

TOMASZ GRZEGOREK*

Czy jest jakiś pożytek ze studenckiej ankiety do ewaluacji zajęć?

1. Wprowadzenie

Człowiek uczy się całe życie, a tylko przez pewien czas w szkole według narzuconych zasad. W pierwszych chińskich szkołach zakładanych za dynastii Zhou, w ateńskim gimnazjum, frankońskiej szkole Alkuina, średniowiecznych uniwersytetach, a potem placówkach różnego stopnia powszechnej oświaty, obowiązywała zawsze ta sama zasada: nauczyciele oceniali uczniów. Dopiero niedawno ten tradycyjny hierarchiczny skos zaczęto niwelować, wprowadzając ocenę zwrotną: studencką ewaluację nauczania (*Student Evaluation of Teaching, SET*). Uczniowie zawsze mieli jakieś zdanie na temat swoich profesorów i sposobu przekazywania wiedzy, tworzyli nieformalne rankingi surowości i katalogowali charaktery jak w powieściach Niziurskiego, ale pomysł wystawiania nauczycielom oficjalnych ocen w formie naśladującej urzędową powagę jest w dziejach edukacji nowinką. Pomysł przyjął się zadziwiająco szybko i dzisiaj dyskutowana jest już przeważnie techniczna strona jego realizacji, natomiast celowość i pożytek rzadko. Ankiety studenckie upowszechniły się łatwo, ponieważ wpisały się w szeroki trend kulturowy napędzany oświeceniowym ideałem równości oraz charakterystyczną dla nowoczesnego kapitalizmu ideą profesjonalizmu. Profesjonalizacja nauczania i wymóg kontroli jakości usługi wymusił stosowanie instrumentu pomiarowego dostarczającego „twardych” danych ilościowych, natomiast ideał równości ułatwił przekazanie tego instrumentu w ręce studentów, którzy obiektywnie znajdują się najbliższej mierzonego obiektu.

Ludzie są ciekawi tego, jak są odbierani i jak wypadają na tle innych. Nauczyciele nie stanowią wyjątku i na pewno zdecydowana większość chce wiedzieć, co uczniowie czy studenci o nich sądzą i jakie opinie na ich temat krążą. Najczęściej zdobywa się takie informacje przy okazji, rozmawiając po godzinach ze studentami albo wymieniając się anegdotami ze współpracownikami. Ankiety można potraktować jako bardziej formalny sposób zaspokajania takiej ciekawości. Ale od lat siedemdziesiątych ubiegłego wieku badanie opinii studentów traktowane jest o wiele poważniej, jego wyniki są uznawane za coś w rodzaju mandatu do wykonywania zawodu nauczyciela i brane pod uwagę w decyzjach dotyczących zatrudnienia i zarobków (Chandler 1978). Ankietom przypisuje się coraz ważniejszą rolę, na przełomie wieków wyniesiono je na poziom adminis-

* Dr Tomasz Grzegorek (tomasz.grzegorek@ignatianum.edu.pl), Katedra Psychologii Ogólnej, Instytut Psychologii, Akademia Ignatianum w Krakowie

tracji państwowej, najpierw w 1998 wprowadzono narodową ankietę dla wszystkich szkół wyższych w USA i Kanadzie (*National Survey of Student Engagement*), potem w 2005 roku w Wielkiej Brytanii (*National Student Survey*) i jeszcze kilku innych krajach.

Przeprowadzono niemało badań nad rzetelnością i trafnością kwestionariuszy do ewaluacji zajęć. Ten artykuł nie jest ich wyczerpującą metaanalizą, zawiera odwołania tylko do wybranych wyników, które da się potraktować jako argumenty za taką bądź inną odpowiedzią na tytułowe pytanie. Trzeba mieć na uwadze, że jest to pytanie o ankietę w ogóle, podczas gdy w każdym badaniu analizowany jest konkretny zestaw pytań i prawie w każdym inny. Na dodatek zdecydowana większość badań odnosi się do realiów anglosaskich, dotyczy tamtejszych ankiet, tamtych studentów i wykładowców, natomiast w artykule chodzi ostatecznie o wnioski dla nas. Mimo rzucających się w oczy różnic uniwersytety w krajach Europy i Ameryki Północnej są jednak chyba wystarczająco do siebie podobne i stosowane wszędzie ankiety również, żeby przytaczanym wynikom badań ostrożnie przypisać walor względnej uniwersalności. Niestety polskich badań jest stosunkowo niewiele, niektóre istotne dla tematu zostały w tekście uwzględnione.

2. Jakich informacji ma dostarczać ankietą?

Ankieta studencka to rodzaj badania z użyciem kwestionariusza podawanego studentom. Na ankietę można spojrzeć jak na procedurę pomiarową, ponieważ dostarcza wyników liczbowych, żeby umożliwić policzenie przynajmniej wartości przeciętnej w zbiorze (np. średniej arytmetycznej na wydziale). Nie jest jednak oczywiste poziom *czego* wyraża taka liczba, *co* jest w gruncie rzeczy mierzone. Najczęściej mówi się, że typowy kwestionariusz z pytaniami w rodzaju „Czy zajęcia podniosły moje kompetencje w zakresie objętym kursem?” mierzy skuteczność nauczania (*teaching effectiveness*). Jest to pewien konstrukt teoretyczny, którego jasne zdefiniowanie sprawia niejaki problem (Estelami 2019; Hornstein 2017; Marsh 1984). Skuteczność (sprawność) w technice oznacza stosunek mocy pobranej do mocy oddanej, więc z dwóch silników elektrycznych bardziej skuteczny jest ten, który zużywa mniej prądu do wykonania tej samej pracy. Analogicznie, bardziej skuteczna metoda nauczania zapewnia uczniom więcej wiedzy i umiejętności przy identycznym wysiłku włożonym w uczenie się i trening. Niełatwo znaleźć ilościowe wskaźniki dla takiego wejścia i wyjścia: ani wysiłku nie można sprowadzić do ilości poświęconego czasu (uczniowie mogą siedzieć nad książką równie długo, ale włożyć w naukę różny wysiłek), ani rezultatu uczenia zamknąć w liczbie posiadanych informacji (uczniowie mogą mieć tyle samo informacji, ale rozumieć sens płycej lub głębiej, lepiej lub gorzej zintegrować wiedzę nabytą z już posiadaną). Rozsądną miarą wysiłku wkładanego w uczenie się może być liczba i poziom trudności wykonanych zadań, a miarą efektów uczenia wyniki zróżnicowanych sprawdzianów.

