



EWALUACJA

jakości dydaktyki w szkolnictwie wyższym

metody | narzędzia | dobre praktyki

EWALUACJA

jakości dydaktyki w szkolnictwie wyższym
metody| narzędzia | dobre praktyki

EWALUACJA

jakości dydaktyki w szkolnictwie wyższym
metody| narzędzia | dobre praktyki

Redakcja

Wojciech Przybylski, Seweryn Rudnicki, Anna Szwed

egzemplarz bezpłatny

KRAKÓW 2010

Publikacja została przygotowana w ramach Projektu Rozwojowego „Kompetentnie ku przyszłości” współfinansowanego przez Unię Europejską w ramach Europejskiego Funduszu Społecznego Działanie 4.1. Wzmocnienie i rozwój potencjału dydaktycznego uczelni oraz zwiększenie liczby absolwentów kierunków o kluczowym znaczeniu dla gospodarki opartej na wiedzy, Poddziałanie 4.1.1 Wzmocnienie potencjału dydaktycznego uczelni.

Tytuł:

Ewaluacja dydaktyki w szkolnictwie wyższym.

Metody, narzędzia, dobre praktyki

ISBN 978-83-60005-14-9

Redakcja:

Wojciech Przybylski

Seweryn Rudnicki

Anna Szwed

Recenzentka:

prof. Renata Grochowska

Spis treści

Wstęp _____	7
<i>Wojciech Przybylski, Seweryn Rudnicki, Anna Szwed</i>	
Jakość kształcenia w szkolnictwie wyższym – uwagi wstępne _____	11
<i>Magdalena Jelonek, Joanna Skrzyńska</i>	
Ocenianie oceniających. Trzy pytania o ewaluację zajęć dydaktycznych _____	35
<i>Seweryn Rudnicki, Anna Szwed</i>	
Jak skonstruować dobre narzędzie do ewaluacji jakości zajęć dydaktycznych? Refleksje praktyczne _____	51
<i>Grzegorz Bryda, Magdalena Jelonek, Barbara Worek</i>	
System Ewaluacji Jakości Kształcenia na Uniwersytecie Warszawskim jako przykład rozwoju kultury ewaluacyjnej na dużej uczelni wyższej _____	77
<i>Mikołaj Jasiński</i>	
System ewaluacji jakości zajęć dydaktycznych w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera _____	99
<i>Wojciech Przybylski, Seweryn Rudnicki, Anna Szwed</i>	
Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia na Uniwersytecie Warszawskim – przykład dobrych praktyk ewaluacji jakości kształcenia wyższego. Przegląd kluczowych aspektów badania _____	123
<i>Albert Izdebski</i>	
Wykorzystanie danych z Uniwersyteckiego Systemu Obsługi Studiów oraz systemu Internetowej Rejestracji Kandydatów w procesie ewaluacji jakości kształcenia _____	143
<i>Tomasz Zajac</i>	
Biogramy _____	165

Wojciech Przybylski

Seweryn Rudnicki

Anna Szwed

Wstęp

Niezależnie od tego, czy ewaluacja jest, czy nie jest administracyjną i menadżerską modą, dbanie o jakość dydaktyki staje się palącą koniecznością. Polskie szkolnictwo wyższe, mimo imponujących wskaźników ilościowych, wciąż nie może pochwalić się wysoką jakością, jakkolwiek byśmy go zdefiniowali. Stąd też presja na podnoszenie jakości pracy dydaktyków, która, rzecz jasna, ma źródła zarówno wewnętrzne (wielość uczelni i niż demograficzny), jak i zewnętrzne (konkurencja globalna, polityki Unii Europejskiej). Wzrost zainteresowania ewaluacją jakości dydaktyki pozostaje jednak w tym zakresie w rażącej asymetrii względem większości uczelni. Tylko nieliczne jednostki akademickie podjęły się systematycznej adaptacji praktyk ewaluacyjnych na własne potrzeby. Niniejsza publikacja ma w założeniach wypełnić próżnię, wynikającą z niewielkiej liczby opracowań z zakresu ewaluacji jakości dydaktyki w języku polskim, prezentując doświadczenia najbardziej zaawansowanych w tej dziedzinie polskich uczelni.

