedzy uporządkowanej i zhierarchizowanej według ścisłych reguł, przez odwołanie się do algorytmów, częstego korzystania z gotowych wzorców, a czasem sztywnych i jednolitych rozwiązań metodycznych, przy jednoczesnej rezygnacji z dialogu. Respektowanie postulatów konstruktywistów oznacza odchodzenie

od modelu transmisyjnego wiedzy na rzecz modelu interaktywnego, opartego na aktywności badawczej młodzieży. W modelu tym większego znaczenia nabiera wiedza ucznia zdobyta poza szkołą” (Kowalik, 2005, s. 308).

Zdobywanie wiedzy poza szkołą, czy odrzucenie kanonu uporządkowanej wiedzy – odniesione do poglądów B. Niemierki (1994b) – pozostaje w pewnej opozycji do koncepcji diagnostyki edukacyjnej jako subdyscypliny badającej „uczenie się, przebiegające w zorganizowanym systemie edukacyjnym, będącym celowym układem sytuacji dydaktyczno-wychowawczych” (tamże, s. 12).

Z kolei przykładem postulowania całościowego traktowania szkolnej wiedzy, jako ludzkiego sposobu poznania, są teksty zredagowane przez współautorów: Stanisława Jakubowicza, Stanisława Plebańskiego, Kornelię Rybicką i Beatę Udzik (2005, 2006). Udowadniają oni, jak zmienia się jakość wiedzy przyrodniczej, jeśli zestawiona zostanie z tekstami literackimi, i jak pogłębiona zostaje interpretacja tekstów literackich, gdy studenci polonistyki zrozumieją istotę zjawisk przyrodniczych. Autorzy zaprzeczają tym samym potocznym poglądom o rozdzielności wiedzy przyrodniczej i humanistycznej. Zbliżają się do idei konsilencji nauki, definiowanej jako „poszukiwanie wzajemnych zależności i związków między zjawiskami badanymi przez specjalistów z różnych dyscyplin naukowych, wykorzystujących różne, ale właściwe tym dyscyplinom metody i narzędzia badawcze” (Encyklopedia PWN, <http://encyklopedia.pwn.pl>). Stwierdzają także, że: „W obiegu opinii odbiór tekstów naukowych jest prawie w całości intelektualny, natomiast poezji emocjonalny. Czy można jednak założyć, że przy poezji narzędziem czytelnika będzie wyłącznie interpretacja, a wyjaśnianie zostanie wyłącznie przy tekstach naukowych?” (Jakubowicz i in., 2006, s. 77). Pośrednio wskazują zatem na zubażanie rozumienia znaczeń przez zatomizowanie szkolnej wiedzy i określanie „dozwolonych” znaczeń w obrębie danej dziedziny.

Holistyczne ujęcie zostało jednak zakłócone przez odwołanie się w analizowanym tekście do piramidy faktów, która, tworzona przez wiedzących „lepiej”, zawsze narzuca sposób postrzegania świata, zakreśla granice wiedzy prawomocnej i – mówiąc o faktach potrzebnych „wykształconemu” człowiekowi – nie mówi o wiedzy jako całości, atomizuje ją. W tym kontekście wydaje się, że globalna piramida faktów, a szczególnie jej wierzchołek (informacje niezbędne wykształconemu człowiekowi) możliwy jest do tworzenia w czasie nieustającej dyskusji. Można założyć, że elementy wiedzy przyrodniczej występujące w literaturze pięknej (np. lektury szkolne, uznane arcydzieła literatury, literatura noblowska) tworzą ten szczyt globalnej piramidy faktów. Ich znajomość może stać się cechą człowieka uważającego się za wykształconego” (tamże). Jakkolwiek mówi się w artykule o dialogu, nieustającej dyskusji nad faktami, zawsze są to decyzje podejmowane przez kogoś za kogoś, dlatego nie wynikają z potrzeb tych, którym wiedza ma służyć. Wiedza zaś, według założeń konstruktywizmu, na który autorzy się powołują, nie jest zbiorem faktów, ale sposobem ich funkcjonowania w jednostkowym umyśle.

Zupełnie inaczej całościowe ujęcia wiedzy przyrodniczej traktuje Elżbieta Tyrańska-Wojtyca, lokując je w ustanowieniu jednakowych standardów wymagań egzaminacyjnych części matematyczno-przyrodniczej egzaminu gimnazjalnego (obowiązujących od roku 1999 do 2008). Autorka pisze: „Standardy wymagań egzaminacyjnych były absolutnym *novum* i postępem w opracowywaniu i interpretacji wyników egzaminacyjnych uczniów w polskiej szkole. Dzięki nim nastąpił rozwój, wdrożyliśmy się do całościowego – choć nie zawsze wygodnego – postrzegania i analizowania

umiejętności i wiadomości uczniów w grupie przedmiotów matematyczno-przyrodniczych” (Tyralska-Wojtycza, 2011, s. 497).

Symptomatyczne w tym przypadku jest rozdzielenie wiadomości i umiejętności, co samo w sobie jest zaprzeczeniem jedności wiedzy odnoszącej się do całości doświadczenia. Zwracał na to uwagę Krzysztof Konarzewski, krytykując sformułowania standardów egzaminacyjnych w obszarze przedmiotów matematyczno-przyrodniczych: „Pisząc «wiedza i umiejętności», autorzy błędnie sugerują, że umiejętność nie jest wiedzą” (Konarzewski, 2003, s. 44). Tymczasem E. Tyralska-Wojtycza odnotowuje odchodzenie w szkole od specyficznego pojmowanego holizmu: „Wkrótce stało się oczywiste, że badanie wyników egzaminu gimnazjalnego przez pryzmat przedmiotowy jest bliższe praktyce szkolnej i znacznie wygodniejsze w analizie osiągnięć edukacyjnych uczniów” (Tyralska-Wojtycza, 2011, s. 497). W tej perspektywie wiedza jako kategoria semantyczna i jako sensowna z punktu widzenia uczenia się ucznia (zob.: Jakubowicz i in., 2005, 2006) wydaje się nieużyteczna z punktu widzenia osoby analizującej osiągnięcia uczniów, dla której ważny jest pragmatyzm i takie dostosowanie zadań, by ułatwić pracę sprawdzającym i analizującym osiągnięcia uczniów.

Dyskurs „uczenia się”

Zadziwiająco, jak nieproporcjonalnie rzadko na konferencjach diagnostycznych pojawiało się zagadnienie procesu uczenia się przedmiotów przyrodniczych, choć powinno ono stanowić, oprócz wyników tego procesu, przedmiot analiz istotny dla diagnostyki edukacyjnej. Źródeł tego stanu rzeczy można upatrywać w fakcie, że z punktu widzenia technologicznej wersji diagnostyki łatwiej zmierzyć efekt końcowy – wynik nauczania się zadanych treści, niż dynamiczny, podlegający wielu czynnikom proces uczenia się.

Niemniej niektóre teksty, podane wcześniej jako przykłady w charakteryzowaniu dyskursu wiedzy, mówią też o uczeniu się. Często jawnie lub w zawołany sposób nawiązują one do konstruktywistycznej wizji uczenia się. I tak, E. Kowalik diagnozuje przebieg uczenia się chemii w liceach z punktu widzenia konstruktywizmu. Podaje przykłady zadań chemicznych, których rozwiązanie mogłoby być efektem samodzielnych poszukiwań i rozmów z nauczycielem. Jednocześnie ma świadomość, że „ocenie osiągnięć ucznia sprawdzanych powyższym zadaniem oznacza podejście holistyczne, wyłamujące się z kanonu przewidywalnego instrukcją punktowania, na rzecz dostrzegania uczniowskich zachowań samodzielnych, kreatywnych, słowem: konstruktywistycznych” (Kowalik, 2005, s. 309). Autorka zwraca uwagę na przyciasny gorset formalnych warunków, powodujących że „na każdej lekcji uczniowie wyposażeni są w wiedzę faktograficzną, której opanowanie jest warunkiem koniecznym nadszycia za wartkim tempem podstawy programowej” (tamże, s. 310).

Podejście konstruktywistyczne przejawia się też pośrednio poprzez pokazanie innej, brytyjskiej kultury edukacyjnej, dla której ilustracją są zadania, jakie rozwiązują brytyjscy uczniowie (Ochendusko, 1999).

Przeciwstawiając przebieg uczenia się w dwóch szkołach, reprezentujących dwa modele edukacji, Maria Groenwald (1999) pokazała, jak sposób podejścia do wiedzy, do uczniów może wpływać na późniejsze mierzalne efekty edukacyjne. W szko-

le realizującej pierwszy z tych modeli, opartym na rygorze, „atmosferze strachu, uczniowie reagowali swoistą «bezradnością umysłową»” (Groenwald, 1999, s. 81, 82) i realizacją wymogu pamiętania wszystkich faktów. Z kolei w drugim – opartym na wspólnej pracy uczniów i nauczyciela, pozwalano na niekonwencjonalne rozwiązania, rozbudzano motywację: „na lekcjach geografii panowała atmosfera autentycznej, dobrej pracy” (tamże).

Jednocześnie warto zwrócić uwagę, że nie wszystkie czynniki związane z procesem kształcenia da się wymierzyć, a czasem bardziej adekwatny jest jakościowy opis zdarzeń podczas lekcji. Taka koncepcja badawcza nawiązuje do „potrzeby przełamania redukcji w pojmowaniu badań ewaluacyjnych w edukacji, (...) [co] niemal od początku eliminuje tendencje do redukcyjnego ich pojmowania ze względu na potrzebę określenia nie tylko przedmiotu działania, a więc tego «Co jest czynione?», ale także jego podmiotu, a więc «Kto czyni?», i niemal automatycznie: «Po co czyni?», «Dla kogo czyni?», «Z kim czyni?», «Gdzie czyni?», «W jakim środowisku czyni?» itd., a więc prowokuje do ujęcia zintegrowanego, wieloczynnikowego, intencjonalnego, wkomponowanego w związku z innymi działaniami” (Korporowicz, 2012, s. 24, 25).

W niewielu tekstach pojawia się zainteresowanie samym uczniem jako podmiotem, a nie narzędziem do produkcji pożądanych efektów edukacyjnych. Jeden z wniosków z badań dotyczących motywacji uczenia się botaniki wskazuje na regres motywacji do uczenia się jako częste zjawisko „związane ze zdobywaniem doświadczenia i uzależnieniem się od przyjętych schematów” (Suska-Wróbel, 1999, s. 76). Koresponduje to z wnioskami dotyczącymi uczenia się chemii jako zbioru faktów i algorytmów.

Jeszcze mniej tekstów zajmuje się problemami etycznymi na lekcjach przedmiotów przyrodniczych. Bodaj jedyne to artykuły Anny Kossobuckiej, a wśród nich jeden pod znaczącym tytułem: *Wychowanie moralne na lekcjach biologii – zniewolenie czy wyzwolenie?* (2007).

Zakończenie

Przedstawiane dyskursy wiedzy i uczenia się przedmiotów przyrodniczych nie są bardzo zauważalne na konferencjach diagnostyki edukacyjnej. Dominuje bowiem podczas nich dyskurs, który można nazwać: dyskursem efektów oraz dyskursem technicznym, charakterystycznym dla diagnostyki, która sytuuje się w technologicznym nurcie pedagogiki. Sama idea diagnostyki wydaje się także ewoluować, zgodnie z charakterystyką programu jednej z konferencji: „Fundamentalną warstwę konferencji stanowią rozprawy i doniesienia charakteryzujące obecne poglądy na ocenianie oraz na przebieg egzaminów o różnym zasięgu” (Szmigielski, Niemierko, 2010, s. 11). Skupienie się na tych dwóch aspektach ponownie odsuwa w cień procesy uczenia się ucznia i namysł nad charakterem wiedzy, którą ma się oceniać na egzaminach.

Warto też przypomnieć, że zmiany to nie zawsze postęp, to tylko zmiany, wbrew entuzjastycznym deklaracjom: „Tak oto na naszych oczach zachodzi kolejny proces rozwoju szkoły, a tym samym systemu egzaminów zewnętrznych” (Tyrańska-Wojtyca, 2011, s. 497). Każde to poddawać zmiany refleksji, nie czekając na bodziec z zewnątrz, który może uwalniać od myślenia. Jeśli chcemy wychować samodzielnych, myślących – czasem niezgodnie z naszymi oczekiwaniami – uczniów, sami nie

pozbawiamy się wątpliwości. W tym kontekście problematyczne staje się czekanie na odgórne zalecenia czy bodźce płynące ze strony władzy, którą to postawę nauczycieli przybliża następująca wypowiedź: „By móc kontynuować rozważania na ten temat, zastanawiać się, na czym polega postęp/rozwój egzaminu, w tej kwestii, należy jeszcze poczekać do chwili, aż będziemy wiedzieć, jak planuje to rozwiązać Centralna Komisja Egzaminacyjna. Są jeszcze inne pytania, dotyczące choćby organizacji tego egzaminu, w tym zakresu przyrodniczego, czasu trwania poszczególnych zakresów, dostosowań, interpretacji wyników itd. Wszystkie one z czasem przestaną być pytaniami i będziemy mogli dokonać ewaluacji nowej formuły egzaminu gimnazjalnego i w całej rozciągłości stwierdzić, na czym polega jego rozwój” (tamże, s. 504). Na ograniczenia tak zarysowanej koncepcji wskazuje inne podejście, zgodne z którym „specyfika licznych analiz prowadzonych w naukach społecznych (...) polega na interdyscyplinarnym otwarciu i traktowaniu analizowanego zagadnienia jako złożonego interpretacyjnie” (Klus-Stańska, 2010, s.13).

Moim zamierzeniem było przedstawienie nieciągłości, zjawisk ważnych, ale niedostrzeganych, sprzeczności, które pojawiają się w pozornej zgodności. Zapewne bardziej eleganckie byłoby wydzielenie rozłącznych kategorii, do których można by przyporządkować poszczególne teksty. Ale posługiwanie się archeologią wiedzy zakłada też, że żadna kategoryzacja nie jest kategoryzacją naturalną. Jest założona z góry. Nawet w obrębie wydzielonych przez mnie kategorii pojawiła się niejednorodność. Można by ją nazwać pluralizmem, który oznacza „otwarcie na wielość i różnice, na tolerancję, ale zarazem stawia problem «końca tożsamości», wytwarza zgiełk «nowej wieży Babel»; nostalgię za przejrzystą przeszłością, tęsknotę do prostych jednoznacznych wykładni normatywnych...” (Kwieciński, 2010, s. 25).

Literatura

- Foucault M. (1977): *Archeologia wiedzy*, tłum. A. Siemek. PIW, Warszawa.
- Garfinkel H. (2007): *Studia z etnometodologii*, tłum. A. Szulżycka. Wyd. Naukowe PWN, Warszawa.
- Groenwald M. (1999): *Co się zdarzyło w dwóch szkołach czyli rzecz o ewaluacji modułowego programu kształcenia*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Wyd. ODESiP, Legnica.
- Grzymała-Kazłowska A. (2004): *Socjologicznie zorientowana analiza dyskursu na tle współczesnych badań nad dyskursem*. „Kultura i Społeczeństwo” nr 48(1).
- Encyklopedia PWN*, <http://encyklopedia.pwn.pl/haslo/3925190/konsiliencja.html> (2.03.2014).
- Jakubowicz S. i in., (2005): *Dwie perspektywy – jedna droga. Fizyka w tekstach literackich*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej – perspektywy informacyjne egzaminów szkolnych*. Wyd. Fundacja Rozwoju Uniwersytetu Gdańskiego, Gdańsk.
- Jakubowicz S. i in. (2006): *Komplementarność wiedzy wyjaśniającej i interpretacyjnej w przygotowaniu uczniów do egzaminów maturalnych*. W: B. Niemierko, M.K. Szmigel (red.): *O wyższą jakość egzaminów szkolnych. Etyka egzaminacyjna i zagadnienia ogólne*. Wyd. GRUPA TOMAMI, Kraków.
- Janowski A. (2007): *Poznanie uczniów jako podstawa indywidualizacji kształcenia*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Wyd. GRUPA TOMAMI, Kraków.
- Klus-Stańska D. (2010): *Dydaktyka wobec chaosu pojęć i znaczeń*. Wyd. Akademickie „Żak”, Warszawa.
- Konarzewski K. (2003): *Egzamin 2002: rzetelność i trafność testu matematyczno-przyrodniczego*. W: B. Niemierko (red.): *Trafność pomiaru jako podstawa obiektywizacji egzaminów szkolnych*. Wyd. Wyższej Szkoły Humanistyczno-Ekonomicznej w Łodzi, Łódź.
- Korporowicz L. (2012): *Refleksja jako działanie. O stałej potrzebie przelamywania redukcji w pojmowaniu badań ewaluacyjnych w edukacji*. W: B. Niemierko, M. K. Szmigel (red.): *Regionalne i lokalne diagnozy edukacyjne*. Wyd. GRUPA TOMAMI, Kraków.

Kosobucka A. (2007): *Wychowanie moralne na lekcjach biologii – zniewolenie czy wyzwolenie?* W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach uczniów*. Wyd. GRUPA TOMAMI, Kraków.

Kowalik E. (2001): *Trafność testu laboratoryjnego wobec podstawy programowej i standardów wymagań egzaminacyjnych z chemii w liceum ogólnokształcącym*. W: B. Niemierko, K. Małecki (red.): *Dawne i nowe formy egzaminowania*. Wyd. Dolnośląskiej Szkoły Wyższej, Wrocław.

Kowalik E. (2004): *Standardy edukacyjne w kształceniu chemicznym*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. Wyd. GRUPA TOMAMI, Kraków

Kowalik E. (2005): *Ocenianie osiągnięć uczniów a konstruktywizm w kształceniu chemicznym*. W: B. Niemierko, G. Szyling (red.): *Holistyczne i analityczne metody diagnostyki edukacyjnej – perspektywy informacyjne egzaminów szkolnych*. Wyd. Fundacja Rozwoju Uniwersytetu Gdańskiego, Gdańsk.

Kwieciński Z. (2010): *Zmiany w edukacji – próba bilansu dwóch dekad*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.

Niemierko B. (1994a): *Międzynarodowe seminarium na temat „Diagnostyka edukacyjna”*. Informacja ogólna i przegląd dorobku. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Niemierko B. (1994b): *Diagnostyka edukacyjna*. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Ochenduszo J. (1999): *Zróżnicowanie zadań biologicznych w angielskich i walijskich badaniach osiągnięć przyrodniczych (SCAA)*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Wyd. ODESiP, Legnica.

Scheurich J.J., McKenzie K.B. (2009): *Metodologie Foucaulta. Archeologia i genealogia*, tłum. P. Pluciński. W: N. K. Denzin, Y. S. Lincoln (red.): *Metody badań jakościowych*. Tom 2. Wyd. Naukowe PWN, Warszawa.

Suska-Wróbel R. (1999): *Diagnoza motywacji uczenia się przedmiotu szkolnego uczniów szkół średnich ogólnokształcących*. W: B. Niemierko, B. Machowska (red.): *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*. Wyd. ODESiP, Legnica.

Szaleniec H. (1994): *Umiejętności uczniów szkół podstawowych w dziedzinie posługiwania się wykresem na lekcjach fizyki*. W: B. Niemierko (red.): *Diagnostyka edukacyjna*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.

Szmigel M. K., Niemierko B. (2010): *Przedmowa*. W: B. Niemierko, M. K. Szmigel (red.): *Teraźniejszość i przyszłość oceniania szkolnego*. Wyd. GRUPA TOMAMI, Kraków.

Tyralska-Wojtyca E. (2011): *Nowa formuła egzaminu gimnazjalnego, czyli wiele hałasu o tak niewiele – na przykładzie przedmiotów przyrodniczych*. W: B. Niemierko, M. K. Szmigel (red.): *Ewaluacja w edukacji. Koncepcje, metody, perspektywy*. Wyd. GRUPA TOMAMI, Kraków.

Dydaktyczne dyskursy w diagnostycznych rozważaniach o przedmiotach przyrodniczych na Konferencjach Diagnostyki Edukacyjnej

W artykule autorka analizuje teksty odnoszące się do edukacji w zakresie szkolnych przedmiotów „przyrodniczych” (biologii, chemii, fizyki, geografii) będące efektem Konferencji Diagnostyki Edukacyjnej w ciągu ostatnich dziesięciu lat. Podejmuje próbę pewnych kategoryzacji odnoszących się do treści publikacji, podejść badawczych (świadomych lub nie), używanego języka, sposobu prezentacji zagadnień, odniesień do innych tekstów. Tło stanowią inne artykuły publikowane w materiałach Konferencji Diagnostyki Edukacyjnej i problemy przez nie poruszane, potraktowane jako odpowiadające czasowi, w jakich powstawały. Podjęta „analiza lokalnej dyskursywności” może pozwolić na zrekonstruowanie nie do końca świadomych przesłanek w diagnostyce, czy szerzej – w edukacji przedmiotów przyrodniczych. W założeniu mogą to być badania „gier prawdy” w ramach przemieszczenia teoretycznego dotyczącego przede wszystkim „pytania o formy praktyk dyskursywnych, przez które wypowiedzi się wiedza (*savoir*)”, będąca czymś więcej niż tylko „znajomością” zagadnienia (*connaissance*, za: M. Foucaultem). W ten sposób ujęta analiza dyskursu zbliżona jest do podejścia etnometodologicznego.

Słowa kluczowe: dyskursy edukacyjne, przedmioty przyrodnicze, diagnostyka edukacyjna

Didactic discourse in diagnostic rumination on natural sciences at the Conference of Educational Diagnostics

In that paper the author analyses articles on science education at the school (in school subjects: biology, chemistry, physics and geography), which are effects of the Conferences of Educational Diagnostics from last ten years. The author attempts some categorizations concerning contents of publications, research approaches (conscious or not), usage of language, ways of problem presentations, relations with other texts. Research problems in other articles (e.g. on education in general) published in materials of Conferences Educational Diagnostics and suiting to the specific time provide a background for the analysis. This “analysis of local discursion” can enable reconstruction of not so conscious circumstances of diagnostics of science education or even of school education of science in general. This can be research of “games of truth” in frames of theoretical dislocation. This theoretical dislocation applies predominantly to the question on the form of discursive practices by which knowledge (*savoir* by Foucault) is expressed. It is more than only awareness of the problem (*connaissance* by Foucault). This way of analyses is similar to the etnometodological approach.

Keywords: educational discursions, school science’s subjects, educational diagnostics

II. Nowe obszary zainteresowań

RUCH PEDAGOGICZNY 2/2014

Małgorzata Żytko
Uniwersytet Warszawski

GOTOWOŚĆ DZIECI I GOTOWOŚĆ SZKOŁY DO UCZENIA SIĘ Rozważania wokół problematyki diagnozowania jakości wczesnej edukacji

Problematyka jakości pracy szkoły we wczesnej edukacji powraca w debatach naukowych, rzadziej publicznych, zwłaszcza w okresach wprowadzania kolejnych reform i zmian na tym etapie kształcenia. Decyzje podjęte przez MEN, w tym realizowanie od 2014 roku obowiązku szkolnego przez dzieci 6-letnie, a także projektowanie zmian związanych z podręcznikami dla edukacji elementarnej, wzbudzają dyskusje wśród pedagogów i psychologów, ale również wśród rodziców, którzy są coraz intensywniej zainteresowani jakością oferty edukacyjnej proponowanej ich dzieciom na początkowym etapie szkoły.

W takim kontekście doprecyzowania wymaga zagadnienie jakości wczesnej edukacji oraz podjęcie próby określenia tych działań instytucji edukacyjnych, które służą rozwojowi i dobru dziecka. Odnoszę wrażenie, że w debacie publicznej akurat to zagadnienie jest przyjmowane jako oczywiste, choć takim zdecydowanie nie jest. Dyskusję wokół tego problemu podejmują badacze wczesnej edukacji w Polsce, lecz trudno dostrzec związek między ich analizami i rekomendacjami a działaniami władz oświatowych regulujących praktykę szkolną. Wydaje się, że świat współczesnej wiedzy naukowej na temat potrzeb rozwojowych i procesu uczenia się dzieci w wieku wczesnoszkolnym oraz świat działań i praktyk istniejących w szkolnej rzeczywistości są absolutnie rozbieżne. Praktyka rządzi się tradycyjnymi, sformalizowanymi, oderwanymi od współczesnej wiedzy naukowej prawami. Trudno w niej dostrzec gotowość do zmiany, a działania resortu edukacji wzmocniają tylko *status quo* szkoły. Także pojęcie jakości w edukacji jest interpretowane w polskiej praktyce szkolnej w dość jednoznaczny sposób. Ma ono związek z normalizacyjną, zestandaryzowaną koncepcją jakości, którą charakteryzuje dominacja łatwo mierzalnych wskaźników efektywności działania instytucji edukacyjnych.

Inną perspektywę przyjmują Gunilla Dahlberg, Peter Moss i Alan Pence, autorzy znakomitej książki: *Poza dyskursem jakości w instytucjach wczesnej opieki i edukacji* (2013). Ich podejście ilustruje następujące stwierdzenie: „Gdy wyglądasz tylko tego, czego się spodziewasz, gdy cenisz tylko wyznaczone cele, wiele tracisz!” (tamże, s. 24) Pokazują oni odmienny od tradycyjnego paradygmatu analizowania pojęcia jakości,

podając w wątpliwość podejście normalizacyjne, a więc rozumienie jakości w kategoriach oceny działania instytucji pod kątem zgodności z przygotowanymi wcześniej zewnętrznymi wytycznymi, ramą interpretacyjną czy kategoriami, do których trzeba się dopasować. Jakość w takim rozumieniu, zdaniem autorów, „oznacza pracę w ramach cudzej perspektywy, przedstawianej nam jako obiektywna prawda” (Dahlberg i in., 2013, s. 25). Proponowana przez nich nowa orientacja w spojrzeniu na jakość instytucji wczesnej edukacji wiąże się z podejściem określanym jako nadawanie znaczeń. Oznacza ono konieczność wypracowania własnej perspektywy analizowania instytucji edukacyjnej, perspektywy, która będzie miała subiektywny charakter, ale za to będzie usytuowana w konkretnym kontekście praktyki edukacyjnej oraz pozwoli pogłębić jej rozumienie poprzez zadawanie pytań i problematyzowanie jej funkcjonowania. Efektem takiego podejścia powinno stać się uświadomienie sobie przez uczestników działalności pedagogicznej w tej instytucji sensu jej funkcjonowania, koncepcji podejmowanych działań i ich konsekwencji, a także wzięcie za nie odpowiedzialności, przy jednoczesnym troszczeniu się o refleksyjne i świadome nastawienie do codzienności edukacyjnej. Podejście takie wymaga również przyjęcia perspektywy językowej innej niż instrumentalna czy technologiczna, stosowana w normalizacyjnym dyskursie jakości. Orientacja na rozumienie jakości instytucji wczesnej edukacji w kategoriach nadawania znaczeń wiąże się ze zmianą ich funkcji społecznej. To nie są instytucje, w których odbywa się proces przekształcania dzieci w określony produkt, spełniający parametry wyznaczone wcześniej w modelu absolwenta. W tej zmienionej orientacji instytucja edukacyjna staje się miejscem dyskusji, wymiany myśli, zaangażowania się w zmiany wszystkich podmiotów – uczniów, nauczycieli, rodziców oraz przedstawicieli środowiska lokalnego.

Hillew Lenz Taguchi (2010), analizując dominujący dyskurs w podejściu do wczesnej edukacji, zwraca uwagę na swoisty paradoks: liberalny humanizm, który kładzie duży nacisk na wspieranie indywidualnego rozwoju, *de facto* godzi się na podporządkowanie go uniwersalnemu standardom, wspólnym dla wszystkich, normalizującym miarę jakości i ocenianie efektywności edukacji. Podobnie teorie psychologiczne: poznawcze oraz konstruktywistyczne, preferują uniwersalne stadia czy etapy rozwoju i podporządkowują im proces uczenia się. Jak stwierdza H. L. Taguchi, dominujące dotychczas rozumienie socjalizacji jako istotnego czynnika rozwojowego, zostaje zastąpione planową inkulturacją, obejmującą każde dziecko. Brak akceptacji dla tej dominującej perspektywy spojrzenia na wczesną edukację prowokuje do zmiany (tamże, 2010, s. 29–30), wskazania innego dyskursu, innej perspektywy myślenia o procesie uczenia się i rozwoju. Zmiana ta oznacza przejście:

- od celów i wartości z góry zadekretowanych i założonych do działań (*in the middle of thinks*) wymagających współpracy, negocjowania i renegocjowania znaczeń,
- od wertykalnego i hierarchicznego układu relacji między ludźmi i rzeczywistością do horyzontalnej i partnerskiej relacji, w której trudno odseparować nauczającego od uczącego się,
- od podmiotu indywidualnie i samotnie uczącego się oraz rozwijającego do uczenia się we współpracy i tworzenia znaczeń we wspólnym działaniu,
- od reprezentacji świata w języku do tworzenia nowych idei, koncepcji, pomysłów, problemów wartych zbadania i opracowania adekwatnych do nich strategii działania,
- od identyfikowania braków i zaburzeń do afirmacji zróżnicowania i inwencji twórczej,

- od uniwersalnego pomiaru efektywności uczenia się do afirmacji, pojedynczych podmiotów i zdarzeń oraz kontekstu ich materialnego i społecznego rozwoju.

H. Taguchi stwierdza, że przyjęcie właśnie takiej perspektywy wymaga odejścia od myślenia o rozwoju i edukacji dzieci w kategoriach założonych efektów i etapów, a więc niejako od końca. W to miejsce przyjmuje perspektywę eksponującą społeczną naturę konstruowania wiedzy, nacisk na pogłębianie rozumienia tego, co teraz się dzieje w dialogu i refleksyjnej analizie z innymi, z akceptacją tymczasowości i niepewności.

Gotowość szkolna w perspektywie normalizacyjnej

Jak w kontekście tego nowego podejścia do jakości wczesnej edukacji sytuuje się diagnostyka pedagogiczna na poziomie edukacji przedszkolnej i wczesnoszkolnej, a w szczególności na etapie przejścia dziecka z kultury edukacyjnej przedszkola do szkoły? Zgodnie z zapisami podstawy programowej wychowania przedszkolnego (DzU 2012, nr 0, poz. 997) nauczyciele są zobowiązani do prowadzenia i dokumentowania wyników systematycznych obserwacji pedagogicznych dzieci oraz przeprowadzania analizy gotowości szkolnej w roku poprzedzającym rozpoczęcie nauki w szkole. W wyniku tych działań nauczyciel zdobywa informacje niezbędne do wspierania rozwoju dzieci odpowiednio do zdiagnozowanych potrzeb, a rodzice – informacje potrzebne do podjęcia decyzji o rozpoczęciu edukacji dziecka w klasie I. Jest to również wstępny materiał konieczny do pogłębionej diagnozy dzieci ze specjalnymi potrzebami edukacyjnymi w poradni psychologiczno-pedagogicznej.

Stosunkowo najdłuższą tradycję diagnozowania ma mierzenie umiejętności dzieci na starcie szkolnym, a więc badanie dojrzałości, a potem gotowości szkolnej. Barbara Wilgocka-Okoń (1972), autorka ogólnopolskich badań realizowanych w latach 1969–1970, definiowała dojrzałość szkolną jako „osiągnięcie przez dziecko takiego stopnia rozwoju intelektualnego, społecznego i fizycznego, jaki jest niezbędny do sprostania wymaganiom szkoły” (Wilgocka-Okoń, 2003, s.11). Takie podejście zakładało bierność szkoły i aktywność rozwojową dziecka związaną z przystosowaniem się do wymagań instytucjonalnych. Dziecko musiało dojrzeć do szkoły, a tę dojrzałość diagnozowano za pomocą odpowiednich narzędzi („Test dojrzałości szkolnej” B. Wilgockiej-Okoń). Sposób diagnozowania był ściśle ograniczony do obszarów umiejętności, które twórcy testu wydawały się możliwe do osiągnięcia przez dziecko na starcie szkolnym. Pułap możliwości wiązał się zatem z wyobrażeniem dorosłego o umiejętnościach dzieci, a te z kolei wiązały się z wymaganiami programowymi przewidzianymi dla klasy pierwszej szkoły podstawowej. Dzieci wykonywały więc zadania obejmujące diagnozowanie ich rozwoju poznawczego, w tym dotyczące umiejętności porównywania, różnicowania i klasyfikowania przedmiotów, odwzorowywania kształtów, orientacji w przestrzeni, analizy oraz syntezy wzrokowej i słuchowej wyrazów, elementów wiedzy matematycznej, głównie liczenia w zakresie 10, a także tworzenia opowiadań na podstawie historyjki obrazkowej. Celem tych badań diagnostycznych było przede wszystkim określenie braków w wiedzy i umiejętnościach u dzieci rozpoczynających naukę szkolną i udzielenie im odpowiedniego wsparcia. Testy dojrzałości szkolnej były zatem tak skonstruowane, by poziom trudności zadań nie wykraczał poza z góry przyjętą normę. W związku z tym uzyskiwano tylko fragmentaryczną wiedzę, limitowaną poprzez określoną normę, co powodowało,

że prawdziwy potencjał i możliwości rozwojowe dzieci nie były w pełni poznane. Dzieci, które przekraczały w zakresie różnych umiejętności ustaloną normę, nie miały w ogóle szansy się ujawnić, co potem w procesie kształcenia mogło skutkować spadkiem motywacji do uczenia się.

Z czasem rozumienie dojrzałości szkolnej zaczęto rozszerzać i modyfikować, akcentując jego interakcyjny, dynamiczny aspekt, czyli „wynik interakcji między zdatnym do szkoły dzieckiem a środowiskiem rodzinnym i wychowaniem przedszkolnym” (tamże, s.12). Zmiana podejścia do interpretacji tego pojęcia, przynajmniej w zakresie teoretycznym, skutkowałą przyjęciem założenia (mającego podstawy psychologiczne i pedagogiczne), że na osiągnięcie odpowiedniego poziomu dojrzałości szkolnej nie tylko trzeba czekać, ale można także na nie wpływać dzięki różnorodnym oddziaływaniom, czyli kształtować określone kompetencje. Stąd wyłoniły się terminy: „gotowość szkolna dziecka”, oznaczająca jego przygotowanie do aktywności szkolnej, oraz „gotowość szkoły” do przyjęcia nowych uczniów.

Współcześnie więc gotowość dziecka do szkoły traktuje się w kategoriach interakcyjnych, dynamicznych, jako efekt współdziałania jego określonych właściwości oraz cech funkcjonowania szkoły, a konkretnie – prezentowanej przez nią oferty edukacyjnej (Brzezińska, 2002). Takie podejście do pojęcia gotowości szkolnej eksponuje zarówno intensywny wpływ kontekstu społecznego dziecka, jak i jego własnej aktywności w kreowaniu korzystnego rozwojowo środowiska. Uzasadniając to ujęcie koncepcji gotowości szkolnej Anna Brzezińska odwołuje się między innymi do następujących koncepcji rozwoju: ekologicznej Urie Bronfenbrennera (1979), kontekstualnej Richarda Lerner (2002) oraz wzajemnej socjalizacji H. Rudolpha Schaffera (2005). W każdej z tych teorii można odnaleźć interpretacje edukacji i wychowania jako elementu dynamicznej struktury oddziaływań społecznych, kulturowych i jednostkowych, związanych zarówno z przeszłością, jak i przyszłością rozwoju. Z takim podejściem do pojęcia gotowości szkolnej wiąże się też diagnozowanie rozwoju dziecka na dwóch poziomach, uwzględniających rozróżnienie wprowadzone przez Lwa Wygotskiego. Jest to diagnozowanie aktualnego poziomu kompetencji dziecka, tego, co potrafi i umie wykonać samodzielnie, oraz potencjału, przyszłości dziecka, a więc tego, co dopiero się kształtuje, swoistych obszarów wrażliwości, od których zaczyna się kolejny etap rozwoju.

Warto zatem przeanalizować, w jakim zakresie dostępne materiały diagnostyczne do badania gotowości szkolnej i diagnozowania umiejętności dzieci na starcie szkolnym uwzględniają szerszą perspektywę w spojrzeniu na rozwój dziecka. Jednym z tych, które poddam analizie, jest opracowana przez Elżbietę Koźniewską Skala Gotowości Szkolnej (SGS), która powstała w 2006 roku w Centrum Metodycznym Pomocy Psychologiczno-Pedagogicznej w ramach projektu „Badanie gotowości szkolnej sześciolatków”. Skala Gotowości Szkolnej to metoda obserwacyjna przeznaczona dla nauczycieli i dzieci 6-letnich (Frydrychowicz i in., 2006). Składa się ona z pięciu części i dotyczy obszarów umiejętności dzieci w zakresach aktywności: poznawczej, relacji społecznych w grupie rówieśników, samodzielności i radzenia sobie w trudnych sytuacjach, zadaniowej wykonywanej samodzielnie lub pod kierunkiem nauczyciela oraz przygotowania dziecka do nauki czytania, pisania oraz matematyki. Tak więc skalę współtworzą następujące podskale: Umiejętności Szkolnych, Kompetencji Poznawczych, Sprawności Motorycznej, Samodzielności, Niekonfliktowości, Aktywności Społecznej.

