

PIOTR BIELECKI

*Szkoła Główna Handlowa
w Warszawie*

**DYKTAT TESTOCENTRYZMU W POLITYCE EDUKACYJNEJ.
REFLEKSJE NA MARGINESIE KSIĄŻKI
MARKA PIOTROWSKIEGO
OD TQM DO „ŻANDARMA”, CZYLI POD PRĄD**

ABSTRACT. Bielecki Piotr, *Dyktat testocentryzmu w polityce edukacyjnej. Refleksje na marginesie książki Marka Piotrowskiego Od TQM do „żandarma”, czyli pod prąd* [The Dictate of Test-orientation in Education Policy. Reflections on Marek Piotrowski's Book *From TQM to "Gendarme", i.e. Against the Mainstream*]. *Studia Edukacyjne* nr 37, 2015, Poznań 2015, pp. 109-134. Adam Mickiewicz University Press. ISBN 978-83-232-2967-4. ISSN 1233-6688. DOI: 10.14746/se.2015.37.8

The book is an adequate and novel response to the continually growing interest both in the effectiveness of the external evaluation of schools and students and in the improved quality of common education. Analyses of the data derived from an empirical research conducted by the Author of the book reviewed do not support the proposition that external standardized tests, given the reality of Poland's schooling system, are appropriate and satisfactory (fair and objective) measures of students' academic achievements. Moreover, a similarly poor effect has been observed in the field of continuous quality improvement. Research evidence provided by the Author reveals the high importance (also social) and complexity of the negative consequences of standardized high-stake testing. Bearing that in mind, the Author of the book addresses the issue of changing the so called "social paradigms" in education policy and schooling, that is a shift from "classic", old-established patterns of teaching and testing to application of principles and tools of TQM. This primary claim constitutes THE correct base for considering current and future numerous initiatives for the sake of common education improvement. Owing to the positive final evaluation of the value of the topic, the major results of the study, the Author's generalizations and conclusions, and the overall scientific contribution, the review notes are generally complimentary. Though largely favorable, the review contains some significant criticisms and items of concern (*shortcomings and omissions*) with regard to the research aims and objectives of the book, its range, contents, structure, style of writing, arguments used, theoretical approach, methods employed, and the literature used. Comments on some of the limitations of the book, especially related to its theoretical background, methodology and content of the study, provided the reviewer with the opportunity for a more critical discussion *highlighting* the topics in question.

Key words: education policy, education quality, TQM, standardized tests

Ujęty metaforycznie tytuł książki¹ może skłaniać także do jego odczytania *à rebours*, a więc rozumienia nawiązującego do głośnego poetyckiego przesłania Zbigniewa Herberta, inspirowanego znaną jezuicką, ascetyczną maksymą „idź pod prąd” (św. Ignacy Loyola). Owa odmienna interpretacja jest zatem swoistą zapowiedzią trudów, z jakimi zmierzyć się musi każdy autor decydujący się podjąć wyzwanie uporania się („działać przeciwie” – *agere contra*) z poglądami przyjmowanymi przez wielu jako stwierdzenia oczywiste i bezsporne. Poglądy te, zakorzenione od dawna w tradycyjnej kulturze edukacyjnej, w dużej mierze zawdzięczają swoją treść i charakter wieloletnim doświadczeniom praktyki szkolnej i dokumentom polityki edukacyjnej prowadzonej przez „żandarma” zapatrzonego w „dyktaturę liczb”. Istotę kluczowego problemu badawczego książki ujawnia najpełniej i najtrafniej przeświadczenie Autora o użyteczności i stosowalności koncepcji zarządzania przez jakość w edukacji jako alternatywy dla omnipotencji pomiaru edukacyjnego (zewnętrznych testów osiągnięć szkolnych), pełniącego rolę podstawowego instrumentu polskiej polityki edukacyjnej na poziomie oświaty powszechnej. Innymi słowy, rozwiązanie prowadzącego w ostatecznym rachunku do wielu niezamierzonych, negatywnych skutków, przejawiających się dysfunkcyjnością oświaty. Rozprawa, jako jedna z niewielu prac krajowych, przybliży problematykę zagrożeń związanych z wprowadzaniem innowacji edukacyjnych w polskim systemie oświaty, analizując w sposób dogłębny rozstrzygnięcia krajowej polityki edukacyjnej zmierzające do poprawy jakości szkół. Świadomość realności tych zagrożeń w kontekście polskim wzmacniają wyniki badań Autora, dokumentujące istnienie rozdźwięku między dominującymi innowacjami w zakresie ewaluacji zewnętrznej, podejmowanymi z inicjatywy podmiotów polityki edukacyjnej (koncepcja „optymalizacji braków” prowadząca do swoistego dyktatu badań ilościowych wyników nauczania – testów i egzaminów) a oddolnym doskonaleniem systemu oświaty (mającym swoje źródło w koncepcji TQM w edukacji). Ten ostatni, bardzo ważny, rodzaj aktywności innowacyjnej odgrywa dziś, ze względu na prymat oddziaływań państwa, drugoplanową rolę, nie pozwalając na skuteczną kooperację szkół, rozwój współpracy nauczycieli i środowiska akademików, a także na pełne wykorzystanie potencjału społecznego nauczycieli, rodziców i środowisk lokalnych.

Dobór podjętej przez Autora tematyki jest trafnym zamierzeniem badawczym nie tylko ze względu na jej doniosłość i wagę społeczną, ale także oryginalność i nowatorstwo na tle współczesnej wiedzy zawartej w piśmiennictwie z zakresu metodologii badań ilościowych nad edukacją, dia-

¹ M. Piotrowski, *Od TQM do „żandarma”, czyli pod prąd*, Warszawa 2013, s. 180.

gnostyki edukacyjnej, krajowej polityki edukacyjnej oraz pedagogiki porównawczej. Książka wypełnia wyraźną lukę w krajowej literaturze przedmiotu tych dyscyplin.

Na jej treść składają się dwie części poprzedzone wprowadzeniem. Pierwsza część pracy (dwa rozdziały) skupia się na charakterystyce teoretycznych i metodycznych zagadnień pomiaru osiągnięć szkolnych uczniów i efektywności pracy szkół w kontekście oceny skuteczności reformy polskiego systemu oświaty. W ocenie tej ustalenia audytorów firmy konsultingowej o międzynarodowej renomie konfrontowane są z wynikami badań własnych Autora książki. Druga część, pojemniejsza treściowo (trzy rozdziały), podejmuje próbę krytycznej oceny skuteczności reform systemu edukacji, opierając się na wynikach badań przebiegu i rezultatów procesu reformowania oświaty w trzech polskich gminach – „gminach laboratoriach”. Autentyczną wartością dodaną w tej części pracy jest przekonująca argumentacja przemawiająca za redefinicją podstawy metodycznej oceny skuteczności polityki edukacyjnej samorządu, której nie mogą stanowić analizy danych statystycznych ilustrujących poziom osiągnięć szkolnych czy wskaźniki zdawalności egzaminów zewnętrznych. Ogólnie rzecz biorąc, w obydwu częściach pracy dominuje podejście empiryczno-analityczne oparte na indukcji jako metodzie postępowania badawczego, odwołującego się do przykładów odzwierciedlających realia polskiej praktyki oświatowej i polityki edukacyjnej.

Większej logicznej spójności tekstu prezentowanego w obydwu rozdziałach, a także eliminacji odczucia przypadkowości fragmentów tekstu, sprzyjałaby większa precyzyjność postawienia problemu we „Wprowadzeniu”. Mam tu na myśli większą odpowiedniość postawionych celów (konceptyjnych i badawczych) oraz układu pracy, która wymuszałaby na Autorze określenie ważności i sekwencji realizacji tych celów ujętych we wspomnianym wprowadzeniu do rozprawy.

Uchybieniem konstrukcji pracy jest spora liczba dygresji wplatanych w zasadniczy tok wywodów, które nie są, moim zdaniem, niezbędne i pasowałyby raczej do tekstu o charakterze popularyzatorskim. Na przykład, odniesienie do zaangażowanej i wielce kontrowersyjnej (jednostronność opcji ideologicznej) publicystyki Boya-Żeleńskiego można było zastąpić przypisem do tekstu zaczerpniętego z naukowego piśmiennictwa pedagogicznego lub socjologicznego.

Zastrzeżenie czytelnika może budzić także zbyt duża liczba kwestii wyróżnianych boldem lub ujmowanych w tzw. boksach (nie zawsze przy tym mających charakter konkluzji). Swoista inflacja wyróżnień staje się niemile widzianą manierą, w szczególności w tekstach podrozdziałów o niewielkiej

objętości. Ponadto, obiektywnie budzi niezbyt staranna redakcja techniczna tekstu (przykładem wielokrotnie spotykany błędny zapis nazwiska prof. Janusza Czapińskiego).

Za ważne, wartościowe oraz inspirujące należy uznać ustalenia i wnioski Autora przedstawione w części I książki. Dotyczą one głównie inherentnych właściwości egzaminów zewnętrznych – ich wad i niedoskonałości metodologicznych, a także negatywnych konsekwencji stosowania tej metody pomiaru osiągnięć szkolnych jako instrumentu polityki edukacyjnej. Na odnotowanie zasługują najważniejsze ustalenia badawcze, tworzące syntetyczny przegląd niedomagań pomiaru osiągnięć szkolnych oraz niepożądanych skutków odgórnie sterowanego procesu reformowania oświaty. Warto zatem wymienić: znaczące opóźnienie we wprowadzeniu oceniania „za postęp” (od 2002 r.); wady prezentacji danych surowych egzaminów zewnętrznych (wartości średnie) oraz wskaźników EWD: brak pełnego obrazu zróżnicowania wyników oraz wpływu uwarunkowań zewnętrznych; pułapki interpretacji wyników egzaminów związane z trudnością precyzyjnego oszacowania realnej efektywności pracy szkoły; nieprecyzyjne określanie wskaźnika EWD jako edukacyjnej, a nie egzaminacyjnej wartości dodanej (niemożność wyeliminowania wpływu czynników niezależnych od szkoły na wyniki ucznia); zbyt częste i woluntarystyczne zmiany podstawy programowej; ignorowanie znaczenia kształcenia umiejętności heurystycznych; negowanie potrzeby nauczania wielopoziomowego (przykład matematyki); negatywne skutki egzaminów zewnętrznych w sferze segregacji uczniów i szkół; błędy w ocenianiu wewnątrzszkolnym skutkujące nasileniem się segregacji wewnętrznej; niedocenianie inicjatyw reformatorskich na szczeblu lokalnym; nieprzemyślane decyzje odnośnie wyboru priorytetowego obszaru reformy (większa skuteczność reformy skoncentrowanej na edukacji przedszkolnej); niewystarczające przygotowanie reformy obniżenia wieku szkolnego; lekceważenie tworzenia kapitału społecznego jako krytycznego czynnika powodzenia reformy; niedocenianie roli coachingu i instytucji pośredniczących.

