

Florian Plit*

NEGATYWNE STRONY STANDARYZACJI SYSTEMÓW OCENIANIA

Negative side effects of standarization of grading systems

Zarys treści. Celem artykułu jest ukazanie negatywnych skutków towarzyszących standaryzacji systemów oceniania uczniów i studentów, odchodzenia od odpytywania, preferowania testów, najlepiej jednokrotnego wyboru. Zwrócono uwagę na fakt, że testy pisemne wprawdzie zwiększają obiektywizm, ale też nie zapewniają pełnej porównywalności wyników i nie likwidują subiektywizmu oceny.

Słowa kluczowe: ocenianie, standaryzacja, testy.

Key words: evaluation, standardization, tests.

Kolejne reformy systemu edukacji w Polsce, wprowadzane na wszystkich szczeblach, począwszy od zerówki i pierwszych klas szkoły podstawowej aż po studia wyższe, podejmowane po 1989 roku zmierzają w kierunku obiektywizacji i standaryzacji systemów oceniania. Służyć temu mają różne formy egzaminów zewnętrznych, standardy nauczania w szkołach wyższych, punkty ECTS itd. Jest to część znacznie szerszego i nieuniknionego procesu związanego z globalizacją w edukacji, mobilnością uczniów, studentów i pracowników naukowych, potrzebą porównywania dyplomów uzyskiwanych w różnych krajach etc. O ile jednak w innych dziedzinach życia krytyka negatywnych aspektów globalizacji ma charakter powszechny, co pozwala na ich ograniczanie, to w zakresie edukacji w Polsce są to raczej postawy odosobnione. Przeważa bezkrytyczny zachwyty nad standaryzacją i obiektywizacją procesów kształcenia. W tej sytuacji celowe wydaje się zwrócenie uwagi na kilka słabości promowanych systemów na przykładzie edukacji geograficznej.

W oficjalnych dokumentach wprowadzających reformy edukacji w Polsce (i w rozlicznych projektach reform) zwraca się uwagę na konieczność odcho-

* Uniwersytet Warszawski, Wydział Geografii i Studiów Regionalnych, Zakład Dydaktyki Geografii i Krajoznawstwa, e-mail: flplit@uw.edu.pl.

dzenia od zdobywania wiadomości na rzecz kształtowania umiejętności. Pierwsze oficjalne projekty reformy z 1998 roku operowały terminem „kompetencje”. Przykładowo, szkoła nie miała wykształcić ucznia tak, by znał on problemy dotyczące rynku pracy w Polsce, ale by „posiadał kompetencje” w zakresie rozwiązywania problemów rynku pracy. Na szczęście, ta nieudana kalka językowa z angielskiego szybko znikła, nie ma jej już w projekcie podstawy programowej z 27.11.1998, ani też w ostatecznej wersji podstawy dla szkół podstawowych i gimnazjów (*Podstawa programowa...*, 1999) oraz dla liceów (*Podstawa programowa...*, 2002). Zgodnie z tymi dokumentami sprawdzać należy „osiągnięcia” ucznia. Z kolei według obecnie wprowadzanej reformy (*Rozporządzenie Ministra...*, 2009) oceniane mają być „umiejętności”. O umiejętnościach jest też mowa w standardach kształcenia na poszczególnych kierunkach studiów oraz w dokumentach Państwowej Komisji Akredytacyjnej. Przejście od „wiedzy” do „umiejętności” powinno zaowocować zmianą sposobu sprawdzania i oceniania ucznia, zwłaszcza że równocześnie coraz więcej mówi się i pisze o celach operacyjnych kształcenia (operacjonalizacji celów kształcenia).

EGZAMINY ZEWNĘTRZNE JAKO PANACEUM NA SUBIEKTYWIZM OCENIANIA

Na wszystkich poziomach kształcenia odchodzi się, a w zasadzie już odeszło się, od klasycznego sprawdzania wiadomości i umiejętności poprzez odpytywanie (egzamin ustny, odpowiedź przy tablicy itd.), a także od dłuższych prac pisemnych (rozprawek). W przypadku odpowiedzi ustnych w zasadzie nie ma możliwości weryfikacji, czy wystawiona ocena była prawidłowa. Oceniający może się tylko starać, by zminimalizować swoje subiektywne odczucia, nigdy nie może być pewien, że ocenił w pełni sprawiedliwie. S. Zajac (1997, s. 164) ma rację, twierdząc, że: „...przy indywidualnej kontroli ustnej ocenianie jest bardzo subiektywne”. Jeśli po kilku dniach uczeń (student) kwestionuje ocenę, trudno wykazać jej zasadność, nawet w przypadku wypowiedzi publicznych, gdyż świadkowie nie zawsze pamiętają szczegóły, a często są tendencyjni (np. solidarność klasowa). Powszechnie wiadomo, że na ocenę wpływają też takie czynniki, jak samopoczucie oceniającego (humor), jego stan zdrowia, warunki pogodowe itd. Z kolei odchodzenie od pisanych rozprawek uzasadniane jest m.in. czasochłonnością sprawdzania, niechlujnym charakterem pisma, licznymi błędami stylistycznymi, gramatycznymi i ortograficznymi.