W opisie charakteryzującym tę skalę znalazło się też stwierdzenie, że wyodrębnione w niej kategorie umiejętności i zachowania dzieci są zgodne z prezentowanymi w podręcznikach psychologii osiągnięciami rozwojowymi charakterystycznymi dla dzieci w wieku 6–7 lat. To stwierdzenie jest kluczowe dla określenia paradygmatu, w którego ramach powstał analizowany kwestionariusz. Jest ono zgodne z podejściem normalizacyjnym, ponieważ zakłada z góry określony repertuar zachowań i umiejętności dzieci, które powinno ono przejawiać rozpoczynając edukację szkolną. W tym kontekście rodzi się pytanie, czy zakres umiejętności wyodrębniony w SGS jest wystarczający i czy w ten sposób *de facto* nie ogranicza i zawęża pola obserwacji oraz diagnozowania prowadzonego przez nauczyciela skoncentrowanego na rejestrowaniu właśnie tych, a nie innych zachowań dziecka, co wynika z konstrukcji tego narzędzia. Trzeba jednak dodać, że autorzy tego narzędzia zamieścili także informację, iż w skład skali wchodzi umiejętności, które nie są wymagane od dzieci na starcie szkolnym, ale ponieważ rozwijają się one w indywidualnym tempie (jak np. liczenie w zakresie 100), to warto je uchwycić.

Proponuję teraz analizę stwierdzeń, które weszły w skład wybranych podskal SGS, prowadzoną pod kątem obrazu funkcjonowania dziecka, czyli Niekonfliktowość oraz Samodzielność. Na podskalę Niekonfliktowość składają się następujące rodzaje zachowań (Frydrychowicz i in., 2006, s. 9):

„sprząta zabawki i pomoce, pamięta o zasadach bezpiecznego zachowania w grupie, unika sytuacji konfliktowych, próbuje radzić sobie w pokojowy sposób, uwzględnia prawa innych (czeka na swoją kolej), nie skarży w sytuacji konfliktowej, nie wymaga przypominania umowy, nie wywołuje konfliktów, nie obraża się, nie reaguje gniewem, nie wybuchają złością i płaczem, wykonuje polecenia, nie sprzeciwia się”.

Podskalę Samodzielność budują następujące zachowania (tamże):

„sprawnie się ubiera, podejmuje próby poradzenia sobie z trudnością, nie zraża się trudnościami, wykazuje inicjatywę, wypróbować różne sposoby działania, rzadko prosi o pomoc osobę dorosłą, nie naśladuje zachowania i prac innych dzieci, nie unika zadań wymagających samodzielności, nie stara się przebywać blisko nauczyciela lub innej osoby dorosłej, interesuje się wynikiem swoich prac, kończy rozpoczęte zadania, lubi dobrze wykonać prace, poprawia je, nie prosi o dodatkowe wyjaśnienia i informacje, potrafi zapamiętać i wykonać polecenie”.

Pierwsza z przedstawionych podskal charakteryzuje dziecko jako osobę posłuszną, unikającą konfliktów, pamiętającą o umowach, wykonującą polecenia bez sprzeciwu, stonowaną emocjonalnie. Natomiast Podskala Samodzielność buduje obraz dziecka, które radzi sobie z trudnościami, wykazuje inicjatywę w działaniu, ale jednocześnie unika kontaktu z dorosłym, trzyma się od niego z daleka, nie prosi o informacje i wyjaśnienia, ani o żadną pomoc, bo potrafi zapamiętać i wykonać polecenie.

Ta krótka analiza pokazuje, jak sędzę, uproszczony, powierzchowny i chwilami niejasny obraz funkcjonowania dziecka. Stanowi dobry przykład zawężania pola widzenia obserwatora do kategorii zachowań, które tworzą obraz dziecka będący efektem wyobrażeń autora kwestionariusza, popartych oczywiście wiedzą z zakresu psychologii rozwojowej. Nasuwa się pytanie, dlaczego w podskali Niekonfliktowość dominuje unikanie konfliktów, a nie umiejętność radzenia sobie w sytuacjach konfliktowych, i na przykład, rozwiązywania ich ze wsparciem dorosłego. Równie przydatna byłaby umiejętność nawiązywania relacji z rówieśnikami, świadcząca o tym, że dziecko opanowało już sposoby działania pozwalające na w miarę zgod-

ne funkcjonowanie w grupie. Dlaczego w podskali Samodzielność został nakreślony obraz dziecka-samotnika, ograniczającego kontakt z dorosłym, wykonującego w milczeniu polecenia i pokonującego trudności bez zadawania pytań oraz unikającego prezentowania swojego zdania?

I jeszcze jeden przykład, ograniczony do wybranych stwierdzeń tworzących podskalę Kompetencje Poznawcze:

„próbuję wyjaśniać obserwowane zjawiska, przewidyuję zachowania innych dzieci, potrafi trafnie powiedzieć, co wie i umie, identyfikuje nazwy przez odniesienie do ogólniejszej kategorii, dodaje i odejmuje w pamięci, w zabawie wybiera gry liczbowe, chętnie rozwiązuje zagadki matematyczne, liczy w pamięci w zakresie 100, ma umiejętności matematyczne powyżej oczekiwanych”.

W przypadku tej podskali też rodzą się wątpliwości dotyczące wyboru właśnie takich a nie innych stwierdzeń charakteryzujących dziecko. Trudno na przykład zrozumieć, dlaczego tak ważne jest liczenie i to w pamięci, w zakresie 100, a nie eksperymentowanie z liczeniem z wykorzystaniem klocków, tworzenie intuicji matematycznych. Również w zakresie aktywności językowej dziecko powinno – zgodnie z opisem skali – identyfikować nazwy pojęć, a nie – na przykład – konstruować ich rozumienie w zabawie i relacjach z rówieśnikami. Przytoczone przykłady ilustrują dobrze określony rodzaj podejścia do diagnozowania gotowości szkolnej według przyjętego z góry wzorca kandydata do klasy pierwszej szkoły podstawowej.

Gotowość szkolna a umiejętności dzieci

Zaprezentowany kwestionariusz do badania gotowości szkolnej ma charakter szacunkowego pomiaru i nie mierzy rzeczywistych umiejętności dzieci, ale opinie nauczycieli lub rodziców o tych umiejętnościach. W 2012 roku został opracowany przez zespół badawczy Instytutu Badań Edukacyjnych Test Umiejętności na Starcie Szkolnym (TUNSS) (Karwowski, Dziedziewicz, 2012). Jest to test wykonania, mierzący tym razem faktyczne umiejętności dzieci poprzez próbki zachowania. Realizuje się go na tabletach. To test adaptacyjny, który podczas diagnozowania umiejętności dostosowuje trudność zadań do tego, co dziecko już potrafi. Podczas jego przygotowywania korzystano z probabilistycznej teorii testu IRT (*Item Response Theory*) – procedury analitycznej, która powinna sprzyjać wysokiej jakości pomiaru. TUNSS mierzy umiejętności w zakresie matematyki, czytania, pisania. Skala umiejętności matematycznych obejmuje cztery zakresy treści: liczby, pomiar, przestrzeń i kształt oraz związki i zależności. Skala pisania dotyczy sprawności wzrokowo-motorycznej, wzrokowo-przestrzennej, słuchowo-językowej oraz kaligrafii i pisania. Natomiast skala czytania obejmuje sprawność słuchowo-językową, wzrokowo-słuchową oraz umiejętność czytania i rozumienia tekstu. Zadania testowe odwołują się do trzech kontekstów funkcjonowania dzieci: osobistego, szkolnego oraz publicznego. W ten sposób test ma zapewnić równowagę w zakresie tematyki zadań i uwzględnić kontekst osobistych doświadczeń dzieci, a także ich doświadczeń: szkolnych i związanych z przestrzenią publiczną. Autorzy testu nie zaliczają go do grupy narzędzi badających gotowość szkolną, lecz do narzędzi określających zakres umiejętności sześć- i siedmiolatków na starcie szkolnym.

W jego analizie skupię się na sposobie realizacji tych badań oraz typach zadań i rodzajach umiejętności diagnozowanych w poszczególnych obszarach. Przeprowa-

dzenie tego badania z wykorzystaniem tableatów wskazuje na wyjątkowość przedsięwzięcia, ale też wzbudza wątpliwości dotyczące znaczenia owego środka technicznego w diagnozowaniu określonych umiejętności. Czy posługiwanie się tabletem dla wszystkich badanych dzieci było tak atrakcyjne i przyjemne, jak zapewniają autorzy zadań? Czy rzeczywiście sprzyjało ono pozytywnej motywacji i wpływało na to, że badanie miało charakter bardziej zbliżony do zabawy niż do poważnego, stresującego testu? Jaki wpływ na szybkość i poprawność wykonywanych zdań mogły mieć wcześniejsze doświadczenia dzieci z tabletem? Co tak naprawdę mierzy wykonywany na tablecie test umiejętności dzieci?

Analiza skal umiejętności, które autorzy zaplanowali zmierzyć, ponownie skłania do refleksji, że mamy do czynienia z materiałem, który jest nastawiony na z góry założony cel, czyli zakres umiejętności dziecka, które wywodzą się z określonego sposobu interpretowania rozwoju (czasem dość schematycznego czy sformalizowanego) danej umiejętności u dzieci. W teście wyodrębniono trzy skale: umiejętności matematyczne, pisanie i czytanie. Poniżej zaprezentuję sposób rozumienia przez autorów testu umiejętności pisania, w zakresie której oczekują oni od dzieci sześciu-, siedmioletnich przede wszystkim umiejętności odtwórczych, głównie – technicznego aspektu pisania, sprawdzanego poprzez opanowanie: odtwarzania znaków według wzoru, szybkości reakcji na bodziec wizualny i słuchowy, rozumienia pojęć określających kierunki i miejsce na kartce, tworzenia wyrazów z podanych głosek i sylab, wyróżniania sylab w nagłosie, wygłosie i śródgłosie; wyróżniania wyrazów w zdaniu, rozpoznawania liter pisanych, przepisywania liter w izolacji i połączeniu, pisania z pamięci i ze słuchu itp. W ramach badania umiejętności pisania wyodrębniono trzy kategorie: obszar sprawności wzrokowo-motorycznej, słuchowo-językowej oraz kaligrafii i pisania. Z powyższego opisu wynika określony sposób interpretacji rozwoju tej umiejętności, a mianowicie: nie przewiduje on samodzielnych prób pisania, eksperymentowania z pisaniem czy pisania z sensem. Dominuje odtwarzanie znaków języka oraz usłyszanych informacji. Duży nacisk położony jest na poprawność kreślenia znaków graficznych, co na tym etapie rozwoju nie powinno być priorytetem, a wyodrębnienie kategorii kaligrafia na to wskazuje. Tymczasem rozwój umiejętności pisania to proces, w którym dziecko ma prawo do eksperymentowania, samodzielnych prób pisania, korzystania ze wzorów, ale z pozostawieniem elastyczności ich wykorzystania. To nie tylko szybkość, płynność i precyzja ruchów powinny być podstawą diagnozowania tej umiejętności. Poza tym autorzy założyli, że koncentrują się w przypadku tej umiejętności tylko na wzorach liter pisanych, a nie np. drukowanych. A dzieci w naturalny sposób zaczynają pisać literami drukowanymi, bo z nimi przede wszystkim zapoznają się w otoczeniu.

Umiejętność czytania, która też jest badana za pomocą zadań analizowanego testu TUNSS, skupia się na następujących obszarach treści: sprawności słuchowo-językowej, sprawności wzrokowo-słuchowej, umiejętności czytania. Wśród szczegółowych umiejętności uczniów, które są przedmiotem badania w tej kategorii znalazły się: umiejętność syntezy i analizy głoskowej, sylabowej, analiza zdaniowa (wyróżnianie wyrazów w zdaniu), uwaga i pamięć słuchowa (odtwarzanie informacji, szybkość reakcji na bodziec), percepcja wzrokowa, słuchanie ze zrozumieniem (odtworzenie i interpretacja wysłuchanego tekstu), znajomość liter, czytanie (sylab, wyrazów, zdań, tekstu). W przypadku badania umiejętności czytania autorzy zadań podkreślają, iż pomiar „nie uwzględnia tempa oraz techniki czytania. Skupia się na ocenie jego

poprawności oraz pomiarze czytania ze zrozumieniem” (Karwowski, Dziedziewicz, 2012, s.30). W przypadku tej umiejętności większy nacisk został położony, jak można wnioskować na podstawie założeń zaprezentowanych w podręczniku do testu, na rozumienie i interpretowanie. Jednak różnica w badaniu między umiejętnością czytania i pisania nie jest jasna. Ponadto czytanie dość tradycyjnie kojarzy się autorem testu z rozpoznawaniem liter drukowanych.

Założenia testu TUNSS zostały zaprezentowane dość ogólnikowo i powierzchownie. Punktem wyjścia do jego powstania był „przegląd wymogów sformułowanych w podstawie programowej, jak również analiza istniejących testów i skal szacunkowych” (tamże, s. 21), poczynając od materiałów z lat siedemdziesiątych ubiegłego wieku. Skupiono się głównie na podstawie programowej wychowania przedszkolnego oraz na częściowym nawiązaniu do umiejętności rozwijanych i utrwalanych na początku klasy pierwszej. Autorzy podkreślają w podręczniku, że „posługiwanie się w diagnozie kompetencji szkolnych zadaniami zbyt łatwymi, jak i zbyt trudnymi nie wpływa korzystnie na jej rzetelność” (tamże, s. 22). Jednak podstawę konstrukcji testu TUNSS stanowi założenie, że treści i umiejętności zawarte z dokumentach, takich jak podstawy programowe, najlepiej opisują to, co dziecko potrafi na tym etapie rozwoju. To niewątpliwie zawężenie pola obserwacji i zdobywania wiedzy o rozwoju i postępach dzieci, ograniczone zwykle do „sfery aktualnego rozwoju” dziecka. Znajomość zapisów podstawy programowej pozwala stwierdzić, że jej autorzy niejednokrotnie nie doceniają możliwości dzieci, infantylizują zakres ich umiejętności i sytuują je czasem w sferze „minionego”, a nie aktualnego rozwoju.

W świetle przedstawionych argumentów nasuwa się pytanie: czego o rozwoju dziecka dowiadujemy się na podstawie takich badań diagnostycznych? Test został niewątpliwie opracowany bardzo starannie od strony metodologicznej, znajdziemy w podręczniku szereg informacji na temat właściwości psychometrycznych, dokładności pomiaru, trafności pomiaru, normalizacji testu. Pozostaje natomiast wątpliwość dotycząca wartości poznawczej wyników uzyskiwanych za jego pomocą. Prezentowanie wyników wyłącznie w globalnych kategoriach: matematyka, czytanie, pisanie, ma ograniczoną wartość poznawczą, gdyż nie zdobywamy wiedzy szczegółowej o poziomie rozwoju poszczególnych aspektów tworzących daną kategorię umiejętności.

Odmienne od omówionych podejście badawcze wprowadza natomiast koncepcja diagnozowania gotowości szkolnej dzieci pięcio- i sześciolletnich w ujęciu interakcyjnym, opracowana w Zespole Wczesnej Edukacji IBE pod kierunkiem Anny Brzezińskiej (Czub, Wiliński, 2014). Po raz pierwszy bowiem w Polsce pojawia się rozwiązanie diagnostyczne, przeznaczone dla poradni psychologiczno-pedagogicznych, które uwzględnia szerokie spojrzenie na problematykę gotowości szkolnej dzieci i koncentruje się na zasobach dziecka i środowiska, w którym się ono wychowuje, a nie tylko na deficytach. Autorzy tej koncepcji podkreślają, że start szkolny jest uzależniony od wielu czynników, nie tylko od kompetencji dziecka rozpoczynającego naukę, ale także od wsparcia środowiska rodzinnego i lokalnego oraz od zróżnicowania oferty edukacyjnej i elastycznego nastawienia szkoły do przyjmowania nowej grupy uczniów. W takim ujęciu gotowość szkolna jest traktowana jako problem wielowymiarowy, „zorientowany na osobę dziecka i (...) złożone relacje między przyszłym uczniem, jego rodziną i środowiskiem społecznym – sąsiedzkiem oraz lokalnym a szkołą” (Czub, Wiliński, 2014, s. 3). Interakcyjne ujęcie problematyki

gotowości szkolnej dziecka oznacza, że jest ona efektem wzajemnego oddziaływania kilku czynników: biologicznych, psychologicznych i społecznych. Jakość gotowości szkolnej dziecka staje się zadaniem, na którego realizację ma wpływ środowisko rodzinne, ale również doświadczenia zdobyte w instytucjach i placówkach wczesnej opieki i edukacji. Autorzy tej koncepcji podkreślają, że najważniejszym elementem jest poszukiwanie związków między historią życia dziecka a jego funkcjonowaniem, a więc uzyskanie odpowiedzi na pytanie o uwarunkowania sukcesów i trudności dziecka oraz czynników wspierających jego rozwój.

Gotowość szkolną dziecka pięcioletniego i sześciolatniego warunkują w ujęciu interakcyjnym następujące komponenty (tamże, s. 4–5):

- gotowość środowiska lokalnego rozumiana jako możliwość zdobywania różnorodnych doświadczeń w kontaktach z ludźmi, a także dostęp do formalnych i nieformalnych źródeł wsparcia (przedszkola, kluby dziecięce, centra edukacyjne itp.),
- gotowość rodziny, której wskaźnikami są: wspieranie rozwoju dziecka w środowisku rodzinnym, jakość i zakres oddziaływań związanych ze statusem społeczno-ekonomicznym rodziny, kompetencjami, zaangażowaniem,
- gotowość dziecka obejmująca: dotychczasowe doświadczenia rozwojowe dziecka i poziom aktualnego funkcjonowania z uwzględnieniem wieku, płci, cech układu nerwowego i temperamentu, rozwoju w różnych sferach (fizycznej, społecznej, emocjonalno-motywacyjnej, poznawczej),
- gotowość szkoły, czyli jej otwartość i gotowość do przyjęcia dzieci o zróżnicowanych potrzebach edukacyjnych, nastawienie na indywidualizację działań pedagogicznych, wspieranie szkolnego startu i dbanie o pomyślną adaptację w ścisłej współpracy z rodzicami i zaangażowaniu wszystkich pracowników szkoły.

W interakcyjnym podejściu do analizy gotowości szkolnej na uwagę zasługuje to, że porządkuje się w nim zbierane informacje wokół trzech obszarów: dziecko (funkcjonowanie społeczno-emocjonalne i zdrowie), rodzina (funkcjonowanie społeczne, jakość opieki nad dzieckiem, status społeczno-ekonomiczny), środowisko (zaplecze instytucjonalne i społeczne). W ramach wymienionych obszarów zostały opisane czynniki, które mogą wspierać pomyślny start szkolny i utrudniać rozpoczęcie szkolnej edukacji. Obszary gotowości szkolnej dziecka zostały też dokładniej scharakteryzowane i połączone ze wskazaniem kierunków zbierania szczegółowych informacji. Przykładem niech będzie opis gotowości rodziny: dążenie do zaspokajania nowych, związanych z edukacją szkolną potrzeb dziecka, modyfikacja dotychczasowego stylu życia, gotowość do współpracy rodziców ze szkołą, wspieranie dziecka w nowej dla niego sytuacji emocjonalnej i społecznej (tamże, s. 10). Z kolei opis gotowości dziecka wiąże się ze: zmianami rytmu dnia, zaspokajaniem potrzeb, przejściem z kultury domu i przedszkola do kultury szkoły, podejmowaniem nowych zadań, obowiązków, nawiązywaniem relacji społecznych z dorosłymi i kolegami oraz koleżankami w różnym wieku.

Autorzy podkreślają, że w interakcyjnym ujęciu gotowości szkolnej w podejściu długoterminowym próbuje się też odpowiedzieć na pytanie, czy dziecko o określonym profilu kompetencji jest przygotowane do radzenia sobie z wyzwaniami w toku całej edukacji i zmieniającej się dynamicznie rzeczywistości (szczegółowy opis interakcyjnej koncepcji gotowości szkolnej można znaleźć na stronie: www.eduentuzjasci.pl/dziecko-nastolatek).

„Gotowości szkolnej dziecka nie można traktować jako zamkniętego zbioru czy zestawu »gotowych« do użycia w momencie startu szkolnego kompetencji – wiedzy i umiejętności – pisze A. Brzezińska (2014, s. 3). – Przez cały okres dzieciństwa dziecko stopniowo dorasta do podejmowania coraz trudniejszych wyzwań, uczy się działać coraz bardziej samodzielnie, ale także współdziała z innymi dla osiągnięcia ważnych celów – osobistych i grupowych. (...) Takie procesualne i jednocześnie interakcyjne ujęcie gotowości szkolnej dziecka oznacza, że gotowość ta rozwija się dalej, także po rozpoczęciu nauki w szkole, przyjmując postać coraz dojrzalszą, otwartą na nowe wyzwania”.

Opisana koncepcja interakcyjnego ujęcia gotowości szkolnej ujmuje analizę rozwoju z uwzględnieniem kontekstu środowiskowego i znacznie szerzej niż to miało miejsce w przypadku wcześniej charakteryzowanych przykładów skal gotowości szkolnej. Niemniej i tak sytuuje się ona w normalizacyjnym nurcie badań, który uwzględnia – szczególnie w charakterystyce dziecka – normalizacyjne podejście badawcze. Został w niej bowiem przyjęty określony wzorzec funkcjonowania dziecka w danym wieku, zgodnie z założeniami psychologii rozwojowej oraz wiedzy pedagogicznej, które to kryteria pozwalają na przygotowanie opisu. Uwzględnienie szkolnego i środowiskowego kontekstu funkcjonowania pogłębia perspektywę widzenia problemu, nadając mu interakcyjny wymiar i czyni tę koncepcję gotowości szkolnej wielowymiarową, dającą zdecydowanie szerszą perspektywę widzenia podejmowanej problematyki. Rodzi się jednak pytanie, w jaki sposób tak przygotowana diagnoza gotowości szkolnej dziecka, będzie wykorzystywana w szkole, po rozpoczęciu przez dziecko edukacji? Niewątpliwą zaletą tego podejścia jest przygotowanie szczegółowej charakterystyki sytuacji dziecka z uwzględnieniem czynników wspierających i ograniczających pomyślny start szkolny. Ich uwzględnienie wymaga indywidualizacji pracy z uczniami oraz dostosowania oferty edukacyjnej do specyficznych potrzeb rozwojowych. Czy szkoła jest przygotowana do podjęcia takiego wyzwania? Jakie kompetencje w tym zakresie posiadają nauczyciele edukacji elementarnej? Czy potrafią podjąć się takich zadań i czy odczuwają w ogóle taką potrzebę?

Normalizacyjny charakter podstawy programowej i organizacji pracy szkoły zakłada, że proces indywidualizacji pracy z dziećmi nie jest właściwie możliwy ani konieczny. Z tego punktu widzenia ważne jest tylko wspieranie rozwoju dzieci, które założonej normy jeszcze nie osiągnęły. Natomiast dzieci, a jest ich niemało, które tę normę przekroczyły, nie wymagają specjalnego wsparcia. Nauczycielka klasy pierwszej w jednej ze szkół warszawskich, na pytanie matki dziecka zatroskanej faktem, że jej córka dużo potrafi i chciałaby rozwijać w szkole swoje umiejętności oraz zainteresowania, dowiedziała się, że teraz musi ona wyrównać poziom dzieci i nie może zajmować się kimś, kto wie więcej od innych. To jest zadanie rodziców.

Poszukując innej perspektywy analizowania jakości edukacji dla małych dzieci, G. Dalberg, P. Moss i A. Pence (2013, s. 222) proponują wyjście poza tradycyjne schematy normalizacyjne i wypracowanie nowych działań, które będą początkiem zmiany. Twierdzą, że „koniecznym warunkiem tego rodzaju działania jest wypracowanie aktywnego sposobu przeciwstawiania się i stawiania oporu oddziaływaniu spłotu wiedzy-władzy, tym reżimom prawdy, które dążą do określenia za nas, co jest prawdziwe, a co fałszywe, co dobre, a co złe, co możemy myśleć i robić, a czego nie”. Proponują wykorzystanie w diagnozowaniu takiego narzędzia jak dokumentacja pedagogiczna, której pogłębiona, wielostronna analiza prowadzi do refleksyjnego i demo-

kratycznego, ich zdaniem, działania pedagogicznego. Dokumentacja pedagogiczna odgrywa zasadniczą rolę w dyskursie nadawania znaczeń. To nie jest już zewnętrzna, chłodna obserwacja aktywności dziecka, której celem jest gromadzenie fragmentarycznych danych z pozycji badacza, ale wzięcie odpowiedzialności za samodzielne nadawanie znaczeń i interpretację. Taka perspektywa wymaga refleksyjności i samodzielności myślenia nauczyciela, konstruowania wiedzy o dziecku w kontekście społecznym, jego doświadczeniach i postępach rozwojowych, zamiast ulegania wizji zestandaryzowanych pomiarów jakości.

Z przedstawionej perspektywy wydaje się więc uzasadnione stwierdzenie, że diagnozowanie rozwoju dzieci we wczesnej edukacji powinno koncentrować się w zdecydowanie większym zakresie na obserwacji procesu uczenia się i aktywności w różnych sferach działania, a nie wyłącznie na klasyfikowaniu i szukaniu odpowiednich kategorii, do których można podporządkować ich funkcjonowanie w środowisku szkolnym. Czy taka postawa jest możliwa do urzeczywistnienia w praktyce edukacyjnej? Jak przygotować nauczycieli do roli refleksyjnych praktyków odpowiedzialnych za analizę szkolnych doświadczeń i podejmujących świadome i samodzielne decyzje edukacyjne? To są pytania, na które warto zacząć szukać odpowiedzi w dynamicznie zmieniającej się rzeczywistości społecznej i ściśle z nią związanej wczesnej edukacji.

Literatura

- Bronfenbrenner U. (1979): *The Ecology of Human Development: Experiments by Nature and Design.*: Harvard University Press, Cambridge, MA.
- Brzezińska A. (2002): *Współczesne ujęcie gotowości szkolnej*. W: W. Brejnak (red.): *O pomysłny start ucznia w szkole*. „Biuletyn Informacyjny Polskiego Towarzystwa Dysleksji”, wydanie specjalne, Warszawa.
- Brzezińska A. i in. (2014): *Sześciolatki w szkole: rozwój i wspomaganie rozwoju*. Wyd. Humaniora, Poznań.
- Czub M., Wiliński M. (2014): *Diagnoza gotowości szkolnej w ujęciu interakcyjnym*. Broszura 1. IBE, Warszawa.
- Dahlberg G., Moss P, Pence A. (2013): *Poza dyskursem jakości w instytucjach wczesnej edukacji i opieki*, tłum. K. Gawlicz. Wyd. Naukowe Dolnośląskiej Szkoły Wyższej, Wrocław.
- Frydrychowicz A. i in. (2006): *Skala Gotowości Szkolnej*. Centrum Metodyczne Pomocy Psychologicznej, Warszawa.
- Karwowski M., Dziedziewicz D. (2012): *Test Umiejętności na Starcie Szkolnym. Podręcznik*. Instytut Badań Edukacyjnych, Warszawa.
- Lerner R. (2002): *Concepts and Theories of Human Development* L.Erlbaum Associates, Mahwah, NJ.
- Schaffer H.R. (2005): *Psychologia dziecka*. Wyd. Naukowe PWN, Warszawa.
- Taguchi H., L. (2010): *Rethinking pedagogical practices in early childhood education: a multidimensional approach to learning and inclusion*. W: N. Yelland (red.): *Contemporary Perspectives on Early Childhood Education*. Open University Press, Maidenhead.
- Wilgocka-Okoń B. (1972): *Dojrzałość szkolna dziecka a środowisko*. PWN, Warszawa.
- Wilgocka-Okoń B. (2003): *Gotowość szkolna dzieci sześciolatków*. Wyd. Akademickie „Żak”, Warszawa.

Gotowość dzieci i gotowość szkoły do uczenia się. Rozważania wokół problematyki diagnozowania jakości wczesnej edukacji

Pojęcie jakości w edukacji oraz gotowości szkolnej jest interpretowane w Polsce w jednostronny sposób. Dominuje podejście normalizacyjne, które charakteryzuje eksponowanie wystandaryzowanych, mierzalnych wskaźników efektywności działania instytucji edukacyjnych. W artykule podjęto

dyskusję tego problemu w kontekście innego interpretowania jakości edukacji, zaprezentowanego w książce G. Dahlberg, P. Mossa i A. Pence'a, *Poza dyskursem jakości w instytucjach wczesnej opieki i edukacji*. Autorzy pokazują w niej paradygmat analizowania jakości nazwany – nadawaniem znaczeń. Orientacja na rozumienie jakości instytucji wczesnej edukacji w kategoriach nadawania znaczeń wiąże się ze zmianą ich funkcji społecznej. W tej orientacji instytucja edukacyjna staje się miejscem dyskusji, wymiany myśli, zaangażowania się wszystkich podmiotów – uczniów, nauczycieli, rodziców, przedstawicieli środowiska lokalnego w zmiany. W artykule podjęto dyskusję tego zagadnienia w kontekście dotychczasowych doświadczeń dotyczących diagnozowania jakości wczesnej edukacji w Polsce.

Słowa kluczowe: gotowość szkolna, jakość wczesnej edukacji, diagnozowanie jakości, testy gotowości szkolnej

Children's school readiness and school readiness to learn. Discussion on the quality of early childhood education

The notion of quality in education is interpreted in a one-sided way in Polish school practice. It follows the concept of normalised and standardised quality characterised by dominance of easily measurable effectiveness indicators concerning educational institutions' activities. Another perspective and paradigm of quality analysis is presented by G. Dahlberg, P. Mossa and A. Pence'a in their book *Beyond Quality in Early Childhood Education and Care. Languages of Evaluation*. The authors question the normalisation approach based on imposed criteria to be followed by schools. They suggest elaborating subjective/individual perspective to analyse the functioning of a given educational institution. Such an approach results in realisation by all those involved in educational activity in that institution what the sense and consequences of their activities are. It entails being ready to take responsibility for them. The paper focuses on the quality of early education in Poland taking into consideration the tools of school readiness diagnosis as well as changes to the idea of diagnosis.

Keywords: children's school readiness, quality of elementary education, school readiness diagnosis, educational institution, school readiness tests

Barbara Ciżkowicz

Uniwersytet Kazimierza Wielkiego w Bydgoszczy

O POTRZEBIE EWALUACJI EFEKTÓW KSZTAŁCENIA W SZKOŁACH WYŻSZYCH: STAN, PROBLEMY, PERSPEKTYWY

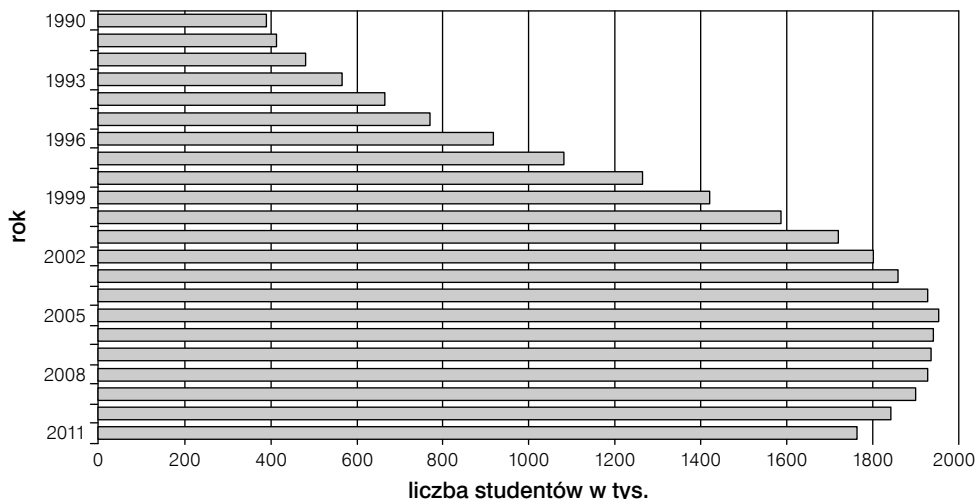
Zmiany społeczno-polityczne zachodzące w Europie oraz rozwój technologiczny spowodowały konieczność dostosowania systemów edukacji krajów integrujących się do nowej sytuacji. Otwieranie się rynków pracy stworzyło pilną potrzebę zapewnienia uznawalności kwalifikacji, a dynamiczne zmiany struktury zatrudnienia wymagały przygotowywania osób uczących się do uczenia się przez całe życie. W działaniach innowacyjnych, mających zapewnić krajom UE konkurencyjność na światowych rynkach, konieczna stała się integracja wiedzy, umiejętności i wartości, które wykraczały często poza dyscypliny naukowe (OECD, 2010). Wszystko to zmusiło kraje Unii do wypracowania ram systemowych zmian w obrębie całego szkolnictwa, w tym szkolnictwa wyższego. W efekcie wprowadzenia odpowiednich regulacji prawnych uczelnie w Polsce zyskały znaczną autonomię w tworzeniu nowych kierunków studiów. Jednak – w zamyśle ustawodawcy – gwarantem wysokiej jakości kształcenia w szkołach wyższych ma być sprawnie funkcjonujący wewnętrzny system zapewniania jakości kształcenia. Jednym z jego istotnych elementów jest kontrolowanie i dokumentowanie osiągnięcia przez studentów efektów kształcenia zapisanych w programach kształcenia.

Spojrzenie wstecz

Przez wiele lat edukacja na poziomie wyższym była w Polsce edukacją elitarną. W okresie 1945–1989 uczelnie polskie zapewniały wysokiej jakości kształcenie naukowe i techniczne pomimo licznych ograniczeń organizacyjnych i ideologicznych. W tym przedziale czasowym liczba instytucji szkolnictwa wyższego zwiększyła się z 54 do 97, czemu towarzyszył wzrost liczby studentów z 86,5 do 378 tysięcy. Ponieważ liczba chętnych znacznie przekraczała liczbę miejsc dostępnych w uczelniach, gwarantowało to selekcję wśród kandydatów na studia (Rószkiewicz, 2009, s. 27).

Transformacja ustrojowa w Polsce doprowadziła do istotnych przemian w funkcjonowaniu szkolnictwa wyższego. Obserwowane od lat dziewięćdziesiątych zmiany w strukturze zatrudnienia zwiększyły zapotrzebowanie na wysoko wykwalifikowaną kadrę pracowników. Spowodowało to znaczący wzrost zainteresowania kształceniem wyższym zarówno wśród młodzieży, jak i wśród osób dorosłych funkcjonujących już na rynku pracy. Począwszy od lat dziewięćdziesiątych w szkolnictwie wyższym

zachodzą dynamiczne zmiany ilościowe i jakościowe. *Ustawa o szkolnictwie wyższym* (DzU 2008, nr 65, poz. 385, z późn. zm.) zezwoliła na tworzenie prywatnych szkół wyższych (GUS, 2013a). Już w 1991 roku powstaje pierwsza taka uczelnia (POLON, 2013). Od roku 1990 nastąpił intensywny wzrost liczby studentów (rys.1).



Rys. 1. Liczba studentów w Polsce w kolejnych latach

Źródło: opracowanie własne na podstawie danych statystycznych (GUS, 2013 a).

W roku akademickim 2005/06 liczba osób studiujących w polskich uczelniach osiągnęła wartość najwyższą (1,954 mln) i w ciągu kolejnych siedmiu lat zmniejszyła się o 277 tys. W roku 2012 uczelnie niepubliczne stanowiły 70,1% uczelni wyższych w Polsce, kształcąc 24,2% ogółu studentów (GUS, 2013a). Współczynnik skolaryzacji, utrzymujący się w początkach lat osiemdziesiątych ubiegłego wieku na poziomie nieco powyżej 4%, wzrósł w sposób istotny (tabela 1).

Tabela 1. Współczynniki skolaryzacji

Współczynnik skolaryzacji	1990/1991	1995/1996	2000/2001	2005/2006	2010/2011	2011/2012	2012/2013
Brutto*	12,9	22,3	40,7	48,9	53,8	53,1	51,8
Netto**	9,8	17,2	30,6	38,0	40,8	40,6	40,2

* Jest to relacja liczby wszystkich osób uczących się na danym poziomie kształcenia do liczby ludności w grupie wieku określonej jako odpowiadająca temu poziomowi nauczania.

** Jest to relacja liczby osób w danej grupie wieku uczących się na danym poziomie kształcenia do liczby ludności w grupie wieku określonej jako odpowiadająca temu poziomowi nauczania.

Źródło: http://www.stat.gov.pl/cps/rde/xbcr/gus/E_szkoly_wyzsze_2012.pdf

Znaczna różnica między wartością współczynnika netto i brutto wskazuje na duży udział osób powyżej 24 roku życia wśród podejmujących studia wyższe. W roku akademickim 2012/13 na studiach niestacjonarnych studiowało 42,1% wszystkich uczących się w szkołach wyższych. Oznacza to najczęściej brak barier na drodze do podję-

cia studiów zaocznych, znacznie mniejszą liczbę godzin zajęć niż na studiach stacjonarnych i – często – gorszą jakość kształcenia (Rószkiewicz, 2009).

