

Łukasz Remisiewicz 

Uniwersytet Gdański

TESTYZOWANIE EDUKACJI I JEGO EFEKTY

W niniejszym esejju autor proponuje reteoretyzację głównych idei zawartych w książce *Testy edukacyjne. Studium dynamiki selekcji i socjalizacji* (2017) autorstwa Agnieszki Gromkowskiej-Melosik. Z przykładów szczegółowo opisanych w książce, autor wyodrębnia uniwersalne wzorce zmian społecznych spowodowanych wprowadzeniem testów edukacyjnych, w ramach których podmioty uwikłane na różnych poziomach w testy edukacyjne reorientują swoje zachowania. Wszystkie te zmiany nazywane są testyzacją. Testyzacja składa się z trzech elementów: testokracji (poziom makrospołeczny), testolatrii (poziom przekonań kulturowych) i uczenia pod testy (poziom mikrospołeczny). Testokracja jako „rządy testów” polega na wykorzystaniu systemów abstrakcyjnych jako narzędzia selekcji. Testokracja jako „zarządzanie poprzez testy” polega na porównawczym ocenianiu stanu systemów edukacji na świecie. Testolatria obejmuje zespół przekonań kulturowych odnoszących się do obiektywności i niezależności testów. Przekonania te zapewniają stabilny i społecznie uprawomocniony porządek testowania. Wreszcie, uczenie pod testy modyfikuje techniki przekazywania i przyswajania wiedzy, przypisując wyższe wartości tym, które z większym prawdopodobieństwem poprawiają wyniki testu.

Słowa kluczowe: socjalizacja; selekcja; egzaminy; stratyfikacja edukacyjna; nierówności w edukacji

Łukasz Remisiewicz, University of Gdańsk

Testizing Education and Its Outcomes

Abstract

In the essay, the author offers to re-theorize the main ideas of the book *Testy edukacyjne. Studium dynamiki selekcji i socjalizacji* [Educational Tests. The Study in the Dynamics of the Selection and Socialization] (2017) by Agnieszka Gromkowska-Melosik. Using the examples meticulously described in the book, the author isolates the universal patterns of social changes affected by educational testing whereby the agents involved in the testing at various levels reorient their behaviour. These changes are called testization and are seen to unfold at three levels: testocracy (macro-social level), testolatry (cultural values level) and teaching to the test (micro-social level). Testocracy understood as „governing through tests” is the use of abstract systems as a selection tool. Testocracy understood as „managing through tests” allows for comparative evaluation of educational systems worldwide. Testolatry refers to the cultural beliefs about objectivity and independence of tests. These beliefs secure a stable and socially legitimized order of testing. Finally, teaching to the test

Uniwersytet Gdański, lukremisiewicz@gmail.com, ORCID 0000-0001-7710-6880.

* Pragnę podziękować autorom anonimowych recenzji za wiele trafnych uwag do pierwszej wersji tego esejju.

modifies the techniques of knowledge transmission and ascribes a higher value to such techniques that are more likely to improve test results.

Key words: socialization; selection; examinations; educational stratification; educational inequalities

Wprowadzenie globalnego systemu testów standaryzowanych niewątpliwie zupełnie odmieniło systemy edukacji oraz rynki pracy i wprowadziło zmiany na wielu płaszczyznach, między innymi biografii jednostek, sposobów uczenia się i nauczania, rywalizacji międzyszkolnej czy polityki edukacyjnej. Przerformowało także problemy podejmowane w dyskursach dotyczących celów edukacji, wartości wiedzy oraz sukcesu osobistego. Do Polski system egzaminów zewnętrznych dotarł na początku XXI wieku, wywołując podobną rewolucję. Choć na wyrost byłoby powiedzieć, że jest to rewolucja „prześlona” (używając terminu Andrzeja Ledera [2014]), to jednak z pewnością nie została ona w pełni urefleksyjniona w polskiej literaturze, zwłaszcza socjologicznej. Wszakże egzaminy jako zjawisko społeczne samo w sobie nigdy nie cieszyły się szczególnym zainteresowaniem socjologów, niemal zawsze ujmowane były jako elementy większej całości (Remisiewicz 2013). Krytyka testów standaryzowanych znana jest polskiemu czytelnikowi głównie z punktu widzenia przeciwników neoliberalnego podejścia w edukacji (np. Melosik 2009; Potulicka i Rutkowiak 2012; Potulicka 2014). I choć akcenty krytyczne wobec neoliberalnych tendencji obecne są także w recenzowanej książce Agnieszki Gromkowskiej-Melosik (2017), to ogólna perspektywa, którą oferuje, jest dużo szersza.

W niniejszym artykule chciałbym rozpatrzeć procesy, które zachodzą pod wpływem wprowadzenia i nadania egzaminom zewnętrznym znaczącej roli w edukacji. Zbiorczo będę je, zgodnie z tytułem, nazywał testyzowaniem różnych obszarów życia społecznego. Posiłkując się przykładami z książki spróbuję wskazać, jak testy wpływają na przeorientowanie polityk edukacyjnych, systemów edukacji, praktyk zarządzania instytucjami oświatowymi (poziom makrospołeczny), jak zmieniają społeczne przekonania na temat natury wiedzy, uczenia się oraz oceniania (poziom kulturowy), a także w jaki sposób zmieniają praktyki na poziomie mikrospołecznym.

W tym sensie pojęcie testyzacji jest analogonem pojęć takich jak medykalizacja (Shilling 2010) czy genetyzacja. Jej podstawą jest dostrzeżenie tego, jak wprowadzenie jednej zmiany całkowicie reorganizuje pewien analizowany wycinek rzeczywistości społecznej. Na przykład, choć badania genetyczne są po prostu pewną laboratoryjną biochemiczną procedurą, to ich rozpowszechnienie w społeczeństwie generuje skutki dalece wykraczające poza laboratorium, i w tym sensie jest *stricte* społecznym obiektem dociekań. Autorka cytuje w swojej książce badania dotyczące międzykulturowych aspektów funkcjonowania

systemu testów, co pozwala adekwatnie wyjaśnić, w jaki sposób zwykle mierzenie wiedzy urosło do rangi zjawiska zmieniającego społeczeństwo.

Amerykanie na początku XX wieku byli pionierami wdrażania naiwnego, pozytywistycznego przekonania, że umiejętności i zdolności ludzkie da się mierzyć tak, jak w fizyce mierzy się prędkość ciała stałego – wystarczy tylko skonstruować odpowiednio rzetelne narzędzie pomiarowe. Co prawda później odkrywano stopniowo rozmaite uwikłania testów, w szczególności skupiono się na czynnikach kulturowych, które już na starcie wpływały na obniżenie ocen niektórym grupom zdających. Potrzeba było kilkadziesiąt lat, aby w końcu zakwitła nieśmiało myśl o tym, że być może testy mają zasadnicze, a nie tylko przygodne wady, i nie da się usunąć poprzez tworzenie coraz bardziej neutralnych pytań i coraz bardziej precyzyjnych schematów punktowania odpowiedzi (por. np. Jencks i Phillips 1998; Owen i Doerr 1999: 180–181). I że być może skupienie na doskonaleniu narzędzi pomiarowych, choć rozwiązuje stare problemy, prowadzi jednocześnie do wielu nowych.

Cały błąd pozytywistycznego stanowiska w edukacji polega na tym, że prędkość ciała stałego można mierzyć bez ingerowania w to ciało. Udaje się stworzyć coraz doskonalsze narzędzia i dysponować dzięki temu coraz lepszą precyzją pomiarów. Każdy sposób wprowadzenia jakichkolwiek pomiarów do tkanki społecznej spowoduje natomiast cały szereg efektów, który socjologowie nazywają niezamierzonymi konsekwencjami działań celowych (Merton 1936).

System testów standaryzowanych nie tylko mierzy wiedzę. Jest biurokratycznym narzędziem polityk publicznych, służącym dystrybucji instytucjonalnych nagród i kar. Dzieje się tak dlatego, że wyniki egzaminu („skwantyfikowana wiedza jednostki”) agregowane są w statyczne makrorezultaty, w taki sposób by jako liczby mogły stać się łatwo porównywalne na różnych poziomach analizy.

Jednostki poddawane testom także nie poprzestaną na tym, żeby ich wiedza była mierzona obiektywnie, ale będą starały się możliwie zmaksymalizować uzyskiwane wyniki, chociażby za sprawą przyswajania strategii rozwiązywania testów (*test-wiseness*; Niemierko 1991: 44). Nie ma w tym nic dziwnego, skoro ich wyniki wyznaczają horyzont możliwości w zakresie wyboru dalszej drogi edukacyjnej lub zawodowej.

Podobnie rozpowszechnienie opartych na wynikach egzaminów praktyk ewaluacyjnych w zakresie oceny efektywności nauczycieli i szkół będzie prowadziło do prób maksymalizacji takich wyników kosztem tych umiejętności i obszarów wiedzy, które nie pojawiają się na egzaminach i tym samym nie są objęte działaniami ewaluacyjnymi. W związku z tym nauczyciele pod groźbą sankcji wdrażają program szkolny narzucony przez Ministerstwo Edukacji Narodowej. Szkoły mogą porównać się z innymi placówkami w rankingach osiągnięć, licząc że wysokie wyniki testów zachęcą w przyszłości lepszych kandydatów, by wziąć udział w ich rekrutacji (podnosząc poziom w następnych latach). Te zaś,

które osiągają niskie wyniki, narażone są na kontrole zewnętrzne prowadzone przez kuratora oświaty.

Tym sposobem niepostrzeżenie przechodzimy od „pomiaru wiedzy uczniów” do złożonych procesów społecznych, które zachodzą na poziomach instytucjonalnym, kulturowym i jednostkowym.

W związku z tym praca Gromkowskiej-Melosik nie tylko może służyć za doskonałe opracowanie myśli (głównie) krytyków testów standaryzowanych, którzy zjawisko testyzacji widzą i badają na co dzień, ale jednocześnie powinna być punktem wyjścia do namysłu dla tych, którzy naiwnie podchodzą do rozpowszechnienia testów i nadawania im wysokiej rangi w kształtowaniu polityki oświatowej.

Słowo o książce

Książka Gromkowskiej-Melosik skupia się na dwóch konkretnych aspektach testyzacji, a mianowicie socjalizacji oraz selekcji. Wpisuje się tym samym w dotychczasowy dorobek autorki, która oba aspekty badała już wcześniej jako osobne problemy. Tematowi selekcji poświęciła trzy książki. Pierwsza z nich, napisana z Tomaszem Gmerkiem, była zbiorem esejów na temat stratyfikacji i nierówności społecznych (Gromkowska-Melosik, Gmerek 2008), druga ściśle dotyczyła nierówności w dostępie edukacji ze względu na płeć (Gromkowska-Melosik 2011), natomiast trzecia koncentrowała się na zjawisku elitarnego szkolnictwa średniego, szczególnie w wymiarze ruchliwości oraz reprodukcji społecznej i kulturowej (Gromkowska-Melosik 2015). Oprócz tego, autorka ma w swoim dorobku także pracę dotyczącą nieuczciwych praktyk w edukacji, w tym między innymi ściągania i fałszowania dyplomów (Gromkowska-Melosik 2007). Recenzowana publikacja jest więc naturalnym wynikiem wieloletnich studiów poświęconych problemowi selekcji oraz innym, pokrewnym, problemom.