Wszystkim tym mankamentom powinien zapobiec pisemny egzamin testowy¹: jednolity dla wszystkich, a zatem w pełni porównywalny. W przypadku

¹ Termin test (testowy) jest obecnie stosowany w tak wielu znaczeniach, że podobnie jak w przypadku terminów ekologia, ekologiczny, regionalny oznaczać może niemal wszystko i każ-

odpowiedzi ustnych, np. maturalnych, nawet na pozór jednoznaczne wymaganie typu „Uczeń potrafi pokazać na mapie ściennej Wisłę” nie jest w pełni porównywalne. Należałoby sprecyzować, ile ma na to czasu, zapewnić we wszystkich szkołach identyczne mapy ścienne (z upływem czasu jednakowo wyblakłe), sprecyzować warunki oświetlenia, dokładność pokazywania, zunifikować wskaźnik, którym posługuje się uczeń itd. Kryteria pełnej porównywalności spełniać mają natomiast egzaminy zewnętrzne kończące gimnazjum oraz tzw. nowa matura. Nad ich przeprowadzeniem czuwa Centralna Komisja Egzaminacyjna i Okręgowe Komisje Egzaminacyjne, a zakres ich działania oraz szczegółowy przebieg egzaminów reguluje szereg zarządzeń Ministra Edukacji Narodowej (i Sportu), wydanych po uchwaleniu przez Sejm *Ustawy o systemie oświaty* z dnia 7 września 1991 roku. By zapewnić pełną porównywalność, egzaminy zewnętrzne przeprowadzane są w tym samym czasie w całym kraju. Wszyscy uczniowie mają takie same arkusze egzaminacyjne. Przepisy określają, między innymi, kto może przebywać w sali, jakie są kompetencje obserwatorów, egzaminatorów i weryfikatorów, jakie są odległości między stolikami oraz co może znajdować się na stoliku. Jednak mimo starań, identyczność warunków pozostaje zapewne fikcją, gdyż warunki oświetlenia i natężenie ulicznego hałasu w salach nie są identyczne, nadto warunki pogodowe w momencie egzaminu różnić się mogą istotnie w poszczególnych regionach kraju, co rzutuje chociażby na zdolność koncentracji. Odpowiednich badań jednak nie przeprowadzono.

Jednakowy sprawdzian wymaga identycznego, a co najmniej podobnego przygotowywania uczniów przez wszystkich nauczycieli. W oczywisty sposób grozi to zanikiem indywidualizmu kształcenia. Wcześniej powinien też być ustalony, w sposób jak najbardziej precyzyjny i jednoznaczny, zakres wymagań. Służą temu z jednej strony sylabusy (np. *Informator maturalny...*, 2003) oraz zapisywanie szczegółowych wymagań za pomocą czasowników operacyjnych (potrafi, rozumie, pokazuje na mapie, umie naszkicować). Na studiach identyczną rolę pełnią standardy kształcenia, sylabusy poszczególnych przedmiotów, dążenie do identycznego zapisu sylabusów itd.

TRUDNOŚCI ZWIĄZANE Z NIEJEDNOZNACZNYM FORMUŁOWANIEM POLECEŃ

Warunkiem koniecznym poprawnie przeprowadzonego egzaminu jest jednoznaczność poleceń. Przy egzaminach ustnych istnieje możliwość korekty

dorazowo wymaga sprecyzowania. W niniejszym tekście oznaczać będzie czynność: metodę pomiaru (pomiar sprawdzający), a nie formularz testu „rozdam wam testy”). Metoda ta przeprowadzana jest pisemnie, zawiera szereg poleceń, przy czym nie muszą one ograniczać się do wyboru odpowiedzi na pytania, ale dotyczyć też mogą np. obliczeń, wykonania prostych rysunków itd.