Na podkreślenie zasługują celne spostrzeżenia Autora związane z krytyczną oceną skutków polskiej reformy oświaty, która nie uporała się z nierównościami edukacyjnymi wyrażającymi się selekcją uczniów i segregacją szkół. Formalna analiza i interpretacja zarówno wyników egzaminów zewnętrznych, jak i wskaźników EWD (nieuwzględnianie zróżnicowania potencjału egzaminacyjnego uczniów, biorącego swój początek we wcześniejszych etapach edukacji), prowadząca do uproszczonej, iluzorycznej i *de facto* umownej kategoryzacji szkół, upoważnia do mówienia o porażce działań

reformatorskich. Stosowanie obydwóch miar skuteczności nauczania na potrzeby oceny pracy szkół przyczyniło się do występowania procesów segregacji i selekcji, a więc do wzrostu nierówności edukacyjnych.

Naszkiecowany przez Autora model alternatywnych rozwiązań polityki edukacyjnej wobec dzisiejszej postaci egzaminów zewnętrznych (podrozdział 5.3: *Walczyć z żandarmem czy odwołać żandarma* – pełniący rolę podsumowania i konkluzji pracy) generalnie rzecz biorąc jest udaną próbą systematyzacji wiedzy o pożądanym kierunkach doskonalenia organizacji i metodyki nauczania. Przemyślenia Autora dowodzą umiejętności krytycznej analizy materiału empirycznego i literatury przedmiotu, a także mają dużą wartość poznawczą. Szkoda, że Autor w swoich wnioskach (patrz s. 162-168) nie uwzględnił w należyтым stopniu potrzeby „nowej wizji oceniania”, czyli propozycji odnoszących się do sposobów oceniania uczniów oraz obiektywności innowacyjnych metod i technik egzaminacyjnych. Mam tu na myśli np. takie rozwiązania, jak przyjęte przez IBO: zapewnienie systemu jednoznacznych i ściśle sformułowanych kryteriów oceny, opisy zróżnicowanych poziomów osiągnięć szkolnych, wewnętrzna i zewnętrzna (recenzenci) ocena prac kursowych (semestralnych/wielosemestralnych), a także lokalnie rozwijane i administrowane systemy oceniania uczniów koncentrujące się np. na prezentacji projektów i prac studenckich, pracach pisemnych uczniów, uczniowskiej dokumentacji nauki własnej (portfolio), technikach oceniania kształtującego. W postulowanych przez Autora rozwiązaniach metodyki nauczania jako alternatywy dla wyników ewaluacji zewnętrznej uczniów i szkół nie wybrzmiewają w wyraźny sposób także dwie inne kwestie. Pierwsza z nich to sugestia uwzględnienia w podstawie programowej zajęć, które roboczo można nazwać „modułem przygotowania ucznia do prowadzenia badań” (analogia do zajęć z zakresu „teorii poznania” w szkołach IBO), obejmujących podstawową wiedzę z zakresu metodyki badań, umiejętność formułowania i analizy problemów badawczych, dobór metod i narzędzi badania, opracowanie i prezentację wyników badania. Druga kwestia dotyczy odwołania do idei samoregulowanego uczenia się, opartego na samokontroli treści, przebiegu i wyników procesu uczenia się, który to sposób nauki własnej powinien być wspierany przez nauczycieli kształtujących u uczniów ich przekonania motywujące.

W głównym nurcie rozważań Autora, skoncentrowanym na analizie swoistych „kosztów i korzyści” stosowania zewnętrznego pomiaru jakości kształcenia, szczególne miejsce zajmują częste odniesienia bibliograficzne i merytoryczne do dwóch ważnych pozycji literaturowych: J.J. Bonstingl

(1999)²; M. Mourshed, Ch. Chijioke, M. Barber (2012)³. Wspomniane książki stały się z jednej strony inspiracją do Jego twórczych przemyśleń i dociekań, z drugiej zaś dały asumpt do krytycznej opinii i podważenia nazbyt optymistycznych i nie zawsze prawidłowych ustaleń badań empirycznych firmy konsultingowej McKinsey & Company, odnoszących się do krajowej polityki edukacyjnej i rezultatów reformy oświaty w Polsce.

Złożoność, wielostronność i interdyscyplinarność podjętej problematyki sprawiły, że Autor zastosował, obok „standardowej” metody studiów literaturowych, zarówno metody ilościowe (dominujące podejście metodyczne), jak też jakościowe, których prawidłowy dobór i wykorzystanie zasługują na pozytywną ocenę. Szkoda tylko, że Autor nie wykazał większej systematyczności w posługiwaniu się metodą historyczno-porównawczą, pozwalającą na dokładną ocenę zmienności/stabilności zjawiska w czasie (interwencje państwa – żandarma, w tym zmiany podstawy programowej, relacja między zakresem regulacji i autonomii szkół od, powiedzmy, 1999 r.).

Z uznaniem należy odnotować fakt, iż Autor dowodząc supremacji kompleksowego zarządzania jakością nad wskaźnikami pomiaru skuteczności nauczania, sięgnął także po metodę benchmarkingu, wskazując cztery przykłady „dobrej praktyki”. Dotyczą one zarówno metodyki nauczania, jak i pomiaru edukacyjnego oraz projektów polityki edukacyjnej. Należą do nich: nowatorskie rozwiązania metodyczne w szkołach objętych systemem matury międzynarodowej (IBO), reforma edukacji w polskiej gminie „Laboratorium” zapoczątkowana w latach 90. XX wieku, której realizacja i ocena były wspierane pracami badawczymi i eksperckimi Autora, a także dwa polskie projekty służące ciągłej poprawie edukacji powszechnej – „Szkoła z klasą” oraz „Akademia uczniowska”.

Dostrzeżone potknięcia Autora książki, niemające przy tym charakteru poważnych warsztatowych uchybień i błędów merytorycznych/terminologicznych, nie wpływają w istotny sposób na całościową, pozytywną ocenę pracy. Po stronie „winien”, obciążającej bilans dokonań autorskich, należy natomiast odnotować szereg zaniechań i niedopowiedzeń. Kwestie pominięte lub nieprzedstawione w wyczerpujący sposób mogłyby być ujęte chociażby w podsumowaniu rozdziałów w formie listy kwestii nierozstrzygniętych czy budzących kontrowersje, a także rekomendacji wskazujących dalsze kierunki badań nad podjętym przez Autora tematem. Umiarkowana obszerność tekstu pracy wskazuje, że wspomniana możliwość swobodnego

² J.J. Bonstingl, *Szkoły jakości. Wprowadzenie do Total Quality Management w edukacji*, Warszawa 1999.

³ M. Mourshed, C. Chijioke, M. Barber, *Jak najlepiej doskonalone systemy szkolne na świecie stają się jeszcze lepsze*, Warszawa 2012.

wypowiedzenia się Autora nie napotkałaby nie tyle sztywnych, co autor-
skich ograniczeń objętości (praca liczy wraz z bibliografią 180 stron).

Fascynacja nową wizją oceny osiągnięć szkolnych i argumentacja Autora
książki na rzecz korzyści płynących z doktryny TQM jako jednego z para-
dygmatów reformy oświatowej polskiego systemu oświaty nie pozbawiona
jest jednak słabości. Wśród nich należy wskazać kilka niedociągnięć. Nie
sposób nie wspomnieć o przyjęciu milczącego założenia o homogeniczności
założeń filozofii TQM w edukacji oraz niewystarczającym uwypukleniu
kwestii zróżnicowania głównych zagrożeń i barier implementacji w odnie-
sieniu do różnych wariantów doktryny TQM. Należy pamiętać, że wybór
zróżnicowanych interpretacji tej doktryny (np. model Ph. Crosby'ego <zero
braków> w opozycji do modelu W.E. Deminga <permanentne doskonalenie
jakości>) skutkuje odmiennymi implikacjami dla polityki edukacyjnej, przy-
bierającymi postać zróżnicowanych interwencji. Na przykład, według
Crosby'ego warunki eliminacji niepowodzeń egzaminacyjnych uczniów
obejmują m.in. zapewnienie instruktażu i coachingu w zakresie technik
zdawania egzaminów, dążenie do nagradzania nauczycieli i studentów
opartego na kryterium osiągnięć szkolnych uczniów, czy podkreślanie wagi
i znaczenia osiągnięcia przez wszystkich uczniów założonych standardów
edukacji. Natomiast, przeciwstawne tezy Deminga postulują m.in. pomniej-
szenie roli egzaminowania końcowego w procesie zapewniania jakości, eli-
minację wszelkich standardów edukacyjnych oraz ilościowych wskaźników
osiągnięć szkolnych, zastąpienie tradycyjnego mistrzostwa (nauczyciela)
nauczania, a także uczenia się „przez wkuwanie” kształceniem zorientowa-
nym na ucznia, współokreślającego efekty kształcenia i potrafiącego oceniać
jakość swojej nauki.