Artykuł wprowadzający – *Jakość kształcenia w szkolnictwie wyższym...* – autorstwa Magdaleny Jelonek i Joanny Skrzyńskiej rysuje horyzont rozważań nad jakością kształcenia na uczelniach wyższych. Autorki starają się ją pokazać w szerszym kontekście roli uczelni wyższych w kształtowaniu rozwoju na poziomach regionalnym, krajowym i unijnym. Ponieważ polskie szkolnictwo wyższe okupiło sukces ilościowy obniżeniem jakości, koncentracja na jakości dydaktyki wydaje się oczywistością i koniecznością zarazem. Ten prosty wniosek wsparty jest analizą Strategii Lizbońskiej i innych dokumentów dotyczących gospodarki opartej na wiedzy. Autorki podkreślają, że jakość w szkolnictwie wyższym może być definiowana

na wiele różnych sposobów, zależnie od punktu widzenia danej grupy interesariuszy czy też misji konkretnej uczelni. Jelonek i Skrzyńska podejmują udaną próbę uporządkowania tych definicji oraz budowanych na ich podstawie modeli oceny funkcjonowania uczelni. Praktyczny wniosek jest następujący: niemożliwe jest opracowanie uniwersalnego systemu ewaluacji jakości kształcenia – każda jednostka musi wypracować model dopasowany do swoich potrzeb.

W kolejnym tekście – *Ocenianie oceniających* – Anna Szwed i Seweryn Rudnicki, w oparciu o analizę dobrych praktyk uczelni anglosaskich, dyskutują trzy fundamentalne kwestie związane z ewaluacją zajęć dydaktycznych. Po pierwsze, wskazują na konieczność triangulacji danych ewaluacyjnych, czyli nieopierania się wyłącznie na jednym ich źródle. Tak pozyskane informacje muszą być następnie interpretowane zależnie od formatywnego lub sumatywnego celu ewaluacji. Z tej konstatacji wynika druga kwestia i rekomendacja zarazem: ewaluacja musi mieć charakter systemowy. Nie może być narzędziem lub nawet zbiorem narzędzi, a funkcjonalnie powiązaną całością, wliczając w to system oceny pracowniczej i planowanie strategiczne uczelni. Ponieważ jednak trudno sobie wyobrazić system ewaluacji niezawierający ocen studentów, odrębna część artykułu Rudnickiego i Szwed poświęcona jest wiarygodności tego źródła danych. W oparciu o bogatą literaturę przedmiotu autorzy uznają ankiety studenckie za zasadniczo trafne i rzetelne narzędzie do uzyskiwania informacji o jakości sposobu prowadzenia zajęć.

Po artykułach wprowadzających proponujemy serię tekstów opisujących konkretne rozwiązania, od systemów ewaluacji po szczegółową analizę narzędzi badawczych. Tekst Grzegorza Brydy, Magdaleny Jelonek i Barbary Worek *Jak skonstruować dobre narzędzie do ewaluacji jakości zajęć dydaktycznych?* prezentuje etapy tworzenia dobrego narzędzia ewaluacyjnego oraz zawiera bardzo praktyczną, ilustrowaną przykładami, listę najczęściej popełnianych błędów w tym zakresie. Autorzy zapoznają nas również z elementami systemu ewaluacji wdrażanego przez Uniwersytet Jagielloński, który jako jedyna z prezentowanych z tym zbiorze uczelni, zdecydował się na elektroniczną formę ankietyzacji studentów w obszarze jakości prowadzenia zajęć.

Tekst Mikołaja Jasińskiego *System Ewaluacji Jakości Kształcenia na Uniwersytecie Warszawskim...* jest syntetycznym opisem systemu ewaluacji jakości kształcenia wdrażanego od kilku lat na Uniwersytecie Warszawskim. Pokazuje zalety powołania niezależnej, dedyko-

wanej temu celowi jednostki, co może być rozwiązaniem wartym polecenia szczególnie dla największych polskich uczelni. Poza ankietami oceniającymi jakość zajęć, realizowanymi w formie papierowej (lecz zbieranymi i analizowanymi centralnie, co udaje się mimo wielkości uczelni), artykuł przedstawia m.in. internetowe narzędzie oceniające różne aspekty działalności UW czy program „Trajektoria studiowania”, pozwalający śledzić losy studentów od kandydata po absolwenta. Narzędzia te mają być użyteczne zarówno do definiowania i analizy lokalnych problemów, jak i planowania strategicznego w skali całego uniwersytetu. Analiza Jasińskiego dotyka też kulturowego kontekstu wprowadzania systemu ewaluacji na uczelni publicznej.