Książka składa się z kilku rozdziałów, z których każdy można traktować jako osobny esej. Autorka rozpoczyna od przedstawienia selekcyjnej i socjalizacyjnej roli testów, poświęcając szczególną uwagę dwóm konsekwencjom ich rozpowszechnienia, a mianowicie redukowaniu tożsamości oraz rzeczywistości. Następnie pochyła się nad problemem międzynarodowych testów PISA porównujących osiągnięcia edukacyjne uczniów konkretnych państw. Kolejnym tematem są międzynarodowe testy z języka angielskiego, których rozpowszechnienie autorka analizuje w sferze kulturowej dominacji państw zachodnich. Czwartym tematem jest państwowy chiński egzamin gaokao, którego wyniki mają najważniejszy wpływ na biografię chińskich nastolatków. Ostatni esej rozszerza nieco perspektywę i odnosi się do roli mierzenia osiągnięć edukacyjnych w wyłanianiu się brytyjskiej elity społecznej.

Niniejszy artykuł recenzyjny nie będzie typowy, ponieważ zrezygnowałem w nim z organizacji wywodu w sposób zbliżony do wywodu autorki, to znaczy analizując konkretne rodzaje egzaminów (PISA, certyfikaty z języka angielskiego, gaokao). Chciałbym raczej przeanalizować jej treść lokując ją w pewnych socjologicznych ramach teoretycznych, pomocnych by wyekstrahować uniwersalne tendencje czy zjawiska społeczne, które wspólnie prowadzą do aberracji czy wręcz patologii w dziedzinie sprawdzania wiedzy, a których szczegółowe aspekty zostały opisane w książce. Na testyzowanie edukacji składa się w moim przekonaniu co najmniej kilka społecznych lub instytucjonalnych efektów. Dla celów narracji podzieliłem je na trzy rodzaje – efekty związane z reorganizacją instytucji społecznych, efekty związane z przemianami kulturowymi, a także efekty związane z przemianami zachowań społecznych na poziomie mikro. Książka autorki pełna jest interesujących, a często także mało znanych przykładów funkcjonowania testyzacji na wszystkich tych poziomach. Niniejszy artykuł recenzyjny może stać się czymś w rodzaju mapy problemowej, wskazującej ogólne trendy rozwoju testyzacji, które warte są dalszych pogłębionych studiów.

Testokracja (poziom makrosocjalno-instytucjonalny)

Za pomocą terminu testokracja¹ można opisać system, który czyni testy edukacyjne głównym lub jedynym instrumentem mechanizmów selekcyjnych. W tym sensie testy „rządzą” procesami społecznymi. W pierwszej kolejności zarządzają procesami selekcji powodując, że możliwości kandydowania na określone miejsca w instytucjonalnych strukturach systemu edukacji lub rynku pracy wyznacza interakcja wyników osób zdających z ilością dostępnych miejsc. W drugiej kolejności wyniki testów standaryzowanych mogą być traktowane jako wskaźniki stanu rzeczywistości edukacyjnej, wykorzystywane przez decydentów na różnych szczeblach we wprowadzaniu zmian w zakresie sposobów, celów i programów kształcenia.

Egzaminy w roli narzędzia sprawdzania wiedzy używane są właściwie od samych początków najszerzej rozumianej, sformalizowanej edukacji; najdawniejsze świadectwa pochodzą ze starożytnej Mezopotamii. Przez wieki jednak egzaminy były właśnie jedynie i n s t r u m e n t e m, za pomocą którego to konkretni nauczyciele badali stan wiedzy swoich uczniów (Remisiewicz 2016: 21–44). Sama relacja pomiędzy obiema stronami zakładała bezpośredni lub pośredni

¹ Termin jest dość rozpowszechniony w dzisiejszym dyskursie, trudno ustalić, kto jest jego autorem. Jedno z wcześniejszych użycí znajdziemy w pracy Hermana J. Petersa (1961). Używa go też w jednym miejscu omawianej książki autorka (Gromkowska-Melosik (2017: 32).

kontakt. W wiekach średnich egzaminy miały charakter widowiskowych dysput, podczas których wykorzystywano zarówno wiedzę akademicką, jak i umiejętności z zakresu retoryki (Hoskin 1979). W przypadku wielu egzaminów oceny były binarne, egzaminatorzy decydowali jedynie o tym, czy kandydat zdał, czy też nie. Posiadali także znacznie większą swobodę w zakresie podejmowanych podczas egzaminu działań. Sytuacja podczas egzaminu była w pewnym sensie samoregulująca się – egzaminatorzy mogli między sobą negocjować kryteria oceny, ustalając w toku dyskusji i narad właściwy werdykt. Jak wskazuje Gerard Giordano (2005: 23–26), taki sposób oceniania też posiada wady, a czasami nawet, według słów cytowanego przezeń Cubberleya, niewiele różni się od osobistej opinii.

Dopóki egzaminy posiadają w systemie edukacyjnym względnie ograniczoną rolę, samo ich istnienie nie świadczy o zjawisku testokracji. Ta pojawia się dopiero wraz z zaistnieniem powiązań między różnymi systemami instytucjonalnymi wykorzystującymi wyniki testów, w którym to procesie formuje się ich wysoka stawka. Egzaminy miały dawniej ograniczone zastosowanie – organ przeprowadzający je archiwizował, a zdający się nimi posługiwali. Cechą testokracji jest poszerzenie pola instytucjonalnego obiegu wyników, a tym samym także rozszerzenie spektrum podmiotów nimi zainteresowanych. Stają się one nie tylko obowiązkowo honorowane na wyższym szczeblu kształcenia lub na rynku pracy, ale również użyteczne statystycznie – pozwalają na analizowanie wyników zdających w odniesieniu do innych zmiennych, a jak słusznie podkreśla autorka, wpływa to także na pozycję zdającego, który nagle staje się „odpowiedzialny” za wynik każdej z „reprezentowanych” przez siebie kategorii – klasy, szkoły, miasta, regionu, rocznika itd.

System testokratyczny jest w różnym stopniu scentralizowany, co oznacza, że instytucje rządowe/agendy/podmioty tworzące testy, odpowiadają za ustalenie kanonów będących podstawą egzaminów (w tym także na przykład programów szkolnych, gdy mowa o testach szkolnych), a także za dobór odpowiednich zadań egzaminacyjnych, zapewniających określoną zdawalność. Mogą także zarządzać samym procesem przeprowadzania egzaminu lub zlecić to innym podmiotom. Tworzą one także agendy odpowiedzialne za wytwarzanie oficjalnych narracji dotyczących testów (przede wszystkim posługując się narzędziami statystycznymi), które mogą być wykorzystywane w celu kształtowania polityk edukacyjnych.

Rzeczywistość testokratyczna ufundowana jest na przekonaniach dotyczących obiektywnego charakteru egzaminów – o zasadności tego przekonania będę pisał w następnym rozdziale, natomiast w tym miejscu opiszę jego wytwarzanie. Polega ono przede wszystkim na technicznym wyizolowaniu egzaminów w zakresie ról społecznych oraz interakcji w taki sposób, by sprawdzić wrażenie, że wszystkie czysto „ludzkie”, nieobiektywne, arbitralne czynniki są

skutecznie wyeliminowane i nie mają wpływu na proces oceny, i tym samym powinien uchodzić za całkowicie bezosobowy, niemal automatyczny.

W przypadku ról społecznych wszyscy uczestnicy testów stają się dla siebie obcy. Cała sytuacja wymaga odegrania oficjalnego przedstawienia, które ma zaświadczać o tym, że formalnie zachowane są neutralne stosunki, dochowane zostały ściśle procedury, a tym samym uprawomocnić pogląd, że cała sytuacja zapewnia sprawiedliwość i bezstronność. Zdającym odbierana jest ich codzienna tożsamość, zastąpiona przez identyfikator, który uniemożliwia egzaminatorom rozpoznanie ich nazwiska, płci czy rasy. Egzaminatorzy stają się urzędnikami, którzy mają wykonać jedynie swoją pracę, polegającą na zastosowaniu przewidzianych przez zewnętrzne regulacje procedur. Zarówno treść zadań, jak i kryteria oceny narzucone są przez inne, wyższe hierarchicznie, podmioty, silnie ograniczając swobodę egzaminatora w zakresie formułowania pytań oraz sposobów i kryteriów oceniania odpowiedzi.

Kolejnym aspektem jest wyabstrahowanie egzaminów z interakcyjnego kontekstu. Ideałem jest, by egzaminatorzy oceniali jedynie pracę, nie poznając samego zdającego. Firma ETS, zarządzająca kilkoma egzaminami o zasięgu globalnym, zaimplementowała system umożliwiający ocenianie w ten sposób nawet wypowiedzi ustnych na egzaminie TOEFL (Xi, Mollaun 2006), ale w przypadku innych egzaminów jest to możliwe poprzez zanonimizowanie pisemnych prac egzaminacyjnych i rozlosowanie ich wśród konkretnych grup egzaminatorów.

Wszystko to tworzy odpowiednie warunki do istnienia systemu testów, który w dużej mierze automatycznie administruje procesami selekcji społecznych. Jego działanie polega na wykorzystaniu narzędzi biurokratycznych, abstrakcyjnych reguł transformowania pojedynczych wyników w makrokonfiguracje, przyporządkowując zdających do odpowiednich kategorii. Czysto ludzka działalność sprowadza się tutaj jedynie do ustalania warunków brzegowych funkcjonowania tego skomplikowanego „urządzenia społecznego”. Wśród nich można wymienić ustalenie skali ocen, progów przyjęć, liczby miejsc, kryteriów rekrutacji i tym podobnych.

Tak funkcjonujące urządzenie społeczne rokrocznie generuje wyselekcjonowane na podstawie swoich wyników egzaminacyjnych „pakiety absolwentów”. Ci, którzy osiągnęli lepsze wyniki, mają większą swobodę w wyborze szkół lub miejsc pracy. Natomiast tym, którym poszło gorzej, pozostaje czekać na zwolnione przez prymusów miejsca lub wybranie mniej prestiżowych miejsc kontynuowania nauki albo rozpoczęcia kariery. Zatem istotnym aspektem systemu testokratycznego jest nadanie wynikom egzaminacyjnym przeważającej roli w procesach selekcji. Wydane na ich podstawie dyplomy stają się walutą, która w zależności od sytuacji rynkowej może posiadać większą lub mniejszą wartość (Collins 1979). W przypadku spadku ich wartości na przestrzeni lat, jednostki,

aby utrzymać swoją pozycję, muszą stosować różne strategie ograniczające negatywne skutki tej inflacji².

Polityka testów jest polityką łatwego zarządzania ogromnymi masami ludzimi: sortowania³, selekcjonowania, monitorowania i ewaluowania. Łatwość tych procedur polega przede wszystkim na tym, że obszar odpowiedzialności głównej instytucji nadzorczej (na przykład ministerstwa) kończy się na rozliczeniu instytucji jej podporządkowanych, którym zleca przeprowadzenie egzaminu. Drugim aspektem ułatwiającym zarządzanie jest możliwość rozporządzania punktami węzłowymi testów, ustawiając dość precyzyjnie odpowiedni wskaźnik zdawalności. Wówczas rządzący mają pewność, że w danym roku na rynek pracy lub do szkół wyższego szczebla wejdzie określony odsetek kandydatów. Testy mają w tym sensie bardzo istotny z politycznego punktu widzenia aspekt predykcyjny.