Odczytany przez czytelnika większy stopień dopasowania modelu De-
minga do hipotezy pracy nie oznacza jednakże, że jej czytelnik ma jasność,
który z 14 punktów koncepcji Deminga nie zdaje testu wykonalności
w przypadku zastosowania doktryny TQM w szkołach i innych instytucjach
świadczących usługi edukacyjne. Rodzi się np. pytanie: czy zalecenie De-
minga kwestionujące praktykę odwoływania się do haseł, sloganów i meta-
for w zarządzaniu jakością procesu nauczania i uczenia się jest słuszne. Bo-
wiem, jak wykazuje praktyka edukacyjna, w procesie reformowania
edukacji wykorzystuje się często nośne hasła, jak np. „szkoła ocenia spra-
wiedliwie” czy „<nacabez> - na co będę zwracać uwagę”, które to slogany
Autor stosuje zresztą na potrzeby swoich przemyśleń w recenzowanej pracy.

Lektura wstępnych rozważań Autora w części I pracy skłania do refleksji,
że zabrakło w nich próby syntezy, polegającej na sformułowaniu kilku
twierdzeń natury ogólnej dotyczącej roli filozofii TQM jako teoretycznej

podbudowy reformatorskiej praktyki edukacyjnej oraz wytycznych dla interwencji państwowego regulatora edukacji. Przykładem oczekiwanych generalizacji mogłoby być np. poniższe stwierdzenie. Dostosowanie założeń TQM zrodzonego w sferze produkcji dóbr materialnych do wymagań i działań pro jakościowych w procesie dydaktycznym i polityce edukacyjnej oznacza, że: „produktem” staje się zarówno wykształcenie (kapitał ludzki), jak i kapitał społeczny (czynnik koordynujący działania na rzecz rozwoju i poprawy jakości edukacji); nauczanie wykazuje raczej podobieństwo do zarządzania niż do szczegółowej kontroli czynności ucznia; proces uczenia się nabiera cech właściwych procesowi badań i rozwoju (R + D), a rola ucznia nie ogranicza się wyłącznie do roli konsumenta; uczeń zaczyna także pełnić rolę aktywnego partnera nauczyciela, czyli swoistego „współpracownika” współzarządzającego procesem kształcenia oraz będącego „współproducentem” wykształcenia.

Istotę TQM w edukacji można byłoby także opisać i analizować posługując się szczegółowymi kategoriami nauk o zarządzaniu, takimi jak: rozproszone przywództwo kładące nacisk na przywództwo nauczycielskie, rola kooperacji i pracy zespołowej w procesie podejmowania decyzji, wysoki stopień decentralizacji i delegowania zadań, koncentracja na spełnianiu oczekiwań interesariuszy szkoły (zmiana i ciągle doskonalenie dostosowane do potrzeb interesariuszy), podejmowanie decyzji opartych na analizie danych i myśleniu systemowym (podejście holistyczne), przywództwo edukacyjne odwołujące się do zasad etycznych i społecznej odpowiedzialności szkoły.

Interesującym wątkiem rozważań nad TQM byłoby także omówienie możliwości zastosowania w oświacie jednego ze znanych modeli doskonałości edukacji, jaką jest amerykańska Narodowa Nagroda Jakości im. Malcolma Baldrige'a (*Baldrige Education Criteria*), przyznawana od 1999 roku. Jej wdrożenie, w opinii wielu specjalistów zarządzania oświatą, przynosi nie tylko pożytki prestiżowe i wizerunkowe, ale także realne korzyści, takie jak wzrost jakości kształcenia. Stosowanie kryteriów doskonałości Baldrige'a (m.in. proces projektowania i realizacji nauczania) nie wyklucza korzystania z innych pomocniczych instrumentów (standardów i technik zarządzania), takich jak: ISO 9000, szcuple zarządzanie, sześć sigma, zrównoważona karta wyników, „planowanie poprawy szkoły”.

Pożytecznym dopełnieniem prezentacji koncepcji TQM i modelu Deminga, którego zabrakło w książce, byłoby nawiązanie do osiągnięć myśli pedagogicznej wspierającej podejście jakościowe w edukacji. Na przykład, kierunek i główny nurt przemysłów Deminga zbieżny jest z koncepcją znanego izraelskiego psychologa i pedagoga R. Feuersteina, rozwijaną od wielu

lat w ramach prac Międzynarodowego Centrum Podnoszenia Potencjału Ucznienia się (ICELP)⁴. Zdaniem obydwu badaczy, priorytetowe traktowanie klasyfikacji uczniów na podstawie końcowych egzaminów, podobnie jak innych zewnętrznych bodźców, zapewnia jedynie krótkotrwałe efekty. Nie służy natomiast zwiększeniu zdolności ucznienia się, a także nie jest gwarancją obudzenia potencjalnej pasji i radości z ucznienia się. Zaletą metody ucznienia się Feuersteina, noszącej nazwę „uczenie się przez doświadczenie w mediacji”, jest jej uniwersalne zastosowanie (dzieci z deficytem poznawczym, dzieci utalentowane). Metoda ta, eksponując znaczenie możliwości poznawczych ucznia, roli nauczyciela jako mediatora, a także interakcji dziecko – mediator, ukierunkowana jest na dostosowanie zadań do aktualnych możliwości dziecka i jego potrzeb rozwojowych. Kompetencje rozwijane dzięki tej metodzie obejmują przede wszystkim autonomię i odpowiedzialność ucznia za naukę własną. Również napomknięcie o właściwościach rozkładu Gaussa byłoby dobrą okazją, aby poddać krytyce przekonanie R. Herrnsteina i C. Murraya (autorów *The Bell Curve*)⁵, że zdolności poznawcze uczniów determinowane są głównie przez inteligencję (IQ), atrybut mający w znacznym stopniu charakter dziedziczny oraz posiadający wysoką wartość prognostyczną dla wyników szkolnych i późniejszych losów, jak też osiągnięć życiowych uczniów. Tego rodzaju poglądy stoją w sprzeczności z optymistycznym przekonaniem Feuersteina o modyfikowalności jednostki uzdolnionej do usprawniania procesu ucznienia się.

Warunkiem sukcesu polityki edukacyjnej promującej TQM jako przeciwstawienie dla modelu oświaty opartej na standardach edukacyjnych (SBE) czy modelu podnoszenia ilościowych wskaźników upowszechnienia oświaty (też przedszkolnej) są wysiłki na rzecz wyrównania szans edukacyjnych. Przy czym, w myśl zaleceń amerykańskiego noblisty J.J. Heckmana, wczesne interwencje (w okresie wczesnoszkolnym i przedszkolnym) są o wiele skuteczniejsze dla późniejszych osiągnięć edukacyjnych, zawodowych i życiowych (swoista profilaktyka), niż korekcja w późniejszym okresie (*vide* tzw. równanie Heckmana)⁶.

W obydwu rozdziałach pracy Autor wykazał się sporą erudycją oraz umiejętnością prezentacji licznych i złożonych kwestii, ujęć oraz stanowisk badawczych wobec egzaminów zewnętrznych, ze szczególnym wyeksponowaniem oceny zalet i wad ilościowych metod pomiaru efektywności

⁴ Zob. zasoby internetowe ICELP (Instytutu Feuersteina): <http://www.icelp.info/>

⁵ R.J. Herrnstein, R.C. Murray, *The Bell Curve. Intelligence and Class Structure in American Life*, New York 1994.

⁶ Zob. np. stronę [www prezentującą projekt badawczy noblisty: http://heckmanequation.org/heckman-equation](http://heckmanequation.org/heckman-equation)

kształcenia. Z jednej strony wieloaspektowość obrazu blasków oraz cieni krajowych i międzynarodowych rozwiązań systemowych pomiaru edukacyjnego, będącą naturalną konsekwencją interdyscyplinarnego charakteru wiedzy o ewaluacji osiągnięć szkolnych (konglomerat wiedzy pedagogicznej, statystycznej, socjologicznej, ekonomicznej i psychologicznej), należy uznać za niezaprzeczalny walor metodyczny pracy. Z drugiej strony, wspomnianej wielości kategorii analitycznych współtworzących przedmiot rozprawy oraz związanej z nią wielowątkowości rozważań nie zawsze towarzyszyła dbałość o spójność wywodu, zmuszająca czytelnika do nader uważnej i skrupulatnej lektury tekstu książki. Istotnym uchybieniem tekstu, aczkolwiek nie ciężącym na całościowej ocenie pracy, jest brak głębszej autorskiej interpretacji ustaleń badawczych będących podstawą rekomendacji odnośnie do przydatności poszczególnych kategorii egzaminów zewnętrznych objętych analizą: egzaminów państwowych oraz egzaminów międzynarodowych. Na przykład, w części zamykającej pracę trudno znaleźć wyrażone *explicite* stwierdzenie podważające sens istnienia egzaminów zewnętrznych jako takich, propozycje ukierunkowujące modyfikację tych egzaminów, czy wskazanie określające hierarchię ważności ich funkcji z perspektywy wymagań polityki edukacyjnej zorientowanej na poprawę efektów pracy szkoły i zapewnienie równości edukacyjnej.

We wstępnej części pracy odczuwa się brak krótkiego zarysowania szerszego tła rozważań w sensie podstaw teoretycznych reform „egalitarnej” polityki edukacyjnej, ukierunkowanej równocześnie na wzrost skuteczności i jakości kształcenia oraz respektowanie zasad równości edukacji. Podstawy te stanowią także podbudowę wciąż ewoluującego modelu społecznej odpowiedzialności szkoły ześrodkowanej na uczniu, czyli wzorca będącego przeciwważą dla stosowania testów „wysokiego ryzyka”. Mam tu na myśli np. socjologiczno-filozoficzne koncepcje prezentowane przez J. Deweya (pragmatyzm społeczny, idea szkoły aktywnej, „uczenie się przez działanie”, progresywizm), poglądy C. Rogersa (psychologia humanistyczna, nauczanie skoncentrowane na osobie), czy koncepcje kwestionujące zasadę przymusu szkolnego narzuconą przez państwowego regulatora, ergo ideę państwowych egzaminów zewnętrznych (A.S. Neill). Lepszemu zobrazowaniu problemu służyłoby również przeglądowe przedstawienie poglądów takich przedstawicieli myśli pedagogicznej i społecznej, jak: J.F. Herbart (kształtowanie woli i charakteru oraz budzenie w uczniach szerokich zainteresowań), G. Kerschensteiner (pedagogika kultury, idea szkoły pracy, kształcenie cnót obywatelskich), R.W. Tyler (uczenie się przez działanie), H. Mann (kształtowanie umiejętności społecznych, cnót obywatelskich oraz charakteru uczniów), J. Coleman (decydujący wpływ środowiska społeczne-

go na wyniki szkolne), D. Brandes, P. Ginnis (nauczanie skoncentrowane na uczniu), Ch.E. Silberman (kształcenie alternatywne). Argumentację autor z pewnością również odniesie do głównych pozycji literaturowych ekonomii eksponujących koncepcję całościowego cyklu formacji kapitału ludzkiego (m.in. prace noblisty J.J. Heckmana czy F. Cunha'ya).