I właśnie tak wygląda od wieków szkolna edukacja: przez określony czas, zazwyczaj semestr, uczniowie muszą radzić sobie z metodycznie dobieranymi zadaniami, żeby opanować treści programu nauczania, z czego na końcu są egzaminowani, a uzyskane stopnie świadczą o skuteczności zastosowanych metod nauczania.

Egzaminy na pewno nie dostarczają idealnego wskaźnika skuteczności nauczania, mają ograniczoną rzetelność i nie uwzględniają poziomu wiedzy, z jaką studenci zaczynają naukę. Racją dla przeprowadzania ankiety studenckiej może być to, że dostarcza lepszego wskaźnika, ale wcześniej trzeba udowodnić, że jej wynik w ogóle jest wskaźnikiem, czyli wiarygodną przesłanką do wywnioskowania obiektywnej jakości nauczania. Można to zrobić bezpośrednio, po prostu obserwując na etapie konstruowania kwestionariusza, jak na poszczególne pytania odpowiadają osoby, o których wiadomo skądinąd, że były nauczane mniej lub bardziej skutecznie. Tak buduje się czasem testy psychologiczne, wiadomo na przykład, że w klinicznym teście osobowości MMPI pozycja „Bardzo rzadko kłócę się z kimś z mojej rodziny” wbrew pozorom nie mówi o ugodowości respondenta, ponieważ takie osoby wcale nie reagują tu zgodnie (za to reagują zgodnie osoby z rysem psychopatycznym). Równie dobrze może się okazać, że odpowiedź na pozycję ankiety „Wiedza była przekazywana w sposób uporządkowany” nie informuje o sposobie przekazywania wiedzy na kursie, lecz na przykład o inteligencji studenta albo nie niesie żadnej informacji. Żeby zbudować ankietę studencką metodą empiryczną, trzeba dysponować niezależnym kryterium skuteczności nauczania, ale innym niż oceny z egzaminów, skoro ankietą ma przewyższać egzaminy pod względem wiarygodności. Problem w tym, że takie kryterium nie istnieje. Związku między spostrzeżaną i rzeczywistą skutecznością nauczania można też dowodzić pośrednio przez wykazanie, że studentów stać na bezstronną refleksję i kompetentny osąd. Nie można tego przyjąć na wiarę, jednak weryfikacja empiryczna posiadania takich zdolności znowu wymaga odwołania do jakiegoś widzialnego kryterium skuteczności nauczania, a dostępne są tylko stopnie ze śródkresowych kolokwii i końcowych egzaminów, ewentualnie opisowe oceny aktywności na zajęciach. Tak czy inaczej podstawową miarą skuteczności nauczania muszą być oceny formułowane przez nauczycieli, więc stosowanie w tym celu ankiety ma sens tylko w szkołach wykluczających z jakiegoś powodu stawianie stopni. Taka konkluzja powinna być oczywista: gdy jedna klasa wypada na maturze z polskiego dużo słabiej niż inna, już nie trzeba zbierać prywatnych opinii uczniów na temat skuteczności różnych polonistów.

Żeby usprawiedliwić ankietę studencką, należy więc przypisać jej cel szerszy niż pomiar skuteczności nauczania albo zgoła inny. Faktycznie, rzut oka na pierwszy lepszy kwestionariusz ewaluacyjny każe powątpiewać, czy w praktyce zawsze chodzi o pomiar skuteczności uczenia. Umieszczane na portalach uniwersyteckich pod koniec semestru zestawy pytań często przypominają marketingowe narzędzia do sondowania zadowo-

lenia konsumentów zbudowane na założeniu, że udzielone odpowiedzi można potraktować od razu jak gotowe informacje, a nie surowy materiał wymagający ostrożnej interpretacji. Bedgood i Pollard (2001) przejrzyli ponad 800 pytań używanych w prawdziwych kwestionariuszach ewaluacyjnych i stwierdzili, że sprawdzają one tak naprawdę zadowolenie studentów i nie pozwalają dowiedzieć się wiele o czynnikach istotnych dla kształcenia. Monitorowanie satysfakcji studentów jest ważne, gdy od ich liczebności zależy sytuacja finansowa uczelni, wszelako związek satysfakcji z jakością nauczania nie jest wcale oczywisty. To prawda, że satysfakcja z zajęć nie jest redukowalna do prostej przyjemności z zaspokojenia potrzeb poznawczych czy emocjonalnych i wcale nie wyklucza przeżywania dyskomfortu związanego z wymaganiami. Ale zestresowane wysokim poziomem i jednocześnie zadowolone będą tylko takie osoby, dla których głównym celem jest zdobycie wykształcenia, a nie świadectwa wykształcenia. Proporcja studentów nastawionych na zdobycie wiedzy do studentów nastawionych na zdobycie uprawnień do wykonywania zawodu zależy od kierunku studiów, za którym kryje się wizja przyszłej pracy. W niektórych zawodach tempo kariery i zarobki są powiązane z realnymi umiejętnościami i wiedzą (np. informatyk, architekt), a w innych ten związek jest luźny (np. nauczyciel, psycholog). Zrozumiałe, że wysokie wymagania wykładowcy prędzej popsują zadowolenie osobom, które kształcą się z myślą o pracy drugiego rodzaju.

Stanley Fish wskazuje na jeszcze inną możliwą sytuację, gdy poziom zadowolenia studentów będzie mijał się z poziomem jakości zajęć: „Studenci lubią wszystko, co jest starannie rozplanowane, chcą wiedzieć dokładnie, gdzie się znajdują i z niezadowoleniem przyjmują wprowadzenie wielorakich perspektyw, zwłaszcza gdy nie ma jakiejś nadrzędnej perspektywy, która wszystkie godzi; po prostu pragną odpowiedzi. Ale czasami (choć nie zawsze) skuteczne nauczanie obejmuje celowe wywoływanie zamętu, brak przejrzystości, odmowę udzielenia odpowiedzi (...) Czasami to rozczarowanie, choć najpierw niezwykle irytujące, powinno być dla studenta oznaką, że właśnie brał udział we wspaniałym kursie, chociaż długo może sobie z tego nie zdawać sprawy. Nie trzeba dodawać, że jest mało prawdopodobne, aby tego rodzaju nauczanie otrzymywało wysokie oceny w kwestionariuszu, który nagradza dostarczanie informacji i karze pedagogikę, która bada, powoduje dyskomfort i nie prowadzi do finału.” (za: Davidovitch i Soen 2011: 33–34). W zwykłym handlu zadowolenie konsumenta prawie zawsze idzie w parze z jakością towaru, ponieważ łatwo ją ocenić: po prostu widać, czy kupiona rzecz spełnia swoją funkcję i nie jest zepsuta. Wykształcenie nie jest tego rodzaju towarem, ocena jego jakości nie jest taka łatwa, dlatego uczeń czasem może być zadowolony z zajęć marnej jakości i niezadowolony z obiektywnie solidnego nauczania. Satysfakcja studentów z tego, jak są nauczani, nie jest silnie powiązana z faktycznie wynoszoną wiedzą, współczynnik korelacji między wynikami na egzaminach a zadowoleniem z kursu przyjmuje wartości od 0,3 do 0,5 w zależności od ocenianego aspektu zajęć (Cashin 1995).