Proces boloński

Wdrażanie postanowień Deklaracji Bolońskiej (1999) miało zbliżyć systemy kształcenia i doprowadzić do utworzenia Europejskiego Obszaru Szkolnictwa Wyższego – EOSW (*European Higher Education Area* – EHEA). Sygnatariusze zobowiązali się do koordynowania polityki zwiększającej międzynarodową konkurencyjność europejskiego szkolnictwa wyższego, wzmacniającej współpracę między uczelniami w celu utworzenia sprzyjających warunków rozwijania ich potencjału, zwiększającej mobilność studentów, poprawiającej dostęp do różnego rodzaju programów studiów (Malan, 2004, s. 289–297). Uznano potrzebę zmian w kształceniu ze zorientowanego na treści na zorientowane na uczenie się przez całe życie. Uczelnie wprowadziły Europejski System Transferu Punktów (ECTS) oraz ułatwienia w modyfikowaniu programów kształcenia, zaczęły stosować mechanizmy poprawy jakości kształcenia (Gvaramadze, 2008, s. 443–455). Przyjęto, że akty prawne tworzone przez sygnatariuszy muszą być zgodne z globalnymi wymaganiami i elastycznie do nich dostosowane (Karseth, Sivesind, 2010, s. 103–120; Serrico i in., 2010, s. 35–54). Proces boloński zaczął wywierać też wpływ na szkolnictwo wyższe poza obszarem europejskim, szczególnie w Ameryce Łacińskiej (Ferrer, 2010, s. 601–611).

W integrującej się Europie konieczne okazało się zwiększenie przejrzystości kompetencji i kwalifikacji zdobywanych w różnych krajach. Pod pierwszym z pojęć rozumie się udowodnioną zdolność stosowania wiedzy, umiejętności i zdolności osobistych, społecznych lub metodologicznych wykazywaną w pracy lub nauce oraz w karierze zawodowej i osobistej (ERK, 2009, s. 11). Kwalifikacje natomiast obejmują formalny wynik procesu oceny i walidacji uzyskany w sytuacji, w której właściwy organ stwierdza, że dana osoba osiągnęła efekty uczenia się zgodne z określonymi standardami (tamże).

By osiągnąć efekt spójności, w szkolnictwie europejskim opracowano Europejskie Ramy Kwalifikacji (ERK), jako układu odniesienia, który pozwala porównywać poziomy kwalifikacji uzyskiwanych w różnych systemach kształcenia i szkoleń w całej UE, wspierając mobilność obywateli oraz ułatwiając uczenie się przez całe życie (DzU 2008, C 111, 6.5). Opracowany w ramach procesu bolońskiego dokument (Framework for Qualifications of EHEA, 2005) – przyjęty w Bergen – stanowi podstawę kształtowania krajowych systemów szkolnictwa wyższego. Jest on jednak jedynie politycznym zobowiązaniem rządów krajów-sygnatariuszy Deklaracji Bolońskiej, natomiast nie tworzy stanu prawnego (EUA, 2006).

Polska w procesie bolońskim – tworzenie studenckiej różnorodności

Wejście w życie zmian ustawy o szkolnictwie wyższym dostosowało pod względem formalno-prawnym polskie szkolnictwo wyższe do ERK (DzU 2011, nr 84, poz. 455). Warunkiem koniecznym prowadzenia studiów na kierunku stało się opracowanie programu kształcenia rozumianego, jako „opis określonych przez uczelnię spójnych

efektów kształcenia, zgodny z Krajowymi Ramami Kwalifikacji dla Szkolnictwa Wyższego (KRK), oraz opis procesu kształcenia, prowadzącego do osiągnięcia tych efektów, wraz z przypisanymi do poszczególnych modułów tego procesu punktami ECTS” (DzU 2011, nr 84, poz. 455, art. 2 ust 1 pkt 14b). Zlikwidowano obowiązujące dotąd standardy kształcenia, wyjątek czyniąc dla kierunków, na których trzeba było uwzględnić wymagania określone w przepisach UE lub wymagania rynku pracy. Wszystkie te decyzje zwiększają możliwości uczelni w elastycznym dostosowaniu oferty kształcenia do potrzeb i możliwości osób uczących się i rynku pracy.

W ramach nowych regulacji prawnych instytucje szkolnictwa wyższego, jako podmioty autonomiczne ponoszą odpowiedzialność za jakość kształcenia. W miejsce dotychczasowych dyplomów państwowych absolwenci otrzymują uczelniane dyplomy ukończenia studiów, potwierdzające uzyskanie odpowiedniego tytułu zawodowego i suplementy do dyplomów (DzU 2011, nr 84, poz. 455, art.167 ust. 1). Na uczelni nałożono obowiązek wprowadzenia wewnętrznego systemu zapewniania jakości, który dzięki odpowiednim procedurom pozwala monitorować i doskonalić proces kształcenia (DzU Nr 243, poz. 1445, z późn. zm.). W ramach tych działań jednostka organizacyjna może dokonywać znacznych zmian w realizowanych zajęciach dydaktycznych, sięgających nawet do 50% punktów ECTS, jeśli nie skutkuje to zmianami efektów kształcenia.

W obowiązujących uwarunkowaniach prawnych ewaluację wewnętrzną należy traktować, jako badania w działaniu z udziałem wszystkich pracowników i beneficjentów w podstawowych jednostkach organizacyjnych uczelni (Bazargan i in., 2007; Spencer, 2001). Polska Komisja Akredytacyjna, jako zewnętrzny organ nadzorujący, działający na rzecz doskonalenia jakości kształcenia, zobowiązana została do oceniania między innymi funkcjonowania i doskonalenia systemów zapewnienia jakości kształcenia (DzU 2011, nr 207, poz 1232).

Od kilku lat na sytuację w polskich uczelniach coraz wyraźniejszy wpływ wywiera niż demograficzny. Liczba osób w wieku 18–24 zmniejszyła się z 4,2 mln w roku 2008 do 3,6 mln w roku 2012. Zgodnie z prognozą tendencja ta utrzyma się, osiągając 3,2 mln w 2015 i 2,7 mln w roku 2020 (GUS, 2013b), co powodować może pogarszanie się sytuacji ekonomicznej uczelni (Jurek, 2013). W sytuacji, gdy duża liczba uczelni konkuruje o malejącą liczbę studentów, obniżają się progi punktowe przyjęć, zwłaszcza na kierunkach humanistycznych i społecznych, czyli maleje utrzymująca się przez wiele lat popularność niskonakładowych kierunków studiów (Woźnicki, 2013; GUS, 2013b). Młodzież dokonuje coraz bardziej racjonalnych wyborów, które stanowią wypadkową wielu czynników, takich jak: zainteresowania kandydatów, szacowanie przez nich realnych możliwości podjęcia studiów na określonym kierunku, jakość oferty dydaktycznej szkół wyższych oraz chłonność rynku pracy.

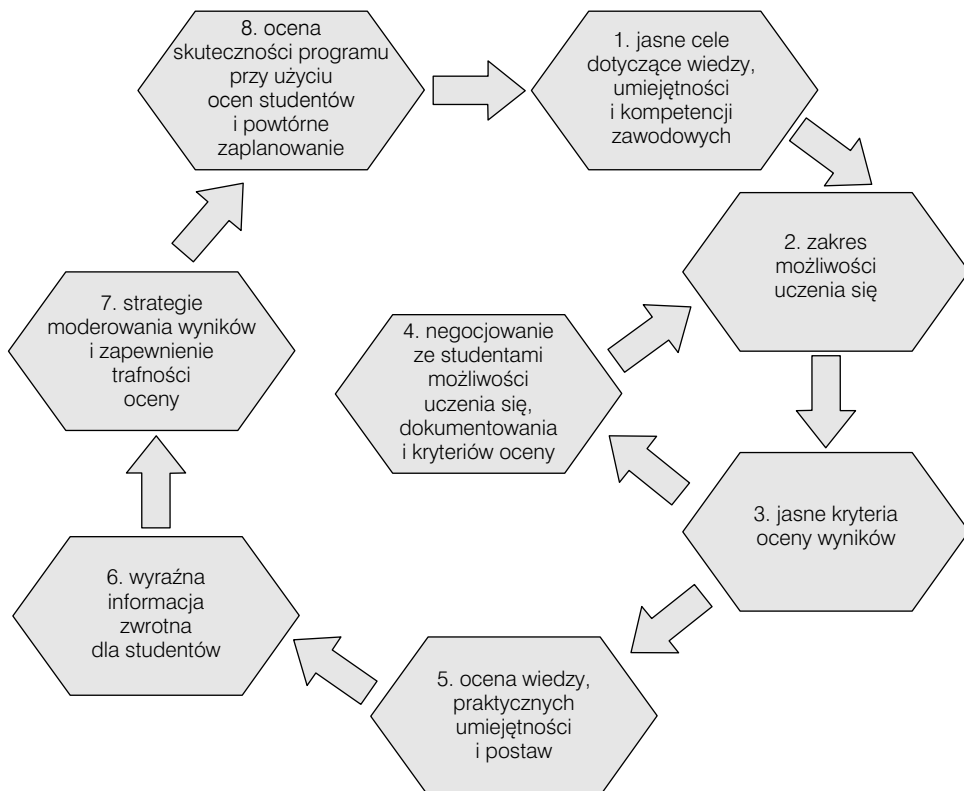
W pogarszającej się kondycji finansowej większego znaczenia nabiera pytanie, jak zapewnić wysoką jakość kształcenia w uczelniach i wyeliminować nierzetelne działania w tym sektorze?

Ewaluacja efektów kształcenia

Znowelizowane prawo zobligowało jednostki organizacyjne uczelni, prowadzące studia na danym kierunku, do określenia sposobu weryfikacji osiągniętych przez studentów efektów kształcenia (DzU 2012, poz. 983). PKA w raporcie z wizytacji

ocenia, czy „jednostka stosuje przejrzysty system oceny efektów kształcenia, umożliwiając weryfikację zakładanych celów i ocenę osiągania efektów kształcenia na każdym etapie kształcenia” (zał. do Statutu PKA, 2011, cz. I, pkt. 2,3). Duże zainteresowanie środowiska akademickiego licznymi konferencjami i szkoleniami, które poświęcone są zagadnieniom oceniania, może wskazywać na problemy występujące w tym obszarze.

Ocenianie osiągnięć studentów powinno pełnić – oprócz najbardziej tradycyjnych funkcji: kształtującej i sumującej – także te, które są związane z doskonaleniem programów kształcenia (Niemierko, 1999, 2007). Zaproponowany przez Susanne Owen i Gary’ego Davisa (2011) kołowy model oceniania w kształceniu wyższym (rys. 2) podkreśla cykliczność działań i porządkuje funkcje, jakie spełnia prawidłowo przeprowadzane ocenianie. Podkreśla się w nim potrzebę jasnego określenia ocenianej wiedzy, a także umiejętności i kompetencji zawodowych, a przy tym uwzględnienie potrzeb i możliwości studentów. Stanowi to bazę do ustalania kryteriów oceny, w którym to procesie powinni uczestniczyć również studenci. Model uwzględnia też znaczenie czytelnej dla studentów informacji zwrotnej. Wyniki oceniania muszą też być wykorzystane do ewaluacji skuteczności programu/modułu, a w konsekwencji do jego doskonalenia poprzez zmiany.



Rys. 2. Model oceniania osiągnięć przez studentów efektów kształcenia

Źródło: modyfikacja własna modelu Owen, Davis (2011)

W licznych publikacjach, poświęconych oddziaływaniu oceniania na uczenie się i motywację studentów (Biggs 2003; Hernander, 2012; Rust, 2002, s. 145–158), autorzy podkreślają, że prawidłowo przebiegający proces oceniania powinien wykształcić u studentów umiejętność samooceny, która jest niezbędna w przygotowaniu do uczenia się przez całe życie, natomiast włączenie studentów w proces oceniania powinno powodować wzrost motywacji do uczenia się (Boud, Associates, 2010; Glyn i in., 2011).

W szkolnictwie wyższym dostrzega się potrzebę zróżnicowania metod oceniania, aby zwiększyć jego szeroko rozumianą trafność. Wyrazem tych dążeń jest wprowadzenie do kształcenia akademickiego portfolio, samooceny i oceny koleżeńskiej (Sambell i in., 1997, s. 349–347). Autorzy tekstów poświęconych tym zagadnieniom podkreślają, że dobieranie stosownych metod oceniania do sprawdzanych efektów uczenia się i udzielanie rzetelnej informacji zwrotnej rozwijają samodzielność i odpowiedzialność studentów za monitorowanie swojego uczenia się i rozwój oraz zarządzanie nim (Calress, 2007, s. 79–89; Ramsden, 2003). Rosario Hernandez (2012, s. 489–502) upatruje skuteczności ciągłego oceniania nie tyle w zwiększaniu ilości i jakości informacji zwrotnej, co w zaangażowaniu studentów w proces oceniania. Wydaje się jednak, że zasady oceniania w większości uczelni nie nadążają za rozległymi zmianami zachodzącymi w szkolnictwie wyższym nie tylko w Polsce.

Z badań poświęconych poglądom nauczycieli akademickich i studentów na rolę oceniania w nauczaniu-uczeniu się wynika, że zasady oceniania nie satysfakcjonują ani jednych, ani drugich. Studenci wyrażają obawy co do wiarygodności kryteriów oceniania, uznając oceny za niesprawiedliwe. Sprawia to, że nie traktują jej jako istotnej informacji zwrotnej w podejmowaniu decyzji dotyczących ich kształcenia (Struyven i in., 2005, s. 325–334). Formułują też liczne zarzuty dotyczące poświęcania w ocenianiu zbyt wielkiej uwagi sprawdzaniu wiedzy (McLellan, 2004, s. 19–33). Z badań wynika też, że studenci akcentują potrzebę kontaktu z nauczycielem, by lepiej zrozumieć ocenę i uzyskać informację zwrotną o swoich osiągnięciach, ponieważ napotykają przy tym na istotne trudności, do których przyczynia się na przykład lakoniczny komentarz lub mało czytelny zapis. Podkreślają też znaczenie wspólnego ustalania kryteriów oceny ze studentami (Carless, 2006, s. 219–233).

W licznych badaniach wskazuje się też na zróżnicowanie opinii studentów i ich nauczycieli akademickich na temat jakości oceniania. Przyczyną tych odmienności może być brak jasnej polityki instytucjonalnej, podstaw empirycznych oraz przejrzystości praktyki oceniania (Price, 2005, s. 215–230). W tym kontekście formułuje się postulaty prowadzenia systematycznych badań w szkolnictwie wyższym w celu wypracowania ram polityki i praktyki oceniania (Meyer i in., 2010, s. 331–350). Uważa się, że brak regulacji tego typu jest konsekwencją historycznie ukształtowanego zaufania do „opinii ekspertów”, którymi są w swoich dziedzinach nauczyciele akademicy (Rust, 2007, s. 229–237). Wśród argumentów za zewnętrzną regulacją oceniania akademickiego znajduje się spostrzeżenie, że w elitarnym kształceniu kontakt „uczeń-mistrz” pozwalał na ciągłą obserwację, natomiast we współczesnych, masowo kształcących uczelniach, bezpośredniość tych kontaktów jest obiektywnie utrudniona. Masowa edukacja wygenerowała warunki kształcenia, w których zmieniły się oczekiwania względem kadry naukowej oraz studentów i absolwentów (Fletcher i in., 2012, s. 119–133).

W związku z tym pojawia się pytanie: czy w nowych realiach tradycyjne podejście do oceniania powinno pozostawać niezmiennie? Opinie na ten temat są podzielone, ale można je przyporządkować do trzech umownych kategorii: (a) przeciwników jakichkolwiek zmian, (b) akceptujących poszerzenie metod sprawdzania, (c) zwolenników standaryzacji procedur w ocenianiu sumującym.

Przeciwnicy wprowadzania zmian podkreślają, że w szkolnictwie wyższym naukowcy wiedzą, co należy cenić u ich studentów oraz jakie przedstawić im problemy i zadania, aby rzetelnie stwierdzić i ocenić osiągnięte przez nich efekty (Hattie, Timperley, 2007, s. 81–112), a sytuacja sprawdzania, postrzegana przez studenta jako zagrożenie, wywołuje lęk, który może powodować powierzchowne, bezrefleksyjne uczenie się (Struyven i in., 2005). Ich zdaniem, kompetencje, których oczekuje się na akademickim poziomie kształcenia, nie poddają się standaryzowanym procedurom oceniania, a rozwój kontroli zewnętrznej zagraża niezależności uczelni wyższych (Carey, 2006; Harvey, 2008; Crossouard, 2010, 247–258; QAA, 2013).

Zwolennicy wzbogacania metod oceniania o nowe rozwiązania argumentują, że kompetencje rozumiane, jako połączenie wiedzy, umiejętności i postaw, mogą być integrowane tylko w działaniu, a w związku z tym mogą być też oceniane tylko przez aktywność, która angażuje je w realizację. Dlatego też w programach studiów uwzględniane powinny być projekty, które pozwalają ocenić kompetencje studenta, jak na przykład kończąca studia obowiązkowa praca dyplomowa. Przy jej realizacji student zdobywa doświadczenie w samodzielnej pracy nad gromadzeniem informacji potrzebnych do rozwiązania problemu, ich krytycznej ocenie, prezentowaniu rezultatów we właściwy sposób oraz w zarządzaniu własnymi zasobami. Ta forma kontroli pozwala oceniać określone kompetencje studentów, w przeciwieństwie do tradycyjnych egzaminów badających treści (Becker, 1997, s. 1347–1373; Mateo i in., 2012, s. 435–441; Popham, 1999, s. 8–16).

Dla niektórych subiektywizm i nierzetelność tradycyjnego oceniania jest argumentem za koniecznością wprowadzenia ustawowego obowiązku dokumentowania efektów kształcenia osiągniętych przez studentów (Goldstein i in., 2014, s. 231–235; Frederiksen, Collins, 1989, s. 27–32).

Bez względu na prezentowane podejście istnieje duża zgodność wśród naukowców, że zbudowanie obiektywnych wskaźników do pomiaru wyników kształcenia w szkołach wyższych jest bardzo trudne, jeśli nie niemożliwe. Duża różnorodność programów kształcenia w połączeniu z malejącą liczbą studentów stanowi poważne utrudnienie w przeprowadzaniu standaryzowanych egzaminów. Podkreśla się jednak, że zmodernizowany system kształcenia wyższego skupia ocenianie na kompetencjach predestynujących do samokształcenia i uczenia się przez całe życie, podając w wątpliwość ważność samego programu nauczania rozumianego jako zakres treści merytorycznych.

Zdarza się, że zewnętrzne egzaminy są przeprowadzane wśród studentów i absolwentów szkół wyższych, ale nie jest to praktyka powszechna. Na przykład w Polsce warunkiem koniecznym uzyskania prawa wykonywania zawodu lekarza jest zdanie z pozytywnym wynikiem Lekarskiego Egzaminu Końcowego (LEK) (Internet1). Należy przy tym podkreślić, że na studiach medycznych nadal obowiązują standardy kształcenia, co w znacznym stopniu uściśla treści kształcenia.

Zewnętrzny, ale lokalny charakter mają też niektóre egzaminy w polskich uczelniach (Internet2). Zdają je wspólnie wszyscy studenci realizujący przedmiot w danym

semestrze bez względu na osobę, z którą odbywali zajęcia (np. ekonometria, matematyka SGH; języki obce SGH i UEK w Krakowie). Natomiast w innych krajach egzaminy standaryzowane wprowadziły na przykład władze Kolumbii, które w celu poprawy jakości kształcenia uznały, że wynik takiego egzaminu jest jednym ze wskaźników określających usytuowanie uczelni w rankingach (Gómez Soler, 2013, s. 201–239).

Dobre praktyki?

W wielu krajach w szkolnictwie wyższym dostrzega się potrzebę poprawy praktyki oceniania, biorąc pod uwagę rosnące wymagania płacących studentów, nasilającą się presję finansową na instytucje szkolnictwa wyższego oraz konieczność utrzymania wysokich standardów akademickich. W Wielkiej Brytanii specjalnie powołana grupa ekspertów przedstawiła obszerny raport na temat pomiaru i dokumentowania osiągnięć studentów, sporządzony po wcześniejszych analizach systemów oceniania wielu krajów (nie tylko UE). Sformułowane w nim wskazania są ostrożne i potwierdzają dużą trudność w wypracowaniu koncepcji oceniania w szkołach wyższych (UUK, 2007), jednak i tak są one krytykowane przez nauczycieli akademickich za zbyt instrumentalne traktowanie oceniania (Crossouard, 2010, s. 247–258).

W brytyjskim szkolnictwie wyższym po konsultacjach prowadzonych w środowisku akademickim opracowano Kodeks Jakości, który zawiera grupę wskaźników precyzujących zagadnienia kluczowe dla oceniania studentów. Wskaźniki te stanowią jedynie podstawę opracowania własnych procedur, uwzględniających specyfikę kształcenia w danej jednostce, a w szczególności efekty kształcenia określone zarówno w odniesieniu do całego programu, jak i do modułów na niego się składających (QAA, 2013). W raporcie dostrzega się konieczność zmian metod oceniania, jednocześnie podkreślając pracochłonność i koszty z tym związane. Wśród elementów, które powinny być uwzględniane w konstruowaniu systemu oceniania, wymieniane są:

1. Większy nacisk na ocenianie dla uczenia się:
 - a) zapewnienie równowagi między oceną kształtującą i sumującą;
 - b) angażowanie studentów w proces oceniania, co ma im zapewnić zrozumienie roli oceny w uczeniu się oraz zwiększyć zaufanie do oceniających i świadomość studentów (Ball i in., 2012; Carless, 2009, s. 57–66).
2. Zapewnienie oceny stosowanej w celu:
 - a) troski o rzetelność i trafność oceniania; przy czym pojęcie trafności obejmuje także skutki oceny w kontekście nauczania i uczenia się oraz społeczne skutki wykorzystania informacji z oceniania (Messick, 1995);
 - b) spójności wymagań stawianych przez różnych nauczycieli akademickich;
 - c) weryfikowania ocen, np. poprzez ocenianie przez dwóch egzaminatorów, porównywanie prac egzaminacyjnych, które uzyskały tę samą ocenę;
 - d) nadania priorytetu programowym efektom kształcenia.
3. Dostarczanie studentom, oprócz wyraźnych kryteriów oceny, dodatkowych informacji (np. przykładów, wzorów) na temat standardów oceniania, kształtowanie zachowań etycznych (przeciwdziałanie plagiatom, ściąganiu na egzaminach).
4. Umiejętność korzystania z oceny do projektowania przedmiotów, wsparcie studentów w praktyce monitorowania i nadzorowania własnej nauki.

5. Doskonalenie umiejętności nauczycieli akademickich w zakresie oceniania studentów (QAA, 2013), ponieważ doświadczenie i postawa nauczycieli akademickich wobec oceniania wpływa na sposób, w jaki włączają je oni do procesu nauczania-uczenia się (Tang, Chow, 2007, s. 1066–1085).

Podsumowanie

Działania podejmowane w celu podnoszenia efektywności instytucji, jak nigdy wcześniej, objęły również szkolnictwo wyższe i stały się przedmiotem zewnętrznej oceny jego wyników. Jednym z czynników wywołujących obawy o jakość kształcenia jest przejście od dawnych systemów elitarnych do masowo kształcących szkół wyższych, tworzących bardzo zróżnicowaną grupę uczelni, pracowników i studentów. Drugi czynnik związany jest ze zwiększeniem autonomii uczelni, co zmieniło relacje między instytucjami szkolnictwa wyższego i władzami: model kontroli przez państwo został zastąpiony modelem nadzoru państwa (Weisbrod i in., 2008).

W założeniu ustawodawcy, prawne dostosowanie polskiego szkolnictwa wyższego do ERK ma przyczynić się do głębokich zmian w organizacji całego procesu dydaktycznego, w którego tworzeniu i ocenianiu powinni aktywnie uczestniczyć nauczyciele akademicki, studenci i przyszli pracodawcy. Jasne kryteria weryfikacji efektów kształcenia i respektowanie ich w ocenianiu studentów są warunkiem koniecznym skutecznego wdrażania KRK. Rzetelnie przeprowadzana weryfikacja efektów kształcenia ma dostarczać informacji o uczeniu się studentów, jakości programu kształcenia oraz odpowiedzialności instytucji.

W wielu krajach, w tym w Polsce, trwa debata na temat potrzeby zmian w systemie oceniania w szkolnictwie wyższym. Prezentowane w tekście różne stanowiska badaczy dostarczają argumentów świadczących o konieczności głębokiego namysłu nad kierunkiem projektowanych zmian. Uwzględniając korzyści i zagrożenia oraz sygnalizowane koszty związane z wprowadzaniem zmian (Pruchnicka, 2013), należy zadbać o to, by bilans zysków i strat wypadł po stronie zysków, a jakość kształcenia nie pozostawała jedynie zapisem udokumentowanym na papierze.

Literatura

Ball S. i in. (2012): *A Marked Improvement Transforming assessment in higher education*, dostępny na: http://www.heacademy.ac.uk/assets/documents/assessment/A_Marked_Improvement.pdf (otwarty 26.01.2014).

Becker W. (1997): *Teaching economics to undergraduates*. „Journal of Economic Literature”, 35 (3)

Biggs J. B. (2003): *Teaching for quality learning at university* (2nd ed.). SRHE and Open University Press, Buckingham.

Boud D. & Associates (2010): *Assessment 2020: Seven propositions for assessment reform in higher education*. Australian Learning and Teaching Council, Sydney.

Carey K. (2006): *College rankings reformed: The case for a new order in higher education*. „Education Sector”, dostępny na: <http://www.educationsector.org/publications/college-rankingsreformed> (otwarty 26.01.2014).

Carless D. (2006): *Differing perceptions in the feedback process*. „Studies in Higher Education”, 31, 2.

Carless D. (2007): *Learning-oriented assessment: Conceptual bases and practical implications*. „Innovations in Education and Teaching International”, 44 (1).

- Carless D. (2009): *Trust, distrust and their impact on assessment reform*. „Assessment and Evaluation in Higher Education”, 34 (1).
- Crossouard B. (2010): *Reforms to higher education assessment reporting: opportunities and challenges*. „Teaching in Higher Education”, 15, 3.
- Deklaracja Bolońska (1999), dostępny na: [http://www.cie.gov.pl/HLP/files.nsf/0/B875466A81D7114FC1256E860027D8EA/\\$file/Deklaracja_Bolonska.doc](http://www.cie.gov.pl/HLP/files.nsf/0/B875466A81D7114FC1256E860027D8EA/$file/Deklaracja_Bolonska.doc) (otwarty 15.01.2014)
- Dziennik Urzędowy C 111, 6.5 (2008): Zalecenie Parlamentu Europejskiego i Rady z dnia 23 kwietnia 2008 r. w sprawie ustanowienia europejskich ram kwalifikacji dla uczenia się przez całe życie.
- European Commission, dostępny na: http://ec.europa.eu/education/lifelong-learning-policy/eqf_en.htm (otwarty 22.01.2014).
- ERK (2009), dostępny na: http://ec.europa.eu/education/pub/pdf/general/eqf/broch_pl.pdf (otwarty 6.02.2014).
- ERK (2014), dostępny na: http://europa.eu/legislation_summaries/internal_market/living_and_working_in_the_internal_market/c11104_en.htm (otwarty 26.01.2014)
- European University Association (2006): *Quality Culture in European Universities: A Bottom-Up Approach Report on the Three Rounds of the Quality Culture Project 2002–2006* (Brussels, EUA).
- Fletcher R. i in. (2012): *Faculty and students conceptions of assessment in higher education*. „Higher Education”, 64.
- Frederiksen J. R., & Collins A. (1989): *A systems approach to educational testing*. „Educational Researcher”, 18 (9).
- Framework for Qualifications of EHEA (2005), dostępny na: http://www.bologna-bergen2005.no/Docs/00-Main_doc/050218_QF_EHEA.pdf (otwarty 26.01.2014).
- Glyn T. i in. (2011): *Using Self- and Peer-Assessment to Enhance Students' Future-Learning in Higher Education*. „Journal of University Teaching and Learning Practice”, v8 n1 Article 5.
- Goldstein S. i in. (2014): *Faculty and resident evaluations of medical students on a surgery clerkship correlate poorly with standardized exam scores*. „American Journal of Surgery”, 207, 2.
- Gómez Soler S. (2013): *Comparing means can be mean: quantifying the performance of students at standardized exit exams*. „Cuadernos de Administración” (01203592), 26, 46.
- GUS (2013a): *Rocznik branżowy*, dostępny na: http://www.stat.gov.pl/cps/rde/xbcr/gus/rs_rocznik_demograficzny_2013.pdf (otwarty 19.01.2014).
- GUS (2013b): *Szkoly wyższe i ich finanse w 2012*. Zakład Wydawnictw Statystycznych, Warszawa, dostępny na: http://www.stat.gov.pl/cps/rde/xbcr/gus/E_szkoly_wyzsze_2012.pdf (otwarty 26.01.2014).
- Gvaramadze I. (2008): *From quality assurance to quality enhancement in the european higher education area*. „European Journal of Education”, 43, 4.
- Hernandez R. (2012): *Does continuous assessment in higher education support student learning?* „High Education”, 64.
- Internet1, dostępny na: <http://isap.sejm.gov.pl/DetailsServlet?id=WDU20111130658> (otwarty 10.02.2014).
- Internet2, dostępny na: <http://akson.sgh.waw.pl/samorzad/forum/viewtopic.php?f=38&t=15306&start=510> (otwarty 11.02.2014).
- Jurek W. (2013): *Some Comments On Financing Higher Education*. W: J. Woźnicki (red.): *Financing and deregulation in higher education*. Institute Of Knowledge Society Polish Rectors Foundation, Warszawa.
- Karseth B. & Sivesind K. (2010): *Conceptualising curriculum knowledge within and beyond the national context*. „European Journal of Education”, 45.
- Malan T. (2004): *Implementing the Bologna Process in France*. „European Journal of Education”, 39.
- Mateo J. i in. (2012): *Evaluation Tools in the European Higher Education Area (EHEA): an assessment for evaluating the competences of the Final Year Project in the social sciences*. „European Journal of Education”, 47, 3.
- Meyer L. i in. (2010): *An investigation of tertiary assessment policy and practice: Alignment and contradictions*. „Higher Education Quarterly”, 64.
- McLellan E. (2004): *Authenticity in assessment tasks: A heuristic exploration of academics' perceptions*. „UK Higher Education Research & Development”, 23(1).University of Strathclyde, Glasgow.
- Messick S. (1995): *Validity of Psychological Assessment. Validation of Inferences From Persons' Responses and Performances as Scientific Inquiry Into Score Meaning*. American Psychological Association.
- Niemierko B. (1999): *Pomiar wyników kształcenia*. WSiP, Warszawa.
- Niemierko B. (2007): *Kształcenie szkolne. Podręcznik skutecznej dydaktyki*. Wyd. Akademickie i Profesjonalne, Warszawa.

OECD (2010): *The OECD innovation strategy: Getting head start on tomorrow*, dostępny na: <http://www.oecd.org/dataoecd/3/14/45302349.pdf> (otwarty 19.01.2014).

Owen S., Davis G. (2011): *Catering for Student Diversity: Building Academic Skills in Graduate Attributes Learning and Assessment Opportunities through Collaborative Work*. „Journal of University Teaching and Learning Practice”, 8, 2.

POLON (2013): Załącznik do Statutu Polskiej Komisji Akredytacyjnej, uchwalony 10 listopada 2011 r., dostępny na: <https://polon.nauka.gov.l/opi/aa/rejstry/run;jsessionid=8C6E7C2CC-7D775079E4551DAF25A C565.NwsProdB?execution=e1s1> (otwarty 19.01.2014).

Popham W. (1999): *Why standardized tests don't measure educational quality*. „Educational Leadership”, 56.

Price M. (2005): *Assessment standards: The role of communities of practice and the scholarship of assessment*. „Assessment and Evaluation in Higher Education”, 30b.

Pruchnicka M (2013): *Zeszyt dobrych praktyk dotyczących wewnętrznego zapewniania jakości kształcenia w uczelniach*. Fundacja Rozwoju Systemu Edukacji, Warszawa.

QAA (2013): *Quality Code for Higher Education: Assessment of students and the recognition of prior learning*, dostępny na: <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Pages/quality-code-B6.aspx> (otwarty 19.01.2014).

Ramsden P. (2003): *Learning to teach in higher education*. Routledge, London.

Rozporządzenie MNiSW z dnia 23 sierpnia 2012 r. zmieniające rozporządzenie w sprawie warunków prowadzenia studiów na określonym kierunku i poziomie kształcenia (DzU 2012, poz. 983), dostępny na: <http://isap.sejm.gov.pl/DetailsServlet?id=WDU20120000983> (otwarty 19.01.2014).

Rozporządzenie MNiSW z dnia 29 września 2011 r. w sprawie warunków oceny programowej i oceny instytucjonalnej (DzU 2011 nr 207, poz 1232), dostępny na: <http://isap.sejm.gov.pl/DetailsServlet?id=WDU20112071232> (otwarty 19.01.2014).

Rószkiewicz M. (2009): *Diagnoza stanu szkolnictwa wyższego w Polsce. Wyzwania w obszarach strategicznych*. W: *Diagnoza stanu szkolnictwa wyższego w Polsce*. Wyd. Uniwersytetu Warszawskiego, Warszawa, dostępny na: www.krasp.org.pl/pliki/07d3834a2204e573048f07df6f757297.pdf (otwarty 19.01.2014).

Rust C. (2002): *The impact of assessment on student learning*. Active Learning in Higher Education, 3(2).

Rust C. (2007): *Towards a scholarship of assessment*. Assessment and Evaluation in Higher Education, 32.

Sambell K. i in. (1997): *But is it fair?: An exploratory study of student perceptions of the consequential validity of assessment*. Studies in Educational Evaluation, 23, 4.

Sarrico C., i in. (2010): *Assessing Quality and Evaluating Performance in Higher Education: Worlds Apart or Complementary Views?* Minerva, 48.

Spencer M. S. (2001): *Enforced cultural change in Academe: A practical Case study*. Implementing Quality Management system in Higher Education, vol. 26.,n.1.

Struyven K. i in. (2005): *Students' perceptions about evaluation and assessment in higher education: a review*. Assessment & Evaluation in Higher Education, 30, 4.

Tang S. & Chow A. (2007): *Communicating feedback in teaching practice supervision in a learning-oriented field experience assessment framework*. Teaching and Teacher Education, 23(7).

Tiana Ferrer A. (2010): *The impact of the Bologna Process in Ibero-America: prospects and challenges*. European Journal of Education, 45.

Ustawa z dnia 12 września 1990 r. o szkolnictwie wyższym (DzU Nr 65, poz. 385, z późn. zm.), dostępny na: <http://isap.sejm.gov.pl/DetailsServlet?id=WDU19900650385> (otwarty 19.01.2014).

Ustawa z dnia 18 marca 2011 r. o zmianie ustawy – Prawo o szkolnictwie wyższym (DzU 2011, nr 84, poz. 455), dostępny na: <http://isap.sejm.gov.pl/DetailsServlet?id=WDU20110840455> (otwarty 19.01.2014).

UUK, (2007): *Beyond the honours degree classification*. Burgess Group Final Report. Universities UK, London, dostępny na: <http://www.universitiesuk.ac.uk/highereducation/Documents/2007/Burgess-BeyondHonours.pdf> (otwarty 19.01.2014).

Weisbrod Burton A. i in. (2008): *Mission and money – Understanding the university*. Cambridge University Press, New York.

Woźnicki J. (2013): *2010–2020 Development Strategy for the Higher Education System The Academic Community's Proposal*. W: J. Woźnicki (red.): *Financing and deregulation in higher education*. Institute Of Knowledge Society Polish Rectors Foundation, Warszawa.

Zał. do Statutu PKA uchwalonego w dniu 10 listopada 2011 r., dostępny na: http://pka.edu.pl/Dokumenty/Uchwały/statut_final_10.11.2011.pdf (otwarty 19.01.2014).

O potrzebie ewaluacji efektów kształcenia w szkołach wyższych: stan, problemy, perspektywy

Tekst ten jest refleksją nad zmianami zachodzącymi w polskim szkolnictwie wyższym oraz ich konsekwencjami dla organizacji procesu kształcenia. Przedstawiono w nim uwarunkowania społeczno-polityczne związane z procesem globalizacji oraz przyjęte przez Polskę zobowiązania dotyczące obszaru szkolnictwa wyższego. Zmiany legislacyjne w Polsce będące konsekwencją tych zobowiązań wprowadziły nową strukturę kształcenia na poziomie wyższym oraz znacznie zwiększyły autonomię uczelni w tworzeniu oferty edukacyjnej. Jednak warunkiem koniecznym prowadzenia działalności dydaktycznej jest funkcjonowanie wewnętrznego systemu zapewniania jakości kształcenia. Uczelnie mają obowiązek przeprowadzania samooceny szerokiego spektrum uwarunkowań procesu kształcenia, w tym osiągnięcia przez studentów założonych efektów kształcenia. W artykule wskazano na trudności w realizacji tego ostatniego zobowiązania. Ukształtowane przez wieloletnie tradycje elitarnego kształcenia wyższego ocenianie na zasadzie obserwacji w kontaktach „uczeń–mistrz” ani nie sprzyjało wypracowaniu koncepcji badań osiągnięć studentów.