W pracy zabrakło, moim zdaniem, także syntetycznego przeglądu ogólnopoznawczych nawiązań i odniesień do pojęć i teorii uzasadniających konieczność interwencji regulatorów (państwa i jego agend sektorowych) jako podmiotów polityki edukacyjnej. Biorąc pod uwagę znaczenie i rolę zewnętrznych testów osiągnięć szkolnych jako instrumentu zarządzania systemem oświaty, mam tu na myśli przede wszystkim teoretyczne podstawy polityki publicznej, takie jak koncepcja Nowego Zarządzania Publicznego (NZP), koncepcja Analizy Polityki Publicznej (*Policy Analysis*), pojmowana w sensie analiz i badań ewaluacyjnych, koncepcja Partycypacji Obywatelskiej – *Civic Engagement* (konsultacja i uzgodnienie polityki publicznej, w tym edukacyjnej ze wszystkimi interesariuszami oświaty powszechnej).

Istotne jest także teoretyczne umocowanie odpowiedzialności szkół i zewnętrznego pomiaru edukacyjnego w ekonomii, którego należy poszukiwać zarówno w teorii agencji (model pryncypała-agenta), jak i teorii kapitału ludzkiego (najważniejszy zasób każdej organizacji oraz czynnik rozwoju gospodarczego). Zgodnie z argumentacją ekonomiczną, w ocenie polityki edukacyjnej i funkcjonowania systemu edukacyjnego pierwszoplanową rolę odgrywa podejście ukierunkowane na rezultaty, a standaryzowane testy dostarczające informacji o poziomie efektów kształcenia urzeczywistniają postulat odpowiedzialności szkół wobec podmiotów polityki edukacyjnej i społeczeństwa jako całości. Egzaminowanie zewnętrzne w myśl powyższych założeń pełni rolę mechanizmu zarządzania stosowanego na potrzeby reform edukacyjnych ukierunkowanych przez cele związane z decentralizacją, wzrostem autonomii szkół oraz wolnym wyborem edukacyjnym. Inny wymiar argumentacji ekonomicznej, przemawiającej za stosowaniem egzaminów zewnętrznych (głównie międzynarodowych), wiąże się z ich rolą jako pośredniej (przybliżonej) metody oceny zasobów kapitału ludzkiego. Podejście oparte na parametrach edukacji stosowane w szacowaniu kapitału ludzkiego wykorzystuje, oprócz takich podstawowych zmiennych, jak alfabetyzacja, skolaryzacja, przeciętny czas kształcenia, także wyniki (głównie międzynarodowych) testów kompetencyjnych (PISA, TIMSS, PIRLS, IALS), traktowane jako (niedoskonałe) miary jakości edukacji. Informacje o efektywności nauczania postrzegane są przez ekonomistów także jako determinanta wyboru edukacyjnego sprzyjającego efektywności alokacyjnej kon-

sumpcji usług edukacyjnych, konkurencji oraz lokalnych (szkoła, samorząd) zachęt stymulujących większą efektywność produkcyjną edukacji szkolnej.

Jeśli chodzi o eksplanacyjną funkcję nauk pedagogicznych, która jest typową funkcją teorii wszystkich, bez wyjątku, dyscyplin społecznych, należy wskazać na brak jednolitej, wyrazistej teorii, której pojęcia i sądy ogólne dostarczałyby racji i argumentów uzasadniających wykorzystanie standaryzowanych egzaminów w praktyce oświatowej. Jak dotychczas, w wyniku braku konsensu, nauki pedagogiczne oferują podejście dychotomiczne, zestawiając dwa przeciwstawne modele testowania zewnętrznego. Istotą pierwszego z nich jest bycie odpowiedzialnym za skuteczność uczenia się i nauczania (*soft accountability model*) opatrzone raczej symbolicznymi sankcjami. Drugi model (*hard accountability model*) zakłada zaś bycie odpowiedzialnym połączone z dotkliwymi konsekwencjami. Egzemplifikacją praktyczną tych modeli jest odmienna polityka testowania zewnętrznego, przyjęta w anglosaskich (głównie w USA i Wlk. Brytanii) i europejskich systemach oświatowych (jak np. Niemcy, Francja, Austria). Według studium N. Mons (2009)⁷, opartego na przeglądzie krajowych i międzynarodowych badań edukacyjnych, różnorodność wyników badań empirycznych nie pozwala na sformułowanie jednoznacznej opinii odnośnie korzyści stosowania standaryzowanych testów, zarówno w warstwie efektywności, jak i równości edukacyjnej. Innymi słowy, związek pomiędzy testowaniem zewnętrznym a poziomem i przyrostem osiągnięć szkolnych nie ma charakteru regularnej zależności przyczynowo-skutkowej.

Wyniki badań własnych Autora, mające znaczący walor poznawczy i dokumentujące przebieg procesów reformowania polskiego systemu edukacji, zyskałyby z pewnością na wartości i klarowności, gdyby towarzyszyła im krótka krytyczna analiza dotychczasowych krajowych osiągnięć teoretycznych i badawczych, w tym prognostycznych w zakresie podjętej problematyki. Jako przykład można by wskazać nawiązania i odniesienia do prac powstałych w ramach Zespołu Badawczego EWD czy prac z zakresu diagnostyki edukacyjnej prezentowanych w ramach corocznych konferencji PTDE. Analogiczną uwagę można odnieść do problemów (wyboru progresywnej/wspólnotowej *versus* neoliberalnej) polityki edukacyjnej *sensu stricto* (np. nawiązanie do raportu Zespołu Doradców Strategicznych Prezesa RM „Polska 2030”)⁸. Szkoda także, że Autor nie pokusił się o wzbogacenie studiów literaturowych i krytycznej oceny autorskiej o pozycje poświęcone teoretycznym odniesieniom problematyki pomiaru dydaktycznego w posta-

⁷ N. Mons, *Theoretical and Real Effects of Standardised Assessment*, Eurydice, 2009.

⁸ *Polska 2030. Wyzwania rozwojowe*, Kancelaria Prezesa Rady Ministrów, Warszawa 2009.

ci egzaminów zewnętrznych, takich jak np. wspomniana wcześniej praca N. Mons (2009). Słowem, zakres wykorzystanej literatury obcojęzycznej może budzić uczucie niedosytu.

Istotnym rozwinięciem treści książki byłyby uwagi zestawiające wady i niedostatki testów zewnętrznych, w tym także egzaminów typu *high-stake* (określanych jako egzaminy doniosłe) z alternatywnymi propozycjami rozwiązań dydaktyczno-organizacyjnych dla praktyki szkolnej. Na przykład, stwierdzeniu, że wyniki testów są jedynie częściowym odzwierciedleniem postępów ucznia powinno towarzyszyć wskazanie konkretnej alternatywy. Jest nią np. sięgnięcie do wielu źródeł wiedzy o tendencjach zmian osiągnięć szkolnych, a także metod ich oceny oraz podkreślenie pierwszoplanowej roli nauczyciela mającego wiedzę o uczniu opartą na własnym doświadczeniu. Równie ważna jest ocena szerokiego spektrum umiejętności uczniów, obejmującego także atrybuty charakteru, określane też jako umiejętności psychologiczno-społeczne, a także tworzenie przez szkoły inwentarza tego typu umiejętności. Według noblisty J.J. Heckmana – badacza zajmującego się ekonomią rozwoju człowieka, tradycyjnie pojmowane cechy osobowościowe (charakter) należałoby traktować jako umiejętności, które można rozwijać i kształtować w szkole, a także poddawać pomiarowi („kapitał ludzki to o wiele więcej niż iloraz inteligencji mierzony w testach”)⁹. Testy, jak dotąd, nie mierzą umiejętności afektywnych, które wyjaśniają sporą część zmienności wyników testów. Ich zakres można określić posługując się np. powszechnie akceptowaną przez psychologów osobowości taksonomią, obejmującą tzw. wielką piątkę. Należą do niej: otwartość na doświadczenia, sumiennność, ekstrawertyzm, przychylnie nastawienie, zrównoważenie emocjonalne. Swojego rodzaju substytutem pomiaru kompetencji emocjonalnych czy tzw. ilorazu emocjonalności (EI) – kategorii analitycznej psychologii zarządzania, stosowanym w polskiej praktyce edukacyjnej – jest element oceny opisowej, określanej mianem dojrzałości szkolnej (poszanowanie własności, przygotowanie do lekcji, punktualność, samodzielność, wytrwałość, umiejętność koncentracji na czynności).

Rodzi się zatem pytanie: czy testy zewnętrzne mierzą w sposób wystarczający transferowalne umiejętności poznawcze, takie jak krytycyzm myślenia, analiza zjawisk i procesów w kategoriach przyczynowo-skutkowych, samodzielne wyciąganie wniosków w sytuacjach nietypowych i problemowych, znajomość głównych koncepcji oraz idei nauczanego przedmiotu, jak też analiza i interpretacja danych.

⁹ J.J. Heckman, T. Kautz, *Fostering and Measuring Skills: Interventions That Improve Character and Cognition*, „IZA Discussion Paper”, 7750, November 2013, Forschungsinstitut zur Zukunft der Arbeit (IZA).