Zupełnie inna rzeczywistość akademicka jest właściwa dla Wyższej Szkoły Europejskiej im. ks. Józefa Tischnera – uczelni niepublicznej i z założenia kameralnej. Autorzy kolejnego tekstu *System ewaluacji jakości zajęć dydaktycznych w Wyższej Szkole Europejskiej...* i zarazem redaktorzy tego tomu pokazują, że również w takim kontekście możliwe jest opracowanie kompleksowego systemu ewaluacji, odpowiadającego na potrzeby szkoły, która wybrała jakość dydaktyki jako główny czynnik swojej przewagi konkurencyjnej na rynku edukacyjnym. Równocześnie, ze względu na wspólnotowy charakter uczelni, duże znaczenie odgrywają jakościowe elementy systemu, spełniające przede wszystkim funkcję formatywną. Oba wyżej wymienione artykuły malują ciekawy pejzaż systemów ewaluacji w skrajnie różnych jednostkach, mających mimo to wiele wspólnych cech.

Artykuł Alberta Izdebskiego *Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia na Uniwersytecie Warszawskim...* jest pogłębioną prezentacją i analizą wspomnianej już wcześniej Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia na Uniwersytecie Warszawskim. Autor opisuje szczegółowo proces przygotowania oraz realizacji badania, które – dzięki części wspólnej pytań powtarzanych co rok – stwarza możliwość analizy trendów. Równocześnie każda edycja ma swoje tematy szczegółowe, pozwalające dokładniej przyjrzeć się wybranym aspektom funkcjonowania uczelni. Warta uwagi jest szeroko zakrojona promocja tego przedsięwzięcia, której efektem jest wysoka, jak na ankietę internetową, stopa realizacji.

Przykładu potencjalnie interesującej i konstruktywnej analizy danych zastanych dostarcza Tomasz Zająć, autor ostatniego tekstu w zbiorze *Wykorzystanie danych z Uniwersyteckiego Systemu Obsługi Studiów...* Korzysta on z faktu, że elektroniczne bazy danych kumulują

olbrzymią ilość informacji o ścieżkach edukacyjnych poszczególnych osób. Próbką możliwości połączenia łatwego pozyskiwania olbrzymiej ilości danych oraz zaawansowanych metod statystycznych jest analiza procesu rekrutacji na Uniwersytet Warszawski, uwzględniająca „rzeczywistą” popularność poszczególnych kierunków i ich konkurencyjność względem siebie.

Tom *Ewaluacja jakości dydaktyki w szkolnictwie wyższym. Metody, narzędzia, dobre praktyki* został stworzony z myślą o wszystkich badaczach i praktykach zajmujących się jakością kształcenia w jednostkach akademickich. Tym, którzy chcieliby szybko i w prosty sposób zaadaptować przedstawiane tu rozwiązania, dedykujemy fragment jednego z artykułów: „Kultury ewaluacyjnej, którą zdefiniować można jako formalne i nieformalne zakorzenienie (również na poziomie zwyczajów, praktyk i sposobów myślenia) zasad wzajemnego i opartego na jasnych kryteriach oceniania oraz nie mniej ważnej otwartości na wyniki tej oceny, nie buduje się ani łatwo, ani szybko”. Tym bardziej jednak warto zacząć jak najwcześniej, do czego serdecznie zachęcamy.

Magdalena Jelonek

Joanna Skrzyńska

Jakość kształcenia w szkolnictwie wyższym – uwagi wstępne

Niniejszy artykuł, mający charakter teoretycznego wprowadzenia do debaty praktyków nad problemem mierzenia, oceny i modyfikacji jakości kształcenia w szkołach wyższych, poświęcony został dwóm fundamentalnym dla rozumienia pojęcia jakości w szkolnictwie wyższym kwestiom, tj. definiowaniu pojęcia jakości oraz decyzji, jaką musimy podjąć, wybierając odpowiedni model oceny funkcjonowania szkół wyższych. Artykuł rozpoczyna się od zarysowania szerokiego kontekstu prowadzonych rozważań poprzez próbę przedstawienia zmian, jakie projektowane są w europejskim i polskim obszarze szkolnictwa wyższego. Zmiany te implikują konieczność przedefiniowania pojęcia „jakość w szkolnictwie wyższym”, a także podjęcie wyzwania modyfikacji podstawowych modeli jej oceny. Znaczna część artykułu poświęcona została problemom definicyjnym; scharakteryzowane są w nim relacje pomiędzy pojęciami, takimi jak: jakość, skuteczność i efektywność, a także podkreślona została konieczność szerokiego ujęcia koncepcyjnego i operacyjnego charakteryzowanego pojęcia.