polecenia: gdy egzaminator widzi, że egzaminowany nie rozumie pytania, bądź też rozumie je opacznie, z reguły precyzuje, o co mu chodzi. W klasycznych podręcznikach dydaktyki (Czekańska 1959; Zajac 1997) podawany jest przykład źle sformułowanego pytania: „Jakie mamy góry w Afryce?”. Otóż, po pierwsze – „my” (Polacy, uczniowie, studenci) nie mamy w Afryce żadnych gór na własność, po drugie – pytanie „jakie?” dotyczy cech przymiotnikowych, zatem poprawną odpowiedzią mogłoby być: ładne, wysokie, niskie, denudowane etc. Jeszcze bardziej pouczający jest przykład bardzo niejednoznacznego pytania: „Nad jaką rzeką leży Warszawa?”. Oczekiwana przez nauczyciela odpowiedź to, oczywiście, „nad Wisłą”, ale i w tym przypadku pytanie dotyczy cechy, zatem poprawniej byłoby na nie odpowiedzieć: „nad długą, nad brudną, nad głęboką” itd. Przeformułowanie pytania na „Nad którą rzeką leży Warszawa?” tylko częściowo rozwiązuje problem. Po pierwsze, Wisła nie jest jedyną rzeką przepływającą przez Warszawę, dlatego też odpowiedź „nad Wilanówką” byłaby również poprawna. A jak ocenić odpowiedź „nad Potokiem Służewieckim”? Czy ciek ten uznać można za rzekę? Po drugie, wątpliwości budzi położenie „nad rzeką”, gdyż sugeruje ono, iż chodzi o rzekę płynącą „pod” miastem. Dlatego też najbardziej poprawną odpowiedzią na pytanie: Nad którą rzeką leży Warszawa? byłoby... „nad Drną (Drzesną, Drzęsną)”. Tak nazywała się bowiem rzeczka wypływająca niegdyś z podmokłych terenów koło Czystego (część dzisiejszej Woli), płynąca Okopową i Stawkami (nazwa ulicy pochodzi od licznych podpiętrzeń na rzeczce, której energię wykorzystywały liczne niewielkie zakłady, głównie młyny), przez dzisiejszy Plac Inwalidów i uchodząca do Wisły koło Cytadeli. W XIX w. niemal cały jej bieg został włączony do systemu kanalizacyjnego, znalazł się pod ziemią i rzeczywiście, dzisiejsza Warszawa leży „nad” Drną. Jednoznacznym poleceniem wymienienia Wisły byłoby natomiast: „Podaj nazwę największej rzeki przepływającej przez Warszawę, stolicę Polski?”.

Przy poleceniach pisemnych możliwości wyjaśnienia polecenia w zasadzie nie ma, stąd na zewnętrznych egzaminach gimnazjalnych i na maturze, mimo że zadania opracowywane są przez specjalistów i przeprowadzane jest próbne testowanie (ze względu na potrzebę zachowania tajności w ograniczonym zakresie), co roku pojawia się wiele niejasności. Nie wszystkie są wychwytywane i zgłaszane do ekspertów w celu rozstrzygnięcia. Najczęściej stosowana jest zasada, że „wątpliwości przemawiają na korzyść ucznia”. Skoro nie można doprecyzować polecenia, skutkuje to zaliczeniem części odpowiedzi przypadkowych i zaniżeniem poziomu wymagań.

Jednak inaczej przedstawia się sprawa uznawania odpowiedzi w testach opracowywanych przez nauczycieli szkolnych i akademickich. Z reguły ocenia je ta sama osoba, która przygotowywała test, wewnętrznie przeświadczona, że zrobiła to poprawnie i najczęściej oczekująca jednego tylko rodzaju odpowie-

dzi. Odpowiedzi nietypowe są wtedy uznawane za błędne, zwykle nie dopuszcza się nawet możliwości istnienia innej (innych) poprawnej (poprawnych). Jest to sprawa o tyle istotna, że uczeń (student) ma niewielkie tylko możliwości uzasadnienia swojej odpowiedzi i rozwiania wątpliwości. Tymczasem w geografii operujemy problematyką tak złożoną, że jednoznaczne odpowiedzi należą do rzadkości. Przykładem może być polecenie dawane studentom na zajęciach z dydaktyki geografii na Uniwersytecie Warszawskim: „Podaj nazwę liczącego najwięcej mieszkańców miasta, przez które przepływa rzeka Don”. Dotyczy ono faktografii i zmiennej łatwo mierzalnej (liczba mieszkańców), zatem odpowiedź powinna być jednoznaczna. Najczęściej studenci wymieniają Donieck, który nie leży nad Donem. Jednak poprawną odpowiedzią nie jest również Rostow nad Donem, ale Toronto, gdyż przez tę kanadyjską metropolię przepływa niewielka rzeka też nazywająca się Don. Jednak i do tej odpowiedzi trzeba podchodzić ostrożnie, gdyż autor niniejszego tekstu nie sprawdził, czy nie ma identycznie nazwanej rzeki płynącej np. przez Kalkutę, Szanghaj czy Teheran. Uczniowie i studenci czerpią wiedzę nie tylko z podręczników i wykładów, ale też z telewizji, internetu, własnych podróży, magazynów ilustrowanych, tabloidów, a nawet gier komputerowych. Jest ona przez to bardzo wycinkowa i nieuporządkowana, często błędna, ale zdarza się, że fragmentarycznie zaskakująco szczegółowa.