3. Czy można wierzyć wynikom ankiety?

Szkoła powinna przede wszystkim dobrze uczyć, zadowolenie studentów ze sposobu prowadzenia zajęć jest miłym bonusem, wcale niekoniecznym korelatem zdobywania wiedzy. Wynika to z powinności szkoły wobec społeczeństwa, które jako wspólnota cywilizacyjna jest jej projektodawcą, a jako organizacja podatników zwykle też materialnym fundatorem: ludzie mają prawo oczekiwać, że most, po którym idą, zaprojektował porządnie wykształcony architekt, albo że psycholog, któremu powierzają swoje zdrowie, naprawę zna się na rzeczy. Ta powinność stoi nad innymi zobowiązaniami szkoły, także nad zobowiązaniami wobec studentów opłacających czesne. Jeśli ankieta studencka ma być w pierwszym rzędzie narzędziem ewaluacji nauczania, a nie zbierania informacji o nastrojach studentów, to musi opierać się na optymistycznym założeniu, że uczniowie są w stanie sprawiedliwie ocenić jakość pracy nauczyciela. Taka ocena wymaga, po pierwsze, znajomości standardów, a po drugie, umiejętności porównania z tymi standardami obserwowanych faktów. Niektórych standardów akademickich studenci mogą zwyczajnie nie znać, na przykład mogą oczekiwać dyskusji podczas wykładu albo domagać się negocjacji wystawionych ocen i negatywnie ocenić wykładowcę, który tego nie dopuszcza. Mogą znać standardy, ale nie umieć przyłożyć ich do rzeczywistości, na przykład oceniając punktualność nauczyciela po kilku miesiącach wypełnionych różnymi kursami, student może nie pamiętać, czy ów rozpoczynał zajęcia z dziesięćminutowym poślizgiem zazwyczaj czy tylko pod koniec semestru, a jeśli nawet student skrupulatnie zanotował trzy spóźnienia, może się wahać, czy takie zachowanie już zaliczyć do ogólnej kategorii „niepunktualny”. Studenci różnią się tego rodzaju presupozycjami, a jeszcze bardziej pod względem wrażliwości emocjonalnej i nawykowych oczekiwań interpersonalnych mimowolnie ujawnianych w odpowiedzi na pytania typu „Czy zajęcia przebiegały w życzliwej i zachęcającej do aktywności atmosferze?” albo „Czy prowadzący odnosił się do studentów z szacunkiem?”. Wiadomo na przykład, że pewną rolę odgrywa tu płeć, ponieważ studentki częściej wchodzi w negatywne interakcje z wykładowcami niż studenci (Jones i Dindia 2004). Wiecznie poważny ton wykładowcy pozbawionego drygu do zabawiania słuchaczy może niektórym w ogóle nie przeszkadzać, a dla innych być objawem bufonady i nieuprzejmości. Nawet jasność i zrozumiałość przekazu nie jest doświadczeniem w takim sensie obiektywnym i odpornym na różnice indywidualne. Zdolność do śledzenia wyводу nie jest sprawą samego intelektu, lecz zależy również od czynników motywacyjnych i osobowościowych, takich jak: wpływająca na utrzymanie uwagi emocjonalna stabilność, sumienność, potrzeba domknięcia poznawczego czy samoocena. Studentka, która zbyt wcześnie zdefiniuje siebie samą jako „umysł humanistyczny”, przez złe nastawienie i brak motywacji nie będzie miała szansy zrozumieć obiektywnie uporządkowanego wykładu ze statystyki. I przeciwnie, bełkotliwa treść ubrana w zręczną elokwencję może u słuchacza wywołać euforię oświecenia tym łatwiej,

im bardziej jest zaciekawiony tematem, nastawiony na porozumienie, nieskory do krytycyzmu i niepewny siebie. Korelacja między ocenami z egzaminów a poczuciem jasności przekazu nauczyciela jest słaba i wynosi 0,35 (Feldman 1989). Studenci jednego kierunku mniej różnią się między sobą poziomem inteligencji niż pod względem upodobań, lęków, stylu bycia, marzeń i przyzwyczajęń wynikających z różnego wychowania, odmiennych charakterów i niepowtarzalnych doświadczeń. Nic dziwnego, że te same zajęcia bywają różnie oceniane przez różnych studentów. To oczywiście stawia pod znakiem zapytania sensowność porównywania wyników ankiet z różnych kursów, ponieważ różnice w ocenach mogą być skutkiem składu osobowego grupy, a nie lepiej lub gorzej prowadzonych zajęć.

W założeniu wynik ankiety studenckiej ma być funkcją wyłącznie jakości dydaktyki, tymczasem zależy on w niemałej mierze od rozmaitych innych czynników znajdujących się po stronie studenta, a także od czynników znajdujących się po stronie nauczyciela, ale innych niż dydaktyka. Bardzo duże znaczenie ma osobista charyzma wykładowcy, która nie gwarantuje przecież wysokiej jakości nauczania. Banyard, Davis i Griffiths (2000) poprosili grupę studentów o ocenę charyzmy nauczyciela oraz wypełnienie kwestionariusza do pomiaru efektywności uczenia rozbitej na dwa osobne czynniki: umiejętność nauczania oraz konstrukcja kursu. Okazało się, że oceny w dużym stopniu się pokrywały, charyzmą nauczyciela dało się wyjaśnić 69% wariacji pierwszego czynnika i o dziwo aż 37% wariacji drugiego. Tak zwany efekt doktora Foa został zademonstrowany w 1973 roku, ale przecież od zawsze wiadomo i bez badań, że forma może przykryć treść i nawet merytorycznie nędzną przemową można oczarować słuchaczy, jeśli będzie wygłoszona odpowiednim tonem i z przekonującą gestykulacją. Wyniki badań są zgodne, lepsze oceny otrzymują wykładowcy energiczni, posługujący się żywą ekspresją i pełni entuzjazmu (Bedggod i Donovan 2012). W grę wchodzić może dodatkowo jeszcze inny pospolity odruch psychologiczny, że większą sympatię budzi na ogół osoba skromna i bliska niż pewna siebie i emocjonalnie odległa. Można przypuszczać, że bardziej lubiani są nauczyciele skłonni do bratania się z uczniami niż twardo pilnujący pozycji autorytetu, zaś osobie lubianej trudniej wyrządzić przykrość niską oceną w ankiecie. Tymczasem przekazywaniu wiedzy sprzyja właściwie pojęty autorytaryzm, nie przypadkiem w badaniu IBE lepsze wyniki z egzaminu gimnazjalnego uzyskali podopieczni tych nauczycieli, którzy wyrazili zdecydowaną akceptację dla twierdzenia „Nauczyciele wiedzą znacznie więcej niż uczniowie i nie powinni pozwalać uczniom na formułowanie odpowiedzi, ponieważ mogą być niepoprawne” (Piwowski 2011). Autorytet jest narzędziem nauczyciela niezbędnym do skutecznego nauczania i nie należy się go wyrzekać tylko dlatego, że czasami bywa deformowany w irytującą feudalną pańskość „profesora”.