Wprowadzenie

W ciągu ostatnich 20 lat szkolnictwo wyższe w Polsce było jednym z najbardziej dynamicznie rozwijających się obszarów, dynamika zmian miała jednak głównie charakter ilościowy, w mniejszym zaś zakresie jakościowy. Ilościowemu rozwojowi szkolnictwa towarzyszy wiele poważnych problemów, których rozwiązanie wymaga pilnego podjęcia środków zaradczych. Przekrojowe zestawienie problematycznych kwestii znaleźć można w stosunkowo krytycznym dla Polski raporcie OECD (Fulton i inni 2007), w *Diagnozie stanu szkolnictwa wyższego w Polsce* (2009), w *Strategii rozwoju szkolnictwa wyższego 2010–2020* (2009) czy w analizie potencjału regionalnego szkół wyższych, zatytułowanej *Ocena wpływu projektów zrealizowanych dzięki funduszom europejskim przez uczelnie wyższe z terenu województwa małopolskiego na wzrost konkurencyjności gospodarczej regionu oraz wzmocnienie potencjału sektora badawczo-naukowego w województwie* (Górniak i inni 2009). Na zestaw sygnalizowanych w publikacjach problemów składają się m.in.: małe zaangażowanie uczelni w usługi zewnętrzne, szkolenia, doradztwo lub kształcenie ustawiczne, słabe powiązanie dydaktyki z rynkiem pracy, wieloletowość pracowników naukowych oraz brak planów strategicznych.

Uogólniając, przebieg zachodzących w ostatnich latach przemian, tj. zaistniały boom akademicki, bez wątplenia określać można w kategoriach sukcesu ilościowego. Niemniej w toku zachodzących przeobrażeń zabrakło działań, które gwarantowałyby jakościowy rozwój polskich uczelni (por. Brzeziński, Elias 2003; Chmielecka 2003).

Podejmowany w licznych opracowaniach problem kryzysu uniwersytetu widoczny jest, zdaniem autorów, m.in. w obniżeniu jakości kształcenia i degradacji znaczenia dyplomów jako mierników jakości kapitału ludzkiego (por. Winclawski 1994; Gieysztor 1998; Bauman 2003). Nie oznacza to jednak, że aktywność teoretyków ogranicza się jedynie do krytyki stanu zastanego. Coraz częściej podejmowane są dysputy na temat naprawy czy reformy

systemu kształcenia w Polsce. Jeden z wątków tej dyskusji skupia się na problemie pomiaru, monitorowania, ewaluacji i ewentualnej modyfikacji jakości kształcenia, a niniejsza publikacja, na co mamy nadzieję, będzie ważnym głosem praktycznym w tej dyskusji.

Artykuł potraktować należy jako wprowadzenie do debaty praktyków nad problemem mierzenia, oceny i modyfikacji jakości kształcenia w szkołach wyższych. Jest to wprowadzenie o charakterze teoretycznym, mające na celu nakreślenie szerokiego kontekstu, w którym temat ten jest osadzony. Jego znacząca część poświęcona została kwestiom definicyjnym, które w przypadku tak wielowątkowego konceptu, jakim jest jakość kształcenia, zazwyczaj sprawiają wiele problemów. Scharakteryzowane zostały w nim relacje pomiędzy pojęciami jakości, skuteczności i efektywności, a także podkreślona została konieczność szerokiego ujęcia koncepcyjnego i operacyjnego jakości nauczania. Tłem dla prowadzonych rozważań są zmiany projektowane w europejskim i polskim obszarze szkolnictwa wyższego. Implikują one konieczność przededefiniowania pojęcia jakości w szkolnictwie wyższym, a także zmodyfikowania podstawowych modeli jej oceny.