Negatywne konsekwencje stosowania egzaminów zewnętrznych ujawniają się także w zmienionym zachowaniu nauczycieli, którzy mogą zacząć traktować w odmienny sposób sklasyfikowane przez siebie trzy grupy uczniów: „przypadki bezpieczne” (uczniowie z najlepszymi wynikami), osoby nadające się do pedagogicznej „terapii” (najliczniejsza grupa uczniów o średnich wynikach) oraz „przypadki beznadziejne” (uczniowie o najgorszych wynikach). Wspomniana segregacja uczniów powoduje, że oddziaływania dydaktyczne nauczyciela będą najpewniej koncentrować się na najliczniejszej grupie uczniów kosztem wyników najsłabszych uczniów, a także rozwoju zdolności uczniów utalentowanych. Stosowanie testów i tzw. tabel ligowych szkół może osłabić równość w edukacji, ponieważ nauczyciele mają tendencję do zajmowania się uczniami mogącymi podwyższyć pozycję szkół w rankingu (uczniowie na pograniczu zdania i „oblania” testu, uczniowie przeciętni). Natomiast, uczniowie o specjalnych potrzebach edukacyjnych mogą nie otrzymywać podobnej pomocy (jako wymagający więcej zabiegów i zasobów oraz osiągający słabe wyniki egzaminacyjne obniżają ocenę rezultatów pracy szkoły).

W analizie „poziomu wykonania” szeroko rozumianego zadania ucznia (którym mogą być interesujące nas wyniki kształcenia oceniane przez nauczyciela bądź wyniki zewnętrznych testów umiejętności) należałoby odnieść się do kwestii opisu kluczowych uwarunkowań objaśniających poziom osiągnięć szkolnych, takich jak bodźce (w tym także konsekwencje testu), wysiłek (starania) ucznia oraz dwie podstawowe kategorie umiejętności: umiejętności poznawcze i często pomijane umiejętności charakterologiczne.

Reasumując, lektura pracy nasuwa spostrzeżenie o nie dość mocnym zaakcentowaniu i usystematyzowaniu kosztów pedagogicznych polityki testowania, czyli zagrożeń i niezamierzonych konsekwencji stosowania testów zewnętrznych dla pracy nauczyciela i funkcjonowania szkół. Wśród nich należy wymienić: wyłączenie z pomiaru ważnych cech charakteru (powtórzmy za J.J. Heckmanem – umiejętności charakterologicznych, które można również, jego zdaniem, w miarę obiektywnie zmierzyć i ocenić); profilowanie nauczania „pod testy”; zawężenie treści nauczania; eksponowanie umiejętności i technik zdawania egzaminów; priorytet dla nauczania umiejętności niższego rzędu kosztem kompetencji złożonych i umiejętności rozwiązywania problemów; sprawdzanie jedynie wiedzy deklaratywnej, redukcję zainteresowań poznawczych uczniów; ograniczenie motywacji nauczycieli i uczniów; stres wywoływany u uczniów; zbyt duży nakład czasu poświęconego rozpoznawaniu przez uczniów dystraktorów, czyli odpowiedzi błędnych w testach.

Kategoria „kapitału rodzinnego”, jako czynnik różnicowania wyników egzaminów zewnętrznych, wykorzystywana w pracy na potrzeby ilościowych analiz osiągnięć szkolnych, jest *de facto* wskaźnikiem poziomu wykształcenia rodziców ucznia (wybór Autora podyktowany dostępnością danych). Innymi słowy, jest uproszczoną miarą oddziaływania czynników społecznych na wyniki uczniów. W książce zabrakło zapoznania czytelnika z bardziej miarodajnymi wskaźnikami kapitału społecznego i kulturowego rodziny, chociażby takimi jak indeks statusu ekonomicznego, społecznego i kulturowego rodziny (ESCS), stosowany w badaniach programu PISA (PISA – *index of economic, social and cultural status*).

Indeks ten agreguje cztery szczegółowe czynniki „zewnętrznego oddziaływania” na wyniki uczniów: status zawodowy rodziców, poziom wykształcenia rodziców mierzony liczbą lat nauki, oraz trzy częściowe indeksy PISA: indeks zasobności materialnej rodziny, indeks zasobów edukacyjnych rodziny, indeks zasobów „klasycznej” kultury w rodzinie¹⁰.

Analiza ważnego czynnika wyników szkolnych, w postaci potencjału rodzinnego i społecznego uczniów, mierzonego poziomem wykształcenia rodziców, wskazuje w sposób jednoznaczny na jedną z zasadniczych przyczyn nierówności edukacyjnych wyrażających się znacznym zróżnicowaniem osiągnięć szkolnych. Należy jednak pamiętać o innym czynniku, nie zawsze nieodłącznie związanym z edukacją rodziców czy ich poziomem zamożności. Czynnikiem tym, pomijanym częstokroć w analizie z uwagi na niedostępność lub nieprecyzyjność danych, jest niewątpliwie jakość wychowywania (dbałość o potomstwo), której miernikami jest stymulowanie osiągnięć szkolnych, wspieranie dziecka, doradzanie, okazywanie przywiązania, czy wreszcie wspólne posiłki i właściwe żywienie. Szczególna rola i obowiązki w tej mierze, czyli we współtworzeniu kapitału ludzkiego, przypadają matkom, o czym np. pisał już dawno jeden z najwybitniejszych ekonomistów brytyjskich – Alfred Marshall. System szkolny napotykać wielkie trudności i ponosząc równie niemałe koszty, może zaledwie niwelować środowiskowe zaniedbania powstałe w okresie wczesnego dzieciństwa, które wynikają z niewystarczającej opieki rodzicielskiej. Również zmiany tradycyjnego modelu rodziny, a w szczególności jej funkcji socjalizacyjnej i emocjonalnej (wzrastający udział rodzin niepełnych kosztem udziału rodzin monogamicznych) wpływają na zwiększenie się deficytu umiejętności poznawczych i kompetencji osobowościowych dzieci (ergo nierówności edukacyjnych). Badania analizujące częstość stymulacji poznawczej oraz wsparcia emocjonalnego dziecka w przekroju typu rodziny (pełna, rozbita,

¹⁰ *Glossary of Statistical Terms*, OECD, 2008.

niezamężne samotne matki) potwierdzają, że najwyższy poziom tego rodzaju swoistej „inwestycji” występuje w pełnych, tradycyjnych rodzinach, najniższy zaś w ostatniej kategorii: niepełnej rodzinie¹¹.

Ewaluacja zewnętrzna, której przedmiotem są przede wszystkim wyniki sprawdzianów i egzaminów zewnętrznych oraz inne efekty pracy szkoły, może zdominować znaczenie autoewaluacji szkolnej, pomniejszając jej rolę i wskazując na rzekomy brak obiektywizmu ewaluacji wewnętrznej oraz niechęć do rzeczywistej lustracji szkoły ujawniającej jej dobre i złe strony. Sytuacja taka może skutkować umniejszeniem autonomii szkolnej, premiovaniem nauczania zorientowanego na umiejętność rozwiązywania testów, brakiem innowacyjności i kreatywności nauczycieli i uczniów, niezadowalającym poziomem jakości kształcenia.

Argumentacja wspierająca tezę o zagrożeniach związanych ze stosowaniem standaryzowanych testów zewnętrznych nie może także abstrahować od kwestii dwuznaczności i nieprzewidywalności związku między stosowaniem testów a skutecznością nauczania (poziomem osiągnięć szkolnych) i redukcją nierówności edukacyjnych. Pełna i obiektywna ocena funkcji ewaluacyjnej stosowanych w kraju standaryzowanych egzaminów zewnętrznych (sprawdzian szóstoklasisty, egzaminy: gimnazjalny i maturalny), czyli wykorzystania wyników testów na potrzeby ewaluacji wewnątrzszkolnej i zewnętrznej, wymagałaby zatem weryfikacji hipotezy o wpływie testów zewnętrznych na osiągnięcia szkolne uczniów. Wspomniany wpływ można rozumieć w dwojaki sposób. Po pierwsze, jako hipotezę dotyczącą skutków testocentryzmu, czyli nauczania „pod testy”. Po drugie zaś, co ważniejsze, jako hipotezę zakładającą związek przyczynowy między stosowaniem testów zewnętrznych (zmienna niezależna) a poziomem osiągnięć szkolnych (zmienna zależna). Nawiasem mówiąc, sprawdzenie tych hipotez wykracza poza zakres celów, jakie postawił sobie Autor książki.

Budząca wiele kontrowersyjnych ocen i spekulacji teza o pozytywnym związku między wzrostem osiągnięć poznawczych i łagodzeniem nierówności edukacyjnych a stosowaniem standaryzowanych egzaminów zewnętrznych skłania do szerszego ujęcia tej współzależności, którego zabrakło w książce. Z pewnością interesujących ustaleń na temat rzetelności i trafności egzaminów zewnętrznych dostarczyłaby analiza długookresowych tendencji zmian (od 2000 r.) poziomu osiągnięć uczniów polskich szkół, wykorzystująca ustalenia badań PISA oraz wyniki krajowych egzaminów zewnętrznych (ogólnokrajowe dane PWE), a także testów OBUT. Rów-

¹¹ J.J. Heckman, *The Economics of Inequality. The Value of Early Childhood Education*, American Educator, Spring 2011.

nie pouczających wniosków należałoby oczekiwać od analizy porównawczej współzależności/rozbieżności tendencji wyników egzaminów krajowych oraz egzaminów międzynarodowych w określonym przedziale czasu. Rygory rzetelności naukowej nakazywałyby, oczywiście, rozpatrzenie kwestii wpływu czynników egzogenicznych na jakość kształcenia (m.in. status społeczno-ekonomiczny rodziny, wskaźniki ilustrujące interwencje polityki edukacyjnej, a nawet miary decentralizacji zarządzania oświatą, która może także negatywnie wpływać na poziom osiągnięć uczniów).