Powyższe przykłady wydać się mogą wydumanymi. Rzeczywiście, w zdecydowanej większości przypadków mamy do czynienia z odpowiedziami standardowymi, wręcz z myśleniem stereotypowym. Jednak każdy, kto przeprowadzał egzaminy pisemne, spotykał się z zaskakującymi, ale poprawnymi odpowiedziami. Wcale nie jest tak, że wysoka stawka egzaminu (np. matura) wpływa na zmniejszenie ich liczby. Być może nawet stres stymuluje do poszukiwania najbardziej karkołomnych prób rozwiązań. Niektóre z nich bywają wręcz humorystyczne. Przykładem może być zadanie pisemne na zawodach finałowych jednej z olimpiad geograficznych. Zadaniem finalistów, wyposażonych w mapy topograficzne, było udzielenie odpowiedzi na pytanie, czy z wierzchołka konkretnej góry A widać wierzchołek góry C. Pomiędzy tymi wzniesieniami znajdowała się góra B, a prostą odpowiedź „tak” lub „nie” należało uzasadnić rysunkiem. Zdecydowana większość osób wykonywała (zgodnie z oczekiwaniami jurorów) uproszczone profile topograficzne, ale jedna z uczestniczek narysowała dwa drzewka, między nimi niewielką sylwetkę ludzką i napisała: „Nie można zobaczyć, bo na górze A rośnie las, a ja jestem malutka”. Rzeczywiście, rósł tam las i informację tę można było (należało?) wyczytać z mapy topograficznej, zatem jury uznało odpowiedź za poprawną, chociaż na rysunku uczestniczki były świerczki, a rósł tam las bukowy.

Najłatwiejsze do sprawdzania są zadania testowe zamknięte, przede wszystkim proste zadania jednokrotnego wyboru, które sprawdzać można posługując

się szablonem². W latach 2002-2006 na egzaminach gimnazjalnych liczba zadań zamkniętych w arkuszu egzaminacyjnym w części matematyczno-przyrodniczej wynosiła 26-28, a zadań otwartych tylko 7-8. Najbardziej zadania otwarte dotyczyły biologii i chemii, ale w roku 2006 nie było ani jednego otwartego zadania z geografii. Biorąc pod uwagę liczbę możliwych do uzyskania punktów, dysproporcje nie były już tak duże, ale nadal występowała przewaga punktów za zadania zamknięte (Barwinek, 2009). Zadania zamknięte pozwalają na dużą porównywalność wyników, zwłaszcza przy dużej liczbie osób sprawdzających. Znacznie trudniej sprawdza się zadania otwarte, zwłaszcza zawierające polecenia typu „wymień trzy główne czynniki...”. Po pierwsze, często nawet sprawdzającemu trudno jednoznacznie określić, które czynniki są „główne” (zwl. że główne nie jest synonimem najważniejsze), po drugie - czynniki wymieniane przez sprawdzanych zwykle nie spełniają wymogów podziału rozłącznego, niekiedy nawet mamy ten sam czynnik, ale zapisany na trzy odmienne sposoby, po trzecie – sprawdzający mają wątpliwości, jak postąpić w przypadku, gdy wymieniono np. 7 czynników, ale nie przeprowadzono żadnej ich hierarchizacji. Czy uznać odpowiedź za bardzo dobrą (zna więcej niż wymagamy), czy też za niewystarczającą (nie wie, które są główne), a może przy ocenianiu brać pod uwagę tylko trzy pierwsze z wymienionych czynników? Jednak testy zamknięte najlepiej nadają się do sprawdzania wiedzy o faktach (wiedzy faktograficznej), znacznie trudniej sprawdza się za ich pomocą rozumienie wzajemnych powiązań, umiejętność tłumaczenia zjawisk. Dlatego też, niezależnie od deklarowanego w oficjalnych dokumentach Ministerstwa Edukacji Narodowej (przez kilka lat Ministerstwa Edukacji Narodowej i Sportu) odchodzenia od faktografii, stawiania na kształtowanie umiejętności etc. (por. np. *Podstawa programowa...* 1999; *Podstawa programowa...*, 2002, *Rozporządzenie Ministra...* 2009) egzaminy zewnętrzne w tym zakresie sprzyjają utrwalaniu konserwatywnego modelu kształcenia³.

² Trzeba jednak pamiętać, że w paru ośrodkach akademickich, gdzie takie sprawdziany są szczególnie popularne, a używane szablony – nieprzezroczyste, studenci geografii umiejętnie zakreślają wszystkie możliwe odpowiedzi, uzyskując w ten sposób 100% trafień.