Jakość kształcenia w świetle dokumentów i inicjatyw europejskich

Aktualnie problematyka jakości kształcenia w segmencie szkolnictwa wyższego jest jednym z priorytetów zarówno w polityce europejskiej, jak i polskiej. Ze względu na ograniczoną formułę artykułu nie ukażemy w nim kontekstu w pełni jego złożoności. Zaprezentowane poniżej kwestie z konieczności stanowią więc będą przyczynek do dalszych poszukiwań czytelnika.

W ostatniej dekadzie kierunki rozwoju europejskiego szkolnictwa wyższego w dużej mierze wytyczane były przez dwa równoległe i współzależne procesy – realizację Strategii Lizbońskiej oraz Proces Boloński.

Celem ogłoszonej w roku 2000 Strategii Lizbońskiej było stworzenie efektywnej i konkurencyjnej gospodarki. Tym samym gospodarka oparta na wiedzy (GOW) stała się wielkim

wyzwaniem dla Europy XXI wieku, a jej wdrażanie stanowi polityczny priorytet, wpisywany w plany rozwojowe na szczeblach międzynarodowych, krajowych i regionalnych. W nowoczesnej gospodarce wiedza, stanowiąc strategiczny zasób społeczeństwa, staje się podstawą jej funkcjonowania i rozwoju. Właśnie dlatego szczególną rolę przypisuje się szkołom wyższym – to one stanowią bowiem miejsca tworzenia i rozpowszechniania wiedzy na najwyższym poziomie, a w związku z tym uznać je można za swoiste moderatory wzrostu gospodarczego. Koncepcja GOW wiąże się również nierozdzielnie z pojęciem kapitału ludzkiego. O ile dawniej zasoby regionów w zakresie kapitału ludzkiego definiowane były poprzez mierniki ilościowe odnoszące się do odsetka osób z wyższym wykształceniem, o tyle dziś wiadomym jest, że ogromną wagę przykładać należy do jakości oferowanego kształcenia.

Obok Strategii Lizbońskiej w krajach europejskich realizowane są inicjatywy, w których bezpośrednim centrum znajdują się szkoły wyższe. Podpisana 19 czerwca 1999 roku Deklaracja Bolońska zapoczątkowała proces zmian, którego nadrzędnym celem jest utworzenie Europejskiego Obszaru Szkolnictwa Wyższego (*European Higher Education Area*)¹. Proces Boloński, będąc próbą odpowiedzi na problemy dotyczące wiele krajów europejskich, ma stanowić klucz m.in.: do zwiększenia mobilności obywateli, dostosowania systemu szkolnictwa wyższego do potrzeb rynku pracy, przygotowania studentów do bycia aktywnymi obywatelami, rozwoju kształcenia przez całe życie, zwiększenia atrakcyjności oraz konkurencyjności europejskiego systemu szkolnictwa wyższego. Intencją decydentów projektujących reformy nie jest stworzenie jednego zuniifikowanego systemu szkolnictwa wyższego, ale ponad czterdziestu systemów narodowych, które opierać się będą na wspólnie wypracowanych zasadach. W chwili obecnej jednym z podstawowych instrumentów realizacji tego zamierzenia jest Ramowa Struktura Kwalifikacji dla Europejskiego Obszaru Szkolnictwa Wyższego oraz – pozostające w dużej mierze w fazie projektowania – jej krajowe odpowiedniki.

Ramową Strukturę Kwalifikacji przyjęto w Bergen w 2005 roku. Zbudowano ją w oparciu o cykl kształcenia (I, II i III stopień), deskryptory efektów kształcenia i kompetencji absolwentów oraz przypisane poszczególnym etapom studiów punkty ECTS. Narzędzie to

1 Polska, nie będąc członkiem Unii Europejskiej, była sygnatariuszem Deklaracji Bolońskiej. Warto więc zaznaczyć, że Proces Boloński swym zasięgiem obejmuje szereg państw niebędących członkami Unii Europejskiej.

ma być z założenia mechanizmem łączącym krajowe struktury kwalifikacji. W zamierzeniu powinno ono ułatwić określenie relacji między europejskimi systemami szkolnictwa wyższego, gdyż stwarza wspólną podstawę rozumienia tychże systemów oraz funkcjonujących w ich ramach kwalifikacji. Struktura kwalifikacji zwiększać ma ich przejrzystość i porównywalność, maksymalizować elastyczność systemów edukacji, ułatwiać uznawalność wykształcenia, a tym samym przyczyniać się do rozwoju rynku pracy czy zwiększenia mobilności obywateli.