Szczególnym przypadkiem, który nie tyle rządzi procesami selekcji, ile używany jest jako instrument polityki oświatowej, są globalne, porównawcze testy osiągnięć, takie jak PISA, TIMMS, PIRLS. Są one niejako traktowane jako niezależna zmienna, punkt odniesienia, który pozwala obiektywnie zdiagnozować poziom edukacji danego kraju i porównania jego pozycji z innymi systemami na świecie. Zasadność wykorzystywania ich jako obiektywnych wskaźników stanu edukacji jest często krytykowana, o czym będę wspominał poniżej, tutaj jednak warto zaznaczyć, że mają one duży wpływ na myślenie o edukacji w poszczególnych krajach. Cytowana przez autorkę brazylijska minister edukacji stwierdziła, że wprowadzają one wręcz nowe pojęcia i kategorie w myśleniu o edukacji (Gromkowska-Melosik 2017: 62). Należy więc zdawać sobie sprawę z ich szczególnej roli – tak zmienione, zglobalizowane i zunifikowane myślenie prowadzi do rozpowszechnienia także podobnych trendów ewaluacyjnych, które reprodukują porządek testokratyczny.

Testy PISA koncentrują się ponadto na wybranych aspektach wiedzy, to znaczy matematyce, czytaniu i naukach przyrodniczych. Nie jest zaskakujące, że kraje, które pragną podnieść swoje wyniki w globalnym rankingu, będą podejmować wysiłki, aby poprawić osiągnięcia uczniów w tych właśnie obszarach, co niechybnie będzie się wiązało z zaniedbywaniem innych, nie mniej ważnych sfer edukacji. Skutkować to może oczywiście wypaczeniem obrazu edukacji w danym kraju. Nawet jeśli uczniowie będą osiągać lepsze wyniki w tych właśnie obszarach, lub którymsz z nich, nie oznacza to, że cała edukacja stanie się lepsza. Mamy tu do czynienia z typowym redukcjonizmem biurokratycznym, swoistą ramą epistemologiczną, która za istniejące uważa tylko to, co zapisane

² Jedną z nich jest zdobywanie kolejnych dyplomów. Dla sytuacji w Polsce por. np. Opiłowska 2009.

³ Joel Spring (1976) wprost nazywa je „maszynami sortującymi”.

i ujawnione w dokumentach. Zjawisko to znane jest pod nazwą błędu McNamary i polega na ocenie rzeczywistości wyłącznie przez pryzmat kwantyfikowalnych kryteriów⁴.

Warto wspomnieć, że reguły globalnych testów są przynajmniej częściowo ustalane przez podmioty zewnętrzne względem krajowych systemów edukacji. Uznają one siebie za arbitrow, którzy władni są rozstrzygnąć o tym, jak przyporządkować miejsca w rankingach opracowanych za pomocą obiektywnych miar. Efektem sprawowania przez nich funkcji arbitrow jest wytworzenie specyficznych rankingów reputacji, których reguły ustalane są przez nich samych. Rankingi te przeobrażają organizacyjne środowisko instytucji, które muszą dostosować się do nich jako konkurenci rywalizujący o miejsca w tak wykreowanej rzeczywistości. Zajęte ostatecznie miejsce jest więc skutkiem wzięcia udziału w sterowanej zewnętrznie grze, a osiągnięty dzięki temu status staje się widoczny i obiektywnie porównywalny, osiągany jest bowiem za pomocą formalnych reguł (Gromkowska-Melosik 2017: 56).

Testokracja w swoim czysto technologicznym aspekcie całą rzeczywistość sprowadza do liczb. Zadania egzaminacyjne opracowywane są w taki sposób, by stanowić wskaźnik umiejętności ucznia, wyniki zagregowane na podstawie danego kryterium stanowią wskaźnik poziomu populacji uczniów, ale także jakości nauczania w danej szkole, regionie czy kraju. To, jaki system jest lepszy lub gorszy, kto powinien być wzorcem, a kto powinien z niego czerpać, wnioskuje się po rankingu, w którym dany podmiot zajął konkretną pozycję. Liczby, interpretowane w ten sposób ustanawiają niejako rzeczywistość. Stanowią podstawę do opracowywania i wdrażania zmian edukacyjnych. Ale są też podstawą swoistej „gry ewaluacyjnej”, w której wszystkim podmiotom zależy na tym, by w tej grze osiągnąć jak najlepszy wynik. Wszelkie wysiłki ukierunkowane są więc na podniesienie liczbowego wyniku, wykorzystując również nieuniknione słabości biurokratycznych procedur i sztywnych reguł. Edukacja traci swój pierwotny sens: takie pojęcia jak etos, wychowanie, kształtowanie są odrzucane jako trudno kwantyfikowalne. Ponieważ ostatecznym celem jest raportowanie, zdobywanie punktów i porównywanie różnych podmiotów, rzeczywistość testów zaczyna „żyć własnym życiem”, staje się grą o osobnych regułach.

Krytycy testokracji powołując się na liczne badania inspirowane teoriami Bourdieu i Passerona ([1970]2006), Bernsteina (1990) czy Bowlesa i Gintisa (1976) uważają także, że testy przyczyniają się do odtwarzania struktury społecznej, a także promują nierówności już na wczesnych etapach edukacji. Pozornie obiektywne kryteria ocen premiuja osoby, które od urodzenia dysponują

⁴ Autorka wskazuje ten sposób myślenia na stronie 58, cytując Charlesa Handy'ego, jednak nie przytacza oryginalnej nazwy tego błędu.

większym kapitałem kulturowym, co pozwala im znacznie szybciej osiągać postępy w nauce. Testy wspierają sposoby myślenia klasy dominującej, która komponuje program szkolny, a także znacząco wpływa na przekaz ukrytych edukacyjnych treści. Dzieci, które szybciej mają dostęp do tego sposobu myślenia, mogą go przyswoić, są na pozycji uprzywilejowanej, przez co nie tylko łatwiej radzą sobie w szkole, ale także znając sposoby myślenia promowane w testach wysokich stawek mogą je łatwiej odtwarzać.

Istnieją co prawda badania, które pokazują, że egzaminy centralne przyczyniają się do wygaszania nierówności (np. Bol i in. 2014), jednak są one oparte na nieco problematycznej definicji egzaminu centralnego, stąd choćby egzaminy w Danii uważane są za egzaminy centralne (dlatego że centralny jest program nauczania), natomiast egzaminy w Stanach Zjednoczonych, które przecież uważane są za wysoce wystandaryzowane, tutaj uważane są egzaminy rozproszone prawdopodobnie (każdy stan decyduje o programie nauczania).

Autorka w kontekście nierówności społecznych powołuje się także na efekt kodu pocztowego wprowadzony przez Wayne Au (Gromkowska-Melosik 2017: 33). Ten sam wynik z testu czytania, osiągnięty przez ucznia z klasy nieuprzywilejowanej, a dodatkowo pochodzącego z biednego środowiska sąsiedzkiego, oraz o innej niż biała rasie, będzie miał zupełnie inny wydźwięk, niż ten sam wynik osiągnięty przez ucznia o przeciwnych kategoriach, ponieważ będzie osiągnięty zupełnie innym nakładem pracy.

Tym bardziej można zapytać, czy idealna równość w edukacji jest w ogóle możliwa? Skoro w szkole procedowany jest program, który zawiera wiedzę i umiejętności, które już w momencie rozpoczęcia nauki uczniowie opanowali w różnym stopniu – część uczniów otrzymuje te informacje wcześniej w swoich domach, a część nie. Uczniowie już na początku swojej drogi edukacyjnej startują więc z innych pozycji, część z nich przynajmniej w pewnej mierze przygotowana na to, czego będzie się uczyć, dla części nauka szkolna będzie zupełną nowością. Różnice te będą miały wpływ na dalsze losy, ponieważ nawet najlepsi uczniowie z rodzin o niskim statusie socjoekonomicznym będą musieli spędzić o wiele więcej czasu i włożyć o wiele więcej wysiłku, by pozostać na poziomie uczniów, którzy opanowali te umiejętności i wiedzę zupełnie nieświadomie. Nie ma efektywnego sposobu na to, by zjawisko to odwrócić, ponieważ współczesna edukacja polega na przekazywaniu i przyswajaniu określonych idei. O ile program szkolny będzie zawierał materiał stworzony przez przedstawicieli grup dominujących, o tyle dzieci wychowane w rodzinach reprezentujących te grupy będą naturalnie miały szansę się z nimi szybciej zapoznać.

Wracając do kwestii testów, można sobie zadać pytanie, czy testokracja ma jakieś pozytywne strony? Autorka, przytaczając słuszne głosy krytyczne, nie próbuje dostrzec jej pozytywów. W tym kontekście przychodzi na myśl koncepcja

makdonaldyzacji George'a Ritzera (2003). Makdonaldyzacja, mimo wielu swoich wad, jest odpowiedzią na niepewność współczesności. Unifikacja produktów sieci McDonald's prowadząca do tego, by pod każdą szerokością geograficzną zaistniała szansa otrzymania wystandaryzowanej i niemal identycznej kanapki, jest w pewien sposób opanowaniem rzeczywistości i może być psychologicznie korzystne. Podobnie konieczna współpraca zagraniczna, ściśle powiązanie podmiotów na rynkach światowych, otwarcie granic, a co za tym idzie – zglobalizowanie edukacji, z jednej strony przyczynia się do konieczności zaistnienia podmiotów, które będą zapewniały równie globalną selekcję, możliwości porównawcze oraz – co nie mniej ważne – przynajmniej psychologiczne uzasadnienie dla poglądu o równości szans. Z drugiej strony – co może się na pozór wydawać paradoksalne – testokracje poszczególnych krajów mogą wykorzystywać instrumenty testowania właśnie do akcentowania odrębności, swoistości systemów edukacji czy promowanych wartości. Historycznym przykładem niech będzie pierwszy przypadek testokracji, którą był system egzaminów urzędniczych w Chinach. Egzaminowani, mimo że kandydowali na stanowiska administracyjne, musieli wykazać się biegłością w zakresie konfucjańskich i neokonfucjańskich kodów kulturowych (Elman 2013).

Testokrację można scharakteryzować z jednej strony jako „rządy testów” – to jest niezależny, wprawiony jedynie w ruch system abstrakcyjny⁵, który selekcjonuje zdających pod względem wyników, mających przełożenie na ich szanse w dalszych procesach selekcyjnych. Z drugiej strony, testokracja jest „rządzeniem p o p r z e z testy”, gdy wyniki wykorzystywane są przez decydentów politycznych zarówno do bieżącego zarządzania rzeczywistością edukacyjną, jak i do formowania lub ewaluowania jej długofalowych skutków. Jednak system ten nie przetrwałby bez korpusu przekonań, które go legitymizują, aktualizują i wzmacniają.

Testolatria (poziom kulturowy)

Banesh Hoffmann w swojej książce *Tyrania testów* wprowadził termin testolatria (Hoffmann [1962]2003). Choć autor nie podał jego definicji *explicite*, z powodu swojej semantycznej konstrukcji, posłuży mi on doskonale do opisu pewnego zjawiska społecznego. Testolatria jest to przesadna (i, być może, przesadna!) wiara w moc testów. Obejmuje ona szereg przekonań dotyczących zarówno adekwatności konstrukcji testów (tj. że testy rzeczywiście mierzą to, o czym ich twórcy zapewniają, że mierzą), prawidłowości osiągniętych za jej

⁵ Używam tego pojęcia w rozumieniu Anthony'ego Giddensa (2012).

pomocą wyników (tj., że ocena z testu jest wiarygodnym miernikiem określonej umiejętności), a także zawiera w sobie schematy przypisywania uczniom tożsamości kategorialnych, wiążące przedstawiony ilościowo wynik testu jednostki z kategoriami jakościowymi zewnętrznymi wobec testowych schematów pojęciowych (dobry–słaby, mądry–niemądry, kompetentny–niekompetentny)⁶. O tym ostatnim zjawisku pisał zresztą Michel Foucault ([1975]1993), a jeszcze więcej jego komentatorzy (Hoskin 1979, 1990), poszerzając refleksję dotyczącą roli egzaminów w konstruowaniu obrazu danej jednostki. We współczesnych dyskursach edukacyjnych nadzieję pokładaną w testach uzasadnia się poprzez odwołanie do ich cech opisywanych metodami statystycznymi, w szczególności coraz większej trafności i rzetelności⁷.