³ Ta sama trudność pojawia się przy zawodach olimpiady geograficznej. Próbowano ją rozwiązać w ten sposób, że na II etapie, gdy zawody przeprowadzane są i prace sprawdzane równocześnie w 12 okręgach, przeważają polecenia, na które oczekuje się jednoznacznych odpowiedzi. Często dotyczą one właśnie wiedzy geograficznej. Pozwala to na dużą porównywalność wyników, ale i tak pracująca później w Warszawie komisja weryfikacyjna stwierdza szereg różnic. Są np. okręgi „surowe”, gdzie jeden błąd literowy w nazwie geograficznej dyskwalifikuje odpowiedź, są „łagodne”, akceptujące podanie nazwy w ogólnych zarysach zbliżonej do poprawnej. Dopiero na trzecim etapie, gdy wszystkie prace sprawdzane są przez same osoby, główny nacisk kładziony jest na wyjaśnianie zjawisk i procesów.

EGZAMINY ZEWNĘTRZNE A EDUKACJA REGIONALNA

Poważny mankament centralnych egzaminów z geografii stanowi fakt, że nie uwzględniają one specyfiki regionalnej obszaru, na którym żyje uczniów. W edukacji geograficznej, co najmniej od czasów Amosa Komensky'ego, obowiązuje paradygmat, że zaczynamy od poznawania tego, co nam najbliższe, stopniowo przechodząc do coraz to dalszych kręgów poznania (por. np. Hibszer, 1996; Pulinowa, 1993, 2005, Plit, 2008 i inni). Po 1989 roku większy nacisk na edukację lokalną i regionalną uznano za szczególnie ważny zarówno ze względu na potrzebę odbudowania tożsamości regionalnej, jak i podmiotowego traktowania ucznia. Miał to być kres komunistycznej uniformizacji i uśrednionki. Temu służyć miała m.in. wprowadzona w poprzednich podstawach programowych (*Podstawa programowa...*, 1999; *Podstawa programowa...*, 2002) międzyprzedmiotowa ścieżka edukacji regionalnej, a także idea ścieżek dydaktycznych realizowanych w pobliżu szkoły. W taki właśnie sposób skonstruowana też została nowa podstawa programowa nauczania geografii w gimnazjum (*Rozporządzenie Ministra...*, 2009). Uczniów powinien zatem nie tylko lepiej znać obszar terytorialnie mu bliższy, ale też wyjaśniać procesy geograficzne, nawet te o zasięgu globalnym, odwołując się do swojego otoczenia, np. migracje ze wsi do miast na przykładzie najbliższego miasta powiatowego, zanieczyszczenie wód – na przykładzie rzeki przepływającej przez jego miejscowość itp. Ponieważ w egzaminie centralnym obowiązuje zasada identyczności i pełnej porównywalności zadań, trudno je odnosić do skali lokalnej. Dominują zadania dotyczące skali krajowej (np. porównanie gospodarki poszczególnych województw) lub globalnej, a na maturze szczegółowe zagadnienia analizowane na mapach dotyczą obszarów, które można uznać za najbardziej znane w Polsce (Babia Góra, Pieniński Park Narodowy, Pobrzeże Kaszubskie, Mazurski Park Krajobrazowy – por. Wiczorek, Soja, 2008). Nowe tendencje w sprawdzaniu osiągnięć i ocenianiu ucznia idą zatem w przeciwnym kierunku, niż w jego kształceniu. Chcemy wykształcić umiejętności, których – jak dotychczas – nie potrafimy skutecznie sprawdzać na egzaminach centralnych. Wcześniej praktykowane matury ustne, a także dłuższe pisemne rozprawki pozwalały na lepsze ich sprawdzanie.

INNE SŁABE STRONY ZADAŃ TESTOWYCH

Wbrew pozorom, ułożenie dobrego zestawu zadań testowych nie jest wcale łatwe. Wielu nauczycieli, a przede wszystkim pracowników naukowych, którzy nie mają za sobą zajęć z bloku dydaktycznego, nie orientuje się nawet, jakie kryteria powinien taki zestaw spełniać. Do najczęstszych błędów należy nie-

przestrzeganie niezależności zadań. „Żadne zadanie nie powinno bazować na wyniku innego zadania” (Niemierko, 1975, s. 28; patrz też inne prace tego autora, np. Niemierko, 1990), tymczasem często konstruowane są zestawy, w których błąd popełniony na początku oznacza całą serię kolejnych błędów (dotyczy to np. obliczeń różnicy czasu słonecznego między miejscowościami, obliczania rzeczywistej wielkości obiektów przedstawionych na mapach, analizy zmian struktury użytkowania ziemi na podstawie wykresów itd.). Niestety, zdarza się, że tego typu postępowanie wynika z wygody sprawdzających, bo już po znalezieniu pierwszego błędu wiadomo, że dalsze zadania nie mogą być wykonane poprawnie. Innym częstym błędem jest zbyt częste umieszczanie w zadaniach jednokrotnego wyboru poprawnej odpowiedzi na drugim miejscu (szczególnie przy czterech wariantach odpowiedzi najczęściej zdarza się, że należy wybrać b). Taki dobór odpowiedzi ma uwarunkowania psychologiczne, ale warto pamiętać, że część studentów i uczniów doskonale o tym wie.