Wielu wykładowców uważa, że studenci wyżej oceniają łatwe kursy, na których uzyskują dobre stopnie, a niżej kursy kończące się trudnym egzaminem. Są dowody na

istnienie takiej zależności, na przykład w badaniu Maurera (2006) studenci spodziewający się słabej oceny D oceniali nauczyciela jako gorszego niż studenci spodziewający się otrzymać stopień A, B lub C. Psychologicznym podłożem obserwowanej prawidłowości raczej nie może być żądza odwetu, ponieważ ewaluacja najczęściej odbywa się przed egzaminami, ani szukanie sympatii u wykładowcy, ponieważ ankiety są anonimowe (opinie wyrażane nieanonimowo są zazwyczaj lepsze). Chodzi raczej o coś w rodzaju efektu halo wokół zajęć, gdy przypisanie jednej ważnej cechy automatycznie pociąga przypisywanie podobnych. Ewaluacja zajęć odbywa się *ex post*, dotyczy minionych zdarzeń z okresu kilkunastu tygodni, więc akt oceniania musi angażować pamięć i siłą rzeczy podlegać zniekształceniu wspomnień. Mechanizm można opisać w uproszczeniu następująco: wymagania wywołują stres, trudniej lubić zajęcia kojarzące się ze stresem, ogólne i niejednoznaczne pytania ułatwiają przywołanie wspomnień zgodnych z emocjonalnym nastawieniem, więc w tym wypadku – wspomnień negatywnych. Istnieje ryzyko swego rodzaju korupcji, gdy nauczyciel celowo zaniża wymagania, żeby zbudować reputację „przyjaznych” zajęć i uzyskać lepsze oceny w ankietach. Owocność takiej strategii nie jest jednak zagwarantowana, w większości badań zaobserwowana korelacja między stopniami z egzaminu i notami wystawianymi wykładowcy była słaba i nie przekroczyła wartości $r = 0,30$ (Davidovitch i Soen 2009). Niemniej kwerenda ogromnego zbioru danych z nieformalnego ratingu wykładowców w Internecie (www.ratemyprofessors.com) pod kątem związku oceny łatwości ukończenia kursu z oceną jego ogólnej jakości ujawniła o wiele silniejszą korelację na poziomie $r = 0,61$ (Felton, Mitchell i Stinson 2004). Należy przy tym rozróżnić pomiędzy trudnością zaliczenia kursu a nakładem pracy podczas zajęć. Nuda męczy, dlatego studenci nieco wyżej oceniają kursy, które ich aktywizują i proponują jakieś wyzwania, aczkolwiek tylko do pewnego poziomu trudności (Rudnicki i Szwed 2009).

Ważnym czynnikiem, który zniekształca wynik pomiaru jakości nauczania jest nastawienie studentów do samej procedury. Wiadomo, że ankietowanie online jest o wiele mniej efektywne od papierowego, choć nawet w tym drugim duża część studentów odmawia udziału. Na polskich uczelniach w ewaluacji zajęć bierze udział mniej niż połowa studium, zwłaszcza po semestrze letnim zwrot ankiet jest mizerny, przykładowo w roku przed pandemią na Uniwersytecie Marii Curie-Skłodowskiej ankiety online po semestrze zimowym wypełniło 18,4%, a po letnim 13,2% studentów¹. Można mieć uzasadnione wątpliwości, czy zebrane w takiej grupie opinie są reprezentatywne. Być może niezadowolone z zajęć silniej motywuje do wzięcia udziału w ankiecie, więc jej wynik będzie ostatecznie niedoszacowany, może też być na odwrót. Ewaluacja zajęć ma z reguły charakter dobrowolny, włączenie jej do obowiązków studenckich z pewnością podniosłoby odsetek zwrotów, ale wymuszona frekwencja raczej nie podniosłaby

¹ <https://www.umcs.pl/pl/raport-dla-roku-akademickiego-2018-2019,17226.htm>

ogólnej rzetelności pomiaru. Większość dobrowolnie wypełniających przyznaje, że robi to raczej pobieżnie i nie przywiązuje do odpowiedzi wielkiej wagi, nawet na uniwersytetach amerykańskich, gdzie ankieta ma najdłuższą tradycję i ugruntowany status. W badaniu przeprowadzonym na Lander Univeristy w Południowej Karolinie tylko 20% studentów potwierdziło wypełnianie kwestionariusza z tą samą skrupulatnością od początku do końca, natomiast zdecydowana większość przyznała się do głębszego namysłu tylko nad pojedynczymi pytaniami (Bassett 2017). Warunkiem poważnego podejścia do wypełnienia kwestionariusza jest wiara w to, że ma on jakieś znaczenie i przyniesie konkretny pożytek (Chen 2003). Tymczasem większość z przepytanych studentów miała wątpliwości, czy władze uczelni biorą wyniki ankiet na poważnie, podobnie w badaniu przeprowadzonym trzy lata temu na Politechnice Rzeszowskiej zaledwie 27% studentów wyraziło głębokie przeświadczenie, że ankietowanie realnie wpływa na podnoszenie jakości kształcenia (Ostasz, Woźniak i Pacana 2017).