Testolatria nie jest oczywiście poglądem zupełnie nowym. Edukacja niemal od zawsze cieszyła się prestiżem, a szczególnie wysoko cenione było zdobywanie dyplomów uznanych uczelni czy stopni naukowych, które uchodziło za wyznacznik statusu, umiejętności, możliwości czy nawet mądrości. Znow można wspomnieć w tym miejscu cywilizację chińską, która wraz z rozwojem gradacji dyplomów zdobywanych poprzez zdawanie egzaminów na konkretne pozycje urzędnicze, wprowadzała podobną hierarchię: ten, kto posiadał wyższy dyplom, uważany był za bardziej godnego, by piastować wyższą pozycję w hierarchii biurokratycznej Chin. W tym właśnie znaczeniu Chiny były merytokratyczne, ponieważ obowiązująca narracja, która ściśle wiązała zdanie konkretnego egzaminu z konkretną pozycją biurokratyczną, uzasadniała pogląd, że rządzą najmądrzejsi. Oczywiście kategoria mądrości musi być rozumiana w ramach tego konkretnego systemu światopoglądowego. Warto pamiętać o tym, że egzaminy chińskie nie miały nic wspólnego z wiedzą, która była wykorzystywana do zarządzania krajem. Sprawdzały one jedynie wiedzę z zakresu ksiąg konfucjańskich, chodziło więc nie tyle o zmierzenie technicznych umiejętności i partykularnej wiedzy, ile o diagnozę osobowości kandydata; czy jako człowiek nadaje się na określone stanowisko. Max Weber pisał w tym kontekście o egzaminach na gentlemana (Weber [1922]2002), osoby reprezentującej konkretny, promowany styl myślenia i działania.

Współczesna testolatria ma oczywiście zupełnie inny wymiar, ponieważ nie dotyczy egzaminów urzędniczych, a egzaminów edukacyjnych, i osiã zainteresowania nie jest to, jakiego stopnia dyplom otrzyma dana osoba, ale to, ile punktów zdobędzie na konkretnym egzaminie. Skwantyfikowany i ustratyfikowany prestiż ma dalekosiężne konsekwencje.

⁶ Autorka odwołuje się w tym miejscu do myśli Pauline Lipman. Por. Gromkowska-Melosik 2017: 44.

⁷ Zob. np. Phelps i in. 2005: 31. Podobny styl myślenia można zauważyć w przypadku testów o zasięgu globalnym, takich jak PISA. Por. OECD 2016.

Czasami niewielka różnica w wyniku decyduje o tym, czy dana osoba dostanie się na prestiżowy kierunek studiów lub na prestiżową uczelnię, czy też nie – i ten właśnie fakt zupełnie zmienia dalszą biografię. Już Bourdieu pisał o „błahych ułamkach punktów”, które mogą zmienić życie jednostki⁸. W tym właśnie sensie testolatria wydaje się wiarą zupełnie irracjonalną, ponieważ trudno rozumowo wyjaśnić, w jaki sposób jeden punkt procentowy różnicy w teście, który przecież nie sprawdzał całej wiedzy, ale jedynie jej bardzo niewielki ułamek, może świadczyć o tym, że dana osoba bardziej lub mniej zasługuje na miejsce na prestiżowej uczelni. W takich wypadkach znacznie bardziej prawdopodobne jest to, że jedna z tych osób miała po prostu szczęście, a czynników, które poza poziomem wiedzy mogły mieć tu wpływ, jest bardzo dużo.

Rozpowszechnienie przekonania o obiektywnym charakterze testów zależy od skuteczności podmiotów je organizujących. Instytucje, które tworzą systemy testów standaryzowanych starają się przekonać o tym, że są obiektywnymi arbitrami, którzy w uprawomocniony sposób władni są rozporządzać miejscami zajmowanymi przez konkretne podmioty w rankingach. Stosowane są tutaj rozmaite dyskursywne uzasadnienia, najczęściej dotyczące idei równości: dostępu, szans, kryteriów oceny czy procesu oceny (Gromkowska-Melosik 2017: 50–51).

Wiele firm i uczelni w krajach azjatyckich uważa, że posługiwanie się tym językiem jest niezbędnym elementem wykształcenia, a testy ETS w rodzaju TOEFL lub TOEIC są obiektywnymi wskaźnikami poziomu władania angielskim. Jak zauważa autorka, powołując się na krytyków językowych testów standaryzowanych, traktowanie ich jako wymiernych w celu diagnozy kompetencji językowych, traktując ten drugi termin w sposób zwyczajny, to znaczy kompetencji umożliwiających porozumiewanie się we wszystkich kontekstach życia społecznego, jest zupełnie nieuprawnione (Gromkowska-Melosik 2017: 83–84). Testy standaryzowane mierzą tylko niewielki zakres kompetencji, który polega na wypełnianiu arkusza testowego, biorąc pod uwagę specyficzne zadania o konkretnych profilach. Ponadto, sytuacja egzaminu jest nienaturalna i nietypowa, biorąc ją jako środowisko posługiwania się językiem obcym. W związku z tym kompetencje językowe oceniane są w bardzo szczególnym – klinicznym – układzie odniesienia (Gromkowska-Melosik 2017: 90). Twórcy testów próbują zaszczepić przekonanie, że można odnieść je do znacznie szerszego układu odniesienia, co jak słusznie podkreśla autorka, jest nieuprawnione, a co można uzasadnić biorąc pod uwagę chociażby problemy Azjatów w posługiwaniu się językiem angielskim poza kontekstem egzaminacyjnym. Autorka przytacza tutaj

⁸ „Francuski Uniwersytet dąży zawsze do wyjścia poza funkcję techniczną egzaminu konkursowego, aby pośród określonej liczby kandydatów, których wyboru się odeń żąda, ustalić z pełną powagą hierarchię opartą na błahych ułamkach punktów – śmiesznych, a jednak decydujących.” (Bourdieu i Passeron ([1970]2006: 244–245).

skrajny przypadek japońskich profesorów literatury angielskiej, którzy choć doskonale czytali i pisali w tym języku, to jednak nie potrafili w nim mówić (Gromkowska-Melosik 2017: 91).

Można więc zapytać: co rzeczywiście mierzą testy? Najbardziej skrajna odpowiedź mówi, że testy mierzą umiejętności wypełniania testów, a bardziej wyważona – że mierzą poziom umiejętności w pewnym wyizolowanym układzie, ale odnoszenie wyników do innych układów może być nieuprawnione. Warto w tym miejscu nadmienić, odchodząc od tekstu książki, że niektórzy socjologowie podejmujący incydentalnie temat egzaminów edukacyjnych (Weber [1922]2002; Bourdieu i Passeron [1970]2006; Foucault [1975]1993), wspominali o swoich wnioskach dotyczących „ukrytych” umiejętności, które faktycznie są badane. Bardziej systematyczne i porównawcze badanie (Watanabe 2015), wskazuje na różnorodność wzorców, zależnych od kultury i polityki.

Pojawiają się głosy krytyczne, według których obiektywizm testów standaryzowanych w odniesieniu do treści jest założeniem nieuprawnionym. Jak referuje autorka, są one wynikiem zachodniej myśli społecznej i, jako takie, akcentują wyłącznie zachodnie wartości oraz wiedzę (Gromkowska-Melosik 2017: 57; 95; 120). Autorka powołuje się na liczne badania wskazujące na różnorodność kulturowych naleciałości związanych z szeroko pojętą wiedzą matematyczną, nie mówiąc o lokalnych historiach czy narracjach, które w żaden sposób nie mieszczą się sposobie myślenia świata zachodniego. W tym sensie, na przykład, test PISA jest instrumentem kolonizacji mniejszych krajów, które zmuszone są zaszcześcić obce sobie wzorce w taki sposób, by podnieść wskaźniki osiągniętych wyników (Gromkowska-Melosik 2017: 57). Za kolonialistyczne uważa się także rozpowszechnienie testów kompetencyjnych z języka angielskiego (Gromkowska-Melosik 2017: 94 i n.).

Autorka wskazuje na słowa Andersona o tym, że wynik testów jest często dla rodziców wyznacznikiem wartości co najmniej wiedzy, a często również i wartości dziecka (Gromkowska-Melosik 2017: 31). W kulturach azjatyckich to na dzieciach ciąży obowiązek zaopiekowania się starszym rodzicem, toteż rozbudzanie wysokich ambicji i budowanie tożsamości uzależnionej od wyników testów, prowadzące do zintensyfikowania przygotowań do testów skutkujących możliwością rozpoczęcia obiecującej kariery, są w długoterminowej perspektywie w interesie samych rodziców. Ale także w mniej ekstremalnych przypadkach ocenianie uzdolnień, mądrości i wiedzy dziecka często korelowane jest z wynikami jego egzaminów.

Zjawisko testolatrii bardzo silnie oddziałuje na aspekt tożsamościowy życia społecznego. Silne przekonania o obiektywnej wartości testów uprawomocniają kategoryzowanie zdających podług ich wyników. Wyniki te wykorzystywane są wówczas już nie tylko w procesach selekcji, ale stają się także nieformalnym źródłem uzyskiwania statusu, emocjonalnej gratyfikacji dla wygranych, dumy

dla ich rodzin. Wysoki wynik uzyskany w teście pełni tu rolę najlepszego uwierzytelnienia mądrości, wiedzy i umiejętności. W tym względzie warto zauważyć, że wyniki testów odnoszone są do czegoś innego niż podstawa programowa danego testu, do kategorii dużo bardziej płynnych i dalece ogólniejszych. Kwantyfikujące procedury testów, mające sens w tym danym układzie odniesienia, wykorzystywane są do rozdzielania i kategoryzowania jednostek w ramach zupełnie innego uniwersum. Choć najczęściej mówi się o wygranych, takie nieuprawnione stosowanie kategorii jednego rodzaju do kategorii rodzaju drugiego prowadzi oczywiście jednak i do tej ciemniejszej, rzadziej opisywanej, ale paradoksalnie znacznie częściej występującej sytuacji, w której tłumy „przeegranych” muszą poradzić sobie z myślą, że tych wszystkich pozytywnie rangowanych cech nie posiadli w stopniu wystarczającym do poczucia spełnienia i dumy. W krajach azjatyckich bywa, że kończy się to samobójstwami dzieci (Rohlen 1983).

Testolatria jest rzeczywistością, w której wynik testu przyjmuje się za obiektywny miernik wiedzy zdającego w pewnym zakresie. Karol Marks pisał o tym, że egzamin jest chrztem wiedzy, uprawomocnia ją jako prawdziwą, rzeczywistość – mając na myśli to, że zespół egzaminatorów poprzez uznanie odpowiedzi zdającego nadaje jej specjalny, uświęcony status (Marks 1949). Dziś powiedzielibyśmy raczej, że testy standaryzowane konserwują wierzenia o merytokratycznym porządku rzeczywistości uświęcając rankingi szeregujące uczniów na podstawie „obiektywnych” kryteriów, przyporządkowując miejsca zależnie od posiadanych kompetencji. Aby znaleźć się na jak najwyższym poziomie, jednostki podejmują rozmaite działania, które w założeniu mają im to umożliwić.