Egzaminy zewnętrzne organizowane przez Centralną Komisję Egzaminacyjną poddawane są każdorazowo bardzo starannej analizie statystycznej, obejmującej m.in. ocenę rzetelności i trafności testu, a także mocy różnicującej poszczególnych zadań. Powszechnie przyjmuje się, że dobrze skonstruowane zadania to te, których moc różnicująca wynosi 0,60 lub więcej, natomiast zadania o wskaźniku poniżej 0,30 zostały wadliwie skonstruowane (Niemierko, 1975, 1990, Zajac 1997). Wydaje się, że do wskaźników statystycznych CKE, OKE (Okręgowe Komisje Egzaminacyjne) i dydaktycy geografii przywiązują zbyt dużą wagę. Dotyczy to zwłaszcza mocy różnicującej zadań. Wskaźnik ten ma istotne znaczenie jedynie przy założeniu jednorodności materiału sprawdzanego, tymczasem geografia stanowi dyscyplinę co najmniej dychotomiczną, co znajduje odzwierciedlenie w przygotowaniu uczniów. Dobrzy uczniowie często mają jeden z działów opanowany słabiej. Jeśli np. większość uczniów ma najlepiej opanowaną geografii fizyczną, natomiast słabo polityczną, a na egzaminie większość zadań dotyczyć będzie właśnie geografii fizycznej, a tylko nieliczne politycznej, to te ostatnie zadania automatycznie mieć będą niewielką moc różnicującą, być może nawet osiągną ujemną wartość wskaźnika. Ponieważ dychotomii geografii nie da się (na szczęście) usunąć, a egzaminy zewnętrzne powinny sprawdzać pełen zakres wiadomości i umiejętności, z żadnego działu geografii nie należy rezygnować. Jedynym arbitralnym sposobem usunięcia zadań o niskiej wartości wskaźnika mocy różnicującej byłoby przyznanie jednokrotnej liczby zadań (i punktów) każdemu działowi geografii. Nie wydaje się to jednak słuszne, chociażby ze względu na ilość czasu poświęcanego poszczególnym działom naszej dyscypliny w szkole. Dydaktycy geografii powinni raczej starać się przekonać innych członków CKE i OKE o ograniczonej możliwości stosowania wskaźnika mocy różnicującej zadań przy końcowych egzaminach gimnazjalnych i na maturze z geografii. Wskaźnik ten może być natomiast bar-

dzo przydatny przy sprawdzianach przygotowanych przez nauczycieli szkolnych i akademickich, a dotyczących jednego tylko działu geografii, np. geografii miast czy też klimatu. Niestety, wtedy akurat jest on rzadko obliczany.

Upowszechnienie egzaminów pisemnych powoduje też zanik umiejętności odpowiadania, prezentacji swoich umiejętności tak przed egzaminatorem, jak i przed szerszym audytorium. Coraz częściej studenci proszą, by zamiast egzaminu ustnego mogli wykonać jakąś pracę, bądź też mieć zaliczenie pisemne. Nie wynika to wyłącznie z nadziei na internet bądź ściągnięcie od kolegi, ale i z autentycznego strachu przed odpowiedzią. Jednak w życiu stresów uniknąć się nie da, trzeba nauczyć się je przezwyciężać, a umiejętność prezentacji swych umiejętności i walorów w bezpośredniej rozmowie jest bardzo cenna w dorosłym życiu. Odchodząc od odpytywania, rezygnujemy z kształtowania tej umiejętności. Bardzo groźną pochodną zmiany sposobu oceniania jest też zanik umiejętności logicznej konstrukcji wypowiedzi. Można zaobserwować (także u studentów!), że podczas odpytywania ustnego, zamiast omawiać temat, dana osoba milczy i czeka na serię krótkich pytań, na które mogłaby odpowiadać pojedynczymi słowami. Niejednokrotnie okazuje się potem, że na każde szczegółowe pytanie rzeczywiście potrafi odpowiedzieć, ale nie potrafi skonstruować logicznej całości. Obniżenie umiejętności odpowiadania widać także na kolejnych zawodach olimpiady geograficznej, zwłaszcza jeśli polecenie dotyczy porównania obiektów bądź procesów. Bardzo często zamiast porównywania mamy wtedy dwa niezależne opisy (Plit, 2007). Wynika to m.in. z faktu, że w przypadku porównań trudniej jest dokonać standaryzacji oceny, dlatego polecenia sprawdzające tę umiejętność rzadko spotykamy w zadaniach testowych.