Poruszona w książce kwestia międzynarodowego benchmarkingu (testy PISA, TIMSS) abstrahuje np. od oceny przydatności wyników testu PISA dla polityk edukacyjnych krajów ujętych w omawianym rankingu osiągnięć edukacyjnych¹². Krytycy tego rankingu wskazują, iż wyniki oceny biegłości uczniów w zakresie czytania, matematyki i nauk przyrodniczych są raczej odzworowaniem skumulowanych doświadczeń edukacyjnych ucznia (szkolnych i domowych) na przestrzeni wczesnego dzieciństwa oraz wieku 15 lat, niż rezultatem oddziaływań pedagogicznych szkół danego systemu edukacyjnego. Stąd, mówienie o ewentualnych implikacjach PISA dla działań dotyczących sposobów doskonalenia systemu szkolnego nie są uprawnione. Ponadto, wiarygodność rankingu krajowych systemów edukacji budzi wątpliwości ze względu na niepewność pomiarową. W grę wchodzi m.in.: operowanie danymi średnimi w tablicach wynikowych PISA, a także zróżnicowane liczebności uczniów odpowiadających na niejednorodne zestawy pytań skutkujące znacznym zróżnicowaniem rankingu pozycji poszczególnych krajów. A zatem, prezentacja dużej części danych, poprzedzona skalowaniem wyników mającym na celu uwzględnienie błędów pomiaru (*plausible values*), oparta jest na liczbach losowych. Krytykę wiarygodności wyników badania PISA wzmacniają także wnioski z porównania tendencji zmian wyników badania OECD (przykład pogorszenia wyników umiejętności matematycznych w Anglii, dekada lat 2000) z odmiennymi trendami zmian potwierdzonymi przez inne standaryzowane testy międzynarodowe (przykład badania TIMSS, Anglia, dekada lat 2000, tendencja wzrostowa wyników).

Pozytywna ocena testów OBUT wystawiona przez Autora wspomnianemu instrumentowi badającemu poziom osiągnięć w zakresie umiejętności matematycznych i umiejętności językowych (odbioru tekstów w języku polskim) skłania do postawienia pytania o Jego stanowisko w kwestii celowości stosowania innych dobrowolnych testów mających *stricte* badawczy charak-

¹² Zob. np. B. Froese-Germain, *The OECD, PISA and the Impacts on Educational Policy*, September 2010; P. Mortimore, *Alternative Models for Analysing and Representing Countries' Performance in PISA*, Brussels 2009.

ter. Należy żałować, że w rozprawie nie wspomniano o amerykańskich testach NAEP czy krajowych badaniach diachronicznych, badających ścieżki rozwoju edukacyjnego uczniów na danym etapie nauki/w danej szkole (przykład badania szkół ponadgimnazjalnych, IFiS PAN, 2009-2013). Badania tego typu umożliwiają nie tylko rejestrację przyrostu wiedzy (podobnie jak EWD), ale także, co jest ich dodatkowym atutem, pozwalają poznać przyczyny różnicowań wyników uczenia się w przekroju trzech uwarunkowań: indywidualnych (geny i uzdolnienia), szkolnych (wkład szkoły) oraz środowiskowych (kapitał kulturowy i społeczny).

Doświadczenia sektorowej instytucji amerykańskiej (NCES)¹³, specjalizującej się m.in. w diagnostyce edukacyjnej na poziomie szkół podstawowych i średnich oraz odpowiedzialnej za administrowanie ogólnokrajowymi testami NAEP¹⁴, wskazują, że jej testy mogą być postrzegane jako alternatywa dla obowiązkowych standaryzowanych testów zewnętrznych „wysokiego ryzyka” obowiązujących ogół uczniów. Testy NAEP uwzględniają, obok najczęściej testowanych dyscyplin, takich jak czytanie, matematyka, nauki przyrodnicze, również szereg innych przedmiotów objętych podstawą programową (np. ekonomię i geografę). Obszar tematyczny wiedzy i kompetencji poznawczych weryfikowanych przez testy NAEP (badania oparte na próbie krajowej) świadczy o znacznym potencjale pomiaru dydaktycznego za pośrednictwem wspomnianego instrumentu. Testy NAEP badają poziom osiągnięć m.in. w zakresie następujących umiejętności wyższego rzędu: lokalizacja i wyszukiwanie informacji; integracja i interpretacja pojęć, idei i koncepcji; rozumowanie indukcyjne i dedukcyjne; znajomość i umiejętność stosowania 5 obszarów wiedzy matematycznej w przekroju 3 poziomów złożoności tej wiedzy; opis i objaśnienie relacji między informacją a koncepcjami ekonomicznymi; stosowanie informacji i koncepcji ekonomicznych na potrzeby oceny i rozwiązywania problemów oraz interpretacji sytuacji; identyfikacja i zastosowanie zasad wiedzy przyrodniczej; realizacja badania naukowego; stosowanie projektów technicznych (inżynierskich); krytyczne myślenie i ocena tekstów.

Krytyczne stanowisko oponentów testów „wysokiego ryzyka” wspiera m.in. rekomendacja brytyjskiego renomowanego Instytutu Badań Polityki Publicznej (IPPR). Według IPPR, z uwagi na niedostateczną skuteczność testów zewnętrznych mierzoną przyrostem osiągnięć szkolnych, polityka edukacyjna powinna skłaniać się ku opcji oceniania wewnątrzszkolnego,

¹³ Zasoby internetowe National Center for Education Statistics (NCES): <https://nces.ed.gov/nationsreportcard/>

¹⁴ NAEP – National Assessment of Educational Progress.

a ewaluacyjną funkcję w odniesieniu do szkół zamiast testów obowiązkowych powinny pełnić testy oparte na losowym doborze uczniów.

Zdecydowanie negatywne postrzeganie ewaluacyjnej roli egzaminów zewnętrznych jako narzędzia oceny pracy uczniów, szkół i rejonów szkolnych cechuje np. systemy edukacyjne Finlandii, Szkocji, Nowej Zelandii, częściowo Walii. Gestorzy tych systemów rezygnując z ogólnokrajowych zewnętrznych egzaminów państwowych (z wyjątkiem egzaminu maturalnego), wybrali opcję polityki edukacyjnej przenoszącą w pełni ciężar społecznej odpowiedzialności szkoły na placówki oświatowe, nauczycielstwo i środowisko lokalne.

Wbrew oczekiwaniom, w pracy nie sięgnięto po przykłady miarodajnych, krytycznych opinii o testach zewnętrznych, których autorzy wskazują na nieuprawnione wyciąganie wniosków z wyników testów, skutkujące błędnymi decyzjami dotyczącymi procesu nauczania i uczenia się. Na przykład, źródłem tych opinii mogłyby być wyniki szeroko zakrojonych (25 stanów) i często przytaczanych badań amerykańskich, przeprowadzonych przez Uniwersytet Stanowy w Arizonie¹⁵. Dowodzą one, że stanowe testy zewnętrzne nie tylko pogarszają wskaźniki sprawności kształcenia, a raczej utrudniają uzyskanie dobrych wyników szkolnych, niż przyczyniają się do ich poprawy. Z drugiej strony zabrakło też przykładów badań przynoszących odmienne rezultaty, jak np. wyniki tzw. naturalnego eksperymentu walijskiego¹⁶. Impulsem do jego przeprowadzenia była likwidacja ważnego elementu odpowiedzialności („rozliczalności”) walijskich średnich szkół, a mianowicie zaprzestanie w 2001 r. publikacji tabel z wynikami pracy szkół. Okoliczność ta umożliwiła analizę porównawczą osiągnięć uczniów Walii z osiągnięciami uczniów szkół angielskich, których wyniki nadal były publikowane w „tabelach ligowych” szkół (egzamin końcowy). Według zespołu badaczy z Uniwersytetu w Bristolu, skutkiem reformy walijskiej był znaczący spadek efektywności pracy szkół, wyrażający się obniżeniem średniego poziomu osiągnięć szkolnych i wzrostem nierówności edukacyjnych.

Autor recenzowanej książki posługując się podstawowymi miarami statystyki opisowej na potrzeby analizy struktury zmiennych charakteryzujących osiągnięcia szkolne (miary położenia, zróżnicowania oraz hierarchii ważności zmiennych), wskazuje na użyteczność różnych skal pomiarowych w diagnostyce edukacyjnej w zależności od właściwości zmiennych oraz

¹⁵ A.L. Amrein, D.C. Berliner, *The Impact of High-Stakes Tests on Student Academic Performance: An Analysis of NAEP Results in States With High-Stakes Tests and ACT, SAT, and AP Test Results in States With High School Graduation Exams*, Arizona State University, 2002.

¹⁶ S. Burgess, D. Wilson, J. Worth, *A Natural Experiment in School Accountability: The Impact of School Performance Information on Pupil Progress and Sorting*, „Working Paper”, 2010, 10/246.

celu prezentacji. Czytelnik odczuwa jednak pewien niedosyt, a mianowicie brak reasumpcji w postaci usystematyzowanej wiedzy (wraz z przypisami literaturowymi) na temat uwarunkowań doboru stosowanej skali pomiarowej, czy też metod zrównywania wyników testów (na różnych etapach kształcenia) na wspólnej skali. Utrudnieniem lektury rozprawy, w przypadku czytelnika nieznanego rodzajów i procedur pomiaru dydaktycznego, w tym np. zasad normalizacji i standaryzacji testu, zasad porównywania norm wymagań (wymagań programowych/standardów egzaminacyjnych) i norm empirycznych, jest brak syntetycznego przedstawienia miar pozycyjnych osiągnięć uczniów w badanej grupie i miar rozrzutu wyników (w szkole, dzielnicy, gminie, populacji uczniowskiej ogółem). Wspomniana luka zmusza czytelnika do sięgnięcia po materiały metodologiczne publikowane przez regionalne Komisje Egzaminacyjne lub zespół badawczy realizujący Projekt EWD.