Słowa kluczowe: jakość kształcenia, szkolnictwo wyższe, proces boloński, ocenianie, ewaluacja

On the need for evaluation of learning outcomes in institutions of higher education: state, problems and prospects

This text is a reflection on changes in the Polish higher education and their consequences for the organization of the educational process. It presents the socio-political conditions associated with the globalization process and the commitments adopted by Poland in the area of higher education. Legislative changes in Poland resulting from these commitments have introduced a new structure of higher education and have significantly increased the autonomy of universities in the creation of the educational offer. However, a prerequisite for didactic activities is functioning of the Internal Education Quality Assurance System. Universities have an obligation to carry out self-assessment through a wide range of factors related to the educational process, including the achievement of the stated learning outcomes by students. The article points out difficulties in the implementation of the latter obligation. Shaped by a long tradition of elite higher education, assessment based on observations in „student–master” interactions had not been conducive to developing a concept of student achievement assessment

Keywords: Quality of education, Higher education, Bologna Process, Assessment, Evaluation

RELACJE Z BADAŃ

RUCH PEDAGOGICZNY 2/2014

Beata Udzik

Uniwersytet Adama Mickiewicza w Poznaniu

DYSKURS EDUKACYJNY WPISANY W KOMUNIKOWANIE EWALUACJI (WOKÓŁ ZAGADNIEŃ PRAGMATYCZNO- -LINGWISTYCZNYCH)

Jeśli ewaluacja ma być zaproszeniem szkoły do rozwoju i mieć charakter interakcyjny (Korporowicz, 2010, s. 41), nie może ograniczać się do czysto badawczego, a w konsekwencji deskryptywnego działania ewaluatorów. Ważną rolę w uspołecznianiu procesu ewaluacji odegrać może sposób jej komunikowania. Leszek Korporowicz twierdzi: „rozwojowy model ewaluacji w pełni potwierdza możliwość badania ewaluacyjnego w zakresie stymulowania synergii zasobów, kompetencji i procesów, wkomponowując całość jej społecznych rezultatów w strategię doskonalenia działań każdej organizacji, które łączą tak wiele rozmaitych elementów ich dziedzictwa, współczesnych wyzwań i wizji przyszłości” (tamże, s. 38).

W „Poradniku wizytatora” przygotowującego, zazwyczaj zespołowo, raport z ewaluacji zewnętrznej czytamy: „Założeniem ewaluacji zewnętrznej jest przedstawienie rzetelnych, wiarygodnych informacji, które będą zaproszeniem do rozwoju i rozpoczną pozytywne zmiany” (Gocłowska i in., 2013, s. 77). Dlatego tak ważnym dokumentem pozostaje „Raport z ewaluacji” upubliczniany na oficjalnych stronach Systemu Ewaluacji Oświaty. Trudno jednak skonstruować dokument, który kierowany jest do dwóch zdecydowanie różnych adresatów. Z jednej strony bowiem odbiorcą instytucjonalnym jest odpowiedni kurator oświaty jako reprezentant państwa, które w rozporządzeniu określa wymagania stawiane szkołom, co ma się stać elementem kształtowania przez nie polityki oświatowej. Z drugiej strony jednak odbiorcami znacznie bardziej zainteresowanymi wynikami ewaluacji konkretnej placówki są jej dyrektor i nauczyciele, być może uczniowie i ich rodzice, organy samorządowe czy lokalne media.

Tekst raportu, produktu finalnego przeprowadzonej w szkole ewaluacji, wpisuje się w zjawisko znacznie szersze, choć jeszcze mało opisane, jakim jest dyskurs edukacyjny. Trudno przecenić wartość materiałową ponad trzynastu tysięcy raportów ewaluacyjnych, które w momencie powstawania artykułu zamieszczone były na oficjalnej stronie internetowej Nadzoru Pedagogicznego. Stają się one bowiem bogatym materiałem badawczym nie tylko dla pedagogów czy socjologów, ale także

dla językoznawców, których interesuje przede wszystkim ujętykowanie komunikatu w jego aspekcie pragmatycznym, społecznym i tekstologicznym.

Niniejszy artykuł jest efektem analizy 32 – dobranych losowo – raportów z ewaluacji zewnętrznej przeprowadzonej w liceach ogólnokształcących w 16 województwach (po 2 z każdego województwa). Ograniczenie próby badawczej do liceów podyktowane było chęcią analizy materiału językowego, który skupia się wokół zagadnień podobnych ze względu na typ szkoły. W przeważającej większości były to raporty stworzone już w nowej formule, po zmianie rozporządzenia dotyczącego nadzoru pedagogicznego w maju 2013 roku¹.

Grażyna Habrajska (2012, s. 96), idąc za koncepcją procesu komunikacji werbalnej Michaela Alexandra Kirkwooda Hollidaya, wyróżnia trzy poziomy przekazu informacyjnego. Należą do nich: poziom ideacyjny, interakcyjny oraz organizacji dyskursu. Pierwszy z nich pozwala zaistnieć aktowi komunikacji, określa bowiem to, o czym i co się mówi (pisze). Dalej rozważyć należy, do kogo się mówi, w jakim celu i w jaki sposób. Wszystkie składniki sytuacji komunikacyjnej funkcjonują równocześnie i choć bez poziomu ideacyjnego pozostałe zaistnieć nie mogą, to w różnych sytuacjach różna może być waga poszczególnych poziomów przekazu ze względu na jego cel i intencję komunikacyjną. Oto bowiem w tekstach naukowych najważniejsza będzie warstwa ideacyjna, a w tekstach publicystycznych – treść interakcyjna.

Raporty z ewaluacji zewnętrznej pozostają spójne w swej warstwie przedstawieniowej, a spójność tę gwarantuje ujednoczenie wymagań ewaluacyjnych wynikających z rozporządzenia, do którego autorzy raportu się odwołują. „Odtworzenie warstwy przedstawieniowej sytuacji komunikacyjnej jest w istocie procesem obrazowania świata” (tamże, s. 96). Obraz świata odtwarzanego w raportach tworzy leksyka związana z życiem szkoły: jego głównymi uczestnikami, zasobami oraz procesami, które w nim zachodzą. Stąd powtarzające się rzeczowniki osobowe i wyrażenia nazywające najważniejszych uczestników szkolnej rzeczywistości: *uczniowie, nauczyciele, zespoły nauczycieli, zespoły przedmiotowe, zespoły zadaniowe, dyrektor, rodzice, partnerzy szkoły, pracownicy niepedagogiczni, czasami przedstawiciele samorządu, organu prowadzącego*. Wymieniani są oni jako podmioty działające w szkole lub jej otoczeniu, ale także jako respondenci w prowadzonych badaniach.

Kolejna sfera słownictwa odnosi się do metodologii prowadzonych badań i wykorzystywanych w nich narzędzi, które wymienione zostają już w pierwszej części raportu: *ankieta, wywiad grupowy, wywiad indywidualny, obserwacje lekcji oraz analiza dokumentacji*. Dużą frekwencję wykazuje rzeczownik *szkoła* i jego synonimy: *placówka, liceum*. Dalej wkraczamy już w sferę rzeczowników nazywających dokumenty regulujące pracę szkoły: *koncepcja pracy szkoły* (w wymaganiach zastrzega się, że nie musi być ona zapisanym dokumentem, ale najczęściej jest), *podstawa programowa, program wychowawczy, program profilaktyki, statut szkoły, raport maturalny*. Ważną grupę słownictwa stanowią też wyrazy związane ze sferą dydaktyki (np.: *wiadomości, umiejętności, kompetencje, zainteresowania, uzdolnienia, cele nauczania, metody, arkusze egzaminacyjne, zadania, ćwiczenia, oceny, system oceniania*), wychowania (np.: *normy społeczne, postawy, wartości*) czy organizacji pracy szkoły (np.: *samorząd uczniowski, lekcja, klasa, zajęcia pozalekcyjne, koła zainteresowań*). Formy owych działań szkoły w sferze dydaktyki, wychowania i organizacji mają charakter zindywidualizowany w odniesieniu do konkretnej szkoły.

Tym, co jednak łączy wszystkie raporty w sferze słownictwa, jest duża frekwencja rzeczowników określających podejmowane działania i zachodzące procesy. W raportach najczęściej więc czytamy o: *analizowaniu, planowaniu, przygotowywaniu, rozpoznawaniu, podejmowaniu, organizowaniu, wdrażaniu, wprowadzaniu, modyfikowaniu, formułowaniu, kształtowaniu, wspieraniu, angażowaniu, stwarzaniu, zachęcaniu, eliminowaniu, wzmacnianiu, promowaniu, doskonaleniu...* Tak częste występowanie rzeczowników zawierających treści predykatywne pozwala na konstruowanie rozbudowanych zdań, w których dominują nominalizacje. Jeśli grupa imienna (Topolińska, 1984, s. 353) jest właściwą formalizacją językową dla nazwy przedmiotu i nazywanie przedmiotu jest prymarną funkcją grupy imiennej jako konstrukcji gramatycznej, to wysoka frekwencja konstrukcji z rozbudowanymi grupami znominalizowanymi nie pozostaje bez wpływu na sposób ujęcia rzeczywistości pozajęzykowej w tekście. Ma to oczywiście związek ze stylem, którym posługuje się autor i ku któremu ciąży jego tekst. Nominalizacje, wymieniane jako ważna tendencja we współczesnym języku polskim (Nagórko, 2010, s. 33), stanowią tutaj silny sygnał „urzędowości” tekstu, przyjęcia odpersonalizowanej i przedmiotowej perspektywy opisu.

Od warstwy ideacyjnej komunikatu, jakim jest raport z ewaluacji zewnętrznej, przechodzimy do jego warstwy interakcyjnej. Do kogo kierowany jest ten dokument? Na początku już wspomniano o podwójnej niejako perspektywie odbiorczej: kuratora jako przedstawiciela państwa i pracodawcy ewaluatorów oraz bezpośrednio zainteresowanych jakością pracy w konkretnej szkole i zaangażowanych w jej życie dyrektora, nauczycieli, uczniów, rodziców, organu prowadzącego czy lokalnej społeczności (por. Ciolan, 2010; Udzik, 2019). Czy odbiorcy ci są w jakikolwiek sposób projektowani czy wprost przywoływani w raportach? Formalne zapisy w raportach sytuują kuratora w roli nadawcy, ponieważ odpowiednie kuratorium oświaty wymienione zostaje na pierwszej stronie raportu, a pod nazwiskami ewaluatorów na końcu pozostawione jest miejsce na jego podpis, a więc to on ostatecznie treść raportu potwierdza. No właśnie. Raczej potwierdza, nie tworzy, ponieważ w jego imieniu czynią to wizytatorzy ds. ewaluacji. Ważniejsze jednak okazuje się to, jak w samym tekście istnieje jego odbiorca.

Oto wprowadzona zostaje w nim metatekstowa rama modalna. We wstępie, który w nowej formule raportu po zmianie rozporządzenia jest w każdym dokumencie taki sam, czytamy: *Prezentowany raport jest rezultatem ewaluacji zewnętrznej przeprowadzonej w szkole przez wizytatorów do spraw ewaluacji. Z kolei ostatnie zdanie opisu metodologii brzmi: Na podstawie zebranych danych został sporządzony raport, który obejmuje podstawowe obszary działania szkoły.*

Powyżej zacytowane sformułowania znajdziemy w każdym raporcie. Nadawca dokłada więc starań, aby upewnić odbiorcę, że ten ma do czynienia z tekstem-dokumentem powstałym na podstawie rzetelnie przeprowadzonego badania. Służy temu również przytaczanie każdorazowo wymagań państwa, powoływanie się na treść rozporządzenia, wymienienie dat przeprowadzenia ewaluacji, nazwisk ewaluatorów, zastosowanych metod i narzędzi oraz określenie liczby i typu respondentów objętych badaniem. Z perspektywy nadawczo-odbiorczej czyni to raport dokumentem z po- granicza tekstu urzędowego i naukowego.

Jeśli modalność to „informacja o określonym stosunku nadawcy do opisywanego stanu rzeczy” (Grzegorzczkova, 1998, s. 39), „wyraz postawy osoby mówiącej w stosunku do treści komunikowanych w wypowiedzeniu” (Łapa, 2003, s. 11), „swoista

kalkulacja myślowa nadawcy, tj. akt mentalny wyrażony językowo, który zakłada spełnienie ściśle określonego przez nadawcę celu komunikacyjnego wobec odbiorcy” (Boniecka, 1999, s. 11), to strona tytułowa, wstęp oraz opis metodologii jako części kompozycyjne raportu służą uwypukleniu sytuacji powiadomienia (*chcę, żebyś wiedział, że*), podkreśleniu, że mamy do czynienia z konstatacjami.

W dalszej części raportu zbiorowy nadawca najpierw zapowiada tekst, np.:

Raport, który przedstawiamy dotyczy ewaluacji zewnętrznej problemowej przeprowadzonej w...

by następnie stworzyć ramę perswazyjną całości poprzez bezpośredni zwrot do odbiorców w części „Obraz szkoły”, np.:

Raport, do lektury którego Państwo zachęcamy, dotyczy ewaluacji zewnętrznej problemowej...

Raport, który Państwu przedstawiamy, dotyczy ewaluacji zewnętrznej przeprowadzonej w (...). Poniżej znajdują Państwo najważniejsze informacje o szkole wynikające z przeprowadzonych badań. Wszystkie zawarte w niniejszym tekście dane i wnioski znajdują potwierdzenie w wynikach przeprowadzonych badań.

Raport, który przedstawiamy, dotyczy ewaluacji zewnętrznej problemowej przeprowadzonej w (...). W raporcie znajdziecie Państwo szczegółowe informacje pozyskane od różnych grup respondentów oraz wnioski z nich wynikające.

Zachęcamy wszystkich zainteresowanych do przeczytania raportu z ewaluacji zewnętrznej.

Wśród badanego materiału znalazł się także raport, w którym na początku bardzo rozszerzono kategorię odbiorców:

Zachęcamy Państwa do lektury Raportu przygotowanego dla szerokiego kręgu odbiorców, którym przyświeca idea poprawienia jakości pracy szkoły w naszym systemie oświaty.

Wszystkie przywołane powyżej zwroty do potencjalnych odbiorców mają już formę imperatywów (Grzegorzczkowska, 1998, s. 40, 43), oczywiście o niskim stopniu zobowiązania odbiorcy. Nadawca nie przyjmuje tutaj pozycji nadrzędnej wobec odbiorcy, a jedynie zachęca, składa propozycję. W ostatnim zdaniu wyraźnie jednak sugeruje, że jeśli odbiorca zrezygnuje z lektury tekstu, to nie należy do tych, którym przyświeca idea poprawienia jakości pracy szkoły w naszym systemie oświaty.

Przytaczając przykłady/sposoby zachęcania odbiorcy do zapoznania się z treścią raportu, można zaobserwować intencję i cel zakładane przez nadawcę. Należałoby tutaj oddzielić intencje ewaluatorów przeprowadzających w szkole badanie od intencji, która wyrażona zostaje w samym raporcie. Jeśli bowiem zadaniem procesu ewaluacji zewnętrznej pozostaje sprawdzenie, na ile szkoła sprostała wymaganiom sformułowanym przez państwo, o tyle intencją nadawcy raportu jako komunikatu informującego o wynikach „praktycznego badania oceniającego” (definicja z rozporządzenia) jest zachęcenie odbiorcy do zapoznania się z jego treścią, a celem – „przedstawienie rzetelnych, wiarygodnych informacji, które będą zaproszeniem do rozwoju i rozpoczną pozytywne zmiany” (Goćłowska i in., 2013, s. 77).

Wyjaśniając istotę pojęcia raport, Anna Goćłowska pisze: „**Co to znaczy napisać raport** i czym właściwie jest **raport**? Wywodzi się z nomenklatury wojskowej, nazewnictwo przejęte zostało przez inne dziedziny życia. Raport z badania jest pisemnym opracowaniem, które zawiera prezentację wyników i ma posłużyć udzieleniu

informacji na temat, który jest przedmiotem raportu. (...) Zatem sensem raportu jest opisanie rzeczy, ale takie opisanie, które jak najwięcej mówi o szkole, przedszkolu, placówce i zostawia jak najmniejszy margines niepewności. Zadaniem raportu nie jest jednak wymienianie wszystkich działań szkoły/placówki. Masz podać **wyselekcjonowane**, najciekawsze przykłady. (...) Raport stanowi esencję pracy ewaluatora. Jest opisem działań podejmowanych przez szkołę/placówkę w celu realizacji wymagań sformułowanych przez państwo. **Ma być podstawą do dyskusji na temat pracy szkoły.** Musi więc być **wartościowy merytorycznie**" (tamże, s. 65–66).

W „Poradniku wizytatora” podkreśla się, że raport powinien być zdjęciem, a wizytator nie może ani reklamować szkoły, ani też przedstawiać jej w czarnych barwach. W raporcie nie ma też miejsca na rekomendacje. Tymczasem oprócz konstatacji znajdziemy w nim również wypowiedzi postulatywne (modalność deontyczna). Pojawiają się one najczęściej we wnioskach, np.:

W celu doskonalenia procesów edukacyjnych należałoby w większym zakresie umożliwić uczniom wpływ na sposób organizowania i przebieg procesu uczenia się, kształtując w ten sposób ich odpowiedzialność za własny rozwój.

Nie brakuje w nich też sformułowań w rodzaju:

uzasadnione byłoby, istnieje potrzeba ich doskonalenia w zakresie, należałoby jednak położyć większy nacisk na stworzenie warunków wyzwalania samodzielnych inicjatyw uczniów, uzasadniona byłaby jednak wnikliwa analiza wyników matur.

Poziom interakcyjny wypowiedzi to nie tylko relacja nadawczo-odbiorcza, intencja/cel przekazu, ale także kształt komunikatu. Pytamy więc nie tylko *kto? do kogo? w jakim celu?*, ale także *w jaki sposób?* Zagadnienie kształtu językowo-stylistycznego tekstu pozostaje w bezpośrednim związku z poziomem dyskursu (Habrajska, 2012, s. 103–109) rozumianego jako typ społecznej praktyki komunikacyjnej (Nocoń, 2013, s. 15). Tekst będący bodźcem do interpretacji na poziomie dyskursu ma określoną formę językową, narzuca typ kontaktu (tutaj oficjalny) i rodzaj dyskursu ze względu na kryterium tematyczne (edukacyjny) oraz stylowe (głównie urzędowy).

Poddany analizie materiał wykazuje wiele cech stylu urzędowego, do których należą bezosobowość, szablonowość, dążenie do precyzji. Ekspozycja formalnej więzi między uczestnikami aktu komunikacji odbywa się poprzez formy nieosobowe czasownika z morfemem *się* lub zakończone na *-no* i *-to* (Wojtak, 2001, s. 161; Zdunkiewicz-Jedynak, 2010, s. 165), np.: *w liceum **respektuje się** podstawę programową, w szkole **rozmawia się** z uczniami, **omawia się** tematykę bezpieczeństwa, dzięki czemu **nie odnotowuje się** incydentów; w wyniku prowadzonego badania **zdiagnozowano** następujące najważniejsze potrzeby, na lekcjach **zaobserwowano, zorganizowano** zajęcia dodatkowe.* O ile czasowniki z morfemem *się* dotyczą tego, co bezpośrednio dzieje się w szkole, o tyle czasowniki zakończone na *-no* i *-to* orzekają zarówno o rzeczywistości szkolnej, jak i o działaniach ewaluatorów.

Również konstrukcje składniowe typu: *być + imiesłów bierny* lub *zostać + imiesłów bierny* ekspozują bezosobowy charakter tekstu, np.: *przy tworzeniu oferty zajęć pozalekcyjnych **brane są pod uwagę** zainteresowania i potrzeby uczniów, opinie te **są wykorzystywane** przy podejmowaniu decyzji, procesy edukacyjne **są planowane i realizowane** zgodnie z potrzebami uczniów, uczniowie **są zmotywowani** do nauki, większość licealistów **jest zainteresowanych** oceną swoich osiągnięć.*

Językowy wykładnik standardowości i silnej konwencjonalności raportów ewaluacyjnych to powtarzalność różnych elementów językowych, utartych formuł,

zwłaszcza w partiach kompozycyjnych (na początku i na końcu). Szablonowa pozostaje budowa całości, schemat wypełniany jest danymi dotyczącymi określonej szkoły. W porównaniu z raportami sprzed maja 2013 roku te przygotowywane w nowej formule charakteryzuje dużo większa spójność narracji w opisach spełniania wymagań. Znacznie częściej występują tutaj czasowniki w formie osobowej orzekające o działaniach nauczycieli i uczniów, o stanie ich świadomości na temat zachodzących w szkole procesów. Równocześnie jednak stopień szablonowości zwiększyło wykorzystywanie w raportach wykresów i tabel. Jako dowody potwierdzające dane pochodzące z przeprowadzonego badania podaje się kopiowane z platformy dane przedstawiane właśnie w postaci pionowych lub poziomych wykresów. W poradniku zaleca się wręcz „częste odwoływanie się do danych ilustrowanych wykresami oraz uogólnianie danych z wywiadów” (Goćłowska i in., 2013, s. 65.). Tekst raportu traci na ciągłości, ale zyskuje na wiarygodności.

Szablonowość jednak zwiększa się i z tego powodu, że w badaniach przeprowadzonych w różnych szkołach respondenci wypełniają te same ankiety, w których pojawiają się te same pytania powiązane na platformie z adekwatnymi dla nich kryteriami i wymaganiami. W opisie spełniania wymagania pojawiają się takie same sformułowania, ponieważ są one powtórzone z kwestionariusza ankiety.

Przytoczone przykłady (cytaty z raportów) ilustrujące sposób opisu kryterium z poziomu D: „Planowanie procesów edukacyjnych w szkole lub placówce służy rozwojowi uczniów, a nauczyciele stosują różne metody pracy dostosowane do potrzeb ucznia, grupy i oddziału”, dotyczącego wymagania drugiego: „Procesy edukacyjne są zorganizowane w sposób sprzyjający uczeniu się”. Pokazują one, że nadawca po prostu przepisuje nagłówki i łączy je z danymi liczbowymi z wykresów, dołączonymi poniżej (przykład 1 i 2) lub tworzy spójną narrację, korzystając z danych zarówno ilościowych, jak i pochodzących z wywiadów czy obserwacji, nie dołączając wykresów w ogóle (przykład 3).

Przykład 1. Wymaganie spełnione na poziomie C:

*Większość nauczycieli **potrafi zainteresować uczniów tematem lekcji** (rys. 1j); **rozumiale tłumaczy zagadnienia** (rys. 2j) oraz w taki sposób prowadzi zajęcia, że **uczniowie chcą się uczyć** (rys. 3j). Na ponad połowie zajęć uczniowie współpracują ze sobą (rys. 4j).*

Przykład 2. Wymaganie spełnione na poziomie C:

*Procesy edukacyjne służą rozwojowi uczniów i są dostosowane do ich potrzeb. W opinii młodzieży większość nauczycieli **rozumiale tłumaczy zagadnienia** (rys. 1j), **potrafi zainteresować tematem** (rys. 2j) oraz organizuje pracę w grupach (rys. 3j). Stosowane przez większość nauczycieli metody pracy powodują, że **uczniowie chcą się uczyć** (rys. 4j).*

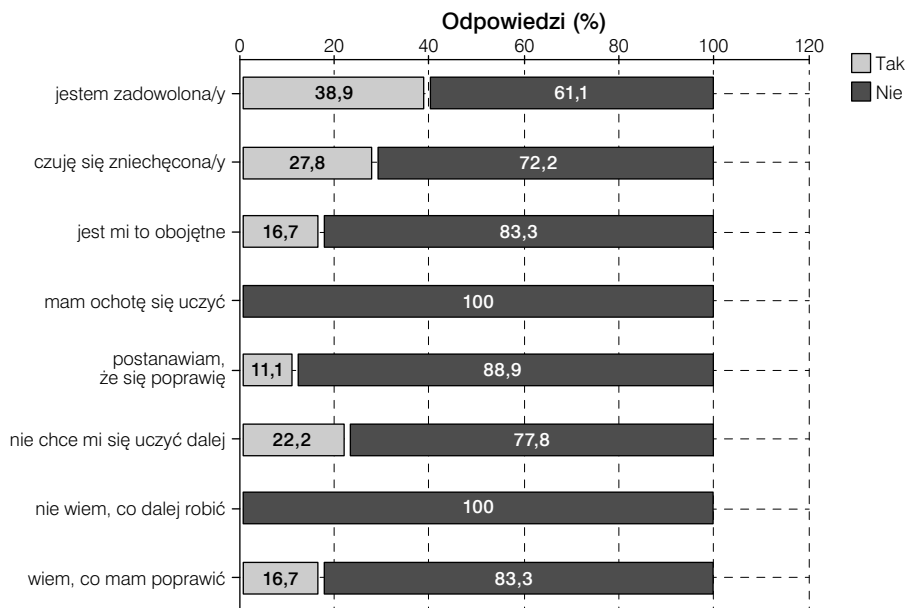
Przykład 3. Wymaganie spełnione na poziomie B:

*Uczniowie w zdecydowanej większości potrafią współpracować ze sobą na zajęciach. Nauczyciele **starają się ich zainteresować tematem, tłumaczą zagadnienia rozumiale** i sprawiają, że **uczniowie chcą się uczyć**. Obserwowane lekcje są bardzo interesujące i prowadzone w podobnym stylu, a uczniowie nie nudzą się. (...). Uczniom podoba się sposób oceniania, praktyczne informacje przydatne w życiu codziennym, rozwiązywanie zadań, osobiste zaangażowanie nauczyciela i żywe tempo lekcji. Podoba się również, że są traktowani indywidualnie, z szacunkiem i czują się bezpiecznie. (...). Za niewłaściwe uważają, gdy na lekcji jest mało wyjaśnień, brak przykładów oraz chaotyczny dobór i rozwiązywanie zadań. (...)*

W raportach można znaleźć także przykłady (por. rysunek) opisywania wykresów: *Kiedy licealiści są oceniani, najczęściej odczuwają zadowolenie lub czują się zniechęceni (rys.1w).*

Jak się zazwyczaj czujesz, kiedy jesteś oceniana/y?
Wybierz opisy, które najlepiej pasują do Ciebie. KADUMS

Liczebność: 18



Zdanie, będące opisem wykresu, jest z punktu widzenia składniowego zdaniem rozłącznym. Ponadto licealiści rzeczywiście: *najczęściej odczuwają zadowolenie lub czują się zniechęceni*, ale przecież dotyczy to około 40% lub 30%, więc choć wydaje się zgodne w wykresem, to sam komunikat pozbawiony rozwinięcia jest dla odbiorcy (niespecjalisty w dziedzinie psychometrii – dyrektor, nauczyciele, uczniowie, rodzice, samorząd) mało czytelny, zwłaszcza w kontekście innych danych (22,1% nie chce się uczyć dalej, a 16,7% jest to obojętne, więc ostatecznie badanie daje informację niepozytywną, podczas gdy jako pierwsza w zdaniu pojawia się pozytywna – *najczęściej odczuwają zadowolenie*). Jest to jakiś rodzaj manipulowania danymi, choć nie ma tutaj formalnie informacji nieprawdziwej, a nadawca zapewne chciał sprostać zadaniu, które w poradniku zostało sformułowane następująco: „**Jeśli chcesz, by Twój raport był wiarygodny, w procesie dowodzenia wskazuj zawsze źródła informacji** (np. dyrektor, nauczyciele, uczniowie) i »odpieraj« je przykładami (ilościowymi i jakościowymi). **Wyniki badań podawaj możliwie precyzyjnie** (dane liczbowe, procentowe), korzystaj z różnych źródeł danych. Nie zapominaj też, że sposób przekazu musi być zrozumiały dla odbiorców” (Goćłowska i in., 2013, s. 66).

Chęć uwiarygodnienia podawanych informacji i opinii powoduje, że ewaluatorzy bezpośrednio powołują się na wypowiedzi respondentów, podają źródła i sposób zdobycia informacji. Stąd duża frekwencja zdań dopełnieniowych, przyłączanych spójnikiem podrzędnym że jako wskaźnikiem zespolenia: *dyrektor szkoły w indywidualnej ankiecie poinformował, że; większość uczniów z klasy II uważa, że; jednak 10 z 37 wyraża*

opinię, że; prawie wszyscy licealiści twierdzą, że; z wywiadu z uczniami po obserwowanych zajęciach wynika, że; rodzice uważają, że. Równie często pojawia się zdanie podrzędne lub jego równoważnik zapowiadające mowę zależną: *jak twierdzi dyrektor szkoły, jak twierdzą rodzice, jak twierdzi większość uczniów, zdaniem wszystkich nauczycieli, w opinii wszystkich ankietowanych rodziców.*

Wymienione powyżej zabiegi kompozycyjne i językowo-stylistyczne, stosowane przez autorów raportów ewaluacyjnych, świadczą o dążeniu do precyzji. To cecha wyróżniająca zarówno teksty urzędowe, jak i naukowe. Napisano już, że urzędowość jest ważną cechą stylu raportów z ewaluacji zewnętrznej. Jeśli zatem do pożądaných cech stylu urzędowego należą – jak chce Dorota Zdunkiewicz-Jedynak (2010) – jednoznaczność i komunikacyjność, ponieważ skierowane są one nie tylko do wąskiego grona specjalistów (w przypadku raportów są to rodzice, przedstawiciele organów samorządowych, społeczności lokalne zainteresowane poziomem oświaty, lokalne media), powinny być zrozumiałe. Tymczasem dążenie do precyzji może powodować różne cele nadawcze, które pozostają w konflikcie. „Długość zdań występująca w służbie precyzji – czyni wiele tekstów urzędowych językowo niezręcznymi, naruszającymi walory estetyczne wypowiedzi, wreszcie trudnymi do zapamiętania i przez to mało komunikatywnymi. Redaktorzy tekstów urzędowych nieuchronnie stają przed trudnym zadaniem znalezienia złotego środka: osiągnięcia każdej z cech (jednoznaczności, precyzji i zrozumiałości) w stopniu optymalnym dla danych okoliczności” (Zdunkiewicz-Jedynak, 2010, s. 165).

Zasadą organizacji tekstów urzędowych pozostaje wyliczenie, które ma służyć precyzji właśnie. Tymczasem w raportach znaleźć można przykłady, w których wyliczenie nie tylko nie przyczynia się do zwiększenia precyzji komunikatu, ale narusza zasady gramatyczne, stylistyczne i logiczne (brak gramatycznego, zwłaszcza składniowego skorelowania z nagłówkiem lub formą fleksyjną innych elementów wyliczenia; brak uporządkowania logicznego wyliczanych cech czy elementów, brak ich uporządkowania i zhierarchizowania), np.:

*W wyniku prowadzonego badania **zdiagnozowano** następujące najważniejsze **potrzeby** psychorozwojowe i społeczne uczniów:*

- **uwzględnianie** w toku nauczania różnic w rozwoju emocjonalnym i edukacyjnym dzieci,
- **potrzeba akceptacji** i wyrażania siebie,
- **poczucia sukcesu**,
- **rozmów** z wychowawcą,
- **indywidualnego traktowania** uczniów o zaburzonych możliwościach intelektualnych,
- **pochwały** pracy każdego dziecka,
- **potrzeba dofinansowania** zakupu podręczników szkolnych, obiadów, wycieczek,
- **zaangażowania się** w życie szkoły poprzez udział w wolontariacie.

Autor (autorzy) komunikatu – rozbudowanego zdania pojedynczego z jednym orzeczeniem (czasownikiem zakończonym na *-no*, czyli formą traktowaną jako wyraz bezosobowości, sposób unikania mówienia czy też pisania bezpośrednio o wykonawcy czynności, podmiocie podejmującym działanie) – zastosował wyliczenie grup nominalnych. Ich członem konstytutywnym pozostaje zawsze formalnie rzeczownik. Czynność czy stan wyrażane rzeczownikiem z pozycji predykatu przechodzą w funkcję argumentu. Ponadto zagubiona zostaje struktura składniowa, brakuje konsekwencji w składniowym kształtowaniu rozwiniętego zdania formalnie ciągle pojedynczego.

Wydaje się, że również dążenie do precyzji i chęć powołania się na informacje pochodzące bezpośrednio od respondentów (dyrektora i nauczycieli) przyświecała autorom poniższego zdania: **W związku z tym trudno jednoznacznie stwierdzić, czy wprowadzone zmiany w pracy, które wymienił dyrektor (zwiększono ilość zajęć dodatkowych dla uczniów słabych w ramach dodatkowej godziny dla n-la; zorganizowano zajęcia dodatkowe dla uczniów startujących w konkursach i olimpiadach przedmiotowych; nauczyciele opracowali plany dydaktyczne uwzględniające możliwości uczniów; systematycznie uzupełniany jest księgozbiór szkolnej biblioteki o lektury, opracowania historyczno-literackie; organizowane są spotkania rodziców klas maturalnych z dyrektorem oraz wychowawcami dotyczące egzaminu maturalnego jego organizacji, ale przede wszystkim poziomu przygotowania uczniów oraz wymagań egzaminacyjnych; organizowane są próbne egzaminy maturalne; systematycznie doposażana jest szkoła w sprzęt multimedialny; wzmocniono nadzór nad pełnieniem dyżurów nauczycieli na korytarzach i przestrzeganiem wewnętrznych zasad oceniania oraz działań zapewniających bezpieczeństwo uczniom w miejscach wskazanych przez nich jako niebezpieczne; wypracowano przez zespół wychowawców zasady postępowania zmierzające do eliminowania liczby godzin nieobecności uczniów) i nauczyciele (rys.20), są efektem monitorowania i analizowania wdrożonych działań wynikających z przyjętych do realizacji wniosków z ewaluacji wewnętrznej i analizy egzaminów zewnętrznych.**

Zdanie wielokrotnie złożone z wtrąceniem składa się ze 153 wyrazów. Jego konstruktor, chcąc przywołać informacje pochodzące z przeprowadzonego badania, nie tylko nie ujednolcił i odpowiednio nie przyporządkował wypowiedzi składowych, ale przede wszystkim nie wziął pod uwagę możliwości percepcyjnych odbiorcy.

W rozbudowanych konstrukcjach składniowych, zarówno rozwiniętych ponad miarę zdaniach pojedynczych, jak i wielokrotnie złożonych, pojawiają się sformułowania pleonastyczne i tautologiczne, które zakłócają zwięzłość i jasność stylu, np.:

Efektywniej wykorzystywać zajęcia dydaktyczne, wykorzystując wiedzę o trudnościach uczniów w nauce...

Podczas lekcji uczniowie mają możliwość odwołania się do innych doświadczeń niż przedmiotowe tj.: wiedzy przedmiotowej, wiedzy z innych przedmiotów, doświadczeń pozaszkolnych.

Dyrektor poinformował, że wspiera nauczycieli wykazujących inicjatywę w zakresie stosowania nowatorskich rozwiązań poprzez wspieranie nauczycieli w ich działaniach...

Pod tym kątem organizuje się zajęcia dydaktyczne oraz tworzy ofertę zajęć dodatkowych umożliwiającą uczniom możliwie wszechstronny rozwój oraz nabycie wiedzy umożliwiającej dalsze kształcenie.

Prowadzona w szkole diagnoza (...) skutkuje zwiększeniem efektywności i skuteczności procesu nauczania.

Wyrazem braku dbałości o formę komunikatu pozostają pojawiające się w raportach błędy frazeologiczne i składniowe, np.:

Dużą rolę w kształtowaniu postaw przywiązuje się do osoby patrona szkoły...

Taki model organizacyjny sprzyja powszechnemu i aktywnemu uczestnictwu kadry pedagogicznej we wzajemnej współpracy.

Także ankietowani uczniowie klas III i II wskazali, że zdecydowana większość nauczycieli wierzą w ich możliwości (zob. rys. 1j) i dają im to odczuć (zob. rys. 2j).

Dokonywane są również analizy wyników próbnego egzaminu maturalnego w celu określenia słabych stron w przygotowaniu do matury, analizę na wejściu – na podstawie danych OKE i analizę wskaźników jakości pracy szkoły, w tym osiągnięć uczniów.

Jeśli „dyskurs to przestrzeń konfiguracji sensu, którą odtwarzamy w procesie interpretacji tekstu”, „wyobrażeniowa przestrzeń mentalna” (Habrajska, 2012, s. 103), a tekst pozostaje fenomenem fizycznym, bodźcem do owej interpretacji, to analizowane raporty z ewaluacji zewnętrznej ze względu na swą wagę instytucjonalną oraz bogactwo zgromadzonego materiału empirycznego wpisują się w dyskurs edukacyjny ostatnich kilku lat.