LEPSZE WYKORZYSTANIE MAPY NA EGZAMINACH PISEMNYCH (ZWŁASZCZA MATURALNYCH)

Zdecydowanie pozytywną stroną sprawdzianów pisemnych jest natomiast fakt ich wykorzystania do oceniania stopnia opanowania wielu innych umiejętności. W nauczaniu geografii są to przede wszystkim ćwiczenia w odczytywaniu, rozpoznawaniu i selekcjonowaniu informacji, wyjaśnianie oraz interpretowanie zjawisk i procesów przedstawionych na mapach, a także wykonywanie obliczeń matematycznych. W tym ostatnim przypadku często zaciera się jednak granica między umiejętnościami geograficznymi a matematycznymi, na poziomie egzaminu zewnętrznego kończącego gimnazjum wyraźnie na korzyść matematyki⁴. Klasycznym przykładem może być... sprawdzanie umiejętności

⁴ Wielokrotnie zwracała na to uwagę M. M. Wilczyńska, m.in. wypowiedź na Zjeździe Polskiego Towarzystwa Geograficznego we wrześniu 2008 w Białowieży.

obliczania procentów w uznawanym za „geograficzne” zadaniu 3. w części matematyczno-przyrodniczej egzaminu gimnazjalnego w 2005 roku.

Sprawdzanie umiejętności posługiwania się mapą najlepiej realizować w terenie. Takie ćwiczenia wymagają jednak pracy z niewielką grupą: przeprowadza je wielu nauczycieli, zwłaszcza na wycieczkach, są one częścią międzynarodowych olimpiad geograficznych. Cały szereg ćwiczeń z mapą (np. pomiary) lepiej realizować w warunkach kameralnych, przy czym są to ćwiczenia czasochłonne, dlatego też na ogół rezygnujemy z nich przy indywidualnej kontroli ustnej. W okresie, gdy w szkole dominowało odpytywanie i krótkie kartkówki, była to dziedzina zdecydowanie zaniedbana, a korzystanie z mapy ograniczało się niejednokrotnie do umiejętności wskazania położenia danego obiektu na mapie ściennej. Sprawdzanie umiejętności korzystania z mapy realizowane było wprawdzie od początku na zawodach finałowych olimpiady geograficznej, ale dotyczyło to jedynie wąskiej, wręcz elitarnej grupy uczniów. Od 2005 roku zadania dotyczące mapy w skali nie mniejszej niż 1:200 000 są stałym i to bardzo ważnym elementem nowej matury z geografii, doczekały się też one sporej liczby opracowań metodycznych (m.in. Soja J., Wieczorek T., 2005; Wieczorek T., Soja J., 2008). Zmusiło to nauczycieli do wprowadzenia na fakultetach odpowiednich ćwiczeń z mapami turystycznymi i topograficznymi, podczas gdy jeszcze niedawno zdarzały się szkoły, w których uczniowie nie wiedzieli, że takie mapy w ogóle istnieją. W efekcie wprowadzenia mapy na nową maturę zdecydowanie poprawił się poziom wiedzy i umiejętności kartograficznych, co zaobserwować można było także na zawodach olimpiady geograficznej.

PODSUMOWANIE

Pisemne sprawdziany, w tym nawet te najprostsze sprawdziany testowe bazujące na zasadzie zerojedynkowej (tak/nie), nie są *a priori* ani złe, ani też dobre. Mają wiele zalet i wiele ograniczeń. Do określonych celów nadają się najlepiej, do innych wcale. Niektóre z tych walorów pokazano w tekście, na temat innych znaleźć można informacje w obszernej literaturze o metodach pomiaru dydaktycznego. Podobnie przedstawia się sprawa z indywidualnym ocenianiem ustnym (odpytywaniem). Traktowanie go jako doskonałego, a już tym bardziej jedynego środka sprawdzania i oceniania stopnia opanowania wiadomości i umiejętności byłoby ogromnym błędem. Ale i ono, prócz licznych wad posiada szereg zalet, zatem i z niego nie należy rezygnować. Wskazane jest raczej stosowanie zróżnicowanych metod sprawdzania i oceniania, dostosowanych do tego, co należy sprawdzić. Należy też ostrożniej traktować zarzut braku obiektywizmu przy indywidualnej kontroli ustnej, bowiem nawet przy doskonałych testach nie da się w pełni subiektywizmu oceny wyeliminować. Trzeba też pa-

miętać, że zadaniem szkoły (i uniwersytetu) jest kształtowanie zaufania do drugiego człowieka, a nie tworzenie atmosfery PPP (państwa powszechnej podejrzliwości) i mnożenie hierarchicznych systemów kontrolnych.