Uczenie pod testy (poziom mikros społeczny)

W mikroskali wysoki poziom testyzacji przejawia się w koncentracji niemal wszystkich wysiłków podejmowanych w szkole na maksymalizowaniu egzaminacyjnych wyników uczniów. Jednym z najważniejszych aspektów jest modyfikacja nauczanych treści oraz sposobów ich uczenia tak, by te pierwsze ściśle odpowiadały organizacji treści na testach, a te drugie rozwijały umiejętności na testach przydatne oraz największy nacisk kładły na te partie, które mogą się pojawić z największym prawdopodobieństwem (Monfils i in. 2004).

Edukacja stestyzowana musi siłą rzeczy socjalizować uczniów do stawiania się zdającymi. Oznacza to nie tylko naukę odpowiednio spreparowanego materiału, ale także sposobów myślenia, które wykorzystuje się wypełniając arkusz egzaminacyjny. Nauka do testu polega najczęściej na przyswojeniu rozproszonych informacji, które następnie mogą być powielone w pewnych abstrakcyjnych

warunkach. Warunki te są całkowicie sztuczne, a wiedza w taki sposób odzwierciedlona pozbawiona jest kontekstu (Gromkowska-Melosik 2017: 22).

Ponadto, ponieważ materiał jest obszerny, uczeń zaczyna myśleć o swojej nauce w kategoriach „ekonomicznych”, zaczyna dozować wysiłek, a przyswajanie wiedzy nieprzydatnej z punktu widzenia testów jest marnowaniem energii, którą można spożytkować na uczenie się rzeczy istotniejszych. Linda McNeil pisze w tym kontekście o uczeniu defensywnym (Gromkowska-Melosik 2017: 19). Prowadzi ona również do zmiany statusu wiedzy, która z wiedzy-o-rzeczywistości staje się sztucznym i sztywnym konstruktem, zbiorem pewnych podanych z góry faktów. Autorka książki wskazuje za McNeil, że wiedza ta nie poddaje się interpretacjom, że uczniowie zachęcani są do tego, by przyswajając wiedzę w gotowej formie, ponieważ na testach nie ma czasu na bardziej pogłębione myślenie. Uczniowie mogą odczuwać wręcz strach przed eksploracją wiedzy, ponieważ ich własna konstrukcja czy sposób myślenia może odbiegać od konstrukcji, która została przyjęta przez autorów testu. W związku z tym – i tutaj pojawia się cała istota uczenia defensywnego – bezpieczniej jest po prostu przyswoić słowa nauczyciela czy autora podręcznika. Zakłada się bowiem, że będą one ściśle odpowiadały promowanemu na testach sposobowi myślenia. W związku z tym uczenie pod testy promuje także przyswajanie wiedzy powierzchniowej w taki sposób, aby uczeń, jak pisze Gromkowska-Melosik, wiedział po trochu o wszystkim, zamiast osiągnąć pogłębioną wiedzę na mniejszą liczbę tematów (Gromkowska-Melosik 2017: 19).

Ten sposób uczenia się jest sprzężony z technikami nauczania wykorzystywanymi w szkołach. Nauczyciele mają motywację do tego, by używać tych technik, które zapewniają efekt maksymalizujący wyniki egzaminacyjne ich uczniów. Jak pisze autorka, wiąże się to z redukowaniem materiału nauczania w taki sposób, by odpowiadał on potencjalnym pytaniom testowym oraz z nauką powtarzania informacji i procedur (Gromkowska-Melosik 2017: 75). W niektórych kulturach o wysokim poziomie testyzacji – na przykład w Chinach – nie ma nawet znaczenia to, jak dalece uczeń rozumie to, co robi. Ważne, by po prostu to robił. Nauka pamięciowa oraz nauka procedur jest postrzegana jako bardziej efektywna, a wręcz redukująca niepewność uczniów, którzy w przeciwnym wypadku nie posiadaliby odpowiednio stabilnego punktu odniesienia.

W związku z tym zanikają wszelkie kreatywne sposoby nauczania, które z jednej strony być może w dłuższej perspektywie okazałyby się bardziej skuteczne w kwestii rozumienia przyswajanego materiału, a z drugiej – rozwijałyby inne kompetencje, w żaden sposób nie premiowane na testach, takie jak umiejętność współpracy, dociekliwość, cierpliwość, kreatywność czy myślenie niesza-blonowe (Gromkowska-Melosik 2017: 16).

Co dotyczy się tego ostatniego, testy wręcz wymagają i zachęcają do myślenia według wypracowanych schematów. Jednym z podstawowych sposobów

przygotowania się do testów jest rozpoznawanie sposobów tworzenia kluczy odpowiedzi tak, by uczeń na teście potrafił nie tyle znaleźć właściwą odpowiedź na zadane pytanie, ile taką odpowiedź, która usatysfakcjonowałaby egzaminatora. Sprawia to szczególnie problem na testach z przedmiotów humanistycznych, które wymagają na przykład interpretacji tekstów. Wiedząc, że spośród różnych możliwych interpretacji istnieje jedna, która z dużym prawdopodobieństwem jest odzwierciedlona przez schemat punktowania odpowiedzi, ostatecznym zadaniem zdającego jest rekapitulacja założeń twórcy arkusza egzaminacyjnego dotyczącego danego tekstu.

Zmiany dotyczą także sposobu organizacji wiedzy oraz programów szkolnych. Program szkolny jest oczywiście zawsze pewnego rodzaju konstruktem, niekiedy zbiorem abstrakcyjnych zasad, niekiedy pewną wizją świata, a bywa że staje się pakietem twierdzeń redukującym złożoność świata. Ale w świecie, w którym dominuje podejście ukierunkowane na testy, jeszcze silniej próbuje się tę wiedzę dostosować do potrzeb tych, którzy będą ją potem sprawdzać. Wiggins i McTighe piszą w swojej książce o „wstecznym projektowaniu” procesu nauczania w taki sposób, by myślenie o ocenie uczniów poprzedzało cel edukacyjny (Wiggins i McTighe 2008: 19). I choć autorzy zastrzegają się, że nie mają na myśli jedynie oceny sformalizowanej, to wszechobecność i zakorzenienie instytucjonalne testów sprawia, że te stają się naturalnym punktem odniesienia. Skutkiem jest tworzenie wiedzy tym silniej atomizowanej i fragmentaryzowanej, podporządkowanej konkretnej wizji, bardzo przewidywalnej i często płytkiej.

Zwraca się przy tym uwagę na fakt, że testy nie sprawdzają wielu cenionych i potrzebnych współcześnie kompetencji, wśród których można wymienić takie, jak postawa obywatelska czy myślenie krytyczne. Nie promują także długotrwałych studiów nad pewnym wycinkiem rzeczywistości, rezygnując z tego na rzecz powierzchownego przyswajania określonych programem informacji i procedur. Cały proces dramatycznie zubaża więc możliwości, jakie daje edukacja, ograniczając pole rozwoju do tych kompetencji, które można łatwo wykazać i tych umiejętności, które dają się skwantyfikować (a przez to ocenić i porównać). Skutkiem tego szkoła nie rozwija dzieci w taki sposób, w jaki byłaby do tego zdolna, gdyby wyniki testów wysokich stawek traktowała jedynie jako instrument pomocniczy, a nie cel sam w sobie.

Niektórzy cytowani przez autorkę badacze uważają, że struktura testów i pytań testowych, poprzez jej kilkunastoletnie praktykowanie, zakorzenia się w umysłach młodzieży jako gotowy skrypt mentalny lub schemat myślenia, który może być wykorzystywany do rozwiązywania rozmaitych problemów. Testy standaryzowane jednak upraszczają rzeczywistość, a w związku z tym traktowanie złożonych, poważnych pytań i problemów życiowych w taki sposób, jakby były pytaniami testowymi, znacznie redukuje ich złożoność

(Gromkowska-Melosik 2017: 7–18). Może także doprowadzić do błędnych wniosków, na przykład takich, że w jakimś przypadku istnieje zawsze najlepsza odpowiedź, choć w rzeczywistości może się zdarzyć, że żadna nie będzie dobra lub kilka będzie na równi satysfakcjonujących.

Wiedza spreparowana pod przygotowania do testów jest rozproszona i zatomizowana, uczniowie mają więc problemy, by zrozumieć naukę jako – używając pojęcia Edwarda O. Wilsona (2011) – z natury konsylientną, czyli składającą się z zespołu naczyń połączonych, które wzajemnie są od siebie zależne. Testy standaryzowane fragmentaryzują wiedzę, każde z pytań może odwoływać się do zupełnie innego jej zbioru czy sposobu myślenia, a w związku z tym pomiędzy kolejnymi zadaniami nie musi istnieć żaden związek logiczny. Zresztą – o czym warto wspomnieć w ramach dygresji – jednym ze sposobów nabywania biegłości testowej jest próba powiązania logicznego kolejnych zadań przez uczniów, odkrywając ukryte założenia czy sposoby myślenia twórcy testu (Niemierko 1991). Rozwiązaniem pozwalającym na pozbycie się tego problemu jest właśnie pełna autonomizacja kolejnych zadań, co w założeniu ma sprawić, że uczeń nie będzie w stanie korzystać ze swoich indukcyjnie wypracowanych rozwiązań, wyszukując nieprzewidziane przez autorów prawidłowości czy wzory.

Jednakże długotrwały trening w rozwiązywaniu takich zatomizowanych zadań może doprowadzić do problemów związanych z umiejętnością wnioskowania. Jeżeli do pytań pojawiających się w życiu przykłada się matrycę zadań testowych, które nie są ze sobą logicznie powiązane, to można rozwiązać jeden problem, nie zwracając uwagi na to, że przyjęte rozwiązanie może stać w logicznej sprzeczności z innym wcześniej powziętym rozwiązaniem.

Institutionalne podporządkowanie edukacji testom, wsparte korpusem praconików dotyczących ich wartości diagnostycznej, prognostycznej i tożsamościowej, prowadzi do zmiany perspektywy na procesy uczenia się. Na przykładzie testów z języka angielskiego autorka przytacza badania dotyczące rzeczywistości japońskiej, w której uczniowie oceniają wartość lekcji biorąc pod uwagę jedynie to, czy przybliży ich ona do osiągnięcia wyższego rezultatu (Gromkowska-Melosik 2017: 91). Ogranicza to zdecydowanie horyzonty szerszego spojrzenia na swoje umiejętności, ale także – co szczególnie istotne – przekierowuje uwagę zdającego, który paradoksalnie w znacznie mniejszym stopniu interesuje się samym przedmiotem, z którego ma właśnie zdawać egzamin.

Nic dziwnego, że taka odhumanizowana i wysoce konkurencyjna rzeczywistość potęguje także negatywne emocje. Z jednej strony pisze się o lęku egzaminacyjnym, który staje się nieodłącznym towarzyszem zdających w tych kulturach, które od wyników testów uzależniają pozycję społeczną zdających. Życie w ciągłym stresie i bycie nieustannie, od najmłodszych lat nękanym powtarzającymi się „ewaluacjami” własnej pracy, może doprowadzić do negatywnych skutków zarówno psychicznych, jak i czysto fizjologicznych. Lęk może się

przejawiać także przed tym, by nie ujawniać własnej niekompetencji, co staje się obsesją szczególnie w krajach azjatyckich. Ta kulturowa cecha w interakcji z koniecznością uczenia się pod testy doprowadzić może do szczególnie trudnego psychicznie sprzężenia.