Rola ankiety studenckiej w uczelnianym systemie kontroli jakości kształcenia musi być adekwatna do jej wiarygodności jako instrumentu pomiarowego, która jest co najmniej dyskusyjna. Interpretacji wyniku ankiety i sformułowanie na jego podstawie ostatecznej opinii o zajęciach i wykładowcy nie należy ograniczyć do porównania liczb: wyniku indywidualnego z wynikiem przeciętnym na wydziale lub na całym uniwersytecie. Trzeba jeszcze brać pod uwagę dodatkowe okoliczności, między innymi tę, że studenci mają tendencję do oceniania wyżej kursów opcjonalnych niż obligatoryjnych, zaawansowanych niż wprowadzających, przedmiotów o tematyce humanistycznej niż matematyczno-przyrodniczych (Cashin 1995). Dobry nauczyciel może uzyskać w ankiecie wynik poniżej średniej tylko dlatego, że na jakimś kierunku humanistycznym prowadził obowiązkowy kurs podstaw logiki, a nie dlatego, że brakuje mu umiejętności albo za mało się przykładał. Znaczenie ma także płeć wykładowcy, badania dowodzą, że zarówno studentki jak i studenci wyżej oceniają zajęcia prowadzone przez mężczyzn (Basow i Silberg 1987). Wyniki te interpretowane są zgodnie jako dowód na tendencyjność oceniających (*bias*), a nie różnicę w kompetencjach pomiędzy nauczycielkami i nauczycielami. Osobowość nie jest aż tak stała jak płeć, niemniej współczesna psychologia doradza raczej zaakceptować u siebie takie cechy jak drobiazgowość albo impulsywność, ponieważ w niewielkim stopniu można je zmienić. A tak się składa, że studenci reagują najbardziej właśnie na osobowość wykładowcy i można tym wyjaśnić według różnych doniesień od 50% do 80% różnic w wynikach ewaluacji zajęć (Clayson 1999). Niewątpliwie pewne typy charakterologiczne nie pasują do pewnych zawodów, niemniej przekazywanie wiedzy jest rzemiosłem czy sztuką, której można się nauczyć i wykonywać lepiej lub gorzej bez względu na to, czy jest się cholerykiem czy flegmatykiem. Skuteczność nauczania na pewno nie jest w decydującym stopniu zdeterminowana osobowością nauczyciela, natomiast postrzeganie tej skuteczności przez studen-

tów już tak. Trzeba się z tym liczyć, jeśli celem ewaluacji ma być poznanie faktycznej jakości zajęć. Skrupulatna analiza wypełnionych ankiet powinna obejmować korektę wyników ze względu na wszystkie wymienione czynniki uboczne, żeby uczciwie porównać wynik nauczyciela przedmiotu obowiązkowego o spokojnym usposobieniu z wynikiem żywiołowej nauczycielki prowadzącej opcjonalny kurs zaawansowany, trzeba ustalić siłę wpływu poszczególnych czynników i oszacować tendencyjność zniekształcającą prawdziwe wyniki. Wymagałoby to przeprowadzenia osobnych badań empirycznych dla konkretnej społeczności studentów, ponieważ siła wpływu tych czynników może być różna na różnych uczelniach ze względu na ich profil i otoczenie kulturowe w jakim funkcjonują. Można przypuszczać, że na przykład stopień tendencyjności w odniesieniu do płci będzie inny na uniwersytecie w Berlinie niż w Białymstoku, na politechnice niż na wydziale filologicznym, na uczelni katolickiej niż świeckiej, inny dzisiaj niż dziesięć lat temu. Dlatego solidne opracowanie ankiety okazuje się zadaniem niezwykle trudnym i pracochłonnym, a i tak nie gwarantuje rzetelnego pomiaru, ponieważ w grę wchodzi jeszcze zmienne zakłócające, których praktycznie nie sposób kontrolować, jak choćby nastrój studenta podczas wypełniania. Zumbach i Funke (2014) przez manipulację nastrojem uczniów skutecznie wpłynęli na wyniki ankiety ewaluacyjnej.

4. Jakie są realne korzyści z ankiety?

Szczytną ideą przyświecającą wprowadzeniu studenckiej ewaluacji zajęć było podniesienie jakości nauczania, natomiast w praktyce akcent położony został na badanie satysfakcji studentów, która wcale nie musi odzwierciedlać naprawdę ważnych parametrów dydaktyki. Dostępne w Internecie doroczne raporty z ewaluacji zajęć na polskich uniwersytetach sprowadzają się do prezentacji przeciętnej oceny uzyskanej z poszczególnych pytań i częstości odpowiedzi, zazwyczaj osobno dla instytutów albo wydziałów, ewentualnie zestawiane są wyniki aktualne z wynikami z lat ubiegłych. Przesunięcie akcentu na inny przedmiot pomiaru wynika z utartej praktyki wykorzystywania ankiety studenckiej: w mniejszym stopniu przez nauczycieli jako źródła informacji zwrotnej o zajęciach, w większym przez władze uczelni jako źródła informacji o pracowniku. Można zaryzykować twierdzenie, że na naszych uczelniach wyniki ankiet realnie mają silniejszy wpływ na politykę kadrową niż na sposób uczenia. Niskie notowania u studentów nierzadko bywają pretekstem do zakończenia współpracy z wykładowcą, chociaż chyba rzadko stanowią wystarczający powód, przynajmniej w szkołach państwowych. Również w innych krajach ankieta odgrywa rolę swego rodzaju kadrowego czujnika z alarmem uruchamianym przez wynik znacznie odbiegający w dół od przeciętnej, zwłaszcza w pytaniach zeszeregowanej i sposobu traktowania studentów (Edström 2008).

Chociaż studencka ewaluacja ma w założeniu pomagać nauczycielom w udoskonalaniu nauczania, to rzadko jest przez nich w taki sposób traktowana, ponieważ rzadko

dostarcza im pożytecznych danych. Badania pokazują (Kember i inni 2002), że ankietowanie studentów w praktyce bardzo słabo przekłada się na wprowadzanie zmian w kursach (jak wcześniej wspomniano, ankietowani to widzą). Informacje płynące z ankiety, przeważnie w postaci pseudoliczbowych odpowiedzi typu Likerta, nie wskazują jasno, co dokładnie przeszkadza studentom w bardziej owocnej nauce i jakie modyfikacje warto wprowadzić, w najlepszym razie mgliście coś sugerują. Dołączana prawie zawsze część przeznaczona na luźne uwagi niewiele pomaga, gdyż – jak pokazuje choćby analiza wyników dziesiątej edycji ankiety na Uniwersytecie Jagiellońskim (Tomanek i Bryda 2015) – studenci wykorzystują ją najczęściej na komentowanie postawy nauczyciela od strony relacji międzyludzkich i tego, „jakim jest człowiekiem”. Lektura tych osobistych uwag i w ogóle konfrontacja z ocenami pozbawionymi niuansów bywa dla wykładowców emocjonalnie trudna, dlatego wielu w ogóle się z nimi nie zapoznaje albo stosuje jakieś inne mechanizmy obronne (Moore i Kuol 2005). W wywiadach na ten temat przeprowadzonych z pracownikami jednej z kanadyjskich uczelni większość przyznała się do skutków czytania ankiet na swój temat w postaci dolegliwości takich jak silne napięcie w karku i szczęce, kłopoty z zasypianiem, poczucie niepewności i przytłoczenia (Moralejo i inni 2019). Wielu wykładowców postrzega ankietę studencką jako dokuczliwość, a nie pomoc w nauczaniu, i nie wierzy w rzetelność narzędzia, które ich zdaniem służy przede wszystkim studentom do odreagowania frustracji (Eckhaus i Davidovitch 2019). Jest coś niesprawiedliwego w honorowaniu subiektywnych przeżyć studentów i jednocześnie lekceważeniu przeżyć nauczycieli.