Literatura

- Barwinek G., 2009, *Osiągnięcia edukacyjne z geografii i ich uwarunkowania (w świetle wyników egzaminów zewnętrznych w gimnazjach województwa świętokrzyskiego)*, rozpr. dokt., Uniwersytet Pedagogiczny im. Komisji Edukacji Narodowej w Krakowie.
- Czekańska M., 1959, *Zarys metodyki geografii*, PZWS, Warszawa.
- Hibszter A., 1996, „Mała Ojczyzna” w nauczaniu geografii, „Czasopismo Geograficzne”, LXVII, z. 1, 93-106, Wrocław.
- Informator maturalny od 2005 roku. Geografia*, 2003, Okręgowa Komisja Egzaminacyjna w Jaworznie/Centralna Komisja Egzaminacyjna, Warszawa.
- Niemierko B., 1975, *Testy osiągnięć szkolnych. Podstawowe pojęcia i techniki obliczeniowe*, WSiP, Warszawa.
- Niemierko B., 1990, *Pomiar sprawdzający w dydaktyce*, PWN, Warszawa
- Plit E., 2007, Trudna sztuka porównywania. Refleksje z XXXIII Olimpiady Geograficznej, „Geografia w Szkole”, z. 2, 24-26, Warszawa.
- Plit E., 2008, Obszar poznawany przez ucznia i jego deformacje. W nawiązaniu do kręgów poznawania przestrzeni Marii Z. Pulinowej, w: Hibszter A. (red.), *Polska dydaktyka geografii. Idee – tradycje – wyzwania*, s. 44-50, Wydział Nauk o Ziemi Uniwersytetu Śląskiego, Sosnowiec.
- Podstawa programowa kształcenia ogólnego dla sześcioletnich szkół podstawowych i gimnazjów*, Rozporządzenie Ministra Edukacji Narodowej i Sportu z dnia 15 lutego 1999 roku.
- Podstawa programowa kształcenia ogólnego dla liceów ogólnokształcących, liceów profilowanych, techników, uzupełniających liceów ogólnokształcących i techników uzupełniających*, 2002, „Dziennik Ustaw” z dn. 26.02.2002, nr 51, poz. 458, Warszawa.
- Pulinowa M.Z., 1993, Międzyprzedmiotowa ścieżka edukacji środowiskowej w polskiej koncepcji kształcenia ogólnego, w: Dudzińska M.R., Pawłowski L. (red.), *Ochrona środowiska w nauczaniu i wychowaniu. Materiały II Ogólnopolskiej Konferencji „Ochrona środowiska w nauczaniu i wychowaniu”*, 159-168, Lublin.
- Pulinowa M. Z., 2005, Zakres wiedzy geograficznej w edukacji szkolnej, w: Jania J., Jankowski A.T. (red.), *Wpływ rozwoju nauk geograficznych na proces kształcenia społeczeństwa oraz promocję wiedzy geograficznej w Polsce*, s. 31-49, Komitet Nauk Geograficznych PAN – Wydział Nauk o Ziemi Uniwersytetu Śląskiego, Sosnowiec.
- Rozporządzenie Ministra Edukacji Narodowej z dnia 23 grudnia 2008 roku w sprawie podstawy programowej wychowania przedszkolnego oraz kształcenia ogólnego w poszczególnych typach szkół*, 2009, „Dziennik Ustaw” z dn. 15.01.2009, nr 4, poz. 17, Warszawa.

- Soja J., Wieczorek T., 2005, Dlaczego Hanka nie weszła na Babią Górę? Trudności uczniów zdających maturę z geografii, „Geografia w Szkole”, z. 6, 25-32.
- Wieczorek T., Soja J., 2008, Co z tą mapą? Trudności na maturze z geografii, „Geografia w Szkole”, z. 6, 5-10.
- Zajac S., 1997, Sprawdzanie osiągnięć i ocena uczniów z geografii, w: Piskorz S. (red.), *Zarys dydaktyki geografii*, wyd. 2, 157-171, Wydawnictwo Naukowe PWN, Warszawa.

Summary

In the Polish education system, a lot of emphasis is placed on standardization and comparability of grading systems, starting from primary school to university studies. This is to be achieved by a departure from oral answers and giving preference to written tests, particularly multiple-choice questions, external examinations and uniformisation of curricula. These measures are producing many positive results. In case of geography, they serve not only to reduce the element of subjectivity in assessment (although ensuring full comparability is not possible), but also allow for practising map-reading skills. However, there are also negative side effects. They include in particular: worse quality of self-presentation skills, worse results in longer oral presentations, lack of regional education content in external lower and upper secondary school examinations (although a lot of emphasis is put on it in the curriculum). The paper also identifies certain frequent methodological errors in preparing tests in geography, and points out that all written tests require that instructions be given in a clear and detailed way; they cannot be clarified as in an oral examination, and therefore students may misunderstand what they are supposed to do – and such answers will also be evaluated.