Odnosnie do kwestii sposobów prezentacji wyników sprawdzianów egzaminów zewnętrznych, które przygotowywane są także z myślą o ich upublicznieniu, nasuwa się pytanie o stanowisko Autora w kwestii minimalizacji zakresu danych i zwiększenia ich przejrzystości. Niewątpliwym kosztem takiej strategii upubliczniania wyników egzaminów w poszczególnych szkołach, dzielnicach i gminach byłaby ich ograniczona interpretowalność i wartość poznawcza dla specjalistów, dyrektorów szkół i organów nadzoru pedagogicznego. Z drugiej strony, większa przystępność prezentowanych danych byłaby rekompensowana przez ich większą klarowność i komunikatywność dla rodziców uczniów i kandydatów do szkół. Jednym z punktów odniesienia dla prezentacji wyników testów może być praktyka przyjęta w amerykańskich raportach NAEP – omawiających wyniki przedmiotowych testów „niskiego ryzyka”. Zamieszczony w nich zestaw informacji obejmuje trzy pozycje: średnie wyniki testu wyrażone w punktach (typowa skala: 0-300 pkt.); procentowy udział uczniów osiągających wyniki na poziomie: poniżej podstawowego, podstawowym, biegłości, zaawansowanym (3 poziomy osiągnięć); wartości wyników wyrażone w punktach, prezentujące osiągnięcia uczniów za pomocą pięciu centyli (najniższe – centyl 10. i 25.), średnie – (centyl 50.) i najwyższe (centyl 75. i 90.) Ostatni sposób prezentacji wyników testu umożliwia analizę tendencji zmian osiągnięć szkolnych na przestrzeni lat.

Słuszny postulat wprowadzenia miary przyrostowej jako podstawowego instrumentu oceny jakości pracy szkół warto byłoby zilustrować przykładem zagranicznej koncepcji lub praktyki zewnętrznego pomiaru przyrostu wiedzy i umiejętności ucznia. Nie zawsze realny kształt tych koncepcji spotyka się z akceptacją użytkowników (nauczycielstwo). Interesującym źród-

dłem wiedzy w tej mierze mogą być np. standardy progowe oraz podnoszone co roku standardy postępu (*floor standards*), opracowywane przez centralnego regulatora – Ministerstwo Edukacji Wlk. Brytanii¹⁷, które są obecnie przedmiotem głębokiej krytyki ze strony brytyjskich związków zawodowych nauczycieli. Mamy tu na myśli zarówno udziały procentowe uczniów osiągających oczekiwany poziom wyników na koniec etapu kształcenia zwieńczonego testem SAT lub egzaminem GCSE (matematyka, czytanie, pisanie lub język angielski), jak i wskaźniki postępu mierzące udział procentowy uczniów w danej szkole (nie mniejszy niż wartość środkowa udziału procentowego uczniów) osiągających oczekiwany postęp na koniec odpowiedniego etapu kształcenia (punktem odniesienia jest przyjęta dla każdego roku mediana tego udziału). Dezaprobatę krytyków rozwiązania brytyjskiego budzi arbitralność wyznaczania standardów progowych (brak uzasadnienia), całkowicie dowolne założenie ustawicznej poprawy osiągnięć szkolnych (ich swoista linearność i jednolitość tempa postępów w nauce), błędne przekonanie o możliwości osiągnięcia „przeciętnych” wyników przez prawie wszystkich uczniów. Standardy progowe, pozwalające na identyfikację słabych szkół, stanowią dogodny instrument uzasadniający podejmowanie różnego rodzaju interwencji wobec źle pracujących szkół: m.in. napomnień i inspekcji nadzoru pedagogicznego, włącznie z wymianą dyrektora lub likwidacją szkoły i nadaniem jej innego statusu. Ta ostatnia sankcja może być także rozumiana jako otwarcie możliwości prywatyzacji. Bardziej zdecydowane i dalej idące reakcje władz oświatowych, oparte na analizie słabych wyników testów „wysokiego ryzyka”, występują w amerykańskim systemie oświaty (*vide* ustawa „Zadbajmy o każde dziecko”, *No Child Left Behind* – NCLB, 2002).

Jednym z alternatywnych rozwiązań wobec swoistej dyktatury testów zewnętrznych jest integracja oceniania formatywnego z codzienną praktyką szkolną. Uwagi Autora wskazujące na ograniczoną użyteczność oceniania kształtującego w reformowaniu edukacji szkolnej (ograniczony zakres przedmiotowy ocenianych efektów kształcenia; trudności z pomiarem skuteczności tego typu oceniania) nie są przekonujące. O możliwościach poprawy jakości funkcjonowania szkół, jakie niesie ze sobą stosowanie oceniania formatywnego świadczy m.in. raport OECD oparty na badaniach empirycznych wśród europejskich szkół średnich (*Formative Assessment...*, 2005)¹⁸. Badania te dowodzą stosunkowo wysokiej skuteczności oceniania formatywnego, przejawiającej się wzrostem osiągnięć szkolnych, łagodze-

¹⁷ Zob. zasoby internetowe Department for Education: <http://www.education.gov.uk/schools/performance/documents.html>

¹⁸ *Formative Assessment. Improving Learning in Secondary Classrooms*, OECD 2005.

niem zróżnicowania wyników w przekroju poszczególnych szkół (nierówności edukacyjnych) oraz poprawą umiejętności uczenia się. Ponadto, doświadczenia stosowania tzw. technik oceniania formatywnego (*classroom assessment techniques* – tzw. techniki CAT) w koledżach i uczelniach krajów anglosaskich upoważniają do stwierdzenia o szerokim zakresie ocenianych efektów uczenia się, nie zawężonym jedynie do prostych kompetencji kognitywnych uczniów, takich jak zapamiętanie i rozumienie uprzednio zdobytej wiedzy. Szeroki wachlarz oceny efektów kształcenia weryfikowanych za pomocą technik CAT obejmuje na przykład: umiejętność analizy i krytycznego myślenia, umiejętność syntezy i twórczego myślenia, umiejętność rozwiązywania problemów, umiejętność zastosowania wiedzy i wykonawstwa, świadomość postaw i wartości, samoświadomość uczniów, umiejętność i strategię uczenia się (m.in. T.A. Angelo, K.P. Cross, 1993)¹⁹.

Metodyka nauczania w oświacie powszechnej, wykorzystująca techniki oceniania formatywnego, rzadko sięga po metody heurystyczne, metody zawierające elementy heurystyki, czy proste wskazówki heurystyczne (np. „bierz pod uwagę zdanie innych”) pomocne w procesie tworzenia i odkrywania. Szkoda, że w ocenianej rozprawie nie uwzględniono związłego przeglądu i omówienia zastosowań najpopularniejszych metod twórczego myślenia, które mogłyby zainspirować nauczycieli oświaty obowiązkowej. Wśród nich na uwagę zasługują metody określane w naukach o zarządzaniu mianem technik organizatorskich, jak: burza mózgów (A.F. Osborn), metoda synektyczna (synektyka J.J. Gordona), metoda delficka, analiza morfologiczna, techniki CERMA (technika adwokata, technika ahistoryczna, technika małych zmian, technika modeli analogicznych, analogia historyczna), wskazówki heurystyczne aktywizujące myślenie kreatywne.

Wspominanej wielokrotnie w książce kwestii „warstwy pośredniczącej”, której roli i znaczeniu poświęca wiele miejsca raport McKinsey & Co., towarzyszy krytyczny komentarz Autora, wskazujący na brak należytej współpracy między szkołami oraz instytucjami odpowiedzialnymi za udzielanie wsparcia placówkom światy, a także na słabości współdziałania szkół ze środowiskiem akademickim. Konstatacje Autora warto byłoby uzupełnić uwagami na temat koncepcji szeroko rozumianego kompleksowego systemu wspomagania pracy szkoły, który obejmowałby np. tworzenie (problemowych i przedmiotowych) sieci współpracy oraz samokształcenia dyrektorów i nauczycieli szkół, pomoc psychologiczno-pedagogiczną oferowaną uczniom, rodzicom i nauczycielom, pomoc w zakresie doradztwa edukacyj-

¹⁹ T.A. Angelo, K.P. Cross, *Classroom Assessment Techniques: A Handbook for College Teachers*, San Francisco 1993.

no-zawodowego dla uczniów, wsparcie we wdrażaniu nowych rozwiązań metodycznych do praktyki szkolnej, pomoc w budowie programu współpracy szkoły ze środowiskiem lokalnym. Istotnym problemem są znaczne różnice w dostępności do instytucji wspierających szkoły, mające swoje źródło w międzyregionalnym zróżnicowaniu zasobów kapitału intelektualnego. Warto też pamiętać o potencjalnym zagrożeniu ze strony ponownej interwencji „żandarma”, gdyby autorstwo założeń programowych takiego systemu zostało przekazane instytucjom związanym ściśle z centralnym regulatorem systemu edukacji.

Podany przez Autora przykład działalności polskiego banku centralnego (NBP) jako „instytucji pośredniczącej”, działającej na rzecz popularyzacji wiedzy i świadomości ekonomicznej wśród dzieci (a także ich rodziców) nie jest przykładem odosobnionym. Dobrych wzorców praktycznych rozwiązań w omawianej kwestii dostarcza także aktywność polskich szkół wyższych na polu upowszechniania wiedzy ekonomicznej. Na przykład, od wielu lat grupa siedmiu polskich uczelni oraz NBP (jako partner strategiczny) są realizatorami dwóch inicjatyw międzyuczelnianej Fundacji FPAKE, takich jak Ekonomiczny Uniwersytet Dziecięcy (uczniowie klas 5. i 6.) oraz Akademia Młodego Ekonomisty (uczniowie gimnazjalni).