Z drugiej strony nie mniej istotne jest także negatywne nastawienie do samych przedmiotów, będących podstawą egzaminu. W tym miejscu można przytoczyć dwa przykłady – otóż wysokie wyniki fińskich uczniów w testach czytania ze zrozumieniem sąsiadują z wynikami badań pokazującymi, że odsetek uczniów czytających dla przyjemności się zmniejsza (Gromkowska-Melosik 2017: 80). Być może, jak słusznie wnioskuje autorka, jest to wynik traktowania czytania jako kompetencji służącej wypełnianiu arkuszy testowych. Z kolei japońscy uczniowie, od których wymaga się podchodzenia do egzaminów TOEIC lub (rzadziej) TOEFL, przyznają się do głębokiej niechęci do języka angielskiego (Gromkowska-Melosik 2017: 93). Może to być pokłosiem nie tylko swoistego imperialistycznego charakteru dyfuzji angielskiego do japońskiej rzeczywistości, ale także sposobu nauczania tego języka, który dostosowany jest do rzeczywistości testowej i polega na zapamiętywaniu słówek i reguł gramatycznych. Nic dziwnego, że taka nauka, która stawia wartość komunikacyjną języka przed jego gramatyczną poprawnością, prowadzi do tego, że kompetencje w zakresie użytkowania języka w codziennych sytuacjach są bardzo niskie.

Można więc potwierdzić, że egzaminy wstępne na studia, wymagające nauki języka angielskiego mają efekt odwrotny od zamierzonego – nie tyle mobilizują do nauki angielskiego, ile zniechęcają do niej. Strategia, jaką przyjmują zdający, polega na nauce tej „wersji” języka, to jest tych konkretnych – a łatwych do wyuczenia – umiejętności, które zmaksymalizują liczbę punktów. Dzieje się to jednak kosztem rzeczywistych kompetencji. Efekt przyrostu umiejętności jest więc w dużej mierze efektem wyłącznie funkcjonowania epistemologii biurokratycznej, która za istniejące uważa tylko to, co jest zapisane w dokumentach. Nie tylko nie przekłada się to na wzrost rzeczywistych kompetencji, ale wręcz mu szkodzi. Zdający tracą w istocie czas i nerwy, ponieważ to, czego się uczą, wykorzystają raz w życiu – na egzaminie.

Istotnym zarzutem podnoszonym w stronę testów o charakterze globalnym, takich jak PISA, ale również międzynarodowe testy z języka angielskiego, jest to, że silnie oddziałują one na kulturowe i jednostkowe wzorce myślenia. W odniesieniu do testów PISA mówi się o promowaniu zachodnich stylów myślowych oraz zachodniej wiedzy. Pomijając spór czysto epistemologiczny, dotyczący jej uprawomocnienia, można wskazać, że wiedza lokalna, na przykład ta dotycząca upraw roślin rosnących na konkretnych obszarach, nawet mimo że zaprzęgnięta w ramy magii lub mistycyzmu dobrze sprawdza się w tych konkretnych warunkach (Gromkowska-Melosik 2017: 57–59). Zmiana wzorców kognitywnych, promujących ścisłą, suchą klasyfikację, fizykalne myślenie przyczynowo-skutkowe,

może okazać się bardzo kłopotliwa i długotrwała; nie warta swojej ceny, którą byłoby zniszczenie dobrze funkcjonujących struktur wiedzy lokalnej.

Globalne testy z języka angielskiego z kolei również zmieniają myślenie jednostek, ale poprzez promowanie kategorii ontologicznych oraz założeń epistemicznych, będących rdzeniem, na którym nabudowany jest ten język. Pomijając już różnicę osobnych języków lokalnych, odzwierciedlających u swoich podstaw rozmaite wzorce kognitywne danej społeczności, istnieją także lokalne wersje języka angielskiego, takie jak nigeryjski pidżyn (Gromkowska-Melosik 2017: 120–121) lub singapurski singlish (Gromkowska-Melosik 2017: 115–116), przystosowane do wspomnianych wcześniej lokalnych warunków. Ponieważ matrycą referencyjną dla kompetencji językowych przewidzianych w testach są umiejętności *native speakerów* ze Stanów Zjednoczonych lub Anglii, osoby biegle posługujące się angielskim w lokalnej wersji, osiągają mimo to niskie wyniki. Biurokratyczna unifikacja kompetencji językowych prowadzi więc w tym sensie do niezamierzonej unifikacji kultury, w tym w szczególności wytwarzania uniwersum wspólnych znaczeń, dla których język angielski jest medium.

W rzeczywistości edukacyjnej o wysokim poziomie testyzacji, tak jak ma to miejsce w Chinach, egzaminy końcowe są podstawowe dla możliwości podjęcia studiów w najbardziej prestiżowych uczelniach, a co za tym idzie – możliwości rozpoczęcia obiecującej kariery. Stają się więc najważniejszym momentem w biografii ucznia, a przygotowanie do tego momentu jest okupione często latami pracy. Jak pisze autorka, kandydaci konkurujący o miejsca na najlepszych egzaminach chińskich (a właściwie ich rodzice), zaczynają myśleć o Gaokao 12 lat wcześniej, podejmując systematyczne przygotowania (Gromkowska-Melosik 2017: 132). Im bliżej terminu, tym bardziej są one zintensyfikowane. Rodzice podejmują rozmaite czynności, by zapewnić przyszłym zdającym komfortowe przygotowania, szkoły zapewniają dodatkowe zajęcia ukierunkowane wyłącznie na ten egzamin, a uczniowie często rezygnują z rodzinnego i towarzyskiego życia, dniami i nocami nie myśląc o niczym innym, poza testami (Gromkowska-Melosik 2017: 133–135). Największym dramatem jest jednak to, że strukturalne i formalne reguły funkcjonowania tego systemu w sposób nieuchronny doprowadzą do wygenerowania milionów młodych chińczyków, którym pomimo takiego poświęcenia się nie udało, i którzy po tylu latach wyrzeczeń będą musieli zadowolić się pracą fizyczną.

Polski kontekst testyzacji

Jeżeli do książki Agnieszki Gromkowskiej-Melosik można mieć jakąś uwagę, to taką, że nie odnosi ona swoich treści do polskiej rzeczywistości. Widzę co najmniej dwa powody takiego stanu rzeczy. Po pierwsze, opisuje ona procesy

zachodzące w innych krajach, i jako takie, ich wpływ na polską rzeczywistość może jawić się jako nieoczywisty. Po drugie, edukacja polska może wyglądać na tle innych opisywanych przypadków jako w niewielkim stopniu dotknięta problemami opisywanymi w książce. Na kilku przykładach chciałbym w tym miejscu krótko wskazać, że badanie testyzowania polskiej edukacji mogłoby być równie interesujące⁹. W skali makro chciałbym opisać przemiany międzynarodowe, które prowadzą do tendencji testyzacyjnych, w skali mikro – fenomen korepetycji, natomiast jako przykład testolatrii – rozwój wskaźników Edukacyjnej Wartości Dodanej.

W Polsce system egzaminów zewnętrznych został wprowadzony za sprawą reformy z 1999 roku, natomiast pierwszy oficjalny egzamin kończący szkołę podstawową odbył się w 2002 roku. W 2005 roku po raz pierwszy w nowej formule przeprowadzono egzamin maturalny. Przemiany polskiej edukacji były sprzężone z ogromną ilością zmian gospodarczych i społecznych wynikających z transformacji ustrojowej po 1989 roku. Wielu autorów widzi je w kategoriach implementacji filozofii neoliberalizmu czy marketyzacji (por. Potulicka i Rutkowiak 2012). Przemiany te stanowią w istocie kilka nakładających się na siebie wzajemnie tendencji. Z jednej strony zmienia się perspektywa kształcenia, funkcja wychowawcza szkoły ustępuje funkcji dydaktycznej (Groenwald 2011: 349), i to biorąc pod uwagę również fakt, że treści przekazywane przez szkołę zdają się coraz bardziej ściśle odpowiadać praktycznym potrzebom, bez koncentracji na zbędnych, „zbyt teoretycznych” zagadnieniach (Lawn 2010). Szkoła w filozofii neoliberalnej ma pełnić rolę służebną wobec rynku pracy, wyłaniając i kształtując osoby, które mogą na nim zająć „odpowiednie”, według ustalonych kryteriów, miejsca. Z drugiej strony szkoły stają się graczami na rynku edukacyjnym, konkurując o najlepszych absolwentów szkół wcześniejszego szczebla. Wyniki egzaminów stają się tutaj po prostu wskaźnikami, które mogą być łatwo porównywalne – jeśli oczywiście zawierzy się przekonaniu o ich trafności i rzetelności.

Wprowadzenie systemu egzaminów zewnętrznych w Polsce wiązało się ze szczególnego rodzaju zmianą myślenia na temat edukacji, którą dobrze oddają słowa z artykułu Henryka Szaleńca i Romana Dolaty, w którym autorzy widzą system egzaminacyjny jako instrument przejścia od zarządzania praktyką nauczania do definiowania efektów kształcenia i kontroli ich osiągnięcia (Szaleniec i Dolata 2012: 37). Autorzy uważają to przejście za konieczne ze względu na wprowadzenie pluralizmu programowego oraz decentralizacji procesu zarządzania szkołami za sprawą podniesienia w tej kwestii roli samorządów. Sednem zmian miało być, jak się wydaje, ustalenie pewnych jasno określonych celów,

⁹ Pragnę podziękować w tym miejscu anonimowemu recenzentowi za zwrócenie uwagi na te kwestie.

z możliwością obrania rozmaitych dróg do tego osiągnięcia przez nauczycieli i uczniów. Egzamin miał w pewnym sensie rozstrzygać to, która droga okazała się najbardziej efektywna. Takie wnioski można wysnuć chociażby biorąc pod uwagę treść przemówienia z 2001 roku, autorstwa Edmunda Wittbrodta, ówczesnego Ministra Edukacji Narodowej (Wittbrodt 2001).

Niestety, niezamierzoną konsekwencją tego działania był fakt, że wysoki wynik na egzaminie, który miał rozstrzygać jedynie o tym, że cel rzeczywisty został przez uczniów i nauczycieli osiągnięty, sam stał się celem. Trudno się temu dziwić, jeśli egzamin zewnętrzny pełni w polskiej rzeczywistości, obok diagnostycznej, również rolę selekcyjną. Można się spodziewać, że twórcze eksperymenty nauczycieli byłyby bardziej powszechne, gdyby nie fakt, że ich ewentualne niepowodzenie może mieć dalekosiężne konsekwencje, w postaci słabych wyników uczniów, które zaważą na ich dalszych biografiach. Odpowiedzialność nauczycieli odgrywa więc tutaj istotną rolę, jednocześnie przyczyniając się do redukcji indywidualności uczniów i nauczycieli na rzecz „człowieka opisanego standardami” (Groenwald 2011: 129). Spełnianie tych standardów w sposób, jaki już wcześniej dał akceptowalne wyniki, wydaje się bezpieczniejsze¹⁰.