Jolanta Nocoń (2013, s. 130), pisząc o stylu dydaktycznym/stylu dyskursu dydaktycznego, zwraca uwagę na polimorficzność szkolnej przestrzeni komunikacyjnej, w której istnieją nie tylko gatunki o funkcjach edukacyjnych, ale także urzędowe, prawno-dyrektywne czy ewaluacyjne właśnie. Raport z ewaluacji zewnętrznej, pozostając tekstem publicznym, kliszowanym, ale i użytkowym, nie może pozostać martwym zapisem, do którego z niechęcią będą sięgać jedynie przymuszeni instytucjonalnie dyrektorzy i nauczyciele. To, czy będzie on dokumentem, który realnie zaistnieje w przestrzeni organizacji, a jego odbiorcy dysponujący wiedzą systemową (kompetencją komunikacyjną), encyklopedyczną (faktograficzną) oraz aksjologiczną (wyobrażeniami o istniejących normach społecznych i wartościach), będą potrafili dopełnić jego sens (Habrajska, 2012, s. 107), zależy w dużej mierze od jego kształtu językowo-stylistycznego, sposobu projektowania odbiorcy i jakości prezentowanych analiz.

Postulatem badawczym pozostaje ciągle postawione przez J. Nocoń pytanie o to, czy przeprowadzona na przełomie XX i XXI wieku reforma polskiej szkoły „wpłynęła na zinstytucjonalizowany styl dydaktyczny”. „Istotną byłaby odpowiedź na pytanie, czy i na ile takie fundamentalne hasła reformy, jak: podmiotowość, antropocentryzm, holizm i integracja zmieniły styl komunikowania się nauczyciela z uczniami na lekcji szkolnej” (Nocoń, 2013, s. 134). W kontekście zaprezentowanego artykułu pytanie powinno rozszerzyć swój zakres o teksty komunikujące diagnozę i ewaluację w ich różnorodnych ujęciach.

Przypisy

¹ Źródłem cytatów i przykładów są raporty pochodzące ze strony: <http://www.npseo.pl/action/raports> (15. 03. 2014). Zapis cytatów zgodny z oryginałem, wszystkie przytoczenia z raportów wyróżnione kursywą, aby podkreślić ich odmiennosc od cytowania tekstów innego rodzaju.

Literatura

- Boniecka B. (1999): *Wykład o modalności*. „Annales UMCS”, Sectio FF, Vol. XVII, Lublin.
- Ciolan L. (2010): *W trosce o wysoką jakość raportu z ewaluacji – kilka uniwersalnych prawd*. W: G. Mazurkiewicz (red.): *Ewaluacja w nadzorze pedagogicznym. Odpowiedzialność*, Wyd. UJ, Kraków.
- Gocłowska A. i in. (2013): *Ewaluacja zewnętrzna. Poradnik wizytatora*. Ośrodek Rozwoju Edukacji, Warszawa.
- Grzegorzczkowska R. (1998): *Wykłady z polskiej składni*. Wyd. Naukowe PWN, Warszawa.
- Habrajska G. (2012): *Wybrane zagadnienia wprowadzające do nauki o komunikowaniu*. Primum Verbum, Łódź.
- Korporowicz L. (2010): *Interakcyjna misja ewaluacji*. W: G. Mazurkiewicz (red.): *Ewaluacja w nadzorze pedagogicznym. Konteksty*. Wyd. UJ, Kraków.
- Łapa R. (2003): *Preedykatywne wyrażenia modalne z bezokolicznikiem we współczesnej polskiej prasie*. Wyd. „Poznańskie Studia Polonistyczne”, Poznań.

Nocoń J. (2013): *Dyskurs edukacyjny i jego społeczny zasięg*. W: J. Nocoń, A. Tabisz (red.): *Tekst edukacyjny. „Język a Edukacja”*, t. 2., Opole.

Nocoń J. (2013): *Styl dydaktyczny – styl dyskursu dydaktycznego*. W: E. Malinowska i in. (red.): *Styl współczesnej polszczyzny. Przewodnik po stylistyce polskiej*. Universitas, Kraków.

Topolińska Z. (1984): *Składnia grupy imiennej*. W: Z. Topolińska (red.): *Gramatyka języka polskiego*. PWN, Warszawa.

Udzik B. (2003): *Raport z ewaluacji zewnętrznej jako komunikat*. W: B. Niemierko, M. K. Szmigel (red.): *Polska edukacja w świetle diagnoz prowadzonych z różnych perspektyw badawczych*. GRUPA TOMAMI, Gniezno–Kraków.

Wojtak M. (2001): *Styl urzędowy*. W: J. Bartmiński (red.): *Współczesny język polski*. Wyd. UMCS, Lublin 2001.

Zdunkiewicz-Jedynak D. (2010): *Wykłady ze stylistyki*. Wyd. Naukowe PWN, Warszawa.

Dyskurs edukacyjny wpisany w komunikowanie ewaluacji (wokół zagadnień pragmatyczno-lingwistycznych)

Artykuł poświęcony jest analizie materiału empirycznego – 32 raportów z ewaluacji zewnętrznej przeprowadzonych w liceach ogólnokształcących. Autorka podjęła próbę scharakteryzowania dyskursu edukacyjnego z perspektywy pragmatyczno-lingwistycznej. Materiał językowy analizowany jest na poziomie leksyki i sposobu obrazowania świata, na poziomie interakcyjnym, stylistycznym i gatunkowym.

Słowa kluczowe: ewaluacja, dyskurs edukacyjny, raport, komunikacja, analiza pragmatyczno-lingwistyczna

Educational discourse enshrined in the communication of evaluation (around pragmatic-linguistic issues)

The article reviews an analysis of experiential material- 32 reports from the external evaluation done in grammar schools. The author tried to describe the educational discourse from the pragmatic-linguistic perspective. The language material is analysed at the level of lexis and the way of depicting the world, at the level of interaction, style and genre.

Keywords: evaluation, educational discourse, report, communication, the pragmatic-linguistic analysis

Alina Kalinowska
Uniwersytet Warmińsko-Mazurski

WYNIKI BADAŃ TRZECIOKLASISTÓW JAKO DIAGNOZA KONTEKSTÓW NAUCZANIA MATEMATYKI W KLASACH NAJMŁODSZYCH

Nauczanie matematyki w klasach najmłodszych wywołuje nieustannie wątpliwości związane z jakością edukacji i niepokojem o jej skutek. Mówi się wiele o podstawie programowej, treściach w niej zawartych i jej podręcznikowych realizacjach¹, co wydaje się zasadne, ponieważ w znacznej mierze wyznaczają one codzienność uczenia się matematyki. Podręczniki i przewodniki metodyczne ukierunkowują działania nauczyciela w klasach najmłodszych, budując określony charakter rzeczywistości poznawczej w klasie szkolnej. Uruhamiana w tej przestrzeni aktywność intelektualna dzieci oraz ich matematyczne doświadczenia przekładają się na jakość wypracowanych intelektualnie procedur postępowania z zadaniami matematycznymi. Uczniowie rozwiązują je, odwołując się do strategii skonstruowanych na lekcjach, dlatego wyniki uzyskane przez najmłodszych uczniów w badaniach umiejętności matematycznych mogą pomóc w identyfikowaniu matematycznej codzienności uczniów i stanowić interesujący materiał dla diagnozowania, w jaki sposób kreowane są szkolne sytuacje poznawcze.

Teoretyczne podstawy testowego sprawdzania wyników nauczania obudowane są obszarem napięć ideologicznych, a pomiar osiągnięć – wkraczający coraz mocniej w systemy oświatowe – budzi liczne kontrowersje. Z jednej strony próbuje się negocjować pojęcie jakości opartej na normatywności, która do tej pory były podstawą myślenia o dziecku i instytucjach dla niego przeznaczonych (Dahlberg i in., 2013). Zakwestionowanie sformułowanych przez psychologię rozwojową założeń i wystandardyzowanych oczekiwań jakościowych zmienia sens i cel badania rozwoju dzieci. Z drugiej strony – w myśleniu o uczniu spotykamy się z nasilaniem koncepcji dziecka „normatywnego”, posiadającego określone kompetencje, możliwe do zbadania ustalonym empirycznie poziomem poprawności odpowiedzi testowych. W tę koncepcję wpisują się obowiązkowe egzaminy zewnętrzne, które w Polsce prowadzone są na prawie każdym etapie nauczania. Bolesław Niemierko zwraca uwagę na rozwojowy charakter tego typu diagnostyki (w odróżnieniu od diagnostyki zaburzeń w uczeniu się), która pozwala wyznaczyć kierunki trwałych zmian zachodzących w uczniach. Celowi temu służy gromadzenie informacji uzyskanych w trzyletnich odstępach czasowych, które po odpowiednim przetworzeniu przyjmują formę edukacyjnej wartości dodanej (Niemierko, 2013, s. 14).

Edukacja przedszkolna i wczesnoszkolna obecnie są jedynymi etapami nieobjętymi ogólnopolską wystandardyzowaną weryfikacją. Niemniej uczniowie z klas

początkowych są poddawani zewnętrznym badaniom poziomu umiejętności, wciąż jeszcze dobrowolnym², bo oficjalnie zależnym od decyzji danej szkoły.

Z poszerzającym się obszarem testowania uczniów współgra język, jakim została napisana podstawa programowa dla nauczania początkowego (DzU 2004, Nr 256, poz. 2572, z późniejszymi zmianami). Zapisane w tym dokumencie cele ogólne nauczania odwołują się – co prawda – do umiejętności bycia we współczesnym, niejednoznacznym i zmieniającym się świecie (Bauman, 2012), jednak treści nauczania często mają charakterystyczny rys, który określiłabym mianem potencjalności testowej. Rozumiem przez to szczególny sposób ich formułowania, w efekcie czego podstawa programowa staje się spisem określonych umiejętności i definicji, sformułowanych w taki sposób, by ich poziom można było często i z łatwością weryfikować. Wydaje się, że wyjątkowo podatna na takiego rodzaju upraszczającą operacjonalizację jest edukacja matematyczna. Jej treści, zawarte w podstawie programowej, można łatwo przekształcić w testowy pomiar, sprowadzony do prostych obliczeń rachunkowych, czasem ukrytych w treści równie nieskomplikowanych zadań tekstowych.

Sprzyjający testowaniu charakter sformułowań mają na przykład następujące zapisy podstawy programowej: „sprawnie liczy obiekty (dostrzega regularności dziesiętkowego systemu liczenia), wymienia kolejne liczebniki od wybranej liczby, także wspak”; „zapisuje rozwiązanie zadania z treścią przedstawionego słownie w konkretnej sytuacji, stosując zapis cyfrowy i znaki działań” (tamże). Wyniki uzyskiwane przez dziecko w takim sprawdzaniu jego umiejętności można odnosić w zerowyjedynekowy sposób do oczekiwanego poziomu osiągnięć ucznia, weryfikując bez większych trudności indywidualny i globalny wymiar osiągniętego efektu nauczania. W procesie kształcenia istotne jest jednak, w jakim zakresie analiza tak uzyskanych wyników „otwiera się” (Szyling, 2011, s. 78) na kontekst budowania umiejętności matematycznych.

Przyjmując to założenie, dokonuję w niniejszym tekście rozpoznania wybranych kontekstów kształcenia, analizując wyniki zewnętrznych badań testowych trzecioklasistów. Szukanie w nich pewnej specyfiki matematycznego kształcenia kierowane będzie podejściem diagnostycznym. U jego podstaw leżą dwa nie w pełni rozłączne założenia: (1) ogląd uczniowskich rozwiązań matematycznych mówi nam wiele o osiągniętych przez nich umiejętnościach; (2) wyniki badań mogą świadczyć o tym, w jakich warunkach poznawczych pojęcia matematyczne były przez uczniów budowane.

Diagnozowanie umiejętności matematycznych – uwagi na marginesie

Diagnoza i jej rozumienie czerpią wprost z pozytywistycznego paradygmatu edukacyjnego, a jej cechą specyficzną jest przeświadczenie o możliwości poznania rzeczywistości. Dorota Klus-Stańska tak określa tę immanentną cechę procesu diagnozowania: „Teoria diagnostyki edukacyjnej jest wpisana w (...) dynamiczny proces naukowej zmiany, a ostatnie lata przyniosły w nich radykalne zwroty, inspirowane paradygmatami partycypacyjnymi (...). Jednak – jakkolwiek by analizować rewolucyjne zmiany w myśleniu w omawianym zakresie – wszystkie one są raczej wpisane w obiektywistyczny paradygmat nauki” (Klus-Stańska, 2013, 77). Mimo to, zdaniem Grażyny Szyling, nowe koncepcje badań edukacyjnych zmieniają przedmiot poznania diagnozy i ewaluacji edukacyjnej, wykraczając poza najczęściej dotychczas

stosowane w szkole badania praktyczne (Szyling, 2013, 58). Można dostrzec zbieżność tej opinii z przekonaniem Krzysztofa Konarzewskiego, że diagnoza nie powinna ograniczać się do podania danych (Konarzewski, 2000, s. 87), a ich interpretacja oraz możliwość teoretycznego osadzenia badanych zjawisk może mieć zróżnicowany charakter, zależnie od perspektywy przyjętej przez badacza.

Rozwinięcie powyższych spostrzeżeń znaleźć można w tekście Dylana Wiliama (2013). Odwołując się do oceniania kształtującego, wskazuje on na trzy stopnie „pogłębienia” analizy danych testowych. Pierwszy z nich, czyli stwierdzenie problemu (np. określenie umiejętności porównywania ułamków zwykłych), będzie tylko monitorowaniem postępów uczniów, natomiast lokalizację trudności (odkrywanie uczniowskich strategii) przypisuje autor diagnozie. Za niezbędne jednak uznaje też uwzględnienie perspektywy teorii umożliwiających doskonalenie uczenia się, ponieważ to one „identyfikują «przepisy» dla przyszłych działań” (Wiliam, 2013, s. 233).

W zarysowane myślenie słabo wpisuje się powszechne przekonanie o matematyce jako dziedzinie naukowej o niezwyklej precyzji definiowania pojęć oraz prowadzenia rozumowania formalnego³. Z kolei szukanie kontekstów uczenia się poprzez analizowanie sposobu rozwiązania zadań przez ucznia ma z założenia charakter intuicyjny, miękki. Pozwala ono na poszerzenie zakresu myślenia o nauczaniu matematyki w klasach najmłodszych i podjęcie próby głębszego wniknięcia w środowisko uczenia się rozumiane jako „ogół warunków, w których zachodzi uczenie się (...), w którym znajdują się uczniowie i odbywają się lekcje” (Dumont, Instance, 2013, s. 52). Jest ono opisane przez Hannę Dumont i Dawida Instance czterema wymiarami, które tworzą: uczniowie, nauczyciele, treści i pomoce dydaktyczne. Tego rodzaju kontekst poznawczy ma szczególne znaczenie w odniesieniu do kształcenia małych dzieci na etapie wczesnoszkolnym, co potwierdzają przytoczone przez Stanisława Dylaka (2013) wyniki badań prowadzonych przez przedstawicieli psychologii poznawczej. Uzasadniają one konieczność rozwijania plastyczności umysłu, która jest największa w okresie dzieciństwa. S. Dylak postuluje radykalną zmianę w myśleniu o nauczaniu w szkole, pisząc: „może nawet ważniejsze jest to, jak się uczy w stonku do tego, czego się uczymy” (Dylak, 2013, s. 45).

Jeśli nie przywiązujemy uwagi do organizacyjnego i kulturowego porządku, w którym zachodzi uczenie się, ograniczamy – twierdzą H. Dumont i D. Instance (2013, s. 53) – wartość naszych badań. W świetle tego stwierdzenia, traktuję podjętą w niniejszym tekście próbę zidentyfikowania wybranych kontekstów poznawczych nauczania matematyki w klasach najmłodszych jako drobny element przyczyniający się do poznania owego porządku.

Wczesnoszkolne konteksty poznawcze konstruowania wiedzy matematycznej – informacje ukryte w liczbach

W latach 2006–2011 zespół przy Centralnej Komisji Egzaminacyjnej zrealizował badania diagnostyczne w ramach projektu EFS. Ich zakres metodologiczny rozszerzono o metody jakościowe, takie jak wywiady i obserwacje zajęć w szkole. Wyniki uzyskane przez dzieci były analizowane i interpretowane z uwzględnieniem prawdopodobnych sposobów pracy nauczycieli (Dąbrowski, Żytko, 2007; 2008; Dąbrowski, 2009; 2011; Dągiel, Żytko, 2009; 2011; Murawska, Żytko, 2012). Uzyskane wówczas dane stały się

podstawą ogólnopolskiej propozycji wprowadzenia testu umiejętności trzecioklasisty OBUT. Kolejne jego edycje objęły lata 2011, 2012, 2013. Polscy uczniowie klas trzecich brali również udział w badaniach PIRLS 2006 i 2011 oraz TIMSS 2011, a raport K. Konarzewskiego (TIMSS i PIRLS 2011) odsłania ponadto możliwości wykorzystania kontekstu w procesie interpretacji wyników uzyskanych przez uczniów.

Podkreślić należy, że identyfikacja kontekstów poznawczych uczenia się matematyki przez uczniów klas najmłodszych może być jedynie pewną interpretacją. Zmienne wpływające na wyniki uzyskane w badaniach są bowiem tak liczne, że można jedynie nakreślić przybliżone kategorie kontekstowe. W tym tekście przyjrze się dwóm z nich, których istnienie w procesie nauczania wczesnoszkolnego nie zawsze jest uświadamiane, a których scharakteryzowanie może poszerzyć rozumienie mechanizmów poznawczych wytwarzanych na lekcjach matematyki.

Kontekst językowo-relacyjny

Znaczenie, jakie ma język potocznego rozumienia dla budowania pojęć matematycznych, ukazywane jest już na poziomie filozofii matematyki. Z perspektywy społecznych konstruktywistów matematyka oparta jest na języku naturalnym: „Reguły logiki i wewnętrzna spójność matematyki uzależnione są od języka naturalnego jako niewzruszonej podstawy, na której spoczywa matematyczna obiektywność. Logiczny system wiedzy uzależniony jest od terminów i twierdzeń pierwotnych, które zaś tworzą naturalny język wiedzy obiektywnej” (Piotrowska, 2013, s. 291).

Zatem doświadczenia językowe, traktowane jak doświadczenia poznawcze najmłodszych uczniów, mogą mieć istotne znaczenie dla odczytywania tekstu zadania matematycznego. Ilustrują to zjawisko wyniki badań trzecioklasistów, które wskazują na istotne różnice w dekodowaniu treści zadania w zależności – chociażby – od kolejności pojawiania się informacji. Na przykład w 2008 roku uczniowie rozwiązywali typowe zadanie:

Janek, Piotr i Michał zbierają modele samochodów. Janek ma już 24 modele. Piotr ma o osiem więcej niż Janek, a Michał o 2 mniej niż Piotr. Ile modeli ma Michał?

Nie było ono żadnym szczególnym wyzwaniem dla 75,1% badanych uczniów, którzy udzielili poprawnych odpowiedzi. Rozwiązanie wymagało wykonania działań na pojawiających się kolejno liczbach. Jednak zmiana szyku zadania oraz zastąpienie zmiennej wyrażonej cyfrą na zmienną nazwaną słowami spowodowała spadek poprawnych rozwiązań do poziomu 51,8% (por. Dąbrowski, 2009, s. 117–118).

Te różnice mogą wskazywać, że podręcznikowy kontekst poznawania, budowania i dekodowania zadań matematycznych jest schematyczny i oparty na jednakowym, kolektywnym konstruowaniu struktury zadania. Uczniowie budują w określony sposób znaczenia dotyczące zadania z treścią, co ujawnia – być może – występowanie nadmiernego wywieranego przez nauczycieli i autorów pakietów edukacyjnych na czytanie podobnie skonstruowanych tekstów zadań oraz wiązanie schematycznych ciągów słownych z równie typowym opisem za pomocą działań arytmetycznych. Natomiast doświadczenia poznawcze uczniów mogą być w mniejszym stopniu związane z budowaniem rozumienia relacji między danymi w określonej sytuacji opisanej treścią zadania.

Kolejny przykład zadania i jego niezwykle wyniki bardziej wyraziście odsłaniają ten kontekst poznawania relacji matematycznych.

Basia i Ola miały po 20 złotych. Basia dała Oli 5 złotych. Teraz Ola ma więcej pieniędzy niż Basia. O ile więcej?

Liczba poprawnych rozwiązań tego zadania była stosunkowo niska i wynosiła 33%, dużo bardziej interesująca natomiast okazała się najczęstsza (50%) błędna odpowiedź: 5 (Dąbrowski, Wiatrak 2012, 33). Ten wynik wskazuje na ograniczanie w procesie kształcenia kontekstu znaczeniowego pojęcia „porównywanie różnicowe”. Poznawanie przez uczniów porównywania dwóch liczb z użyciem określenia: „o ile więcej” lub „o ile mniej” ogranicza się na lekcjach prawdopodobnie do jak najszybszego powiązania w umysłach dzieci słownego hasła i arytmetycznego aspektu tego pojęcia („o ileś więcej = masz dodać”). Znacznie rzadsze są natomiast inne sytuacje poznawcze, w których uczniowie mogą nadawać znaczenia związane z rozumieniem relacji typu: „mam o 5 więcej niż ty, to znaczy, że muszę mieć tyle, co ty, i jeszcze 5”.

Wyniki tego zadania wskazują również na ograniczone możliwości stosowania porównania różnicowego w sytuacjach pozwalających na korzystanie z niego dla matematyzowania sytuacji życiowych. W efekcie kontekst słowny sytuacji opisywanej w zadaniu tekstowym często nie jest rozpoznawany relacyjnie ze względu na ograniczenie sposobów poznawania porównania różnicowego. Sposoby uczniowskiego radzenia sobie z zadaniami typu: „o ile więcej” lub „o ile mniej” mogą świadczyć o słabym rozumieniu przez nauczycieli potrzeby zróżnicowania kontekstu poznawczego niewykluczającego badania relacji i osłabienia nacisku na szybkie opanowanie schematów. Odsłaniają więc ubóstwo sytuacyjne poznawania pojęcia zarówno pod względem jego zastosowania, jak i znaczeń nadawanym tej relacji matematycznej.

Kontekst (para)użytkowy

Używam pojęcia kontekst (para)użytkowy w zakresie znaczeniowym związanym z rozumieniem, na ile wiedza matematyczna ma moc wykorzystywania jej do rozwiązywania problemów z rzeczywistości pozaszkolnej. Knud Illeris (2006, s. 203), charakteryzując przestrzeń szkolnego uczenia się, odwołuje się do koncepcji Pierre’a Bourdieu oraz paradygmatu scholastycznego Steena Wackerhausena. Pokazuje uczenie się w szkole jako koło, w którym cele nauczania są wyznaczone przez oczekiwania społeczne, a te z kolei konstytuują szkołę i edukację. Autor przywołuje badania i opinie Kristena Hanzena dotyczące nauczania matematyki w szkołach zawodowych w Danii, z których wynika, że: „matematyka nie uzyskuje znaczenia z uwagi na zastosowania w przyszłej pracy ani inne czynniki pozaszkolne; matematyka nabiera znaczenia w szkole i przez szkołę jako część szkolnej społeczności praktyki” (Hansen, za: Illeris, 2006, s. 203). Można zatem przyjąć, że jeśli takie znaczenia nadają wiedzy matematycznej także uczniowie najmłodszy, to może oznaczać, że określenie „wykorzystywać wiedzę” odnosi się jedynie do sytuacji szkolnych, a rozwiązywanie problemów poza szkołą podlega innym, bliżej nieuświadomianym sobie strategiom.

W szczegółowej analizie uczniowskich rozwiązań zadań matematycznych kontekst paraużytkowy ujawnia się wówczas, gdy porównamy wyniki zadania rozwiązywanego przez trzecioklasistów w dwóch wersjach testu, stosowanych w odstępie dwóch lat. Pierwsza z nich została wykorzystana w badaniach umiejętności

podstawowych uczniów trzecich klas szkoły podstawowej w 2010 roku⁴, a druga w dwa lata później, w ogólnopolskim teście badania umiejętności trzecioklasistów OBUT 2012.

Jacek kupił czekoladę za 2,20 zł oraz cukierki za 2,60 zł. Ile otrzymał reszty z 10 zł?

W roku 2010 zadanie miało strukturę otwartą, a dane były przedstawione w postaci wyrażen dwumianowanych, czyli cena była zapisana jako 2 zł 20 gr (por. Raport OBUT 2012). Trzecioklasiści udzielili wówczas 51,5% poprawnych odpowiedzi. Dwa lata później to samo zadanie uczniowie rozwiązywali jako zamknięte, z danymi zapisanymi liczbami dziesiętnymi. Tym razem nieco więcej z nich dobrze rozwiązało zadanie: 57,1%. W drugim przypadku musieli wykorzystać wiedzę pozaszkolną (najprawdopodobniej nierozwijaną systematycznie, ale związaną z własnym udziałem w życiu pozaszkolnym). Biorąc pod uwagę, że zadanie w pierwszej wersji (szkolnej) było pewnym sprawdzianem używania wyrażen dwumianowanych nauczanych na lekcjach, musimy zastanowić się, o czym świadczy porównywalny wynik wykorzystania przez uczniów wiedzy pozaszkolnej. Ponieważ poziom użyteczności obu rodzajów wiedzy nie różni się wiele między sobą, zatem można wnioskować, że skuteczność użytkowa poznawanych w szkole pojęć nie jest większa niż sytuacyjne doświadczenia pozaszkolne.

Interesujący dla powyższych rozważań jest też przykład zestawienia w teście dla uczniów klasy trzeciej analogicznych zadań dotyczących obliczeń temperaturowych.

W Olsztynie był mroźny poranek. Termometr wskazywał temperaturę minus 6 stopni. W Warszawie było cieplej o 4 stopnie. Jaka temperatura była w Warszawie? (por. Dąbrowski, 2011, s. 107–112).

Ta wersja zadania zawierała rysunek termometru, na którym uczniowie mogli „policzyć kreski” i wykorzystać tę umiejętność do wykonania poprawnych obliczeń. Jedynie 32,3% badanych trzecioklasistów rozwiązało to zadanie poprawnie.

Druga wersja zadania różniła się niewiele w obszarze danych, ale była pozbawiona modelu termometru.

W Olsztynie był mroźny poranek. Termometr wskazywał temperaturę minus 8 stopni. W Warszawie było cieplej o 2 stopnie. Jaka temperatura była w Warszawie? (por. tamże)

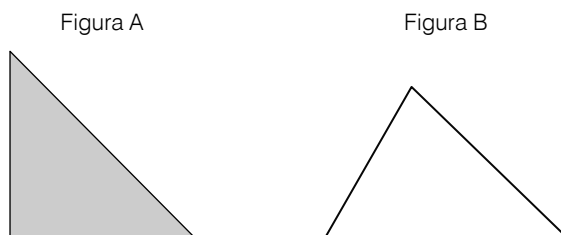
To zadanie poprawnie rozwiązało 32,7% badanych. Okazało się również, że niewielu uczniów sięgało do rysunkowego narzędzia radzenia sobie z sytuacją matematyczną.

Powyższe zestawienie wskazuje na ograniczający kontekst budowania „użyteczności” szkolnej wiedzy matematycznej. Uczniowie klas najmłodszych nie budują narzędzi pomagających im radzić sobie z próbami matematyzowania sytuacji życiowych i prawdopodobnie nie sięgają po nie spontanicznie.

Jeśli przyjąć, że środowisko ubogie w budowanie reprezentacji poznawczych (Bruner, 1978) może być źródłem konstruowania znaczeń o niskich walorach użytkowych, to zasadny wydaje się wniosek, że choć ograniczający kontekst paraużytkowy napędza w pewnym sensie potrzebę uczenia się matematyki, ale służy ono szkole, a nie indywidualnym potrzebom jednostki.

Na koniec analiz chciałabym przytoczyć wyniki rozwiązania zadania, które pokazują ograniczający kontekst poznawczy w zakresie relacji geometrycznych. W 2011 roku dziesięciolatki brały po raz pierwszy udział w międzynarodowych badaniach TIMSS. Oto przykładowe zadanie zastosowane w badaniach:

Rysunek pokazuje dwie figury: A i B. Podaj jedno podobieństwo i jedną różnicę między nimi.



Poprawna odpowiedź brzmiała: A i B to trójkąty, A ma kąt prosty, a B nie ma. Miała ona dowodzić, że uczeń zna pojęcia: trójkąt, kąt, kąt prosty (por. Konarzewski, 2011, s. 19).

Polscy trzecioklasiści uzyskali w zakresie wiedzy geometrycznej poziom wyników poniżej średniej międzynarodowej. Dokładniejsza analiza przykładowego zadania ukazuje, że jego konstrukcja wymaga dostrzegania prawidłowości, porównywania podobieństw i różnic. Mimo zapisu w podstawie programowej o rozpoznawaniu podstawowych figur geometrycznych (a może z powodu jego określonego kształtu), rozwiązanie zadania świadczy, że kontekst poznawczy w tym zakresie może być ograniczany do rozpoznawania kształtów, a nie do budowania pojęć geometrycznych na podstawie rozpoznawania ich własności. Uczniowie prawdopodobnie nie mają wielu okazji do dostrzegania podobieństw i różnic między elementami, badania zmian i odkrywania prawidłowości.

Podsumowanie

Diagnozowanie przez nauczyciela codziennych kontekstów poznawania pojęć matematycznych przez najmłodszych uczniów jest niezbędne dla rozumienia i kształtowania procesu ich uczenia się. Jest to możliwe również dzięki analizowaniu wyników testowania wiedzy matematycznej. Ta relacja ma jednak również kierunek odwrotny, ponieważ wprowadzenie zewnętrznej weryfikacji może zmieniać konteksty nauczania szkolnego. G. Szyling, sięgając do wyników badań Jana Potworowskiego, podaje, że po 2000 roku nastąpiło w Anglii „zamrożenie” wzrostu wyników egzaminów oraz obniżyła się jakość nauczania. Przytaczając z kolei badania K. Konarzewskiego, mówi o podejmowaniu przez nauczycieli zintensyfikowanych wysiłków uczenia pod testy, wbrew ich deklaracjom o wykorzystywaniu wyników testów do doskonalenia własnych kompetencji nauczania (Szyling, 2011, s. 84). Takie podejście do uzyskiwanych przez uczniów efektów nie prowadzi do (re)interpretowania ich w kategoriach własnych działań i znaczeń związanych ze szkolnym środowiskiem uczenia się, a także strategii myślenia tworzonych przez uczniów.

W przedstawionych sygnalnie wynikach trzecioklasistów osiągniętych w testach matematycznych dostrzegamy tendencje ujawniające, że codzienne środowisko szkolnego uczenia się matematyki może okazać się ograniczającym kontekstem

monologicznym⁵. Rozumiem pod tym pojęciem uogólniony zespół czynności lekcyjnych, w których znaczenia nie podlegają społecznym i indywidualnym negocjom. W pewnym sensie pojęcia matematyczne tracą wówczas „społeczny” charakter na rzecz „skończonych” społecznie i mało użytecznych pojęć formalnych. Uczniom proponuje się doświadczenia ubogie, bardziej utrwalające już zdefiniowane pojęcia niż umożliwiające konstruowanie nowych. Uczniowie klas najmłodszych rozwiązują wiele zadań tekstowych, w których jednak rozszerzony jest jedynie kontekst sytuacyjny. Treść zadań zmienia się stosownie do aktualnie omawianych tematów przyrodniczych czy społecznych. Nie zmienia się jednak kontekst relacyjny dotyczący bogactwa „współbycia” ze sobą obiektów matematycznych. Zadania proponowane dzieciom w podręcznikach stanowią powielane w kilku lub kilkunastu sytuacjach te same relacje dodawania lub odejmowania w postaci wykonywania analogicznych rachunków i zapisywania za pomocą znaków matematycznych. Dzieci nie mają okazji do badania innych kontekstów, na przykład sprawdzania, w jaki sposób zmiany składników wpływają na sumę i dlaczego tak się dzieje. Propozycja konstruowania pojęć matematycznych przez najmłodszych uczniów, związana głównie z asocjacyjnymi kontekstami tworzenia treści zadań, stwarza złudne poczucie szerokiego kontekstu poznawczego, jednocześnie ograniczając możliwość badania relacji matematycznych.

Ujawnianie takiego potencjału kontekstowego wymaga podejmowania nieustannych prób rozpoznawania, jak dzieci uczą się wczesnoszkolnej matematyki. Badania nad uczeniem się przedmiotów ścisłych wykazują, że konieczne jest tworzenie kontekstu poznawczego, który można określić jako dialogiczny. Postrzegam to pojęcie jako metaforę zespołu czynności nauczyciela i uczniów zajmujących się nauczaniem-uczeniem się matematyki. W jego skład zaliczyłabym możliwość uzasadniania zauważanych prawidłowości, spontaniczne używanie takich narzędzi budowania znaczeń matematycznych, jak manipulowanie przedmiotami i wykonywanie rysunków przez uczniów. Badania pokazują, że skuteczne nauczanie musi mieć jako podstawę „dynamiczne poszukiwanie, wyznaczone dostrzeganiem samej zmiany, intelektualną płynnością oraz krytycznym myśleniem uczniów” (Dylak, 2013, s. 131). Kontekst rozwijający pojęcia matematyczne pozwala na rozbudowywanie znaczeń relacyjnych i użytkowych. Mamy już bowiem przekonanie, że myślenie polega na konstruowaniu strategii postępowania, natomiast myślenie jest „zadane”, a nie „dane” (tamże, s. 17).

Przypisy

¹ Ta dyskusja nabrała siły w kontekście najnowszego pomysłu stworzenia jednego podręcznika dla klas pierwszych. Wielka liczba krytycznych głosów pokazuje, jak istotne znaczenie ma dla nauczania najmłodszych uczniów zawartość książek szkolnych.

² Owa „dobrowolność” jest dyskusyjna. Po przeprowadzeniu testu OBUT pojawiają się co roku wpisy nauczycieli na forach internetowych, którzy czują się zmuszani przez dyrektorów do wzięcia udziału w badaniach. Możliwości decydowania w tym zakresie nie mają również rodzice, ani tym bardziej uczniowie.

³ Myślenie o matematyce jako o dziedzinie sformalizowanej, indywidualnej i jednoznacznej ma już o kilku dziesięcioleci przed wszystkim charakter sądu potocznego. Przedstawiciele różnych odłamów konstruktywizmu społecznego postulują włączenie również matematyki w ten nurt myślenia o wiedzy, wskazując jej społeczne konotacje (por. Piotrowska, 2008).

⁴ Badania były prowadzone w latach 2006–2011 w ramach projektu realizowanego przez Centralną Komisję Egzaminacyjną, współfinansowanego przez Europejski Fundusz Społeczny.

⁵ Oba konteksty (monologiczny i dialogiczny) traktuję jako kategorie nadbudowane na definicjach monologu i dialogu w klasie szkolnej sformułowanych przez Dorotę Klus-Stańską (2000).

Literatura

- Bruner J. (1978): *Poza dostarczone informacje*. PWN, Warszawa.
- Dąbrowski M., Żytko M. (red.) (2007): *Badanie umiejętności podstawowych uczniów trzecich klas szkoły podstawowej, cz. I: Raport z badań ilościowych*. CKE, Warszawa.
- Dągiel M., Żytko M. (red.) (2009): *Nauczyciel kształcenia zintegrowanego 2008 – wiele różnych światów?* CKE, Warszawa.
- Dąbrowski M. (2009): *Rozwiązywanie zadań tekstowych*. W: M. Dąbrowski (red.): *Badanie umiejętności podstawowych uczniów klas trzecich szkoły podstawowej. Trzecioklasista i jego nauczyciel*. CKE Warszawa.
- Dąbrowski M. (2011): *Wykonywanie obliczeń, w tym w sytuacjach praktycznych*. W: M. Dąbrowski (red.): *Badanie umiejętności podstawowych uczniów klas trzecich szkoły podstawowej. Trzecioklasista 2010*. CKE Warszawa.
- Dąbrowski M., Wiatrak E. (2012): *Rozwiązywanie zadań tekstowych*. W: A. Pregler, E. Wiatrak (red.): *Raport z badania OBUT 2012*. CKE Warszawa.
- Dahlber G. i in. (2013): *Poza dyskursem jakości w instytucjach wczesnej edukacji i opieki*. Wyd. Naukowe DSW, Wrocław.
- Dziennik Urzędowy (2004), nr 256, poz. 2572, ze zmianami, dostępny na: http://bip.men.gov.pl/men_bip/akty_prawne/rozporzadzenie_20081223_v2.pdf (otwarty 2 lutego 2014).
- Dumont H., Instance D. (2013): *Analiza i tworzenie środowisk uczenia się w XXI wieku*. W: H. Dumont, i in. (red.): *Istota uczenia się. Wykorzystanie wyników badań w praktyce*. Wolters Kluwer bussines, Warszawa.
- Dylak S. (2013): *Architektura wiedzy w szkole*. Difin SA, Warszawa.
- Illeris K. (2006): *Trzy wymiary uczenia się*. Wyd. Naukowe Dolnośląskiej Szkoły Wyższej TWP, Wrocław.
- Klus-Stańska D. (2000): *Konstruowanie wiedzy w szkole*. Wydaw. UWM, Olsztyn.
- Klus-Stańska D. (2013): *Myslenie poza diagnozą i ewaluacją. Pytanie o możliwość zmiany logiki tworzącej edukację. „Teraźniejszość–Człowiek–Edukacja”*, nr 4(64).
- Konarzewski K. (2000): *Jak uprawiać badania oświatowe. Metodologia praktyczna*. WSiP, Warszawa.
- Konarzewski K. (2012): *TIMSS i PIRLS 2011. Osiągnięcia szkolne polskich trzecioklasistów w perspektywie międzynarodowej*. CKE Warszawa.
- Niemierko B. (2013): *Czy diagnostyka edukacyjna i ewaluacja w edukacji zachowują odrębność? „Teraźniejszość–Człowiek–Edukacja”*, nr 4 (64).
- Piotrowska E. (2008): *Społeczny konstruktywizm a matematyka*. Wyd. Naukowe UAM, Poznań.
- Szyling G. (2011): *Nauczycielskie praktyki oceniania poza standardami*. OW „Impuls”, Kraków.
- Szyling G. (2013): *Diagnoza i ewaluacja a teoretyczne horyzonty refleksji nad działaniem. „Teraźniejszość–Człowiek–Edukacja”*, nr 4 (64).
- Wiliam D. (2013): *Rola oceniania kształtującego w skutecznych środowiskach uczenia się*. W: H. Dumont i in. (red.): *Istota uczenia się. Wykorzystanie wyników badań w praktyce*. Wolters Kluwer bussines, Warszawa.