Wykorzystywanie wyników ankiety studenckiej w decyzjach administracyjnych i personalnych przez władze uczelni nie jest tylko utartą nieformalną praktyką, lecz ma umocowanie prawne. W Polsce kwestionariusze do oceny zajęć pojawiły się dopiero w latach dziewięćdziesiątych ubiegłego wieku, a upowszechniły po przystąpieniu naszych uniwersytetów do Procesu Bolońskiego. Miejsce dla ankiety studenckiej w krajobrazie życia akademickiego precyzyjnie wskazał dokument podpisany w Bergen w 2005 roku, gdzie pierwszy raz sformułowano wyraźnie standardy kształcenia na europejskich uniwersytetach i wskazówki do podnoszenia jakości (*Standards and Guidelines for Quality Assurance in the European Higher Education Area*). Nowelizacja dokumentu z 2015 musiała być brana pod uwagę przy tworzeniu najnowszej ustawy o nauce i szkolnictwie wyższym zwanej Konstytucją dla Nauki. Ustawa wymaga, żeby każdy wykładowca przynajmniej raz na cztery lata został rozliczony z wypełniania swoich obowiązków. Dokładne kryteria oceny okresowej pracownika ma ustalać rektor, ale ustawa nie pozostawia mu wyboru w kwestii przeprowadzenia ankiety studenckiej: „Uczelnia umożliwia studentom i doktorantom dokonanie co najmniej raz w roku akademickim oceny nauczyciela akademickiego w zakresie wypełniania przez niego obowiązków związanych z kształceniem.” (art. 128, ust. 4). Co więcej, ustawa zobowiązuje do uwzględnienia wy-

ników ankiety przy ocenie pracownika: „Przy dokonywaniu oceny okresowej uwzględnia się ocenę, o której mowa w ust. 4.” (art. 128, ust. 6). Zapisy te ustanawiają wynik ewaluacji studenckiej nieusuwalną składową ocenę okresowej nauczyciela, więc mówiąc bez ogródek, czynią studenta sędzią swojego profesora. Dyrektorowi instytutu albo dziekanowi nie wolno potraktować opinii studentów tylko jako dodatkowego źródła informacji, które może wziąć pod uwagę, ale wcale nie musi, chociaż nie jest na szczęście narzucone, w jakim stopniu musi. Umyślne czy nieumyślne sformułowanie przepisu w taki sposób, że student wystawia ocenę nauczycielowi, a nie, że dostarcza informacji przydatnych władzom uczelni do wystawienia oceny, może okazać się brzemiennie w skutkach, gdy w ministerialnej parametryzacji pojawi się wspólna ankieta dla wszystkich szkół wyższych stanowiąca zunifikowaną miarę jakości dydaktyki. Teraz każdy uniwersytet albo nawet wydział produkuje własną ankietę z innym zestawem pytań i możliwych odpowiedzi, więc nie da się porównywać poziomu jakości dydaktyki między uczelniami, co pewnie już długo nie potrwa. Na marginesie warto odnotować, że taka ogólna państwowa ankieta studencka poza wszystkim jest też elementem szeroko pojętej polityki i instrumentem rządomyślności (od *gouvernementalite* Faucaulta). Na przykład da się zauważyć, jak za pomocą Narodowej Ankiety Studenckiej (NSS) władze Wielkiej Brytanii kontrolują i dyscyplinują instytucje edukacyjne, narzucając dyskurs i wskazując, co ma być dla nich istotne, a co nie (Thiel 2019).

Na razie ankieta nie odgrywa w polskiej rzeczywistości aż tak ważnej roli, niemniej jej wpływ na atmosferę panującą na uniwersytetach jest wyraźnie zauważalny. Choć w niektórych przypadkach wyniki ewaluacji studenckiej okazują się przydatne, to biorąc wszystko pod uwagę, jest z niej więcej szkody niż pożytku. Ta opinia nie może dziwić, zważywszy na opisane wady ankiety jako narzędzia pomiarowego oraz w praktyce wątki związek ze sprawą doskonalenia dydaktyki. To prawda, że ankieta pozwala wyłapać patologie, ale przecież łatwo ją w tej roli zastąpić. Istnieją różne kanały komunikacji, z których studenci mogą skorzystać, gdy na zajęciach dzieje się coś naprawdę niepokojącego: jak nie przez bezpośredni kontakt z zastępcą dyrektora do spraw studenckich, to przez samorząd studencki, który ma swoich przedstawicieli w rozmaitych gremiach, często już w radzie instytutu. W porównaniu ze zwykłymi możliwościami ujawnia się nawet wada ankiety jako zbyt czułego aparatu ostrzegawczego ze skłonnością do fałszywych alarmów. Osobiste złożenie skargi – nawet przy zapewnionej poufności – wymaga od studenta trochę cywilnej odwagi i wzięcia odpowiedzialności za to, co mówi, na pewno w większym stopniu niż impulsywne zaznaczenie skrajnej odpowiedzi w kwestionariuszu czy sformułowanie kąśliwej uwagi, wytrenowane na forach internetowych. Stworzenie specjalnej formalnej drogi anonimowego powiadamiania bywa jedynym sensownym wyjściem wtedy, gdy skala patologii przekracza zwykły poziom i ma charakter systemowy, a jawny akt sprzeciwu nieuchronnie pociąga represje. Sytuacja studenta na polskiej

uczelni na pewno nie jest tak beznadziejna, a bezspornie zdarzające się nadużycia i ekscesy obyczajowe z udziałem wykładowców jeszcze nie pozwalają twierdzić, że środowisko akademickie jest systemowo zdeprawowane i wymaga jakiegoś nadzwyczajnego monitoringu.