Sprawę radykalnej przemiany polskiego systemu edukacyjnego, z centralną rolą egzaminów włącznie, komentuje Piotr Długosz wykorzystując teorię Roberta Mertona (Długosz 2012). Jego zdaniem uczniowie uczęszczający do szkoły w okresie przemian zderzyli się z rodzajem kulturowej anomii, przewartościowania dotychczasowych celów i sposobów ich osiągnięcia. Wywołało to w nich pilną potrzebę poszukiwania strategii adaptacyjnych tym bardziej, że te, proponowane drogami oficjalnymi, wykazywały się niską skutecznością. Jedną z nich jest pobieranie korepetycji, zajęć mających na celu poprawienie wyników szkolnych poprzez pracę w warunkach pozaszkolnych. Choć trudno ustalić dokładnie, jak wielu rodziców zapewnia dzieciom korepetycje, na podstawie badań CBOS z 2012 roku i ankiet wypełnianych podczas testów PISA z 2006 (dane za: Dolata i in. 2013: 169–170), można przyjąć, że w skali ogólnopolskiej około 1/3 rodziców posyłało dzieci na korepetycje co najmniej przez pewien okres, co najmniej z jednego przedmiotu. Podobne wyniki otrzymała grupa badawcza Romana Dolaty (Dolata i in. 2013: 181). Z powodu braku danych porównawczych sprzed ery testów standaryzowanych, trudno jednoznacznie powiedzieć, czy wprowadzenie ich w znaczący sposób powiększyło rynek korepetycji, zresztą dla zagadnienia testyzacji ważniejszym wydaje się raczej ustalenie motywu ich pobierania.

¹⁰ „Hipotezę bezpieczeństwa” można byłoby sprawdzić, badając liczbę programów autorskich w zależności od przedmiotu. Spodziewalibyśmy się, że im więcej uczniów zdaje egzamin z danego przedmiotu, tym mniejsza jest liczba programów autorskich z tego przedmiotu. Niestety próżno szukać danych ogólnopolskich dotyczących liczby programów autorskich.

Również i w tym zakresie nie przeprowadzono dotąd systematycznych badań, dlatego te wzmiankowane dalej mogą stanowić jedynie trop dla dalszych poszukiwań. Na przykład w badaniach Długosza przeprowadzonych na próbie uczniów szkół średnich zdecydowanie najistotniejszym motywem było uzyskanie najlepszego wyniku na maturze (Długosz 2012). Tymczasem badania przeprowadzone w 1984 roku na próbie uczniów kończących szkołę średnią pokazały, że głównym motywem pobierania korepetycji z matematyki było uzupełnienie luk w wiedzy (Rozmus-Chmura i in. 1990: 126). Warto byłoby dogłębniej zbadać ten problem, biorąc pod uwagę, w jakim stopniu testyzacja polskiej edukacji wpłynęła na rozwój rynku korepetycji, a także w jaki sposób go przeobraziła w odniesieniu do celów, jakie chcą za pomocą tych dodatkowych zajęć osiągnąć uczniowie i ich rodzice. Warto byłoby także pochylić się nad problemem praktyk, jakie prowadzi się w ramach korepetycji – jak dalece chodzi o podniesienie rzeczywistego poziomu wiedzy, jak dalece zaś technik i sposobów, które mogą okazać się pomocne w uzyskaniu zadowalającego wyniku.

Jednym z ważnych przejawów testokracji jest uznanie egzaminów zewnętrznych za główne narzędzie ewaluacji pracy szkoły. W Polsce jedną z najbardziej znanych miar wykorzystania wyników egzaminów do oceny pracy szkoły była Edukacyjna Wartość Dodana (EWD), rozwijana przez 10 lat – od 2005 do 2015 roku – ze środków Unii Europejskiej (Dolata i in. 2013; Żółtak 2015). Jej istotą jest ustalenie wpływu szkoły na osiągnięcia uczniów. Aby to zrobić, porównuje się wyniki tych samych uczniów z poprzedniego etapu edukacji (x), z wynikami egzaminów kończących następny etap edukacji ($x+1$). Zakłada się, że jeżeli stosunek wyników egzaminu szczebli $x+1$ do x jest w danej szkole znacząco wyższy od średniej ogólnopolskiej, to miała ona większy wkład w wyniki uczniów. EWD jest odpowiedzią na pewną intuicję, która za niesprawiedliwe uważa ocenianie szkół wyłącznie na podstawie średnich wyników egzaminu końcowego etapu $x+1$, ponieważ taki poziom ewaluacji nie bierze pod uwagę poziomu wiedzy uczniów na wejściu. Tym samym na przykład szkoły prestiżowe, znajdujące się w ośrodkach wielkomiejskich miałyby nieuzasadnioną przewagę. Tymczasem wskaźnik EWD może wykazać, że wysoki poziom danej szkoły nie wynika z wkładu szkoły, ale raczej z wysokiego poziomu uczniów na wejściu, a poziom średni mimo wszystko świadczy o dobrej pracy nauczycieli, skoro trafiają do niej głównie uczniowie słabi. O tendencjach testyzacyjnych w polskiej edukacji świadczyć może oczywiście nie fakt stworzenia takiego wskaźnika, ale zaufanie do niego (poziom testolatryczny) oraz sposób jego wykorzystania do celów ewaluacji (poziom testokratyczny).

Ponieważ wskaźnik EWD, wpisując się w mechanizm urynkowania szkoły, ma za zadanie wypracować płaszczyznę porównawczą dla rynku edukacyjnego oraz ewaluacyjną dla zarządzających oświatą, autorzy, którzy zaangażowani byli w rozwój metod jego szacowania, zdają sobie sprawę z konieczności

zachowania wysokiej trafności i rzetelności egzaminów zewnętrznych, jeśli na ich podstawie ma on być wyliczany. Pewien problem polega na tym, że ustalanie wskaźników trafności i rzetelności zawiera ukryte założenia. Pierwszym z nich jest założenie o istnieniu cechy ukrytej, która manifestuje się w rozwiązaniu określonego zadania (lub rozwiązań zbioru zadań podobnych). Drugim jest założenie o ciągłości tej cechy – w tym znaczeniu z rozwiązania zadania trudnego mierzącego tę cechę wnioskuje się automatycznie o tym, że ta sama osoba rozwiązałaby również zadanie łatwe. „Na podstawie dowodów opanowania bardziej złożonych umiejętności, możemy wyciągnąć wniosek o posiadaniu mniej złożonych kompetencji (nawet bez weryfikowania tego zadaniami w teście). Podobnie, jeśli osoba wykazała się umiejętnością napisania listu formalnego, możemy się spodziewać, że potrafiłaby napisać list do przyjaciela.” (Dolata i in. 2013: 69). O tym, że tak być nie musi, może zaświadczyć znakomicie przykład japońskich profesorów literatury angielskiej, przytaczany przez autorkę recenzowanej książki. Innym argumentem przeciwko testolatrii może być ten podnoszący wartość prognostyczną amerykańskiego testu SAT – jednego z najdokładniej rozwijanych, będących niejako wzorcem pod względem pomiarowym dla systemów egzaminacyjnych w innych krajach. Otóż wyniki SAT mają dla osiągnięć akademickich mniejszą trafność prognostyczną niż oceny w szkole średniej (*high school*) (Geiser i Santelices 2007).

Warto też wspomnieć, że przy stosowaniu wskaźnika EWD istnieje presja ukierunkowana na tworzenie egzaminów, które z większą łatwością można ze sobą porównać. To jednak posiada swoje konsekwencje w postaci coraz większej standaryzacji. Zwolennicy EWD postulują także rozmaite zmiany, które pomogłyby usunąć problematyczne statystycznie aberracje pojawiające się w analizach międzyegzaminacyjnych, spowodowanych na przykład zwalnianiem laureatów olimpiad przedmiotowych z konieczności zdawania egzaminu maturalnego z danego przedmiotu (Żółtak 2015: 65). Należy to odczytać jako efekt tendencji testyzacyjnej, w której sam pomiar wiedzy jest traktowany jako cel nadrzędny, który powinien kształtować inne aspekty edukacyjnej rzeczywistości. W tym przypadku gratyfikacyjna funkcja statusu laureata olimpiady powinna zostać zniesiona, ponieważ utrudnia ona dostosowanie miar egzaminacyjnych dla potrzeb ewaluacyjnych.

Niech ten krótki z konieczności przegląd tematów będzie świadectwem, że w polskiej rzeczywistości również mamy do czynienia z tendencjami testyzacyjnymi, ale oczywiście wektory rozwoju testyzacji są bardzo specyficzne. System egzaminów zewnętrznych wprowadzany był pod wpływem inspiracji zachodnimi metodami diagnostyki i ewaluacji edukacyjnych, lecz specyficzna sytuacja przemian ustrojowych zarysowała dla implementacji tych metod unikatowy kontekst. Dotychczasowe badania skupiają się z jednej strony na analizach czysto funkcjonalnych, które można też nazwać narracjami oficjalnymi. Ich cechą

charakterystyczną jest wysokie zaufanie do narzędzi statystycznych oraz wiara w możliwość obiektywnego pomiaru. Jeśli wskazywane są jakieś słabości egzaminów zewnętrznych, wydają się one autorom możliwe do przewyciężenia. Egzamin jest w tym dyskursie rzeczą oczywistą, a celowość egzaminowania lub dalekosiężne zmiany w procesach oceniania nie są przedmiotem analizy. Drugi typ narracji odwołuje się do krytyki tendencji neoliberalnych, które według autorów mają dalekosiężne, negatywne skutki oddziałujące na cały system edukacyjny. Niestety wciąż brakuje prac, które w sposób krytyczny analizowałyby instytucjonalne aspekty funkcjonowania systemu egzaminacyjnego w Polsce, w sposób systematyczny badały zjawisko testolatrii, a także próbowały zmierzyć się z tematem nauki pod testy (tego problemu dotyczył raport Krzysztofa Konarzewskiego [2008]).

Zakończenie

Cały sens obiektywnego pomiaru w naukach przyrodniczych polega na tym, że osiągnięty wynik nie zależy od badacza, a przeważająca część czynników zewnętrznych, jakie mogłyby na niego wpływać, jest poddana rygorystycznej kontroli. Jak przekonuje Agnieszka Gromkowska-Melosik, w przypadku testów edukacyjnych pomiar taki nie jest możliwy nie tylko ze względów technicznych. Jest on niemożliwy z powodu istoty procesów, jakie wprowadzenie testów standaryzowanych wprawia w ruch. Nawet jeśli z powodu drugorzędnych funkcji spełnianych przez testy w skali makro, takich jak budowanie kanałów transferu uczniów pomiędzy poziomami nauczania, nie da się od nich odejść, to za szczególnie ważny należy uznać dalszy namysł nad ich rzeczywistą rolą na wszystkich poziomach organizacji społeczeństw. Nie mniej istotne jest szerokie informowanie o wynikach tychże badań. Konieczne wydaje się także budowanie świadomości decydentów politycznych na temat globalnych testów PISA, TIMMS, TOEIC, TOEFL, których rola intencjonalna niekoniecznie odpowiada rzeczywistym potrzebom i celom konkretnych państw. Wpisuje się ona w szeroko rozpowszechnioną tendencję znaną jako polityka oparta na dowodach (*evidence-based policy*). Jednak naukowy rygor gromadzenia baz danych umożliwiających przekształcenia statystyczne, nie zawsze idzie w parze z trafnością interpretacji osiągniętych wyników.

Współczesny styl myślenia o edukacji, ewaluacji i roli „beznamiętnych” testów, jak określa to zgrabnie autorka (Gromkowska-Melosik 2017: 176), wydaje się dziś dla osób z „wewnątrz”, czyli ewaluatorów i specjalistów od diagnostyki edukacyjnej – czymś oczywistym. Zmiany przez nich proponowane, za pomocą sugestywnych i obiecujących narracji oddziałują na sposób myślenia polityków,

osób zarządzających oświatą, a dalej nauczycieli, uczniów i ich rodziców. Tymczasem przyjęcie perspektywy historycznej ukazuje, że ten rozliczeniowy styl myślenia o edukacji jest czymś stosunkowo nowym, szczególnie w Europie. Ma oczywiście pewne zalety, ale ma również wady, których wyeksponowanie może stać się w debacie równie silne.