Wyniki badań trzecioklasistów jako diagnoza kontekstów nauczania matematyki w klasach najmłodszych

Raporty z badań dotyczących wiedzy uczniów klas najmłodszych (w Polsce głównie TIMSS i OBUT) od kilku lat stanowią cenny materiał informacji o ich kompetencjach. Rodzaj i treści zadań, a przede wszystkim sposoby ich rozwiązywania są źródłem rozpoznania poziomu uczniowskiej wiedzy matematycznej oraz jej jakości. Analiza rozwiązań dziecięcych może być również źródłem próby stawiania diagnozy, jak przebiegał proces uczenia się, ponieważ intelektualne strategie dziecięce odsłaniają

konteksty uczenia się matematyki w szkole. Badania nad uczeniem się umysłu wskazują na istotność tego, jak jest on konstruowany, ponieważ sposób budowania wiedzy jest predykatorem jej wykorzystania w przyszłości, przede wszystkim w przestrzeni pozaszkolnej. Rozpoznawanie kontekstów tworzenia pojęć matematycznych przez dzieci może stanowić element refleksji nad budowaniem koncepcji nauczania matematyki w klasach najmłodszych.

Słowa kluczowe: wczesnoszkolna wiedza matematyczna, kontekst uczenia się matematyki, zadania matematyczne

Results of research concerning third-formers as a diagnosis of maths-teaching contexts in youngest classes

Reports of research on the youngest pupils' knowledge (in Poland mainly TIMSS and OBUT) have provided valuable information about the pupils' competences. The kind and content of the problems as well as ways of solving them are a source of recognising the level of pupils' knowledge and its quality. The analysis of the solutions can provide material to try and diagnose how the educational process has gone, with the child's intellectual strategies revealing contexts of learning mathematics at school. Research on mind's learning points to the importance of how it is constructed, because the way of building knowledge is a predicator of its use in the future, mainly in the space outside school. Recognising the contexts of creating mathematical concepts by children can cause a reflection on building the concept of teaching mathematics to youngest pupils.

Keywords: Early mathematical knowledge, context, learning math, math

RECENZJE

RUCH PEDAGOGICZNY 2/2014

„TERAŹNIEJSZOŚĆ–CZŁOWIEK–EDUKACJA”, nr 4 (64)/2013

Czwarty numer kwartalnika „Terażniejszość–Człowiek–Edukacja” w całości poświęcony został problematyce szeroko ujmowanej diagnostyki edukacyjnej i składa się z tekstów powstałych i opracowanych w ramach Ogólnopolskiego Seminarium Naukowego „Diagnoza i ewaluacja. Spotkania w szkole”, zorganizowanego w Instytucie Pedagogiki Uniwersytetu Gdańskiego w kwietniu 2013 roku. Powstała dzięki temu możliwość, by prześledzić bieżące problemy diagnostyki edukacyjnej w autorskich interpretacjach badaczy związanych z tą dziedziną.

Pierwszą część: „Studia i rozprawy”, zawierającą siedem tekstów, otwiera artykuł Bolesława Niemierki poświęcony pytaniu o odrębność diagnostyki i ewaluacji: *Czy diagnostyka edukacyjna i ewaluacja w edukacji zachowują odrębność?* Autor dokonuje w nim wprowadzenia w problematykę ewaluacji i diagnostyki edukacyjnej, a zarazem porównuje te subdyscypliny, przyjmując, że jest w nich więcej elementów wspólnych niż różnic. W pierwszej części tekstu przybliży i wyjaśni terminologię, założenia oraz główne kierunki rozwoju ewaluatyki (w dalszej części analiz autor używa też określenia „ewaluacja”), wykazując rodowód tej dziedziny sięgający czasów nauczania programowanego wspieranego behawioryzmem psychologicznym. W części drugiej opisuje znacznie skromniejszy dorobek diagnostyki edukacyjnej, dokonując rozróżnień wynikających z odmiennych założeń skoncentrowanych na procesie wykrywania zaburzeń leżących u podstaw niepowodzeń szkolnych oraz na poszukiwaniu rozwojowych uwarunkowań różnicujących indywidualne osiągnięcia uczniowskie. Te dwa kierunki: diagnostyki zaburzeń i diagnostyki rozwojowej, autor traktuje komplementarnie, co jest istotne w analizach związanych z wyjaśnieniem wzajemnej relacji pomiędzy tymi obszarami. W dalszej części tekstu następuje konkretyzacja sposobów wykorzystywania omawianych kategorii w odniesieniu do procedur zewnętrznego egzaminowania oraz pomiarów EWD. Konkretyzacja ta głębiej wyja-

śnia intencję dobrego odróżnienia diagnostyki zaburzeń od diagnostyki rozwojowej. W końcowej części artykułu autor dokonuje syntetycznego porównania ewaluatyki i diagnostyki, wydobywając wspólne elementy, których opis wyraźnie wskazuje na podobieństwa związane z koncentracją na szeroko rozumianym procesie kształcenia i jego efektach oraz możliwościach indywidualnych dla uczących się.

Kolejny tekst części pierwszej, zatytułowany *Sens spotkań diagnostyki z ewaluacją w szkole*, autorstwa Marii Groenwald, przenosi dyskusję nad ewaluacją i diagnostyką szkolną na inne tory, ujawniając istniejące, ale marginalizowane wątki poświęcone istocie spotkań rozgrywających się pomiędzy osobami zaangażowanymi w omawiane formy aktywności pedagogicznej. Autorka nadaje im status ontologiczny i aksjologiczny zarazem, sięgając do argumentów filozoficznych, wyeksponowanych za Władysławem Stróżewskim, i wskazując tym samym na możliwość prowadzenia analiz w polu semantyki i filozofii. Ujęcie to stanowi punkt wyjścia konstrukcji teoretycznej, w której sens odkrywany jest poprzez koncepcję trafności – przyjętą za Samuelem Messickiem. Powstały „model trafnościowej koncepcji sensu spotkań” ma charakter wielopłaszczyznowy i stanowi konstrukt generujący możliwości analiz w kontekście aksjologicznym, prognostycznym, teoretycznym oraz paradygmatycznym. W odniesieniu do tych dwóch ostatnich można się zastanawiać, czy nie należałoby ich rozpatrywać łącznie, tzn. traktując namysł nad teoretycznymi implikacjami spotkań diagnostyki z ewaluacją jako szansy na wieloparadygmatyczne ujmowanie ich sensu.

Argumentem na rzecz podejścia łączącego problem teoretycznego pola analiz ewaluacji i diagnostyki z konceptualizacją i określeniem ich różnorodności paradygmatycznej może być postawione przez autorkę w tekście pytanie o semantyczny status badanych obszarów, co jest znaczące w przypadku niejednoznaczności badanych kategorii, której źródłem jest zarówno ich interdyscyplinarny charakter, jak i odwoływanie się do zróżnicowanych paradygmatów przez poszczególne badaczy. Uzasadnienie dla takiego

połączenia odnaleźć można na str. 25–26, gdzie autorka pokazuje faktyczne zróżnicowanie metodologiczne badań w obszarze diagnostyki i ewaluatyki, wynikające z odmiennych „subiektywnych ontologii”, a zatem też z odmiennych teoretycznych konceptualizacji przedmiotu badań. Uwaga ta ma charakter w mniejszym stopniu polemiczny, jest raczej próbą dopowiedzenia i wyjaśnienia poruszonego wątku, niemniej doprecyzowanie tej kwestii byłoby znaczącym uzupełnieniem tekstu.

Szczególną wartością tego wywodu jest sproblematyzowanie namysłu nad diagnostyką i ewaluacją edukacyjną za pomocą narzędzi i języka filozofii, wytwarzających refleksyjną metaperspektywę dla badanego obszaru.

Z kolei tekst autorstwa Leszka Korporowicza: *Spoleczna etyka ewaluacji*, zawiera argumentację na rzecz wprowadzenia do debaty pedagogicznej wątku odpowiedzialności społecznej związanej z obszarami instytucjonalnych praktyk społecznych, do jakich zalicza się ewaluacja. Zwracając się w stronę mechanizmów społecznych, będących efektem wprowadzania „idei społeczeństwa wiedzy”, autor opisuje swoisty redukcjonizm zarówno w sferze świadomości podmiotowej jak i na poziomie działań instytucjonalnych, który tym procesom towarzyszy. Mamy do czynienia z zamianą mądrości i refleksyjności na wiedzę, a tej na informację. Dokonuje się też zmiana w procedurach komunikowania społecznego, podlegającego regułom wynikającym z dążenia do efektywności i pragmatycznego dążenia nie tyle do prawdy, co raczej do tworzenia skutecznych procedur działania i wytwarzania skutecznego wizerunku. Zamiast społecznie kontrolowanego procesu wylania się obraz zdehumanizowanych, biurokratycznych i pozbawionych podmiotowej odpowiedzialności procedur, kryjących się pod stechniczonym żargonem. Autor przedstawia schemat opisujący, jak dochodzi do przewartościowania i zamiany mądrości, wiedzy i prawdy na informację, wizerunek i skuteczność, na utylitarne kategorie, które wtargnęły także do obszaru refleksji społecznej.

Dla instytucji administrujących życie gospodarcze oraz warunki rozwoju makroekonomicznego biznesu wymienione kategorie stały się normą w sposób naturalny, zgodny z obecnym porządkiem opartym na logice ekonomii rynku. Odnosząc się do sfery edukacji, logika ta spotyka się nie tyle z oporem, co wręcz z „odporem” (określenie J. Rutkowiak, 2010, s. 212 i kolejne). Wynika to przede wszystkim z samoświadomości pedagogów zaangażowanych w pracę

pedagogiczną – zarówno szkolną jak i pozainstytucjonalną. W tym sensie propozycja nadania ewaluacji walorów wynikających z poczucia odpowiedzialności etycznej w biznesie potwierdza działania podejmowane w sferze edukacji przez krytycznie zaangażowanych pedagogów.

Wątki znaczące dla całości problematyki wnosi tekst *Pedagogika ewaluacji* autorstwa Sylwii Jaskuły. Pełni on rolę wyjaśniającą i porządkującą kwestie terminologiczne, a także wzmacnia pytania o społeczny wymiar działań ewaluacyjnych w obszarze edukacji instytucjonalnej. Autorka pisze o konieczności swoistej „resocjalizacji ewaluacji” i przywrócenia jej zapoznanego sensu wyrażającego się w respektowaniu interesu społecznego, nadrzędnego w stosunku do wymogów administracyjnych i kontrolnych. W dalszej części w syntetyczny i przejrzysty sposób ukazana została ewolucja rozumienia ewaluacji na poziomie jej celów, przedmiotu zainteresowań i roli ewaluatora.

Ewaluacja rozumiana jako „zaproszenie do rozwoju”, „negocjacja” oraz „wspieranie wartości uspołecznionego działania” nabiera szczególnego znaczenia w kontekście zamieszczonego w dalszej części numeru tekstu Doroty Klus-Stańskiej *Myślenie poza diagnozą i ewaluacją*. Zaprezentowana koncepcja ewaluacji jest dla niego swoistym kontrapunktem, dającym Czytelnikowi możliwość zgłębienia istotnych pytań o cele kształcenia i jego wymiary oraz o sens i sposoby rozumienia w tym kontekście kategorii rozwoju.

Szczególną wartość praktyczną zawiera druga część analiz. Autorka wykazuje w nich, na czym polega proceduralny sens ewaluacji, zwłaszcza w odniesieniu do koniecznych etapów tego procesu. Nie można bowiem uznać, iż stanowi o tym raport. Jak podkreśla badaczka, najważniejszy moment wiąże się z upowszechnieniem i praktycznym wykorzystaniem uzyskanych wyników. W tym sensie ewaluacja może włączyć się w nurt praktyki społecznej i związanych z nią doświadczeń oraz stać się elementem wiążącym, a nie atomizującym grupy społeczne skupione wokół procesów edukacyjnych. Jeśli dodamy do tego wyraźne odcięcie się autorki od nadania ewaluacji redukcjonistycznych znaczeń związanych z procedurą akredytacji i biurokratycznych obciążeń, to otrzymujemy znakomity przyczynek do myślenia o ewaluacji w kategoriach humanistyki.

Kolejny tekst: *Diagnoza i ewaluacja a teoretyczne horyzonty refleksji nad działaniem*, napisany przez Grażynę Szyling, wzmacnia obecny w tomie humanistyczny nurt myślenia o ewalu-

acji i zawiera ważny emancypacyjny postulat, zawieszający paradygmat funkcjonalistyczny na rzecz postawy interpretacyjnej. Wyraża się ona w tworzeniu horyzontu samowiedzy podmiotów zaangażowanych w procesualnie rozumianą ewaluację oraz w akceptacji dla włączenia w jej obręb procesów wartościowania. Społeczne zaangażowanie badaczy w trakcie ewaluacji jest faktem, zaś postulat ideologicznej neutralności i całkowitej ateoretyczności nie znajduje uzasadnienia, a nawet jest wewnętrznie sprzeczny. Autorka zauważa, że ewaluacja jest procesem, który w założeniu prowadzić ma do praktycznych zastosowań. Krytyczna analiza przywołanych w dalszej części tekstu przykładów wykazuje, jak trudno jest ustalić miarodajną perspektywę oglądu własnych działań przez osoby zaangażowane w proces edukacji. Autorce nie chodzi jednak o „obiektywizację wyników”, lecz o szeroką płaszczyznę interpretacyjną, w której możliwe stanie się pełne rozpoznanie przez nauczycieli sensu własnych doświadczeń ze wszystkimi składnikami, nadającymi znaczenia niepowtarzalne, subiektywne i każdorazowe. Jednak – wykazuje Grażyna Szyling – takie rozumienie ewaluacji staje się utopią w warunkach zideologizowania i zawłaszczenia jej celów do doraźnych partykularnych potrzeb zewnętrznych instytucji i agend państwa, wymuszających na samodoskonających się nauczycielach postawy schizoidalne i adaptacyjne. Szukanie własnej drogi w tym procesie oznacza nie tylko ustanowienie osobistego horyzontu interpretacyjnego, ale także konieczność porzucenia zarówno perspektywy „z bliska”, jak i tej odległej „spoza” sytuacji, w której znalazł się nauczyciel. Wymaga to jednak spełnienia trudnych warunków, do których z pewnością należy rozwijanie krytycznej samoświadomości odnoszonej do nauczycielskich doświadczeń, które nabierają cech „zdarzeń krytycznych”. Zdaniem Grażyny Szyling stanowią one punkt wyjścia autointerpretacyjnych odczytań i przekształceń własnej praktyki, wykraczających poza dotychczasowy paradygmat nauczycielskiego myślenia i działania. Tak rozumiana autodiagnoza i autoewaluacja jest jednocześnie szansą na wytworzenie alternatywnej perspektywy interpretacyjnej.

Następny tekst, zatytułowany: *Myślenie poza diagnozą i ewaluacją. Pytanie o możliwość zmiany logiki tworzącej edukację*, autorstwa Doroty Klus-Stańskiej wykracza swym zakresem poza klasycznie ujmowaną problematykę ewaluacji i diagnostyki. W artykule w celu uniknięcia nieporozumień terminologicznych i ujednocznienia przyjmowanych założeń przyjęto,

iż paradygmat to „zbiór pojęć i teorii, mających wspólny fundament ontologiczno-epistemologiczny i oparty na podzielanym przez wspólnotę uczonych języku, w którym określonym terminom nadaje się swoiste dla danego paradygmatu znaczenia” (s. 72). W tekście postawiono pytania przewartościowujące i poszerzające sens omawianych dyscyplin. Pytania te jednocześnie generują nowe wątki, w których chodzi o kształt postulowanej i uprawianej praktyki edukacyjnej w kontekście zmiany założeń epistemologicznych i ontologicznych, a w ślad za tym też przesłanek metodologicznych leżących u podstaw przyjętej postawy badawczej. W tak wytworzonej „ekumenicznej perspektywie” oglądu badaczka proponuje swoisty dialog międzyparadygmatyczny, dający szansę na wzbogacenie i rozszerzenie pola analiz badawczych. Kluczowym wątkiem tekstu jest krytyczny namysł nad sposobem uprawiania diagnostyki i ewaluacji edukacyjnej opartej na założeniach obiektywistycznych, wyprowadzonych z pozytywistycznej koncepcji nauk społecznych. Badacze tego nurtu wychodzą z założenia „że realnie istnieje taki byt, jak rozwój jednostki oraz, związana z zapewnianiem tego rozwoju, jakość warunków, w których ów postęp zachodzi” (s. 78). Jednak zamysł autorski jest zarysowany znacznie szerzej – chodzi bowiem nie tyle o wizję międzyparadygmatycznego uprawiania badań edukacyjnych, co o radykalną zmianę znaczeń związanych z używaną terminologią, podstawowymi pojęciami i kategoriami oraz wynikającą z tego zmianą rozumienia celów praktyki edukacyjnej leżących u podłoża przyjętych przesłanek teoretycznych, związanych w szczególności z nadmiernie speyfikowanym pojęciem rozwoju. Z przytoczonych przez autorkę krytycznych argumentów (s. 79–82) wynika, że nieproblematyczność rozumienia tej kluczowej kategorii jest przyczyną ujmowania edukacji w perspektywie włączającej rozwój w procedury wymuszania „znormalizowanego wzrostu” (s. 87–88). Wzrost ów jest jednak dwuznaczny, gdyż progresja występująca w danym obszarze generuje jednocześnie deficyty i ograniczenia w innych sferach. Rozwój ujmowany w kategoriach progresji jakościowej bywa słabo związany z działaniami szkoły, a nawet toczy się wbrew nim, będąc jednocześnie strukturą niekiedy wręcz ambiwalentną, czego przykłady pojawiły się w dalszej partii artykułu. Chodzi o odwołania do społecznego kontekstu współtworzącego zmiany definiowane z jednej strony jako „zdobycz rozwojowa”, a z drugiej jako regres (np. kult pracowitości *versus* zdolność do swobodnej aktywności i zabawy). Autorka

pisze: „Te symptomy określam na razie wstępnie jako anomalie” (s.79) – może to oznaczać nie tyle postulowany, co rzeczywisty przełom paradygmata, świadczący o tym, że edukacja instytucjonalna może być rzeczywistym sprawcą utrudniającego jego zaistnienie.

Waga tych stwierdzeń jest nie do przecenienia, gdyż badaczka dostarcza argumentów pochodzących z niesystematycznych wprowdzeń, ale dobrze udokumentowanych spostrzeżeń badawczych z obszaru psychologii, antropologii, filozofii i geografii społecznej, tworząc interdyscyplinarne uzasadnienie dla przewidywanego przełomu paradygmata. Elementami tak opisywanego przełomu są: zmiana w rozumieniu funkcji rozwoju wraz z jego ambiwalentną specyfiką oraz wynikająca z tego redefinicja samej kategorii edukacji. W tym sensie otrzymujemy rzeczywiste wsparcie dla działań podejmowanych w sferze praktyki, gdzie dominuje odejście od dyskursu opartego na technicznej i adaptacyjnej racjonalności na rzecz podmiotowej partycypacji i gdzie „rozwój” nie stanowi już narzędzia przystosowywania i mierzenia znormalizowanego postępu.

Na tym etapie tekst znakomicie inspiruje do dyskusji nad zmianą paradygmata w edukacji i otwiera pytanie o rozwój i warunkującą go jakość: Czym właściwie są i czy nadal można te wartości traktować autotelicznie, czy raczej wymagają kontekstualizacji historycznej, lokalnej i kulturowej w miarę określania zadań edukacyjnych ważnych społecznie w danym czasie i miejscu? Szczególnie inspirujące jest poszerzenie przez Dorotę Klus-Stańską pola analiz rozumienia kluczowych kategorii pedagogiki oraz usytuowania „rozwój” w kontekście pytań o cele zabiegów edukacyjnych (niekoniecznych instytucjonalnych, czy nawet raczej pozainstytucjonalnych). Z podjętej dyskusji nad rozwojem jako celem wychowania, wyłania się potrzeba poszerzenia/redefinicji omawianej kategorii (por. Kohlberg, Mayer, 1993, s. 59). W kontekście rozważań autorki dyskusyjny staje się pogląd, iż rozwój musi stanowić cel zabiegów edukacyjnych, gdyż wówczas normalizacja może wydawać się nieuniknioną tego konsekwencją. Niezależnie od tego, można traktować tekst jako głos w dyskusji nad zmianą paradygmata rozumienia i opisu istoty normalizujących praktyk związanych z rozwojem.

Tekst kolejny, *Trafność użytkowania wyników egzaminu* Piotra Skorupińskiego, stanowi kłamerę zamykającą całą część i w pewnym sensie dopowiada niektóre wątki pojawiające się w pierwszym artykule. Akcent położony został na traf-

ność użytkowania uzyskanych wyników i – obok ich interpretacji – określone zostały istotne elementy w całym procesie dalszego użytkowania danych uzyskanych podczas egzaminów. W pierwszej części tekstu zaprezentowano pogląd, zgodnie z którym należy szerzej ujmować problematykę trafności, nie ograniczając się do trafności pomiaru testowego, lecz rozszerzając sens tej procedury do dalszych działań podejmowanych przez zainteresowane podmioty (dyrektorzy szkół, kuratorzy, nauczyciele rodzice czy wreszcie sami uczniowie). Wykorzystując model argumentacji unitarnej, autor uzasadnia konieczność włączenia wyjaśnienia sposobu użytkowania testu w całość procedury jego konstruowania. Tym samym o trafności danego testu przesądza nie tylko precyzja pomiaru, ale również „trafność użytkowania” uzyskanych wyników. Przytoczone w dalszej części tekstu na przykładach czterech liceów formy użytkowania wyników maturalnych przybierają zróżnicowaną postać, zależnie od decyzji dyrekcji i grona pedagogicznego, w mniejszym stopniu odpowiadając na oczekiwania samych uczniów i rodziców. Autor wykazuje także brak precyzyjnych uregulowań prawnych (analiza rozporządzeń MEN świadczy o niejednoznaczności w tym zakresie) oraz brak standardów umożliwiających lepsze ich użytkowanie. W tym sensie tekst wpisuje się w oczekiwania tych podmiotów (nauczycieli, rodziców i uczniów, administracji szkolnej), dla których wysoki poziom trafności użytkowania wyników egzaminów stanowić może pomoc w dalszych decyzjach w sferze praktyki edukacyjnej.

W części drugiej numeru „Komunikaty z badań”, znajdują się dwa teksty opisujące wyniki badań OBUT (Ogólnopolskie Badania Umiejętności Trzecioklasistów) oraz badania 15-latków. Pierwszy z nich, Małgorzaty Żytko, zatytułowany *Ogólnopolskie Badanie Umiejętności Trzecioklasistów (OBUT) pedagogiczny eksperyment z „politycznymi konsekwencjami”*, jest rodzajem sprawozdania z kontynuacji wcześniejszych, szeroko zakrojonych badań realizowanych w latach 2006–2011, obejmujących uczniów kończących pierwszy etap edukacji oraz ich nauczycieli. Referowane w tekście badania OBUT są dedykowane w szczególności nauczycielom i dyrektorom szkół jako znaczące narzędzie poznawania uczniowskich umiejętności językowych i matematycznych, umożliwiające wnikliwą ich diagnozę, a także jako narzędzie doskonalące pracę szkoły. Badania te zrealizowano w latach 2011–2012, obejmując blisko 80% szkół. Wykorzystane zadania weryfikowane już w poprzednim etapie,

charakteryzowały się zróżnicowanym stopniem trudności. Wprowadzono w nich także podział na zadania typowe i nietypowe, te ostatnie – nieznajdujące swego odzwierciedlenia w codziennej pracy dydaktycznej, niemniej związane ze sformułowanymi w podstawie programowej celami edukacyjnymi odnoszącymi się do wiedzy i myślenia. Raporty z badań przekazywane szkołom pozwalały zainteresowanym pedagogom na poznanie poziomu wykonania zadań, identyfikację najczęściej popełnianych błędów, strategii podejmowanych przez uczniów oraz na sformułowanie na tej podstawie rekomendacji do pracy dydaktycznej.

W dalszej części tekstu, w wątku poświęconym ewaluacji, autorka zauważa, iż od dyrektorów szkół uzyskano istotne potwierdzenie dla wartości diagnostycznej badania OBUT, a także – co jest szczególnie znaczące – uznali oni, że OBUT pozwoli na lepsze planowanie pracy dydaktycznej na drugim etapie szkolnym oraz zaplanowanie działań wyrównawczych. Ewaluacja ujawniła też często niedemokratyczny sposób podejmowania decyzji w stosunku do nauczycieli (z ankiety wynika, iż 60% nauczycieli nie miało wpływu na przystąpienie do badań). OBUT można postrzegać też jako szansę bardziej refleksyjnego podejścia do własnej praktyki, weryfikacji stylu pracy, podejmowania bardziej świadomych decyzji co do wyboru materiałów dydaktycznych oraz możliwości indywidualizowania pracy z uczniem.

Spostrzeżenia Małgorzaty Żytko pośrednio dotyczą także istniejącego systemu edukacji, z jego obciążeniami biurokratycznymi i przeżytkiem form dydaktycznych. Ta ocena jest tym bardziej istotna, gdyż pokazuje faktyczne przyczyny ujawniające się „poza nauczycielami” – wytworzone przez administrację oświatową i chaotyczną, a właściwie nieobecna politykę oświatową. Widać to wyraźnie w postrzeganiu przez badanych OBUT jako „opresyjnego” w tym sensie, że ujawnia on rozbieżność między rzeczywistym i wykazanim w badaniach potencjałem dzieci a wymuszonym i narzuconym przez podstawę programową stylem pracy dydaktycznej. Opisany przez Małgorzatę Żytko konflikt, pomiędzy wymogami instytucjonalnymi a faktycznymi potrzebami i możliwościami rozwojowymi uczniów, jest głęboką i popartą badaniami diagnozą przyczyn utraty przez szkołę znacznej części dziecięcego potencjału badawczego. Systemowo pozbawieni profesjonalnej autonomii nauczyciele starają się w tej sytuacji pogodzić swe intuicje pedagogiczne wyzwolone przez OBUT z oczekiwaniami wyrażonymi

w podstawie programowej (s. 114–115). W końcowej partii tekstu autorka stwierdza, iż wadą OBUT było utożsamianie go ze sprawdzianem, przez co generować mógł postawy obronne wśród uczniów i samych nauczycieli. Potencjał emancypacyjny OBUT nie mógł być na tym etapie w pełni wykorzystany z powodu systemowej niewydolności przejawiającej się w społecznej fetyszyzacji testów oraz poznawania wartości ewaluacyjnej niektórych z tych narzędzi, tak jak ujawniono to w badaniach OBUT.

Drugi tekst tej części, *Obraz polskiej szkoły w badaniach 15-latków*, autorstwa Barbary Ciżkowicz, stanowi przekrojową interpretację wyników osiągniętych przez polskich uczniów w kolejnych edycjach badań PISA i spojrzenie przez ich pryzmat na polską szkołę. Obok zmian uzyskanych wyników z lat 2000 i 2003 nastąpiła poprawa w testach czytania ze zrozumieniem i w naukach przyrodniczych (2006). Autorka analizowała uzyskane w badaniach opinie uczniów odnoszące się do szkoły: ocenie jej przydatności w przyszłej pracy, przygotowania do dorosłości, ułatwienia podejmowania decyzji. Wynika z nich, że opinie polskich nastolatków są bardziej pesymistyczne w stosunku do uśrednionych opinii uczniów krajów OECD. Negatywna ocena możliwości szkoły w zakresie przygotowania do dorosłości oraz wartości spędzanego tam czasu wyrażona przez polską młodzież wyraźnie kontrastuje z postawami fińskich nastolatków. Lepsze wyniki uzyskano w badaniu relacji z nauczycielami, gdyż pozytywnie oceniana je ponad 80% polskich uczniów, a także – co najciekawsze – oceny te są wyższe niż wyrażone przez uczniów fińskich. Jednak u tych ostatnich występuje związek opinii z wynikami uzyskanymi w testach, czego nie stwierdzono u uczniów polskich. Oznaczałoby to społeczną alienację i częściowy zanik dydaktycznej funkcji szkoły, przy zachowaniu społecznych więzi z nauczycielami. Potwierdzeniem tego może być spostrzeżenie na temat korzystania z kształcenia wyrównawczego, gdyż dotyczy to w Polsce wyższego odsetka 15-latków w porównaniu ze średnią. Obserwowany jest stały proces uzupełniania wiedzy poza szkołą, co świadczy o niespełnieniu przez nią oczekiwanych funkcji. Porównując z fińskimi szkołami zauważyć można niski poziom tradycyjnie rozumianej dyscypliny, co przy dobrych osiągnięciach uzyskanych w tej grupie uczniów oznaczałoby, iż wyniki te nie zależą od środków dyscyplinujących, lecz od umiejętnego wykorzystania społecznych walorów uczenia w grupie i elastycznego stylu pracy dydaktycznej. Wprawdzie Barbara Ciżkowicz

przycacza wyniki mówiące o tym, że 15-latkowie fińscy niżej od polskich oceniają zaangażowanie swych nauczycieli w pracę z tekstem i mobilizację ich do niej, jednak wyższa ocena polskich uczniów w tym zakresie może wynikać z nadmiernego przywiązywania wagi do uzyskiwania wyników w testach, kosztem innych sfer pracy szkolnej. Autorka zauważa: „ilość czasu «na przemyślenie odpowiedzi» okazała się prognostyczna w każdej z rozpatrywanych grup” (s. 127). Ostatecznie niepokoić musi ocena polskiej szkoły jako „miejsca straty czasu” i takiego, gdzie uzyskiwane wyniki są co najwyżej średnie przy wysokim zaangażowaniu w kształcenie dodatkowe.

Zaprezentowany obraz uzupełniony został postulatami wprowadzenia zmian systemowych, jednak wydaje się, że wobec braku spójnej polityki oświatowej może być to trudne. Obraz polskiego systemu edukacji, który wyłania się „spoza” przywołanych badań nie napawa optymizmem, mimo poprawy wyników uzyskiwanych przez nastolatki.

Część trzecia recenzowanego numeru kwartalnika „Opinie” zawiera tekst Ewy Wysockiej: *Standardy etyczne i profesjonalne w pracy pedagoga-diagnosty w kontekście obszarów diagnozy pedagogicznej*. Autorka formułuje w nim rozwiązania kodyfikujące pracę zawodową pedagoga-diagnosty. Rozwiązania te zostały zaczerpnięte z obszaru psychologii, zaś przyjęta koncepcja diagnostyki, jej celów i zadań – koncentruje się na obszarach związanych ze stwierdzeniem właściwości rozwojowych, ocenianiem deficytów i zagrożeń rozwojowych oraz – w konsekwencji – na określeniu standardów niezbędnych pedagogom w pracy diagnostycznej. Standardy te sformułowane zostały w nawiązaniu do założeń aksjologicznych oraz norm etycznych (na ten temat informację zawierają przypisy na s. 136–137). Dalsza część tekstu poświęcona jest omówieniu szczegółowego katalogu zasad pracy pedagoga-diagnosty, co otwiera pytania o możliwości i sposoby ich praktycznej konkretyzacji w codziennej pracy diagnostycznej. Zasady i standardy postępowania w dziedzinie diagnostyki powinny stanowić – zdaniem autorki – ułatwienie jako normy podejmowanych działań diagnostycznych oraz jako ograniczenia związane z przyjętymi wartościami, regulującymi zakres postępowania w sferze profesjonalnej, interwencyjnej i prakseologicznej.

Ostatnie części omawianego numeru to „Kronika”, w której Anna Kossobucka opisuje przebieg seminarium, będącego źródłem zamieszczonych w czasopiśmie tekstów, oraz

sporządzona przez Wojciecha Siegienia recenzja książki *Sprawcy i/lub ofiary działań pozornych w edukacji szkolnej*, pod red. Marii Dudzikowej i Kariny Knasieckiej-Falbińskiej.

Całość numeru ma charakter tematyczny, prezentujący autorskie koncepcje zagadnień dotyczących diagnostyki i ewaluacji. Szczególną wartością zbioru jest przedstawienie zróżnicowanych linii myślenia o edukacji, szerokie osadzenie rozważanych zagadnień w kategoriach humanistyki, a także głosów odrębnych. Włączenie w zbiór raportów badawczych dodaje całości aktualnego, społecznego kontekstu.

Literatura

Kohlberg L., Mayer R. (1993): *Rozwój jako cel wychowania*. W: Z. Kwieciński, L. Witkowski (red.): *Spory o edukację*. IBE, Warszawa.

Potulicka E., Rutkowiak J. (2010): *Neoliberalne uwiłki edukacji*. OW „Impuls”, Kraków.

Jolanta Kruk

Elbląska Uczelnia Humanistyczno-Ekonomiczna

POTENCJAŁ POZNAWCZY WYBRANYCH MONOGRAFII Z KRĘGU DIAGNOSTYKI EDUKACYJNEJ. PERSPEKTYWA JĘZYKOZNAWCZA

Można założyć, iż obraz subdyscypliny naukowej kształtowany jest między innymi przez zakres pojęć, którymi się ona posługuje, a także przez pole odniesień do innych dyscyplin naukowych. W dziedzinach wpisanych w badania o nachyleniu praktycznym ważny może być też zakres pojęć wykorzystywanych przez praktyków, którymi w diagnostyce edukacyjnej są między innymi nauczyciele.

Niniejszy tekst wyrasta z założenia o znaczącej roli spójnego systemu pojęć, który jest warunkiem zrozumienia sensu każdej dyscypliny naukowej, a tym bardziej takiej, która znajduje się w fazie kształtowania swej odrębności. Zmierzam w nim do wyodrębnienia zbioru odniesień oraz kategorii językowych akcentowanych przez wybrane publikacje książkowe, wydane w ostatnich latach i powiązane tematycznie z diagnostyką edukacyjną, konstytuującą się przez ostatnie dwudziestolecie między innymi dzięki cyklicznym konferencjom jej poświęconym. Postawię pytanie o praktyczne aplikacje omawianej dziedziny, uznając, że uzasadnieniem jej istnienia jest możliwość praktycznych zastosowań.

Powyższe założenia określają także strukturę niniejszego tekstu, na którą składają się:

- 1) refleksja nad dyscyplinami naukowymi wyprzedzająca refleksję nad językiem, ponieważ terminologia dziedziny jest często efektem daleko idących zapożyczeń z innych dziedzin;
- 2) identyfikacja zakresu dyscyplin leżących u podstaw diagnostyki edukacyjnej oraz zbioru konkluzji wyprowadzonych przez autorów analizowanych publikacji, co współtworzy pole teoretycznych i praktycznych priorytetów;
- 3) założenie, że o dalszym rozwoju diagnostyki edukacyjnej zdecyduje między innymi to, jakimi kategoriami językowymi będzie się ona systematycznie posługiwać.

Zaplecze teoretyczne, czyli dyscypliny, z których czerpią analizowane teksty

Celem diagnostyki edukacyjnej jest „rozpoznanie przebiegu, wyników i uwarunkowań uczenia się” (Niemierko, 2002, s. 72), jednak – stwierdza Bolesław Niemierko – „jej paradygmat naukowy, tj. zbiór wspólnie uznawanych założeń, pojęć i metod, nie jest jeszcze ustalony” (tenże, 2009). To sprawia, że w podjętych analizach poszukiwać będą punktów wspólnych, łączących pozycje z gruntu różne, czyli kluczowy dla dziedziny podręcznik B. Niemierki (2009), a także monografie poświęcone *etycznym aspektom egzaminów szkolnych* (Groenwald, 2011), *nauczycielskim praktykom oceniania* (Szyling, 2011) oraz *efektowi rówieśników w nauczaniu szkolnym* (Pokropek, 2013).