Główna szkoda wyrządzana przez ankiety polega jednak – zdaniem wielu krytyków – na czymś innym, mianowicie na wprowadzeniu na uniwersytet stosunków znanych z rynku: profesor oferuje towar w postaci wiedzy, a student jest klientem, któremu przysługują wszystkie prawa konsumenckie z reklamacją włącznie (Avador 2006). Jest to fakt niewątpliwy, ale nie należy się na to obrażać, ponieważ z pewnej perspektywy wiedza jest towarem, skoro ma wymiarną cenę i bywa przedmiotem wymiany. Statusu wykładowcy nie rujnuje uwikłanie go w rodzaj transakcji, lecz osłabienie fundamentalnego dla relacji nauczyciel – uczeń elementu zaufania, które w relacji sprzedawca – klient nie jest tak ważne, gdyż oparta jest na wymianie dóbr o namacalnej wartości, a nadto obudowana jest prawnymi gwarancjami. Nauczycielowi trzeba ślepo ufać z tej prostej przyczyny, że przekazuje wiedzę zgromadzoną w postaci zdań, które odnoszą się do rzeczywistości w praktyce niemożliwej do sprawdzenia przez ucznia. Student musi przyjąć na wiarę usłyszane na wykładzie albo przeczytane w podręczniku twierdzenie, że wokół aksonu komórki nerwowej jest izolacja elektryczna z mieliny, gdyż do końca życia prawdopodobnie nie zobaczy tego na własne oczy. Zresztą obserwacje i eksperymenty na ogół są niezrozumiałe bez objaśnienia, więc nawet wtedy potrzebne jest zaufanie do autorytetu. Uczniom zaleca się wprowadzić tak zwany zdrowy krytycyzm, ale ostatecznie podstawą nieufności wobec zdania jednego profesora może być tylko wiara w zdanie innego profesora, które zostało zapisane w książce albo opublikowane na „poważnym” portalu. Decyzja o tym, żeby zaufać wiąże się z oddaniem jakiejś części władzy nad sobą, a władza jest relacją niesymetryczną: jeśli jeden człowiek sprawuje władzę nad drugim, to ten drugi nie może jednocześnie sprawować władzy nad pierwszym. Studencka ewaluacja dydaktyki niszczy zaufanie, niszcząc tę asymetrię, jako że wystawianie oceny obciążonej sankcją jest formą sprawowania władzy. Wystarczy wyobrazić sobie inne hipotetyczne społeczne narzędzia „monitorowania jakości”, żeby stało się jasne, jak to działa: semestralny kwestionariusz oceny współpracy z dyrektorem instytutu wypełniany anonimowo przez jego pracowników do wiadomości rektora albo rządowy formularz dla dzieci do oceny rodziców i tego, jak wywiązują się z rodzicielskich obowiązków. Tego rodzaju ankiety bez wątpienia nadwątląbyby wzajemne zaufanie i na głębokim poziomie zniszczyły relacje między dyrektorem i pracownikami oraz rodzicami i dziećmi. A przecież logika studenckiej ewaluacji nauczania jest właśnie taka. Druga analogia pasuje prawie dokładnie, ponieważ ustawa wśród obowiązków nauczyciela oprócz kształcenia wymienia też wychowywanie studentów (art. 115). W życiu społecznym dzieje się czasem coś podobnego do prawidłowości odkrytej przez Heisenberga

na poziomie mikrofizycznym: akt pomiaru wpływa na mierzony obiekt. Kiedy student mierzy relację nauczyciel – uczeń, to nieuchronnie zmienia tę relację.

Naturalnie każdy pracodawca musi mieć oko na swoich pracowników, czy wywiązują się z obowiązków, za które są opłacani. Trzeba jednak pamiętać, że samo kontrolowanie jakości nie prowadzi automatycznie do jej podnoszenia. Sprawdzonej metodą ułatwiającą wprowadzanie realnych modyfikacji w nauczaniu i udoskonalanie dydaktyki jest obserwacja koleżeńska (*peer observation*) i hospitacja zajęć, która w polskich szkołach z powodu nawyków mentalnych traktowana jest często jako konieczna przykrość i chyba rzadko bywa owocna. Rzecz jasna obserwacje studentów też mogą być bardzo pomocne, ale jak pokazuje choćby doświadczenie Politechniki w Sztokholmie o wiele lepiej robić to metodą swobodnego wywiadu niż za pomocą kwestionariusza z pytaniami zamkniętymi (Edström 2008). Zdobywanie informacji o tym, co studentom sprawia faktyczne kłopoty i przeszkadza w nauce, ma niewiele wspólnego z przymuszaniem ich do wystawiania sformalizowanych ocen na wymyślonych wymiarach. Pora przestać się wygłupiać z kierowaniem do studentów pytań w stylu: Czy sposób, w jaki prowadzone były zajęcia, zachęcał do samodzielnego myślenia? Czy prowadzący umiejętnie wiązał teorię z praktyką? Czy literatura rekomendowana do przedmiotu dobrze uzupełniała wykład². Bo na czym ma polegać „samodzielne myślenie” w przyswajaniu wiedzy z anatomii układu pokarmowego? Co jest „teorią”, a co jest „praktyką” na zajęciach z historii literatury XX wieku i jak ma wyglądać „umiejętne łączenie” jednego z drugim? Do czego ma się odwołać niedawny licealista, recenzując dobór lektur do wykładu z termodynamiki? Studenci udzielają odpowiedzi na takie pytania, wyobrażając sobie najróżniejsze rzeczy, podobnie jak człowiek umiejętnie zagadnięty przez reportera zawsze coś odpowie na pytanie, czy miłość przypomina bardziej jezioro czy las. Zebrane odpowiedzi i policzone średnie albo mediany stanowią jakąś iluzoryczną wiedzę, której nie wolno traktować poważnie w instytucji powołanej przeciwko specjalnie do rozwiewania iluzji. Podsumowując, pożytki z regularnie przeprowadzanej studenckiej ewaluacji nauczania dzielą się na fikcyjne oraz faktyczne, te ostatnie nieliczne i możliwe do uzyskania w inny sposób. Są też szkody, którym można zapobiec jedynie przez zaprzestanie ankietowania.

Literatura

- Avador S. (2006). For whom and for what? Bias factors in student's evaluations at educational colleges and the question of benefit for the evaluated and the organization. *Dapim* 41: 10–37.
- Banyard P., Davies M., Griffiths M. (2000). The Validity of Student Evaluation of Teaching in Higher Education: Love me, love my lectures? *Assessment & Evaluation in Higher Education* 25(4): 397–405.

²To prawdziwe pytania z ankiet przeprowadzonych na Uniwersytecie Medycznym w Katowicach, AGH w Krakowie i Uniwersytecie Ekonomicznym w Poznaniu (przytoczone za: Hońko 2014).