Integralnym elementem składowym tzw. warstwy pośredniczącej, o którym nie wspomina recenzowana praca, powinny być także programy wspierania uzdolnień w zakresie poszczególnych grup przedmiotów oraz funkcjonowanie międzyszkolnych kółek, jak również konkursów i olimpiad przedmiotowych, rozwijających wiedzę i umiejętności wykraczające poza podstawę programową. Programy tego typu, przeznaczone zarówno dla uczniów szkół podstawowych, gimnazjów oraz szkół ponadgimnazjalnych, realizowane są w kraju przez szkoły wyższe i powoływane przez nie fundacje edukacji przedmiotowej. Interesującym przykładem wspomnianych interwencji dokonywanych przez wrocławskie środowisko akademickie jest rozpoczęty w 2011 r. projekt: „Szlifowanie diamentów – innowacyjne programy wsparcia uczniów uzdolnionych w zakresie nauk matematycznych i przyrodniczych”. Projekt ten obejmuje obok głównego zadania – rozwijania talentów dzieci i młodzieży – także działania w zakresie wsparcia psychologiczno-pedagogicznego uczniów uzdolnionych, ich rodzin i nauczycieli. Warto przypomnieć, że wśród organizatorów olimpiad przedmiotowych dla gimnazjalistów i uczniów szkół ponadgimnazjalnych dominującą grupę stanowią szkoły wyższe i stowarzyszenia profesjonalne, których liderzy rekrutują się z reguły ze środowiska akademickiego.

Pochwała zasad i metod dydaktycznych stosowanych w szkołach realizujących program Matury Międzynarodowej (w tym m.in. obiektywność

oceny uczniów i jej porównywalność osiąganą dzięki stosowaniu jednoznacznych i ściśle określonych kryteriów oceniania uczniów, zaznajomieniu uczniów z tymi kryteriami oraz ich opisami dla poszczególnych poziomów osiągnięć) nie wyczerpuje zagadnienia, jakim jest alternatywa dla testocentryzmu. Przykładami metod bieżącego sprawdzania i oceny osiągnięć uczniów są eseje argumentacyjne (rozprawki typu „za” i „przeciw”) czy podsumowania wyników badań uczniów zawierające konkluzje z eksperymentów przeprowadzonych w trakcie zajęć z dyscyplin przyrodniczych („raporty laboratoryjne”), prowadzenie przez samych uczniów tzw. dzienników uczniowskich z zapisami dotyczącymi uczenia się, prowadzenie tzw. portfolio uczniowskiego – „teczki” ucznia (dokumentacji zawierającej osiągnięcia szkolne oraz ilustrującej możliwości rozwojowe – potencjał ucznia), wykazy kryteriów samooceny osiągnięć ucznia (*rubrics*).

Autor przedstawiając scenariusze stopniowego wprowadzania obowiązku szkolnego dla 6-latków oraz edukacji przedszkolnej dla 5-latków nie ustosunkował się do idei partycypacji obywatelskiej w kształtowaniu zasad edukacji wczesnoszkolnej. Częste odwoływanie się do roli kapitału społecznego w recenzowanej rozprawie nasuwa bowiem refleksję, jakie jest stanowisko Autora wobec postulatów i inicjatyw społecznych (w tym akcji organizacji pozarządowych) dotyczących zniesienia ustawowego obowiązku szkolnego dla sześciolatków, sprzeciwu wobec wzrostu uprawnień władczych kuratorów i ograniczeniu wpływu JST na działanie szkół, czy też respektowania aprobaty rodziców dla treści programowych zajęć przedszkolnych i szkolnych.

Bezsporną zaletą pracy jest przedstawienie właściwości dobrze pomyślanego monitorowania jakości pracy szkół oraz przekonująca argumentacja na rzecz opcji alternatywnego modelu monitorowania, zakładającego aktywne uczestnictwo podmiotów tzw. warstwy pośredniczącej w procesie ciągłego, bieżącego monitorowania praktyki szkolnej (przykład programu „Szkoly z klasą”), a także roli rodziców w tym procesie. Pewne zastrzeżenia budzi jednakże brak klarowności w kwestii wskazania priorytetowych obszarów monitorowania (osiągnięcia ucznia to tylko jeden z kilkunastu, jeśli nie kilkudziesięciu, potencjalnych obszarów monitorowania). Pozytywną ocenę wspomnianego wątku pracy dodatkowo wzmacnia fakt, że monitorowanie pracy szkół w wykonaniu organów prowadzących (JST) pozostawia wiele do życzenia (niewystarczająca aktywność JST, ograniczone wykorzystywanie wskaźników edukacyjnych, niski stopień kooperacji międzyszkolnej oraz współpracy z ośrodkami akademickimi, brak konsensusu między JST i dyrektorami szkół w kwestii istoty i zakresu monitorowania).

W podsumowaniu uwag recenzenckich dotyczących monografii Marka Piotrowskiego należy podkreślić zasadnicze walory tej rozprawy:

- kreatywne rozwinięcie tezy badawczej J.J. Bonstingla wskazującej kierunek jakościowej zmiany paradygmatu w oświacie, którą wyraża następująca fraza: od nauczania i testowania do ciągłego uczenia się i doskonalenia;

- poprawność merytoryczną i metodologiczną analizy oraz płynących z niej wniosków dotyczących miarodajności wymiernych wskaźników skuteczności pracy dydaktycznej (jakości kształcenia) opartych na wynikach egzaminów zewnętrznych;

- nowatorstwo, dociekliwość naukową i wnikliwość analizy empirycznej weryfikującej wiarygodność i przydatność istotnego instrumentu polityki edukacyjnej państwa - egzaminów zewnętrznych - mierzących poziom osiągnięć uczniów i ich postępy w nauce;

- interdyscyplinarność podejścia wyrażającą się dobrą znajomością wiedzy i rozwiązań praktycznych dotyczących zróżnicowanych aspektów, kontekstów i instrumentów krajowej polityki edukacyjnej odnoszącej się do oświaty obowiązkowej;

- krytyczny ogląd realiów życia polskich szkół (dydaktyka, zarządzanie jakością kształcenia) w podwójnej perspektywie: deklarowanych celów i zadań zestawionych z działaniami pozornymi obciążającymi skuteczność i jakość kształcenia;

- znaczącą wartość poznawczą i przydatność rozprawy dla praktyki edukacji powszechnej, dającą podstawę do sformułowania rekomendacji zalecającej lekturę książki zarówno dla decydentów krajowej i samorządowej polityki edukacyjnej, jak też dyrektorów szkół i nauczycieli interpretujących i wykorzystujących dane, jakich dostarczają wskaźniki jakości pracy szkoły.

Mając na uwadze zarówno uwagi natury ogólnej, jak też uwagi szczegółowe odnoszące się do mocnych i słabych stron książki Marka Piotrowskiego, można pokusić się o stwierdzenie, że praca ta wnosi wartościowy wkład do badań nad instrumentami polityki edukacyjnej, realizującymi cele związane z permanentną poprawą jakości nauczania i osiągnięć szkolnych w polskim systemie oświaty. Przyczynienie się Autora książki do wzbogacenia wiedzy naukowej w dziedzinie pedagogiki i polityki edukacyjnej polega przede wszystkim na zastosowaniu oryginalnej metodyki badań służących zarówno ocenie procesu reformowania edukacji powszechnej, jak i ocenie użyteczności oraz wiarygodności pomiaru osiągnięć szkolnych za pomocą zewnętrznych egzaminów i testów. Podsumowując, należy również stwierdzić, że o wartości książki Marka Piotrowskiego stanowi nie tylko prezentacja interesującego materiału faktograficznego będącego plonem

badan empirycznych Autora, ale także udana próba całościowego ujęcia, z jednej strony, złożonych uwarunkowań i mechanizmów polityki edukacyjnej, z drugiej zaś – metodyki diagnostyki i pomiaru edukacyjnego oraz metod poprawy efektów kształcenia w krajowej oświacie powszechnej.

BIBLIOGRAFIA

- Amrein A.L., Berliner D.C., *The Impact of High-Stakes Tests on Student Academic Performance: An Analysis of NAEP Results in States With High-Stakes Tests and ACT, SAT, and AP Test Results in States With High School Graduation Exams*, Education Policy Studies Laboratory, Arizona State University, 2002.
- Angelo T.A., Cross K.P., *Classroom Assessment Techniques: A Handbook for College Teachers*, Jossey-Bass Publishers, San Francisco 1993.
- Bonstingl J.J., *Szkoły jakości. Wprowadzenie do Total Quality Management w edukacji*, CODN, Warszawa 1999.
- Burgess S., Wilson D., Worth J., *A Natural Experiment in School Accountability: The Impact of School Performance Information on Pupil Progress and Sorting*, "Working Paper", 2010, 10/246, Bristol Institute of Public Affairs, University of Bristol.
- Department for Education: <http://www.education.gov.uk/schools/performance/documents.html>
- Formative Assessment. Improving Learning in Secondary Classrooms*, Centre for Educational Research and Innovation, OECD, 2005.
- Froese-Germain B., *The OECD, PISA and the Impacts on Educational Policy*, Virtual Research Centre (VRC), Canadian Teachers' Federation, September 2010.
- Glossary of Statistical Terms*, OECD, 2008.
- Heckman J.J., *The Economics of Inequality. The Value of Early Childhood Education*, American Educator, Spring 2011.
- Heckman J.J., Kautz T., *Fostering and Measuring Skills: Interventions That Improve Character and Cognition*, „IZA Discussion Paper”, 7750, November 2013, Forschungsinstitut zur Zukunft der Arbeit (IZA).
- Herrnstein R.J., Murray R.C., *The Bell Curve. Intelligence and Class Structure in American Life*, Free Press, New York 1994.
- <http://heckmanequation.org/heckman-equation>
- ICELP (Instytut Feuersteina): <http://www.icelp.info/>
- Mons N., *Theoretical and Real Effects of Standardised Assessment*, Eurydice, 2009.
- Mortimore P., *Alternative Models for Analysing and Representing Countries' Performance in PISA*, Educational International Research Institute, Brussels 2009.
- Mourshed M., Chijioko C., Barber M., *Jak najlepiej doskonalone systemy szkolne na świecie stają się jeszcze lepsze*, McKinsey & Company, CEO, Warszawa 2012.
- National Center for Education Statistics (NCES): <https://nces.ed.gov/nationsreportcard/>
- Piotrowski M., *Od TQM do „żandarma”, czyli pod prąd*, Wydawnictwo VEDA, Warszawa 2013.
- Polska 2030. Wyzwania rozwojowe*, Kancelaria Prezesa Rady Ministrów, Warszawa 2009.