Przywołane publikacje monograficzne, prezentujące wyniki badań nad edukacją, są efektem teoretycznych i empirycznych eksploracji, a każda z nich stara się uchwycić określony moment rozwoju polskiej szkoły i – sięgając do różnych dziedzin wiedzy – pozwala go lepiej zrozumieć. Nie oznacza to jednak, że mamy do czynienia z kanonem, zresztą autorom nie towarzyszy taki cel, a i sam dobór omawianych w niniejszym tekście książek jest wyznaczony tylko ich tematyczną bliskością z diagnostyką edukacyjną oraz zbliżonym czasem publikacji.

Odniesienia odnalezione w pracy G. Szyling (2011), poświęconej nauczycielskim praktykom oceniania poza standardami, dotyczą dydaktycznych, psychologicznych, społecznych i aksjologicznych aspektów szkolnego oceniania. Te rozliczne osadzenia teoretyczne można uznać za swoiste „przedpole” oceniania, bez którego znajomości wręcz nie da się dobrze oceniać. Równie dobrze można jednak zarzucić autorce niepotrzebne wydłużanie pracy i stwarzanie wrażenia, że ocenianie osiągnięć uczniów wiąże się praktycznie ze wszystkimi problemami edukacji szkolnej, np. poziomami rozwoju

jednostki (tamże, s. 27), zastrzeżeniami wobec ujęć taksonomicznych (tamże, s. 59), zmianami w systemie oświaty (tamże, s. 142). Wbrew różnorodności poruszanych w pracy zagadnień, utrzymana jest jednak ciągłość narracji, co można uznać za dowód uwikłania, interdyscyplinarności i złożoności samego pojęcia oceniania, jak i postrzeganej przez jego pryzmat diagnostyki edukacyjnej.

Nie mniejsze uwikłanie teoretyczne dostrzegam w pracy Marii Groenwald (2011), poświęconej etycznemu wymiarowi egzaminów szkolnych. Dążenie do jego rozpoznania znacząco poszerza wachlarz odniesień diagnostyki edukacyjnej poprzez włączenie w jej obszar perspektywy oglądu zjawisk edukacyjnych przez pryzmat koncepcji etycznych. Wpisana w egzaminy teoretyczna złożoność utrudnia jednak orientację, który z poruszanych w książce aspektów uznać za bardziej znaczący: czy przewijające się przez całą pracę obarczenie egzaminów odpowiedzialnością za „zawieszanie rozwoju wielostronnego uczniów” (tamże, s. 104), czy może na przykład rozbieżność pomiędzy deklaracjami nauczycieli a ich postępowaniem (tamże, s. 119). Pozostaje jedynie uznać, iż są to dwie kwestie wobec siebie komplementarne.

Odmiennej perspektywę metodologiczną odnajdujemy w książce Artura Pokropka (2013), który podjął próbę syntetycznego ujęcia determinantów osiągnięć szkolnych uczniów. Między innymi zestawił wyniki różnych badań dotyczących związku osiągnięć szkolnych ze statusem społeczno-ekonomicznym rodziców uczniów (tamże, s. 20–21), efektywnością pracy szkoły (tamże, s. 24–26), genetycznymi uwarunkowaniami zdolności poznawczych (tamże, s. 42) i osiągnięć szkolnych (tamże, s. 44–45) oraz tytułowego „efektu rówieśników” (tamże, s. 52–53). Nawet pobieżny przegląd tych badań wskazuje na potrzebę szeroko zakrojonej diagnozy efektu rówieśników, co w pełni uzasadnia zrównoważone traktowanie przez autora pracy czynników o charakterze socjologicznym, psychologicznym i pedagogicznym.

Zaplecze terminologiczne, czyli wokół pojęć pojawiających się w analizowanych tekstach

Przystępując do próby wyodrębnienia pojęć szczególnie znaczących dla diagnostyki edukacyjnej w jej obecnym kształcie, trzeba zaznaczyć, iż jest to zadanie szczególnie trudne. Już bowiem sama jej teoretyczna wielowarstwowość narzuca szerokie ramy pojęciowe. Przyczynia się do tego między innymi nadanie przez B. Niemierkę

(2009, s. 33) miana „paradygmatu diagnostyki edukacyjnej” zarówno formalnym, jak i nieformalnym metodom diagnostycznym (tamże, s. 33), co z językoznawczego punktu widzenia można uznać za swoistą zachętę do rozluźniania „gorsetu scjentystycznego rygoryzmu” (tamże) także w zakresie pojęć. Dokonując bowiem niesystematyzowanych diagnoz nieformalnych, przedkładamy opis nad eksperymentowanie, a więc – w przypadku języka – bardziej interesuje nas treść aniżeli forma.

Do podobnej językowej swobody wydaje się również skłaniać diagnostów systematycznie przywoływana przez B. Niemierkę kategoria pojęciowa: „patrzenie w klasie szkolnej”¹, uprawniająca do wyciągania wniosków na bazie codziennych nauczycielskich obserwacji. Intencja widzenia tego, jak przebiega uczenie się uczniów (por. Niemierko, 2009, s. 31) musi nas skłaniać do otwartości terminologicznej, opisowego podejścia do badań jakościowych i powstrzymywania się od oczekiwań, iż nasze obserwacje powtórzą się, czyli do uchylania się od językowo sztywnych uogólnień.

Niemniej możemy podjąć próbę systematyzowania pojęć, którymi uściślamy sam język diagnostyki bez ograniczania przedmiotu naszych spostrzeżeń. Przez pryzmat tego kryterium, traktowanego jak swoisty „filtr”, dokonam oglądu analizowanych książek. Przyjmuję też, że „nowy” język diagnostyki edukacyjnej, zajmującej się zagadnieniami wcześniej nieobecnymi w obszarze jej refleksji, ma charakter opozycyjny w stosunku do ujęć silnie zakorzenionych w tradycji językowo rygorystycznego pomiaru dydaktycznego, który uznają za źródło pojęć na dobre „zasiedziały” w tej dziedzinie. Przywołam zatem niektóre wybrane pojęcia, które uznają za „przybyte” w językowy obszar diagnostyki.

Przegląd terminów wyodrębnionych przez B. Niemierkę (2009) w podrozdziałach jego podręcznika zawierających wykazy najważniejszych pojęć, nasuwa spostrzeżenie, iż autor podejmuje próbę nadawania statusu pojęć określeniom, których chyba większość z nas nie „podejrzewałaby” o jednoznaczny charakter. I tak, np.: *wystarczalność informacji* ma być właściwością stanowiącą jeden z efektów starań nauczyciela realizującego diagnozę nieformalną (tamże, s. 183). Pojęcie to pojawia się równoległe z innymi, nieco ściślej określonymi właściwościami, takimi jak np.: *obiektywizm konwencjonalny* sprowadzający się do zgodności diagnozy z systemem kształcenia (tamże); *rozmowa diagnostyczna* (tamże, s. 184), *normy moralne diagnostyki edukacyjnej* (tamże, s. 47), *systemy odpowiedzialności*

edukacyjnej (s. 266), jak i *instytucja samoucząca się* (s. 288). Każdy z tych terminów wydaje się na tyle pojemny, że mógłby stać się podstawą szerokiej akademickiej dyskusji.

Ta pojemność, czy też otwartość interpretacyjna pojęć diagnostyki edukacyjnej zaznacza się również w analizowanych przeze mnie pracach, które z uwagi na swój metodologiczny charakter mogą rodzić w Czytelniku oczekiwanie znaczeniowo bardziej uściślonej terminologii. Na przykład w pracy G. Szyling (2011), poświęconej nauczycielskim praktykom oceniania (czyli pojęciu również o niedookreślonych granicach), odnotowujemy wyraźnie bardziej jednoznaczna terminologię tam, gdzie autorka sięga do teorii pomiaru dydaktycznego, w tych zaś fragmentach pracy, w których szuka ona uzasadnień tych praktyk, pojawiają się terminy, których konieczność wprowadzenia wynika z uwikłania oceniania w wielopłaszczyznowy kontekst. W obszarze tym stosowane są zatem „miękkie” pojęcia, typu: *mikrokontekst osiągnięć szkolnych* (tamże, s. 80), *sądowy kontekst oceniania* (tamże, s. 106), *niepowodzenia ukryte/jawne* (tamże, s. 136), *okazjonalność obniżania wymagań* (tamże, s. 334), *nakład pracy ucznia* jako składnik przedmiotu oceny (tamże, s. 250), czy też „jawność uzasadnienia oceny” (tamże, s. 227).

Posługiwanie się przez autorów analizowanych tekstów wieloznacznymi kategoriami językowymi odbieram jako utrudnienie w zorientowaniu się, w którym „miejscu” dyskusja prowadzona wokół zagadnień oświatowych zaczyna bądź przestaje wpisywać się w zakres diagnostyki edukacyjnej. Rozpatrując pod tym kątem pracę M. Groenwald (2011), stwierdzam, iż dzieje się tak wówczas, gdy ucieka się ona do języka metafor, mówiąc o: *patrzeniu przez zwierciadła etyki* (tamże, s. 69–73), *zawieszeniu pomiędzy systemami normatywnymi* (s. 95), *świecie egzaminów* (s. 43). W takiej sytuacji próby dyscyplinowania i systematyzowania terminologii – a za taką uznają zabieg wprowadzenia do dyskusji nad egzaminami kategorii *normatywnej trójjedni* (s. 93) – uważam za mocno utrudnione. Proponowane w tym kontekście kategorie jakościowe, które dotyczą egzaminu (np. pozornosc, bezstronność) czy nauczyciela-egzaminatora (np. godność, zaufanie, sprawiedliwość), zdają się mieć tak wiele znaczeń, jak wielu jest użytkowników tych kategorii. Tym samym dyskusja egzaminów przeprowadzona wokół tego rodzaju pojemnych kategorii staje się tylko jedną z wielu możliwych.

Wspomnianej trudności w wytyczeniu precyzyjnej granicy obszaru zainteresowań diagno-

styki edukacyjnej, wyznaczanej terminologią wykorzystywaną przez autorów analizowanych prac, doświadczam również przy lekturze publikacji A. Pokropka (2013), co – jak wspomniałem wcześniej – ma bezpośredni związek z treściowym bogactwem dziedziny. Wykorzystywane w tej książce pojęcia, takie jak, np.: np. *sojalizacja pierwotna i wtórna, grupa społeczna, asymilacja, akomodacja*, są wpisane w kontekst omawianych zagadnień, ale też przynależą do dziedzin ugruntowanych w tradycji nauki, trudno zatem uznać je za kategorie pojęciowe wyłonione na podstawie i na potrzeby badań diagnostycznych. Za ważne uznaję podkreślenia przez A. Pokropka (2013, s. 16) znaczenia dobrze znanej książki pod red. B. Blooma, z 1956 roku, poświęconej taksonomii celów kształcenia. Nazywając ją jedną z najbardziej kluczowych „być może dla dziejów całej pedagogiki” (tamże), autor akcentuje rolę publikacji, w której językowa precyzja ma szczególnie znaczenie w mówieniu o uczniowskich osiągnięciach. Z językoznawczego punktu widzenia istotna wydaje mi się też dbałość o terminologię, która przejawia się przywoływaniem angielskich odpowiedników (czy też raczej pierwowzorów) polskich pojęć.

Wskazania praktyczne, czyli wokół wniosków wyprowadzanych przez edukacyjnych diagnostów

Obok siatki diagnostycznych pojęć, które – jak można stwierdzić na bazie omawianej przeze mnie grupy tekstów – wylaniają się, pozwalając lepiej rozumieć i dyskutować rzeczywistość szkolną, warto zwrócić uwagę również na zjawiska, których empiryczna weryfikacja może przyspieszyć proces kształtowania się „nowej” terminologii, a tym samym przyczynić się do rozwoju diagnostyki edukacyjnej.

W publikacji G. Szyling (2011), odnajduję wiele konkluzji prowadzących do dyskusji, ponieważ ich heterogeniczność ma związek ze złożonością procesu oceniania. Przyglądając się matrycowemu zestawieniu nauczycielskich uzasadnień obniżania kryteriów oceniania (tamże, s. 248–249), trudno mi oczekiwać jednoznaczności wniosków praktycznych. Przywołam tylko zasadnicze kwestie związane z obniżaniem wymagań (i uporządkowanych zgodnie z czterema aspektami trafności teoretycznej S. Messicka, zaadaptowanej do oceniania szkolnego), takie jak: pojmowanie oceny jako nagrody za wysiłek ucznia (tamże, s. 257), nauczycielska intencja wsparcia uczniowskich wysiłków podwyższeniem stopnia (tamże, s. 266), chronienie ucznia przed drugorocznością (tamże, s. 270), czy też

stwarzanie szans na rozwój zgodny z zainteresowaniami ucznia (tamże, s. 272). Są to kwestie znaczące, jednakże każda z nich zasługiwałaby na odrębne studia empiryczne, każda z nich ma swój intuicyjny pierwiastek i na każdą z nich silnie oddziałuje konkretny kontekst.

W tej samej publikacji jednak odnajduję, dzięki zamieszczonym w niej ilościowym analizom empirycznym, wiele dalece bardziej jednoznacznych konkluzji. Najistotniejsza z nich dotyczy – w moim przekonaniu – faktu systematycznego obniżania kryteriów przez nauczycieli, którzy uznaniowo traktują pilność i pracowitość uczniów, a także ulegają presji uczniowskich oczekiwań, szczególnie przy braku wyników w nauce. Zjawiska te występują niezależnie od pozycji szkoły zajmowanej w rankingach egzaminacyjnych i nie tylko w odniesieniu do dolnej części stopni szkolnych, przy czym wiedza uczniów o przesłankach nauczycielskich decyzji w sprawie ich ocen pochodzi głównie ze wzajemnej wymiany nieoficjalnych informacji (tamże, s. 320–355).

Przedstawiona (nie)jednoznaczność wyciąganych wniosków zdaje się wynikać głównie z różnic między badaniami jakościowymi a ilościowymi prowadzonymi w obszarze diagnostyki edukacyjnej, ale nie jest to jedyne źródło odmienności. Na przykład w pracy M. Groenwald (2011), reprezentującej to pierwsze z podejść badawczych i wykorzystującej wyniki analizy wywiadów z uczniami i nauczycielami, znajduje wyraźny brak zgody autorki na: dominację w szkole orientacji na wynik, a nie na proces (tamże, s. 168), przenoszenie cech egzaminu (np. niesprawiedliwości) na osobę nauczyciela (tamże, s. 154), zaniedbanie kwestii moralnych w dyskusjach na temat egzaminów (tamże, s. 161), przedmiotowe traktowanie osób uwikłanych w egzaminy (tamże, s. 166) – co w przypadku uczniów oznacza, iż wykorzystywani są oni do „utrzymywania się biurokratycznej maszyny systemu oświaty” (tamże, s.132). Uważam, że są to niewątpliwie spostrzeżenia bardzo znaczące, które stanowią też dobry punkt wyjścia podjęcia bardziej szczegółowych badań, także ilościowych. To wymagałoby jednak zoperacjonalizowania przywołanych przeze mnie powyżej bardzo pojemnych kategorii pojęciowych.

Z kolei publikacja A. Pokropka (2013) lokuje się na ilościowym biegunie podejść badawczych, co pozwala jej autorowi na dokonanie rozstrzygnięć zawężonych do jasnego zbioru badanych kategorii i przez to zdecydowanie bardziej jednoznacznych. Wśród takowych odnajduję, np. następujące konkluzje: „Zarówno w wypadku

bardzo słabych uczniów, jak i uczniów bardzo dobrych [gimnazjum], związek ich wyników ze składem klasy jest znacznie słabszy niż uczniów średnich” (tamże, s. 230); „[w liceum ogólnokształcącym] wartości efektów indywidualnych zarówno dla języka polskiego, jak i dla matematyki okazują się mieć najwyższą wartość na skrajach rozkładu, najniższą dla uczniów, których wyniki na wejściu lokują się w środkowych 50% wyników” (tamże, s. 231). Uważam, że wnioski te mają charakter praktyczny i dają się przełożyć na codzienną pracę nauczyciela bez podnoszenia poziomu ich szczególności.

Podsumowanie, czyli refleksje językoznawcy

Korzystając z pojęcia „przestrzeni poznawczo niczyjej” (Groenwald, 2011, s. 67), chciałbym w niej ulokować językowe aspekty diagnostyki edukacyjnej. Konstatacja ta nie tyle jest przejawem mojego krytycznego podejścia do omawianych w tym tekście prac, ile wyrazem zdziwienia, że w niewielkim zakresie uwzględniają one rolę języka jako znaczącego składnika diagnozowanej rzeczywistości edukacyjnej.

Uważam, że jeśli dyskusowanie zagadnień związanych z kształceniem szkolnym, silnie zdeterminowanym językowo, jest pozbawione odwołań ściśle lingwistycznych, brakuje w nim ważnego komponentu i tym samym staje się po części ograniczone. Dostrzeżona przeze mnie luka w obszarze diagnostyki edukacyjnej jest ważnym argumentem przemawiającym za uwzględnieniem w dalszym rozwoju tej dziedziny zagadnień lingwistycznych. Uznaję za potrzebne systematyczne włączanie w dyskusje diagnostyczne zarówno pojęć o charakterze mniej specjalistycznym (np. interakcja

językowa, językowe zasoby nauczycieli i egzaminatorów, bariery językowe), jak i tych wąsko zdefiniowanych (np. *interlanguage*, metafora, funkcje języka, argumentacja, socjolekt). Język jest bowiem jednym z tych obszarów, który towarzyszy poznawczemu, emocjonalnemu i społecznemu rozwojowi ucznia, a może też stać się jednym ze źródeł poczucia bezradności uczniów, co ujawnia w swoich badaniach Barbara Ciżkowicz (2009, s. 193), badaczka również powiązana z diagnostyką edukacyjną.

Przypisy

¹ Kategoria pojęciowa wprowadzona przez T. Gooda i J. Brophy'ego (1987)

Literatura

Ciżkowicz B. (2009): *Wyuczona bezradność młodzieży*. Wydawnictwo Uniwersytetu Kazimierza Wielkiego, Bydgoszcz.

Good T. L., Brophy J. E. (1987): *Looking in the classroom*. Harper & Row, New York.

Groenwald M. (2011): *Etyczne aspekty egzaminów szkolnych*. Wydawnictwo UG, Gdańsk.

Niemierko B. (1993): *Diagnostyka dydaktyczno-wychowawcza*. W: W. Pomykało: *Encyklopedia pedagogiczna*, Innowacja, Warszawa.

Niemierko B. (2002): *Ocenianie szkolne bez tajemnic*. WSiP, Warszawa.

Niemierko B. (2009): *Diagnostyka edukacyjna*. PWN, Warszawa.

Pokropek A. (2013): *Efekt rówieśników w nauczaniu szkolnym*. IBE, Warszawa.

Szyling G. (2011): *Nauczycielskie praktyki oceniania poza standardami*. Oficyna Wydawnicza „Impuls”, Kraków.

Michał Daszkiewicz
Uniwersytet Gdański

Anna Kot
Uniwersytet Gdański

PROFESOR BOLESŁAW NIEMIERKO I DIAGNOSTYKA EDUKACYJNA W POLSCE

Diagnoza w najbardziej ogólnym ujęciu oznacza ocenę stanu czegoś przedstawioną na podstawie badań i analiz. Termin ten pochodzi z języka greckiego, w którym *diagnosis* to rozróżnianie, wyróżnianie, osądzanie; *diagnostykos* oznacza umiejącego rozpoznawać, a *diagnostica technē* to sztuka odróżniania oraz umiejętność stawiania diagnozy (Lisowska, 2008, s. 18).

Początkowo pojęcie to związane było przede wszystkim z medycyną, ponieważ postawienie diagnozy uznawano i uznaje się za podstawę działania lekarza. Pierwsze wzmianki o diagnozie spotykamy już w pismach szkoły Hipokratesa. Pierwotnie wiązała się ona wyłącznie ze stanem patologicznym organizmu, jednak z czasem przyjęła się jako określenie stanu zdrowia pacjenta, jego zdolności do pracy itp. (Ziemski, 1973, s. 16). Aktualnie pojęcie diagnozy i związane z nim działania występują niemalże wszędzie. Na jej podstawie podejmują decyzje ekonomiści, urbaniści, sędziowie i adwokaci, jest ona nieodłącznym elementem wszelkiego rodzaju poradnictwa i orzecznictwa. Zajmuje również poczesne miejsce w obszarze nauk społecznych, a zwłaszcza w pedagogice. Sytuacje problemowe występujące w sferze edukacyjno-wychowawczej sprawiły, że stała się ona jedną z podstawowych kategorii badań empirycznych także na gruncie pedagogiki (Kawula, 2001, s. 664).

Diagnostyka pedagogiczna nie jest jednak jednolitą dziedziną wiedzy. W zależności od specyficznego przedmiotu i zakresu diagnozy wyróżnić można poszczególne diagnostyki szczegółowe (tabela 1).

W niniejszym tekście skupię uwagę na jednej z wyróżnionych diagnostyk szczegółowych, jaką jest diagnostyka edukacyjna. Jej celem jest opis i wyjaśnianie „sposobów rozpoznawania sytuacji dydaktycznej i wychowawczej ucznia, w tym zwłaszcza zmian, jakie w nim zachodzą po wpływie oddziaływania pedagogicznego” (Niemierko, 1997, s. 96). Obejmuje ona zatem teorię i praktykę diagnozy edukacyjnej, zmierzając do uściślenia „rozpoznawanie warunków, przebiegu i wyników uczenia się” (Niemierko, 2009, s. 30).

Wskazanie konkretnej daty stanowiącej początek rozwoju diagnostyki edukacyjnej w Polsce jest zadaniem niewykonalnym. W wymiarze praktycznym była ona uprawiana przez nauczycieli „od zawsze”, zaś instytucjonalne „badanie wyników

Tabela 1. Diagnostyka pedagogiczna a diagnostyki szczegółowe (źródło: opracowanie własne)

DIAGNOSTYKA PEDAGOGICZNA	DIAGNOSTYKI SZCZEGÓŁOWE	PRZEDMIOT I ZAKRES DIAGNOZY	TWÓRCY I PRZEDSTAWICIELE
	Diagnostyka kliniczno-wychowawcza	Potrzeby opiekuńcze	Janusz Korczak Stanisław Kawula
	Diagnostyka środowiskowa	Środowisko życia	Halina Radlińska Irena Lepalczyk Ewa Marynowicz-Hetka Edward A. Mazurkiewicz
	Diagnostyka społeczna	Sytuacja społeczna jednostki, jej grupa społeczna i środowisko lokalne	Aleksander Kamiński
	Diagnostyka ratownicza i lecznicza	Jednostki przejawiające trudności wychowawcze, osoby specjalnej troski, niedostosowane społecznie, rodziny dysfunkcyjne	Natalia Han-Ilgiewicz
	Diagnostyka w pracy wychowawczej	Zachowanie i rozwój osobowości wychowanka, organizacja zespołu wychowawczego, wyniki procesu wychowania,	Romana Miller
	Diagnostyka w resocjalizacji	Stan nieprzystosowania społecznego jednostki	Jan Konopnicki Lesław Pytka Bronisław Urban
	Diagnostyka edukacyjna	Sytuacja dydaktyczna; przebieg, wyniki i kontekst uczenia się	Bolesław Niemierko

nauczania” zaczęło rozwijać się po zakończeniu pierwszej wojny światowej w II Rzeczypospolitej (Niemierko, 2002, s. 80). Natomiast o naukowym rozwoju tej dziedziny w Polsce mówić możemy dopiero po II wojnie światowej.

Postacią bezsprzecznie związaną z genezą diagnostyki edukacyjnej w Polsce jest profesor dr hab. Bolesława Niemierko¹. Jego pierwsze zetknięcie z pomiarem dydaktycznym nastąpiło ponad pięćdziesiąt lat temu. Wtedy to, krótko po obronie pracy doktorskiej² otrzymał od profesora Wincentego Okonia propozycję pracy w Instytucie Pedagogiki w Warszawie. W 1968 roku włączył się w prowadzone przez The International Association for the Evaluation of Educational Achievement (IEA)³ badania osiągnięć uczniów z przedmiotów przyrodniczych.

Udział w międzynarodowych projektach badawczych oraz licznych kursach i konferencjach szkoleniowych organizowanych przez IEA pozwolił Profesorowi na zdobycie doświadczenia oraz wymianę poglądów ze światowej sławy naukowcami z zakresie pomiaru dydaktycznego. Jedno z takich spotkań, z amerykańskim psychologiem i pedagogiem Benjaminem S. Bloomem, dało B. Niemierce impuls do rozpoczęcia pracy nad jego własną koncepcją celów nauczania. Już w 1975 roku opublikował książkę *ABC testów*, w której zaprezentował opracowaną przez siebie taksonomię celów poznawczych. Jej nazwa: ABC – na co wskazuje sam autor, podkreśla elementarność dokonanej przez niego klasyfikacji. Dziś ta jedna z najbardziej znanych Polsce taksonomii jest powszechnie wykorzystywana w dydaktyce.

Wydana w tym samym roku, co *ABC testów*, publikacja pt. *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki publiczne* stanowiła podstawę uzyskania przez B. Niemierkę tytułu doktora habilitowanego na Wydziale Pedagogicznym Uniwersytetu Warszawskiego. Jednak jeszcze przed tym wydarzeniem B. Niemierko został kierownikiem Pracowni Metodologii Badań Pedagogicznych w Instytucie Pedagogicznym UW oraz podjął systematyczną współpracę z Ośrodkami Metodycznymi m.in. w Warszawie i we Wrocławiu. W 1977 roku objął B. Niemierko funkcję wicedyrektora Instytutu Badań nad Młodzieżą (1977–1980), a następnie przez trzy lata

(1981–1983) był wicedyrektorem Instytutu Kształcenia Nauczycieli w Warszawie (Kowalik, 2005, s. 13).

Początek lat osiemdziesiątych to dla Profesora okres staży zagranicznych (roczne stypendium Fulbrighta w University of California w Los Angeles, a następnie trzy kolejne staże w USA, w latach: 1985, 1990 oraz 1992), współpracy z zachodnimi autorytetami naukowymi m.in. z Anthonym Nitko oraz czas rozpoczęcia działalności akademickiej. Jej etapy wyznaczają: Wyższa Szkoła Pedagogiczna w Bydgoszczy (od 1983 roku) oraz Uniwersytet Gdański, gdzie od 1989 roku B. Niemierko był najpierw kierownikiem Zakładu Metodologii Badań Pedagogicznych, a później kierownikiem Międzywydziałowego Studium Pedagogicznego UG. W roku akademickim 2003/04 z inicjatywy Profesora powołano w Instytucie Pedagogiki UG pierwszy w Polsce Zakład Diagnostyki Edukacyjnej.

Profesor B. Niemierko koordynował i realizował też liczne projekty ogólnopolskie i międzynarodowe, a wśród nich:

- Badania osiągnięć szkolnych realizowanych według projektu IEA (1983 rok). Ich wyniki zostały zaprezentowane w kilkutomowych zagranicznych i polskich publikacjach.
- II Międzynarodowe Badania Porównawcze z Przedmiotów Przyrodniczych (SISS – Second International Science Study), których koordynatorami w Polsce byli również uczniowie Profesora: Krystyna Czupiał (chemia) oraz Eugeniusz Gabryelski (fizyka).
- Program MOVE – badanie osiągnięć uczniów w szkołach wdrażających zmodernizowane programy kształcenia w przedmiotach fizyka, informatyka, ochrona i kształtowanie środowiska, język angielski oraz wprowadzenia do świata pracy – projekt.
- Program IMPROVE, którego celem było opracowanie i wdrożenie do praktyki szkolnej nowego, dostosowanego do wymogów gospodarki rynkowej modelu kształcenia w szkole zawodowej.
- Program *Nowa Matura* (Niemierko, 1999).

Przykłady narzędzi zbudowanych przez: wykonawców Międzynarodowych Badań Osiągnięć Uczniów, Nauczycieli i Szkół (IEA), realizatorów programów modernizacji kształcenia zawodowego PHARE/IMPROVE i PHARE/MOVE oraz przez zespoły eksperymentatorów Nowej Matury – zostały zaprezentowane m.in. w książce B. Niemierki *Pomiar wyników kształcenia*, opublikowanej w 1999 roku.

Przenoszenie doświadczeń metodologicznych z IEA na grunt Polski, prowadzenie na szeroką skalę badań empirycznych oraz teoretyczne opracowanie koncepcji pomiaru dydaktycznego stanowiło solidny fundament rozwoju diagnostyki edukacyjnej naszym kraju. Obok wspomnianej wcześniej taksonomii celów poznawczych, jako szczególnie znaczące dla powstania tej dziedziny wymieniłem należy zaadaptowanie na grunt Polski pomiaru sprawdzającego wielostopniowego oraz trójwymiarowego modelu treści kształcenia. Koncepcje te nie tylko przyczyniły się do rozwoju diagnostyki edukacyjnej jako dziedziny naukowej i akademickiej, lecz również stanowiły podstawowe narzędzie dla praktyków, nauczycieli różnych przedmiotów szkolnych, tworzących własne testy sprawdzające. Publikacje takie jak: *Pomiar sprawdzający w dydaktyce. Teoria i zastosowania* (Niemierko, 1990), *Między oceną szkolną a dydaktyką* (Niemierko, 1997), *Pomiar wyników kształcenia zawodowego* (Niemierko, 1997), *Pomiar wyników kształcenia* (Niemierko, 1999), *Ocenianie szkolne bez tajemnic*

(Niemierko, 2002), *Diagnostyka edukacyjna* (2009) – stanowią dziś kanon lektur dla nauczycieli oraz wszystkich osób związanych z pomiarem dydaktycznym i egzaminami zewnętrznymi.

Dla rozwoju diagnostyki edukacyjnej w Polsce ogromne znaczenie miało też przygotowanie kadry, która byłaby zdolna w sposób profesjonalny kontynuować pracę nad teoretycznymi i empirycznymi podstawami pomiaru dydaktycznego. B. Niemierko jeszcze jako pracownik Instytutu Pedagogicznego w Warszawie, nawiązał pod koniec lat sześćdziesiątych współpracę z Ośrodkiem Metodycznym we Wrocławiu. W jej wyniku powstała we Wrocławiu nieformalna struktura organizacyjna o charakterze seminarium naukowego. W jej ramach podejmowano próby przenoszenia teorii pomiaru dydaktycznego na grunt poszczególnych dyscyplin kierunkowych, np: do polonistyki (Tadeusz Patrzalek, Leopold Pańtak), chemii (Krystyn Czupiał), fizyki (Eugeniusz Gabryelski). W latach 1970–1989 odbyło się łącznie ponad 50 spotkań tej grupy seminaryjnej (Kowalik, 1998, s. 311). Ponadto pod kierunkiem Profesora zespół ten organizował cyklicznie kursy pomiaru dydaktycznego dla nauczycieli. Pierwszy z nich odbył się w lipcu 1972 roku we Włocławku.

Kolejne seminaria i kursy organizował B. Niemierko już na Uniwersytecie Gdańskim. Na przykład w kursach pomiaru dydaktycznego uczestniczyło blisko stu uczestników – nauczycieli szkolnych i akademickich, którzy po ich ukończeniu mieli okazję doświadczyć pełnego cyklu badań empirycznych: od ich zaplanowania, tworzenia i doskonalenia narzędzi, poprzez organizację badań, aż do analiz i komunikowania wyników (Szaleniec, 2005, s. 36). Ponadto w latach 1996–2004 zrealizowano na Uniwersytecie Gdańskim sześć edycji podyplomowego Studium Pomiaru Dydaktycznego i Podyplomowego Studium Ewaluacji Dydaktycznej, których uczestnikami było blisko tysiąc nauczycieli, metodyków i pracowników administracji szkolnej (Niemierko, 2009, s. 323). Działania te pozwoliły m.in. na przygotowanie przyszłej kadry okręgowych komisji egzaminacyjnych oraz diagnostów praktyków.

Przygotowywaniem kadry specjalistów pomiaru dydaktycznego oraz kadry praktyków towarzyszyła aktywność w obszarze tworzenia środowiska akademickiego i oświatowego przychylnego pomiarowi dydaktycznemu, czego empirycznym dowodem jest wypromowanie przez profesora B. Niemierkę ponad trzydziestu prac doktorskich. Doskonałą płaszczyznę do wymiany doświadczeń, prezentacji doniesień badawczych oraz prowadzenia dyskusji panelowych stanowiły i stanowią konferencje poświęcone diagnostyce edukacyjnej. Pierwsza z nich, pod naukowym kierownictwem Profesora, odbyła się w 1993 roku w Gdańsku. Miejsca kolejnych konferencji to Elbląg (1994), Legnica (1999), Szczecin (2000), Wałbrzych (2000), Kraków (2001), Wrocław (2001), Katowice (2002), Łódź (2003), Kraków (2004), Gdańsk (2005), Lublin (2006), Łomża (2007), Opole (2008), Kielce (2009), Toruń (2010), Kraków (2011), Wrocław (2012), Gniezno (2013) i najbliższa, jubileuszowa – w Gdańsku (2014).

Dorobek tych konferencji to dwadzieścia tomów wydawnictw, w których opublikowane zostały zarówno prace naukowców i specjalistów-diagnostów, jak i teksty początkujących autorów, którzy doskonalą swój warsztat pod okiem Profesora – współredaktora każdej pokonferencyjnej książki. Tematyka tych konferencji była niezwykle rozbudowana i obejmowała m.in. takie zagadnienia, jak: metody diagnozy edukacyjnej, kompetencje nauczycieli, egzaminy wewnątrzszkolne i zewnętrzne, perspektywy informatyczne egzaminów, ocenianie osiągnięć uczniów, egzamin w oczach uczniów,

etykę egzaminacyjną, trafność pomiaru, standardy wymagań i normy, konstrukcję zadań zamkniętych, jakość pracy szkoły (Niemierko, 2009, s. 46).

Kluczowe znaczenie dla rozwoju diagnostyki edukacyjnej w Polsce miało również powstanie Polskiego Towarzystwa Diagnostyki Edukacyjnej z siedzibą w Krakowie. Zostało ono powołane z inicjatywy B. Niemierki podczas odbywającej się w Łodzi Konferencji Diagnostyki Edukacyjnej w 2003 roku.

PTDE stawia sobie za cel m.in. (Statut PTDE):

- wspomaganie rozwoju systemów oceniania wewnątrzszkolnego i egzaminów zewnętrznych,
- promowanie diagnozy edukacyjnej w placówkach oświatowych,
- inspirowanie badań naukowych i dyskusji na temat diagnozy edukacyjnej,
- rozwijanie współpracy z organizacjami i instytucjami działającymi na rzecz diagnostyki w edukacji,
- wspieranie rozwoju kwalifikacji zawodowych członków Towarzystwa.

* * *

Zgodnie z kryteriami klasyfikacji nauk, diagnostyka edukacyjna nie osiągnęła jeszcze statusu subdyscypliny naukowej, nie spełnia bowiem wszystkich wymogów formalnych stawianych subdyscyplinie⁴. Jej status jako subdyscypliny pedagogiki wciąż się konstytuuje. Podkreślić jednak należy, że z inspiracji i pod kierownictwem naukowym profesora Bolesława Niemierki została opracowana w Polsce oryginalna szkoła naukowa diagnostyki edukacyjnej oraz dynamicznie rozwinął się nauczycielski ruch badań pedagogicznych związanych z szeroko rozumianym pomiarem dydaktycznym. Znaczące osiągnięcia naukowe Profesora oraz jego współpracowników i uczniów to z pewnością doskonały punkt wyjścia do stworzenia dojrzałej subdyscypliny naukowej oraz ukształtowania – opartej na wiedzy naukowej – diagnostyki.

Przypisy

¹ B. Niemierko urodził się 17 lipca 1935 roku w Poznaniu. W roku 1952 ukończył szkołę średnią w Olsztynie, a następnie w 1954 roku podjął studia na kierunku pedagogika Uniwersytetu Warszawskiego. W 1960 roku obronił pracę magisterską: *Poglądy na świat współczesnej młodzieży polskiej*.

² Pracę doktorską *Niepowodzenia dydaktyczno-wychowawcze nauczycieli szkół podstawowych* B. Niemierko obronił w 1966 roku, na Wydziale Pedagogicznym Uniwersytetu Warszawskiego

³ IEA to niezależne, międzynarodowe stowarzyszenie instytucji i badawczych agencji rządowych zajmujące się prowadzeniem badań naukowych dotyczących szkolnych osiągnięć uczniów. Powstało w 1959 roku przy Instytucie Pedagogicznym UNESCO w Hamburgu. <http://www.iea.nl/> (otwarty 14 lutego 2014).

⁴ Dyscyplina naukowa stanowi „doniosłą społecznie, ukształtowaną i wyodrębnioną ze względu na przedmiot i cel badań lub kształcenia część nauki w znaczeniu instytucjonalnym, uznaną za podstawową jednostkę jej kwalifikacji” (Krzyżanowski, 1994, s. 44).