- Basow S., Codos S., Martin J. (2013). The effects of professors' race and gender on student evaluations and performance. *College Student Journal* 47(2): 352–363.
- Basow S.A., Silberg N.T. (1987). Student Evaluations of College Professors: Are Female and Male Professors Rated Differently? *Journal of Educational Psychology* 79: 308–314.
- Bassett J., Cleveland A., Acorn D., Nix M., Snyder T., (2017). Are they paying attention? Students' lack of motivation and attention potentially threaten the utility of course evaluations. *Assessment & Evaluation in Higher Education* 42(3): 431–442.
- Bedggood R.E., Donovan J.D. (2012). University performance evaluations: what are we really measuring? *Studies in Higher Education* 37(7): 825–842.
- Bedggood R.E., Pollard R. (2001). Customer satisfaction and teacher evaluation in higher education. In *Proceedings of the tenth biennial World Marketing Congress: Global marketing issues at the turn of the millennium*. Red. H.E. Spotts, H.L. Meadow i S.M. Scott. Cardiff, Wales: Academy of Marketing Science.
- Cashin W.E. (1995). Student Ratings of Teaching: The Research Revisited, *IDEA Paper*. 32, Manhattan, KS: Center for Faculty Evaluation and Development, Division of Continuing Education, Kansas State University
- Chandler T. (1978). The questionable status of student evaluations of teaching. *Teaching of Psychology* 5(3): 150–152.
- Chen Y., Hoshower L.B. (2003). Student Evaluation of Teaching Effectiveness: An Assessment of Student Perception and Motivation. *Assessment and Evaluation in Higher Education* 28: 71–88.
- Clayson D.E. (1999). Students' evaluation of teaching effectiveness: Some implications of stability. *Journal of Marketing Education* 21: 68–75.
- Davidovitch N., Soen I.D. (2009). Myths And Facts About Student Surveys Of Teaching The Links Between Students' Evaluations Of Faculty And Course Grades. *Journal of College Teaching & Learning* 6(7): 41–49.
- Davidovitch N., Soen D. (2011). Student Surveys And Their Applications In Promoting Academic Quality In Higher Education. *Administrator Journal of College Teaching & Learning* 8(6): 31–46.
- Eckhaus E., Davidovitch N. (2019). How do Academic Faculty Members Perceive the Effect of Teaching Surveys Completed by Students on Appointment and Promotion Processes at Academic Institutions? A Case Study. *International Journal of Higher Education* 8(1): 171–180.
- Edström, K. (2008). Doing course evaluation as if learning matters most. *Higher Education Research & Development* 27(2): 95–106.
- Estelami H., (2016). An Exploratory Study Of The Effects Of Online Course Efficiency Perceptions On Student Evaluation Of Teaching (SET) Measures. *American Journal of Business Education* 9(2): 67–82.
- Feldman K.A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multi-section validity studies. *Research in Higher Education* 30: 583–645.
- Felton J., Mitchell J., Stinson M. (2004). Web-based Student Evaluations of Professors: The Relation between Perceived Quality, Easiness and Sexiness. *Assessment & Evaluation in Higher Education* 29: 91–108.
- Hońko S. (2014). Ocena procesu kształcenia na uczelniach – diagnoza i propozycje zmian. *Folia Pomeranae Universitatis Technologiae Stetinensis. Oeconomica* 75: 57–66.

- Hornstein H.A. (2017). Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance. *Cogent Education* 4.
- Jones, S., Dindia, K. (2004.). A meta-analytic perspective on sex equity in the classroom. *Review of Educational Research* 74(4): 443-471.
- Kember D., Leung D.Y.P., Kwan K.P. (2002). Does the use of student feedback questionnaires improve the overall quality of teaching? *Assessment and Evaluation in Higher Education* 27(5): 411–425.
- Marsh H.W. (1984). Students' Evaluations of University Teaching: Dimensionality, Reliability, Validity, Potential Biases, and Utility. *Journal of Educational Psychology* 76(5): 707–754.
- Maurer T.W. (2006). Cognitive Dissonance or Revenge? Student Grades and Course Evaluations. *Teaching of Psychology* 33: 176–179.
- Moore S., Kuol, N. (2005). A punitive bureaucratic tool or a valuable resource? Using student evaluations to enhance your teaching. *Emerging Issues in the Practice of University Learning and Teaching*. Eds. O'Neill, G., Moore, S., McMullin, B. Dublin: AISHE.
- Moralejo L., Andersen E., Hilsmann N., Kennedy L. (2019). Measuring Student Responses in and Instructors' Perceptions of Student Evaluation of Teaching (SETs), Pre and Post Intervention. *The Canadian Journal for the Scholarship of Teaching and Learning* 10(3).
- Ostasz G., Woźniak J., Pacana A. (2017). Badanie satysfakcji studentów z systemu ankietyzacji. *Modern Management Review* 24(4): 91–101.
- Piwowski R. (2011). Nauczyciele a osiągnięcia szkół (na podstawie danych TALIS i wyników egzaminu gimnazjalnego). *Analizy IBE/02/2011*. dostępne na: <http://www.ibe.edu.pl/pl/component/content/category/17-wydawnictwo>
- Rudnicki S., Szwed A. (2009). Ocenianie oceniających. Trzy pytania o ewaluację zajęć dydaktycznych. *Kultura i Polityka* 5: 127–139.
- Thiel J. (2019). The UK National Student Survey: An amalgam of discipline and neo-liberal governmentality. *British Educational Research Journal* 45(3): 538–553.
- Tomanek K., Bryda G. (2015). Odkrywanie postaw dydaktyków zawartych w komentarzach studenckich. Analiza treści z zastosowaniem słownika klasyfikacyjnego. *Przegląd Socjologiczny* 64(4): 51–81.
- Williams J., Kane D. (2009). Assessment and Feedback: Institutional Experiences of Student Feedback, 1996 to 2007. *Higher Education Quarterly* 63(3): 264–286.
- Zumbach J., J. Funke. (2014). Influences of Mood on Academic Course Evaluations. *Practical Assessment, Research & Evaluation* 19(4): 1–12.

Is there any benefit from the student evaluation of teaching survey?

The idea of students giving official grades to teachers is new in the centuries-old history of education. Today, the technical side of its implementation is usually discussed, while the purposefulness and usefulness are rarely in focus. The article provides a justification for the view that the student survey is a poor measuring instrument and it is used contrary to its intended use. First, it is unclear what exactly it measures, but it does so imprecisely as there are many known factors that are potentially sources of bias. Second, only to a small extent is the survey actually used to improve the quality of teaching. A separate argu-

ment was also given for the opinion that student evaluation of classes not only brings little benefit, but also causes some harm.

Key words: Student Evaluation of Teaching, SET, student survey, teaching effectiveness, education quality monitoring