W tym znaczeniu omawianą książkę uważam za bardzo dobre wprowadzenie do tej tematyki, syntetycznie streszczające wielowątkowe zarzuty kierowane w stronę testów standaryzowanych. Biorąc pod uwagę ogólny stan świadomości dotyczący tego tematu, powinna ona stać się lekturą obowiązkową dla polskich decydentów kształtujących system edukacji. Sądzę jednak, że na tym wprowadzeniu środowisko socjologiczne oraz socjopedagogiczne nie powinno poprzestać. Mamy tu bowiem do czynienia z palącymi kwestiami, wpływającymi na kształt całego globalnego myślenia edukacji i, jako takie, domagają się one zastosowania i rozwinięcia również w ramach polskiej rzeczywistości. Fakt, że biorąc pod uwagę skrajne przykłady innych krajów, polski system nie jest jeszcze do końca stetyzowany, tym bardziej powinien mobilizować do tego, by ewentualne zmiany wytyczać z uwzględnieniem wszystkich kosztów, jakie testyzacja mogłaby przynieść. Bo choć, jak podkreśla autorka w zakończeniu, współczesne pokolenia niemal na pewno „skazane” będą na testy, to kształtowanie polityk edukacyjnych również w zakresie przeprowadzania ich może i powinno odbywać się w sposób maksymalnie odpowiedzialny. A to oznacza branie pod uwagę wszystkich dostępnych faktów, wyciągając na światło dzienne wszystkie ukryte założenia.

Bibliografia

- Bernstein, Basil B. 1990. *Odtwarzanie kultury*. Tłum. Z. Bokszański, A. Piotrowski. Warszawa: Państwowy Instytut Wydawniczy.
- Bol, Thijs, Jacqueline Witschge, van de Werfhorst, Herman G. i Jaap Dronkers. 2014. *Curricular Tracking and Central Examinations: Counterbalancing the Impact of Social Background on Student Achievement in 36 Countries*. „Social Forces” 92: 1545–1572. DOI: 10.1093/sf/sou003.
- Bourdieu, Pierre i Jean-Claude Passeron. [1970] 2006. *Reprodukcja: Elementy teorii systemu nauczania*. Tłum. E. Neyman. Warszawa: WNPWN.
- Bowles, Samuel i Herbert Gintis. 1976. *Schooling in capitalist America: Educational reform and the contradictions of economic life*. London: Routledge and Kegan Paul.
- Collins, Randall. 1979. *The Credential Society: An Historical Sociology of Education and Stratification*. New York: Academic Press.
- Długosz, Piotr. 2012. *Korepetycje maturzystów pogranicza w latach 2008–2011*. „Kultura i Edukacja” 88: 88–106.

- Dolata, Roman, Anna Hawrot, Grzegorz Humenny, Aleksandra Jasińska-Maciażek, Maciej Koniewski, Przemysław Majkut i Tomasz Żółtak. 2013. *Trafność metody edukacyjnej wartości dodanej dla gimnazjów*. Warszawa: Instytut Badań Edukacyjnych.
- Elman, Benjamin A. 2013. *Civil examinations and meritocracy in late Imperial China*. Cambridge, Massachusetts, London, England: Harvard University Press.
- Foucault, Michel. [1975]1993. *Nadzorować i karać: Narodziny więzienia*. Tłum. T. Komendant. Warszawa: Wydawnictwo Aletheia.
- Geiser, Saul i Maria V. Santelices. 2007. *Validity Of High-School Grades In Predicting Student Success Beyond The Freshman Year: High-School Record vs. Standardized Tests as Indicators of Four-Year College Outcomes*. „UC Berkeley: Center for Studies in Higher Education”.
- Giddens, Anthony. 2012. *Nowoczesność i tożsamość: „Ja” i społeczeństwo w epoce późnej nowoczesności*. Tłum. A. Szulżycka. Warszawa: WN PWN.
- Giordano, Gerard. 2005. *How testing came to dominate American schools: The history of educational assessment*. New York: P. Lang.
- Groenwald, Maria. 2011. *Etyczne aspekty egzaminów szkolnych*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- Gromkowska-Melosik, Agnieszka. 2007. *Ściągą, plagiaty, fałszywe dyplomy: Studium z socjopatologii edukacji*. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.
- Gromkowska-Melosik, Agnieszka. 2011. *Edukacja i (nie)równość społeczna kobiet: Studium dynamiki dostępu*. Kraków: Oficyna Wydawnicza „Impuls”.
- Gromkowska-Melosik, Agnieszka. 2015. *Elitarne szkolnictwo średnie: Między reprodukcją społeczno-kulturową a ruchliwością konkurencyjną*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Gromkowska-Melosik, Agnieszka. 2017. *Testy edukacyjne: Studium dynamiki selekcji i socjalizacji*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Gromkowska-Melosik, Agnieszka i Tomasz Gmerek. 2008. *Problemy nierówności społecznej w teorii i praktyce edukacyjnej*. Kraków: Oficyna Wydawnicza „Impuls”.
- Hoffmann, Banesh. [1962]2003. *The Tyranny of Testing*. Mineola, N.Y. Dover Publications.
- Hoskin, Keith. 1979. *The Examination, Disciplinary Power and Rational Schooling*. „History of Education” 8: 21–135. DOI: 10.1080/0046760790080205.
- Hoskin, Keith. 1990. *Foucault Under Examination. The Crypto-Educationalist Unmasked*. W: S. J. Ball (ed.). *Foucault and Education. Disciplines and Knowledge*. London/New York: Routledge.
- Jencks, Christopher i Meredith Phillips. 1998. *The black-white test score gap*. Washington, D.C. Brookings Institution Press.
- Konarzewski, Krzysztof. 2008. *Przygotowanie uczniów do egzaminu: pokusa łatwego zysku*. Warszawa.
- Lawn, Martin. 2010. *The ‘Usefulness’ of Learning: The struggle over governance, meaning and the European education space*. „Discourse: Studies in the Cultural Politics of Education” 24: 325–336. DOI: 10.1080/0159630032000172515.

- Leder, Andrzej. 2014. *Prześlona rewolucja: Ćwiczenie z logiki historycznej*. Warszawa: Wydawnictwo Krytyki Politycznej.
- Marks, Karol. 1949. *Przyczynek do krytyki heglowskiej filozofii prawa*. W: *Wybrane pisma filozoficzne 1844–1846*. Warszawa: Książka i Wiedza.
- Melosik, Zbyszko. 2009. *Uniwersytet i społeczeństwo: Dyskursy wolności, wiedzy i władzy*. Kraków: Oficyna Wydawnicza „Impuls”.
- Merton, Robert K. 1936. *The Unanticipated Consequences of Purposive Social Action*. „American Sociological Review” 1: 894. DOI: 10.2307/2084615.
- Monfils, Lora F., Roberta Y. Schorr i William A. Firestone (eds.). 2004. *The ambiguity of teaching to the test: Standards, assessment, and educational reform*. Mahwah, NJ: L. Erlbaum Associates Publishers.
- Niemierko, Bolesław. 1991. *W poszukiwaniu technologii pisania zadań testów sprawdzających osiągnięcia uczniów*. „Neodidagmata” nr XX: 27–52. DOI: (OECD). 2016. *PISA 2015 Results (Volume I)*. OECD Publishing.
- Opiłowska, Marta. 2009. *Współcześni studenci jako stratedzy własnej kariery zawodowej (w świetle teorii gier i kredencjalizmu)*. „Teraźniejszość – Człowiek – Edukacja” 48: 107–126.
- Owen, David i Marilyn Doerr. 1999. *None of the above: The truth behind the SATs*. Lanham, Md. Rowman & Littlefield Publishers.
- Peters, Herman J. 1961. *Testing or Testomania*. „The Clearing House: A Journal of Educational Strategies, Issues and Ideas” 35: 141–144. DOI: 10.1080/00098655.1960.11477595.
- Phelps, Richard, Herbert J. Walberg, Herbert Walberg i J. E. Stone. 2005. *Kill the Messenger: The War on Standardized Testing*. Somerset: Taylor and Francis.
- Potulicka, Eugenia. 2014. *Neoliberalne reformy edukacji w Stanach Zjednoczonych: Od Ronalda Reagana do Baracka Obamy*. Kraków: Oficyna Wydawnicza „Impuls”.
- Potulicka, Eugenia i Joanna Rutkowiak. 2012. *Neoliberalne uwikłania edukacji*. Kraków: Oficyna Wydawnicza „Impuls”.
- Remisiewicz, Łukasz. 2013. *W stronę holistycznej socjologicznej teorii egzaminu*. „Teraźniejszość – Człowiek – Edukacja” 3: 41–54.
- Remisiewicz, Łukasz. 2016. *Egzamin w perspektywie socjologicznej*. Gdańsk: Wydawnictwo Uniwersytetu Gdańskiego.
- Ritzer, George. 2003. *Makdonaldyzacja społeczeństwa: Wydanie na nowy wiek*. Tłum. L. Stawowy. Warszawa: Muza.
- Rohlen, Thomas P. 1983. *Japan's High Schools*. Berkeley, Calif. Univ. of California Press.
- Rozmus-Chmura, Małgorzata, Jadwiga Sajdak, Bogumiła Samoń i Mirosława Zawal-Jarosik. 1990. *Czy korepetycje z matematyki są powszechne?* „Roczniki Polskiego Towarzystwa Matematycznego” XXIX: 123–129.
- Shilling, Chris. 2010. *Socjologia ciała*. Tłum. M. Skowrońska. Warszawa: WN PWN.
- Spring, Joel H. 1976. *The sorting machine: National educational policy since 1945*. New York: Longman.
- Szalaniec, Henryk i Roman Dolata. 2012. *Funkcje krajowych egzaminów w systemie edukacji*. „Polityka Społeczna” nr 1: 37–41.

- Watanabe, Masako E. 2015. *Typology of Abilities Tested in University Entrance Examinations: Comparisons of the United States, Japan, Iran and France*. „Comparative Sociology” 14: 79–101. DOI: 10.1163/15691330-12341339.
- Weber, Max. [1922]2002. *Gospodarka i społeczeństwo: Zarys socjologii rozumiejącej*. Tłum. D. Lachowska. Warszawa: WN PWN.
- Wiggins, Grant P. i Jay McTighe. 2008. *Understanding by design*. Alexandria, Va. Association for Supervision and Curriculum Development.
- Wilson, Edward O. 2011. *Konsiliencja: Jedność wiedzy*. Tłum. J. Mikos. Poznań: Zysk i S-ka Wydawnictwo.
- Wittbrodt, Edmund. 2001. *Informacja rządu o stanie przygotowań i zabezpieczeniu finansowym "Matury 2002"*. Zapis debaty sejmowej 3 Kadencji Sejmu, 107 posiedzenia, 3 dzień (26.04.2001), 25 punkt porządku dziennego. Adres url: <http://orka2.sejm.gov.pl/Debata3.nsf/moin/3550DD92>.
- Xi, Xiaoming i Pam Mollaun. 2006. *Investigating the Utility of Analytic Scoring for the TOEFL Academic Speaking Test (TAST)*. „ETS Research Report Series” 2006: i-71. DOI: 10.1002/j.2333-8504.2006.tb02013.x.
- Żółtak, Tomasz. 2015. *Statystyczne modelowanie wskaźników Edukacji Wartości Dodanej: Podsumowanie polskich doświadczeń 2005–2015*. Warszawa: Instytut Badań Edukacyjnych.