Literatura

Kawula S. (2003): *Diagnostyka pedagogiczna środowiska*. W: T. Pilch (red.): *Encyklopedia pedagogiczna XXI wieku*. Tom I. Wyd. Akademickie „Żak”. Warszawa.

- Kowalik E. (1998): *Profesor Bolesław Niemierko*. W: B. Niemierko, E. Kowalik (red.): *Perspektywy diagnostyki edukacyjnej*. Wyd. Uniwersytetu Gdańskiego, Gdańsk.
- Kowalik E. (2005): *Profesor Bolesław Niemierko*. W: M. Groenwald i in. (red.): *Diagnostyka edukacyjna. Niemierkowskie spotkania i inspiracje*. Wyd. Fundacji Rozwoju UG, Gdańsk.
- Krzyżanowski L. J. (1994): *O podstawach kierowania organizacjami*. Wyd. Naukowe PWN, Warszawa.
- Lisowska E. (2008): *Diagnostyka pedagogiczna w pracy z dzieckiem i rodziną*. Wszechnica Świętokrzyska, Kielce.
- Niemierko B. (1997): *Diagnostyka dydaktyczno-wychowawcza*. W: W. Pomykała (red.): *Encyklopedia pedagogiczna.*, Fundacja Innowacja, Warszawa.
- Niemierko B. (2002): *Ocenianie szkolne bez tajemnic*. WSiP, Warszawa.
- Niemierko B. (2009): *Diagnostyka edukacyjna*. Wyd. Naukowe PWN, Warszawa.
- Statut PTDE, dostępny na: http://www.ptde.org/old/html/statut_PTDE.html (otwarty 17 lutego 2014).
- Szaleniec H. (2005): *Ziarno egzaminów zewnętrznych w Polsce zostało zasiane ćwierć wieku temu*. W: M. Groenwald i in. (red.): *Diagnostyka edukacyjna. Niemierkowskie spotkania i inspiracje*. Wyd. Fundacji Rozwoju UG, Gdańsk.
- Ziemski S. (1973): *Problemy dobrej diagnozy*. Wiedza Powszechna, Warszawa.

Barbara Przychodzeń
Okręgowa Komisja Egzaminacyjna w Gdańsku

CZYNNIKI PODNOSZĄCE EFEKTYWNOŚĆ KSZTAŁCENIA W GIMNAZJACH

Kronikarska relacja z badań Okręgowej Komisji Egzaminacyjnej w Gdańsku

We współczesnej edukacji efektywność kształcenia jest traktowana jak jeden z priorytetów praktyki pedagogicznej. Pojęcie to nie jest jednak jednoznacznie definiowane, do czego przyczynia się zarówno różnorodność sposobów badania i wielość obszarów powiązanych z efektywnością kształcenia, jak też niemożność uzgodnienia uniwersalnych kryteriów jej oceny. W niniejszym tekście przyjmuję za Kazimierzem Denkiem (2005, s. 7), że: „Efektywność realizacji procesu nauczania i uczenia się wyraża zespół pozytywnych cech, przynoszących najlepsze rezultaty w postaci wyników w nauce, obejmujących wiedzę, nawyki, zainteresowania i zdolności poznawcze oraz motywę i przyzwyczajenia do ustawicznego kształcenia się”.

Zainteresowanie problematyką empirycznego ustalania efektywności działań edukacyjnych w szkole wzrosło wraz z wprowadzeniem do polskiej edukacji systemu egzaminów zewnętrznych. Analizując wyniki sprawdzianu i egzaminu gimnazjalnego poszukiwano równocześnie sposobów diagnozowania i szacowania postępu edukacyjnego uczniów (Kutajczyk, Przychodzeń, 2008a, s. 226), co z kolei wywołało zapotrzebowanie na stworzenie metody pozwalającej zmierzyć dydaktyczne osiągnięcia nauczycieli oraz szkół, wykraczające poza proste rangowanie wyników egzaminów. Oczekiwania te zdawała się spełniać miara efektywności nauczania określana wskaźnikiem edukacyjnej wartości dodanej dla szkoły. Śledzenie rozwoju EWD umożliwiają liczne teksty zamieszczone w publikacjach związanych z Konferencjami Diagnostyki Edukacyjnej. W niniejszym tekście jednak przedstawię tylko wąski wycinek tych zagadnień, sytuujący się na pograniczu głównego nurtu poszukiwań, a powiązany z lokalnymi działaniami badawczymi, podejmowanymi przez Okręgową Komisję Egzaminacyjną w Gdańsku.

Edukacyjna wartość dodana, nazywana w skrócie EWD, jest definiowana jako „przyrost osiągnięć w wybranym zakresie programowym w określonym czasie” (Niemierko, 2004, s. 34) oraz jako „przyrost wiedzy uczniów w wyniku procesu edukacyjnego” (Dolata, 2007, s. 9). Proces kształcenia jest jednak różnorodnie uwarunkowany kontekstowo, co znajduje odzwierciedlenie w wynikach osiągnięć uczniów mierzonych egzaminami zewnętrznymi. Poszukiwania tych uwarunkowań podjęli się pracownicy Okręgowej Komisji Egzaminacyjnej w Gdańsku, którzy w latach 2004–2009 prowadzili badania na terenie województw kujawsko-pomorskiego i pomorskiego, korzystając przy tym z merytorycznego wsparcia profesora Bolesława

Niemierki. Niniejszy tekst jest rodzajem syntetycznej relacji z badań już zakończonych, niemniej ważnych dla podmiotów zaangażowanych w ich przeprowadzanie.

Badaniami objęto 126 szkół, w tym: 69 szkół podstawowych, 40 gimnazjów i 17 szkół ponadgimnazjalnych. Łącznie zgromadzono dane z 5334 kwestionariuszy wypełnionych przez uczniów, ich wychowawców i dyrektorów szkół. Wyniki ankiety przeprowadzonej wśród uczniów skonfrontowano z danymi zebranymi od dyrektorów badanych szkół i wychowawców klas oraz z osiągnięciami uczniów mierzonych wynikami egzaminów zewnętrznych.

Założenia metodologiczne badań prowadzonych w gimnazjach

Cele i problemy badań

Głównym celem badań gimnazjalistów było ustalenie relacji między rezultatami egzaminów a wybranymi zmiennymi, które uznaje się za najbardziej charakterystyczne uwarunkowania efektywności uczenia się. Analizie poddano również relacje występujące między zmiennymi a przyrostem osiągnięć uczniów kończących gimnazjum.

W badaniach postawiono następujące pytania badawcze:

1. Jaki związek występuje pomiędzy poziomem osiągnięć uczniów ustalonym w wyniku egzaminu zewnętrznego po szkole podstawowej a ich poziomem ustalonym na podstawie wyniku egzaminu prowadzonego na koniec kształcenia w gimnazjum?
2. Jakie relacje występują między przyrostem osiągnięć uczniów gimnazjów a cechami indywidualnymi tychże uczniów oraz ich środowiskiem rodzinnym?
3. Jaki związek istnieje między przyrostem osiągnięć uczniów w gimnazjach a organizacją pracy szkoły i działaniami edukacyjnymi nauczycieli?

Metody i narzędzia badawcze

Na potrzeby badań skonstruowano kwestionariusz ankiety skierowanej do uczniów. Składały się na niego pytania pogrupowane tematycznie, odpowiadające zoperacjonalizowanym zmiennym, jak klimat społeczny klasy i szkoły, opinia o lekcjach i uczeniu się w gimnazjum, zainteresowania uczniów i zajęcia pozalekcyjne, a także wrażenia z egzaminu gimnazjalnego i plany uczniów po ukończeniu gimnazjum. Dyrektorów szkół i wychowawców klas pytano w badaniu ankietowym przede wszystkim o organizację i przebieg procesu kształcenia gimnazjalistów. Ponadto wychowawcy zostali poproszeni o scharakteryzowanie swoich wychowanków i pozycji ekonomiczno-społecznej ich rodzin oraz o podanie ocen semestralnych w klasie III z przedmiotów objętych egzaminem gimnazjalnym.

Drugą metodą badawczą – równie ważną jak ankieta – była analiza danych zastanych, czyli wyniki egzaminów zewnętrznych uczniów objętych próbą badawczą.

Dobór próby

Próbę objętą badaniami efektywności kształcenia w gimnazjum, prowadzonymi przez OKE w Gdańsku w latach 2006–2009, charakteryzuje tabela 1.

Tabela 1. Próba objęta badaniami uwarunkowań efektywności kształcenia w gimnazjach

Rok	Obszar badania	Uczestnicy badań		Liczba uczestników badań ankietowych		
		Liczba szkół	Liczba uczniów	Uczniowie	Wychowawcy	Dyrektorzy
2006	Powiaty: sztumski, kartuski (6 gmin)	6	730	625	29	6
2007	Powiaty: bytowski, kartuski, słupski (5 gmin)	11	502	418	24	11
2008	Gdańsk, Gdynia (2 gminy)	5	618	480	25	5
2009	Powiaty: sępoleński, rypiński (10 gmin)	18	1 141	959	47	18
Razem		40	2 992	2 482	125	40

W niniejszym opracowaniu przedstawiono wybrane wnioski z analiz danych uzyskanych w badaniu przeprowadzonym w 2007 roku w gimnazjach wiejskich (Kutajczyk, Przychodzeń, 2008b, s. 261), których uczniowie uzyskiwali najniższe wyniki egzaminów zewnętrznych. W takim doborze próby istotne było też to, że zdający z gimnazjów wiejskich stanowili około jedną trzecią populacji zdających w województwie pomorskim.

Założenia dotyczące analizy danych

W badaniach przyjęto, że wyniki egzaminu są wskaźnikiem poziomu opanowania przez uczniów określonych umiejętności poznawczych. Równocześnie jednak uwzględniono, że rozwojowi umiejętności uczniów towarzyszy tzw. efekt wachlarzowy, rozumiany jako dodatnia korelacja między stanem „na wejściu” i przyrostem osiągnięć uczniów (Niemierko, 2009, s. 216–217).

Za miarę efektywności pracy szkoły, czyli przyrostu osiągnięć uczniów w określonym czasie uczenia się, uznano edukacyjną wartość dodaną przedstawioną w skali równoważników klasy (RK), czyli w latach i miesiącach uczenia się (Niemierko, 2009, s. 255–256). Występujące przy RK indeksy dotyczą: G – sumarycznych wyników egzaminu, GH i GM – odpowiednio części humanistycznej i matematyczno-przyrodniczej.

W tej metodzie za miarę wartości dodanej bezwzględnej przyjmuje się efekt standardowy jednej klasy szkolnej równy 0,7. Pozwala to na przedstawienie osiągnięć ucznia w skali RK, czyli w skali lat i miesięcy uczenia się. Do oszacowania równoważników klasy po szkole podstawowej (RK_{sp}) i po gimnazjum (RK_G) oraz bezwzględnej edukacyjnej wartości dodanej w gimnazjum wykorzystano następujące wzory:

$$RK_{sp} = \frac{z_s}{0,7} + 6, \text{ gdzie } z_s - \text{wynik standardowy ucznia ze sprawdzianu,}$$

$$EWD_G = \frac{z_{EG} - r_{EG,S} \cdot z_s}{0,7} + 3,$$

gdzie: z_{EG} – sumaryczny wynik standardowy ucznia z egzaminu gimnazjalnego,
 $r_{EG,S}$ – współczynnik korelacji wyników egzaminu gimnazjalnego i sprawdzianu,
 $RK_G = RK_{sp} + EW_{DG}$.

W niniejszych obliczeniach przyjęto wartości dla populacji zdających z województw kujawsko-pomorskiego i pomorskiego, które dla wyników sprawdzianu 2004 i egzaminu gimnazjalnego 2007 były równe: $r_{EG,S} = 0,82$, $r_{GH,S} = 0,78$, $r_{GM,S} = 0,75$.

W procesie analizy danych dla części humanistycznej i matematyczno-przyrodniczej oszacowano dla każdego ucznia również RK_{GH} i RK_{GM} . Natomiast równoważniki klas RK_{SP} , RK_G i EWD_G dla szkoły, klasy lub innej grupy uczniów ustalono empirycznie, obliczając je jako średnie arytmetyczne poszczególnych uczniów. Przyjęto, że przeciętny równoważnik klasy w kraju wynosi: po szkole podstawowej RK_{SP} 6 lat, a po gimnazjum RK_G – 9 lat, a przeciętna edukacyjna wartość dodana w gimnazjum wynosi 3 lata.

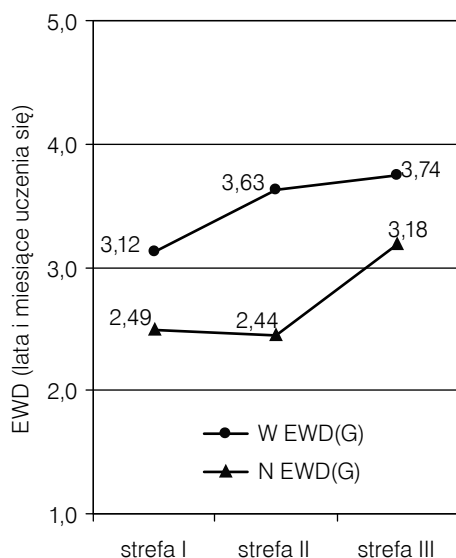
Wybrane wyniki badań

Zróźnicowanie przyrostu osiągnięć uczniów badanych szkół

W poniższej analizie uwzględniono trzy grupy uczniów, dobrane celowo na podstawie wyników sprawdzianu gimnazjalnego: strefa I – uczniowie o wynikach niskich (staniny 1–3), strefa II – uczniowie o wynikach średnich (staniny 4–6), strefa III – uczniowie o wynikach wysokich (staniny 7–9). Porównano także szkoły, do których uczęszczali ci uczniowie, biorąc pod uwagę wartości EWD_G oraz wartości dla obu części egzaminu gimnazjalnego (EWD_{GH} i EWD_{GM}). W analizie przyjęto następujące oznaczenia:

W_{EWD} – szkoły o wskaźniku EWD istotnie statystycznie ($p < 0,05$) większym niż 3,00
 N_{EWD} – szkoły o wskaźniku EWD istotnie statystycznie ($p < 0,05$) mniejszym niż 3,00.

Analizując wyniki badania, zauważono zróźnicowanie wartości wskaźników EWD dla uczniów o różnym potencjale edukacyjnym w dwóch grupach szkół. Ilustrują to profile strefowe, czyli linie łączące punkty odpowiadające średnim wartościom EWD dla uczniów, których RK_{SP} znalazły się w kolejnych strefach (rys. 1).

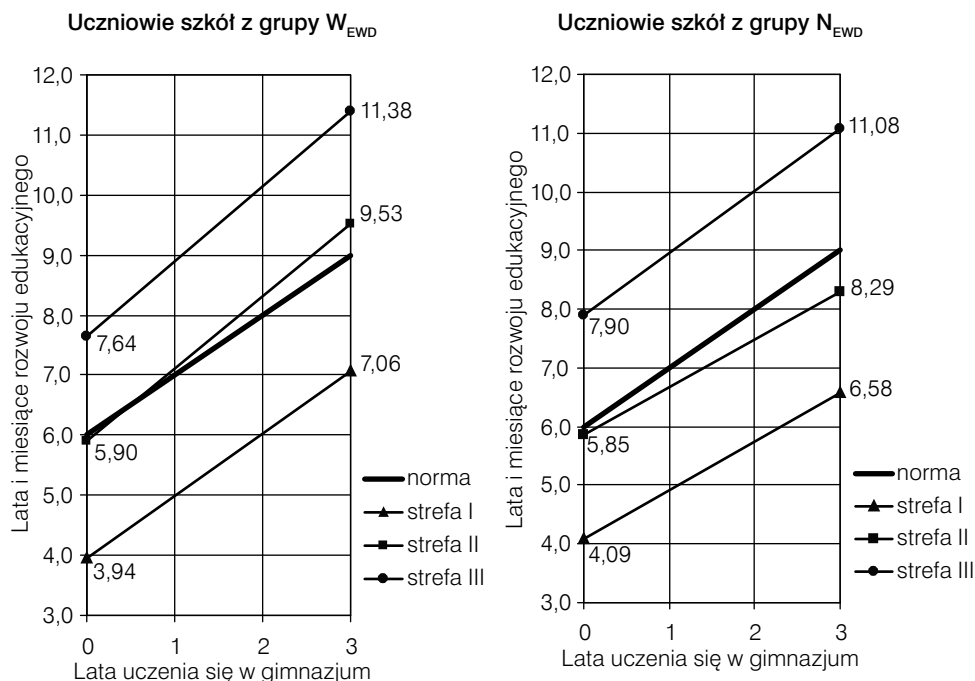


Rys. 1. Profile strefowe dla uczniów szkół z grup W_{EWD} i N_{EWD}

W szkołach z grup W_{EWD} i N_{EWD} zauważyć można, że efektywność działań edukacyjnych była najbardziej zróźnicowana u uczniów ze strefy II (uczniowie „przeciętni”). Przyrost ich osiągnięć w szkołach o wysokiej EWD_G uległ przyspieszeniu, a w szkołach o niskiej EWD_G – spowolnieniu.

Profile na rysunku wskazują, że przyrosty osiągnięć uczniów w tych szkołach współwystępowały z pracą z uczniami „przeciętnymi”.

Łącząc punkty odpowiadające średnim wartościom RK_{SP} i RK_G , wykreślono linie ilustrujące tempo przyrostu osiągnięć uczniów w dwóch grupach szkół, uwzględniając trzy wyróżnione strefy, odniesione do przeciętnego tempa dla populacji w kraju, zilustrowanego tzw. linią normy. Przyjęto przy tym, że tempo przyrostu umiejętności u uczniów o różnych osiągnięciach „na wejściu” może być wskaźnikiem pracy szkoły w zakresie wyrównywania szans edukacyjnych i indywidualizacji procesu uczenia się uczniów (rys. 2.)

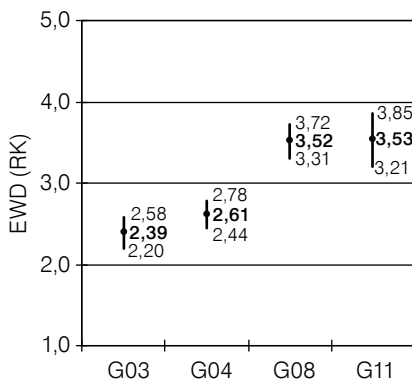


Rys. 2. Linie rozwoju uczniów o różnym poziomie osiągnięć „na wejściu” w grupach szkół wyodrębnionych ze względu na EWD_G

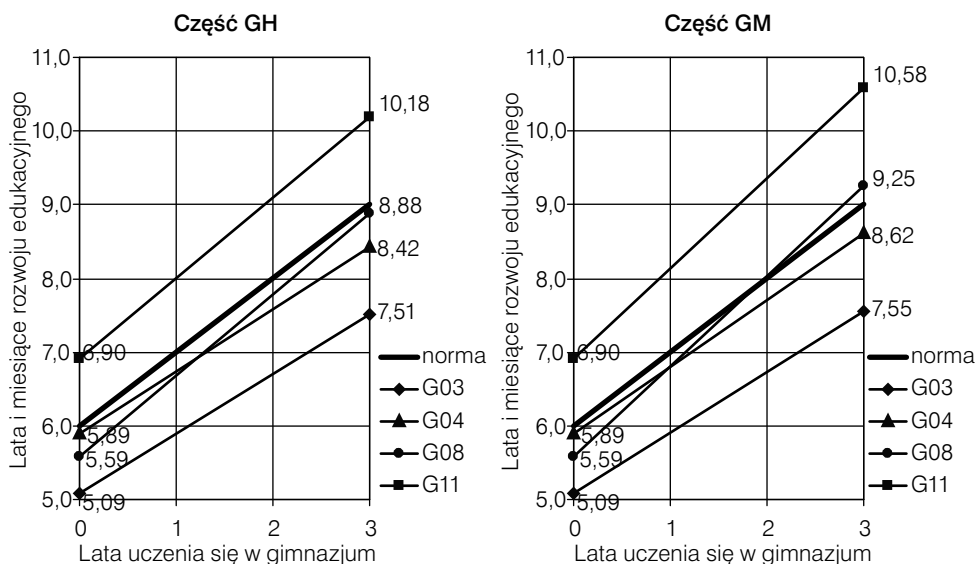
Zróżnicowanie przyrostu osiągnięć uczniów w przedmiotach humanistycznych i matematyczno-przyrodniczych

W niniejszej części zostanie przedstawiona analiza wybranych wyników zgromadzonych w czterech gimnazjach różniących się istotnie EWD . Są to: G08 i G11 – oba z grupy W_{EWD} oraz G03 i G04 – oba z grupy N_{EWD} (rys. 3).

Dla ogółu uczniów EWD_{GH} jest statystycznie mniejsza niż przeciętna dla populacji, a EWD_{GM} – większa (rys. 4.). Ta prawidłowość jest na tyle charakterystyczna dla młodzieży wiejskiej, że wystąpiła nawet w szkołach z grupy W_{EWD} , co ilustruje rys. 2, a jej przyczyny lokować można między innymi w trudniejszym dostępie do źródeł kultury niż w wypadku rówieśników mieszkających w miastach.



Rys. 3. Zróżnicowanie EWD_c w porównywanych szkołach



Rys. 4. Linie przyrostu osiągnięć uczniów szkół istotnie statystycznie różniących się EWD

Związek indywidualnych właściwości uczniów z przyrostem ich osiągnięć

Wyrażone przez wychowawców klas na skali pięciostopniowej opinie na temat cech indywidualnych uczniów (poziom rozwoju intelektualnego, zaangażowanie we własną naukę i w życie klasy, frekwencja na zajęciach i zachowanie) różniły się w porównywanych szkołach. W grupie N_{EWD} średnia ocena była równa 3,0 (w G03 – 2,7), a w grupie W_{EWD} – 3,8 (w G11 – 3,9). Analiza danych pochodzących od wychowawców klas i uczniów wykazała, że badane indywidualne właściwości uczniów (m.in. uzdolnienia, odpowiedzialność za własny rozwój, świadomość potrzeby uczenia się i motywacja wewnętrzna oraz aspiracje) wykazują związek z przyrostem osiągnięć w czasie trzech lat nauki w gimnazjum.

Najsilniej z EWD_G dodatnio korelują uzdolnienia i frekwencja na zajęciach, odrabianie prac domowych, czytanie książek (w tym lektur szkolnych) i nienapotykanie na trudności w uczeniu się. Gimnazjaliści osiągający najlepsze wyniki poświęcali więcej niż ich rówieśnicy czasu na odrabianie lekcji i uczenie się, częściej czytali lektury szkolne i inne książki oraz częściej korzystali z komputera w celach edukacyjnych (tabela 2.).

Tabela 2. Wartości wskaźnika akceptacji stwierdzeń dotyczących zajęć poza szkołą

Stwierdzenie	Grupa W_{EWD}	Grupa N_{EWD}
Odrabiam lekcje/uczę się	4,3	3,9
Czytam lektury szkolne	3,8	3,1
Czytam inne książki niż lektury	4,1	2,8
Korzystam z komputera w celach edukacyjnych	4,0	3,5
Gram w gry komputerowe	2,8	3,2
Uprawiam sport amatorsko	3,3	3,1

Uczniowie z porównywanych grup szkół różnili się także aspiracjami dotyczącymi dalszego kształcenia się po gimnazjum. W szkole G11 największy odsetek gimnazjalistów (75%) zamierzał kontynuować naukę w liceum ogólnokształcącym, a w szkole G03 – 15% (tabela 3). W tabeli 3 zestawiono również wartości odpowiedzi uczniów z grup W_{EWD} i N_{EWD} odpowiadające stopniowi akceptowania przez nich stwierdzeń dotyczących cech indywidualnych uczniów. Uwagę zwraca przy tym, że zainteresowania językiem polskim miały związek z EWD_{GH} , natomiast nie zaobserwowano korelacji między zainteresowaniami matematyką i EWD_{GM} .

Tabela 3. Wartości wskaźnika akceptacji dla grup zmiennych: cechy indywidualne uczniów

Grupa zmiennych	Grupa W_{EWD}	Grupa N_{EWD}
„Pilność” (ocena wychowawcy klasy)	4,1	3,3
Odpowiedzialność za własny rozwój (samoocena)	3,9	3,4
Zainteresowanie językiem polskim (samoocena)	3,7	3,2
Zainteresowanie matematyką (samoocena)	3,3	3,3
Odsetek uczniów, którzy zamierzali uczyć się w LO	55%	30%

Powiązania zbliżone charakterem do omówionych poprzednio wystąpiły również w przypadku czynników związanych z pozycją ekonomiczno-społeczną rodzin, mierzoną wykształceniem rodziców, sytuacją materialną, wydolnością wychowawczą oraz zainteresowaniem nauką dziecka i zaangażowaniem rodzica w życie klasy. Średnia opinii wyrażanych w skali 0–5 wynosiła: w szkołach G03 i G04 – odpowiednio 2,7 i 3,0, a w szkołach G08 i G11 – 3,3 i 3,4. Najsilniej z przyrostem osiągnięć uczniów gimnazjów wiejskich koreluje wykształcenie rodziców, a pozostałe zmienne są pośrednio z nim związane.

Związek organizacji pracy szkoły i działań edukacyjnych nauczycieli z przyrostem osiągnięć gimnazjalistów

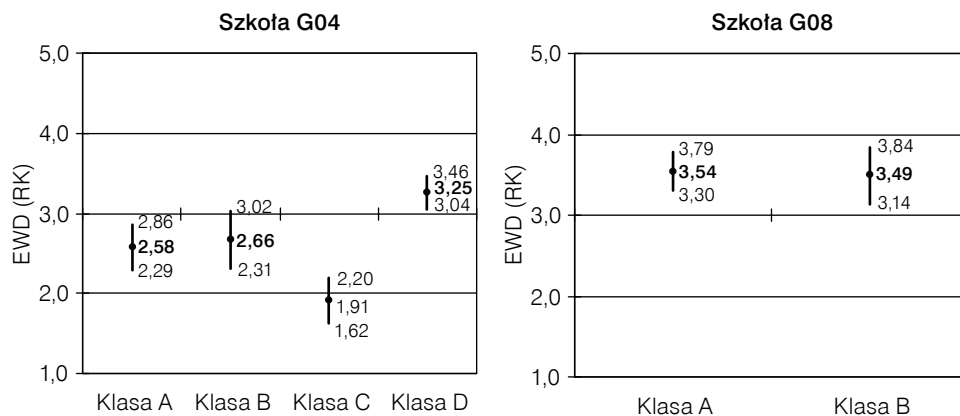
Analiza odpowiedzi uczniów na pytania dotyczące lekcji z przedmiotów wchodzących w zakres egzaminu gimnazjalnego (zwłaszcza języka polskiego i matematyki) pokazała różnice dla szkół o różnych wartościach wskaźnika EWD.

Na podstawie wartości współczynników korelacji między EWD i opiniami uczniów o lekcjach można stwierdzić, że w przypadku języka polskiego najistotniejsze znaczenie dla efektywności kształcenia mają czynniki indywidualne i poznawcze (korelacja wysoka), a w przypadku matematyki – indywidualne (korelacja bardzo wysoka). Wystąpił także silny związek między zainteresowaniem ucznia treściami lekcji i zrozumieniem wyjaśnień nauczyciela. Dotyczy to zwłaszcza przedmiotów matematyczno-przyrodniczych, przy czym uczniowie deklarują, że na lekcjach matematyki uczą się efektywniej, gdy rozwiązują problemy i rzadziej korzystają z pomocy nauczyciela, czyli gdy mają czas na samodzielne rozwiązywanie zadań.

Związek selekcji uczniów przyjmowanych do klas pierwszych z przyrostem osiągnięć gimnazjalistów

W gimnazjach wiejskich jednym z głównych kryteriów przydziału uczniów do klas pierwszych jest miejsce zamieszkania i wynikająca stąd konieczność dowożenia ich do szkoły. Jednak grupowanie uczniów zamiejscowych w jednej klasie nie sprzyja wyrównywaniu ich szans edukacyjnych. Młodzież wiejska na ogół żyje w rodzinach o niskim statusie ekonomiczno-społecznym, co w znacznym stopniu niekorzystnie wpływa na przyrost jej osiągnięć. Jest też rzadziej diagnozowana pod kątem specyficznych trudności w uczeniu się, co – jak wskazują liczne badania – może wynikać ze słabszego zrozumienia przez rodziców potrzeby takich badań i mniejszych możliwości ich przeprowadzenia.

Analiza przyczyn zróżnicowania przyrostu osiągnięć uczniów w oddziałach klasowych wykazała związek z wewnątrzszkolną selekcją uczniów przyjmowanych do klas pierwszych, co ilustrują rysunki 5 i 6.

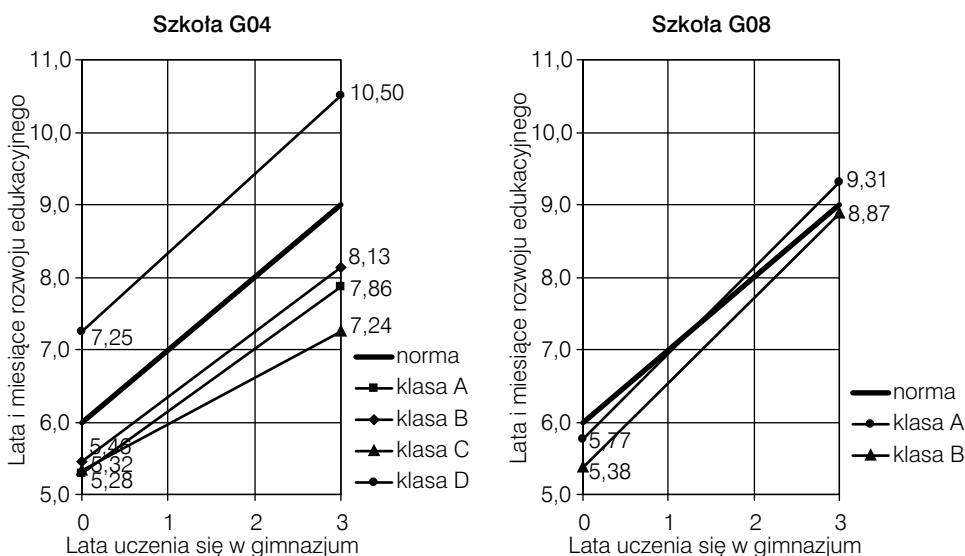


Rys. 5. EWD_G z przedziałami ufności (95%) w oddziałach klasowych szkół G04 i G08

W gimnazjum G04 były cztery klasy (A, B, C i D). Wszyscy uczniowie uczęszczający do klasy B i C byli zamiejscowi, natomiast w klasach A i D ich odsetek wynosił około 50%. Uczniowie klas A, B i C rzadko uczestniczyli w zajęciach kół przedmiotowych, natomiast w klasie D więcej uczniów brało udział zarówno w takich zajęciach, jak i w konkursach przedmiotowych.

W szkole G08 były dwie klasy (A i B). Uczniowie obu mieli podobny potencjał edukacyjny „na wejściu”, natomiast w klasie A panował bardziej sprzyjający kontekst uczenia się niż w klasie B (wyżej wykształceni i pracujący rodzice, znacznie mniejsza grupa uczniów zamiejscowych). Wychowawca klasy B informował: „Większość uczniów pochodzi z rodzin rolniczych, w których dzieci ciężko pracują, poświęcając mało czasu na naukę”. Jednak porównywalny odsetek młodzieży uczestniczył w zajęciach pozalekcyjnych i konkursach przedmiotowych.

Jak można zaobserwować na wykresie (rys. 6), w szkole G04 klasa D ma EWD_G statystycznie wyższą niż średnia dla populacji, klasa B – przeciętną, a klasy A i C mają EWD_G statystycznie niższe niż przeciętna. Natomiast w szkole G08 obie klasy mają EWD_G nieróżniące się od siebie istotnie i statystycznie wyższe niż średnia dla populacji.



Rys. 6. Linie przyrostu osiągnięć uczniów poszczególnych klas w gimnazjach G04 i G08

Patrząc na linie rozwoju edukacyjnego klas szkoły G04 widzimy istotny statystycznie efekt wachlarzowy, który można powiązać z wysokim wskaźnikiem selekcji wewnątrzszkolnej do oddziałów klasowych.

Podsumowanie

Z przeprowadzonych badań (Kutajczyk, Przychodzeń, 2008c, s. 85) wynika – przy zachowaniu ostrożności we wnioskowaniu ze względu na celowy i ograniczony dobór próby – że:

1. Wykorzystując EWD do oceny poziomu pracy szkoły, nie można pominąć kontekstu uczenia się gimnazjalistów.
2. Jest wiele czynników warunkujących efektywność kształcenia w gimnazjach większych. Wyniki badań pozwoliły dostrzec następujące relacje między zmiennymi:

- a) przyrost osiągnięć uczniów po trzech latach uczenia się nie jest zależny wyłącznie od ich poziomu wiedzy „na wejściu”,
- b) młodzież wiejska uzyskuje większy przyrost umiejętności z przedmiotów matematyczno-przyrodniczych niż humanistycznych,
- c) istnieje negatywna relacja między selekcją wewnątrzszkolną w przydziale do klas pierwszych a przyrostem osiągnięć gimnazjalistów,
- d) gimnazja mogą niewiele różnić się średnią efektywnością nauczania, ale statystycznie istotnie rezultatami pracy z uczniami o różnym potencjale edukacyjnym „na wejściu” (w szkołach o wysokiej EWD najbardziej przyspieszają rozwój uczniowie „przeciętni” i „silni” u progu gimnazjum).

W niniejszym opracowaniu przedstawiono wybrane czynniki korelujące z efektywnością kształcenia. Ze względu na małą próbę szkół zaprezentowane wyniki nie są reprezentatywne, choć wykazują zbieżność z ogólnopolskimi badaniami EWD, popularyzowanymi w publikacjach ukazujących się w ramach Konferencji Diagnostyki Edukacyjnej. Istotną mocną stroną tych badań jest ich ułożenie na styku danych ogólnopolskich i lokalnych, pozwalające na uwzględnianie nie tylko potrzeb systemu edukacji, ale także – a może przede wszystkim – specyfiki regionu.

Literatura

Denek K. (2005): *Ocenianie szkolne w kontekście określenia efektywności kształcenia zorientowanego na jego jakość*. „Biuletyn Badawczy CKE”, Nr 5, Warszawa.

Dolata R. (red.) (2007): *Edukacyjna wartość dodana jako metoda oceny efektywności nauczania*. CKE, Warszawa.

Kutajczyk T., Przychodzeń B. (2008a): *Diagnozowanie postępu edukacyjnego uczniów w gimnazjum*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczycieli*. PTDE, Kraków.

Kutajczyk T., Przychodzeń B. (2008b): *Uwarunkowania tempa rozwoju uczniów w gimnazjum wiejskim*. W: B. Niemierko, M. K. Szmigel (red.): *Uczenie się i egzamin w oczach nauczycieli*. PTDE, Kraków.

Kutajczyk T., Przychodzeń B. (2008c): *Czynniki skuteczności kształcenia ogólnego w gimnazjach wiejskich. Raport z badania*. OKE Gdańsk, dostępny na: www.oke.gda.pl (otwarty 5 lutego 2014).

Niemierko B. (2004): *Założone i uzyskane, sprawdzające i różnicujące znaczenie treściowe egzaminacyjnej skali pomiarowej*. W: B. Niemierko, H. Szaleniec (red.): *Standardy wymagań i normy testowe w diagnostyce edukacyjnej*. PTDE, Kraków.

Niemierko B. (2009): *Diagnostyka edukacyjna. Podręcznik akademicki*. Wyd. Naukowe PWN, Warszawa.

PRENUMERATA „RUCH” S.A.

1. Wpłaty na prenumeratę przyjmują jednostki kolportażowe „RUCH” S.A. właściwe dla miejsca zamieszkania lub siedziby prenumeratora.
2. Cena prenumeraty ze zleceniem dostawy za granicę jest powiększona o rzeczywiste koszty wysyłki. Zlecenia na prenumeratę dewizową przyjmowane od osób zamieszkałych za granicą realizowane są odo dowolnego numeru. Wpłaty przyjmuje Oddział Krajowej Dystrybucji Prasy „RUCH” S.A. na konto:
Pekao S.A. IV O/Warszawa 681240–10531111–00000–4430–494
lub kasa Oddziału.
3. Informacji o warunkach prenumeraty i sposobie zamawiania udziela „RUCH” S.A. Oddział Krajowej Dystrybucji Prasy, 00–958 Warszawa, skr. pocztowa 12, ul. Jana Kazimierza 31/33
lub telefonicznie: (22) 5328–731, 5328–820, 5328–816, fax: 5328–732,
Internet: www.ruch.pol.pl, e-mail: prenumerata@okdp.ruch.com.pl
4. Terminy przyjmowania wpłat na prenumeratę krajową i zagraniczną:
 - do 5.12 – na I kwartał roku następnego, na I półrocze i na następny rok,
 - do 5.03 – na II kwartał roku bieżącego,
 - do 5.06 – na III kwartał roku bieżącego,
 - do 5.09 – na IV kwartał roku bieżącego.