W toku realizacji Procesu Bolońskiego wiele uwagi poświęcono bezpośrednio kwestiom jakości kształcenia w jednostkach akademickich, a jednym z priorytetowych celów na najbliższe lata jest zintensyfikowanie współpracy europejskiej w zakresie zapewniania jakości, z uwzględnieniem opracowania porównywalnych kryteriów i metodologii. Realizując powyższe założenie na konferencji ministrów w Bergen przyjęte zostały proponowane przez ENQA (*European Association for Quality Assurance in Higher Education*) standardy oraz wskazówki zapewnienia jakości w Europejskim Obszarze Szkolnictwa Wyższego².

W odniesieniu do kontekstu polskiego podkreślić należy, że kwestia jakości nie zawsze znajdowała swój wyraz w obowiązującym ustawodawstwie. Przeciwnie, niektóre z przyjmowanych rozwiązań wpłynęły na jakość kształcenia w sposób bardzo negatywny. *Ustawa o szkolnictwie wyższym z dnia 12 września 1990 roku*, odchodząc od centralnego planowania szkolnictwa wyższego i dopuszczając do różnorodności instytucjonalnej, w sposób istotny zmodyfikowała PRL-owski ład akademicki (*Diagnoza stanu szkolnictwa wyższego w Polsce 2009*). Jak przyznają autorzy założeń obecnie planowanych reform (*Założenia do nowelizacji ustawy... 2010:6*), *wiele efektów w życiu społecznym wiąże się z decyzjami przypadkowymi. Tak też było w przypadku szkolnictwa wyższego. Nikt w 1991 roku nie miał*

2 W Komunikacie Berlińskim z 2003 roku zwrócono się do ENQA z prośbą o opracowanie zestawu standardów, procedur i wskazówek dotyczących zapewnienia jakości. Powstałe standardy i wskazówki odnoszą się zarówno do wewnętrznych, jak i zewnętrznych, systemów zapewnienia jakości. W odniesieniu do szkół wyższych postuluje się realizację następujących norm:

- istnienie wewnętrznej polityki oraz procedury zapewnienia jakości;
- tworzenie mechanizmów zatwierdzania, monitoringu oraz okresowego przeglądu programów i ich efektów;
- ocenę studentów według opublikowanych i jasnych reguł;
- istnienie metod działania zapewniających wysoką jakość kadry dydaktycznej;
- zagwarantowanie odpowiednich zasobów do nauki oraz środków wsparcia dla studentów;
- funkcjonowanie systemów informacyjnych gromadzących dane dotyczące skutecznego zarządzania oferowanymi programami studiów oraz innymi działaniami;
- publikowanie informacji na temat oferowanych przez siebie programów oraz ich efektów.

wyobrażenia, jakie długofalowe skutki wywoła ustawowy zapis o tworzeniu niepaństwowych szkół wyższych. Rzeczywiście skala umasowienia studiów wyższych przerosła oczekiwania, nie spodziewano się również konsekwencji, które to wywoła. Generalnie, według ekspertów postanowienia ustawy spowodowały, że nastąpił głęboki spadek jakości kształcenia zarówno w odniesieniu do uczelni publicznych, jak i niepublicznych.

Kolejnym istotnym z perspektywy budowania systemu szkolnictwa wyższego dokumentem było przyjęte w roku 2005 *Prawo o szkolnictwie wyższym*. Na marginesie warto dodać, że jest ono uważane za drugą ustawę w Unii Europejskiej, która uwzględnia podstawowe założenia Procesu Bolońskiego (por. Szulc 2005). W kontekście jakości kształcenia i poza prawnym umocowaniem fundamentów Procesu Bolońskiego do najważniejszych zmian wprowadzonych przez *Prawo o szkolnictwie wyższym* należą:

- utworzenie nowych instytucji ponaduczelnianych, w tym Państwowej Komisji Akredytacyjnej (PKA);
- uregulowanie prowadzenia kierunków studiów poprzez m.in. stworzenie listy kierunków i wprowadzenie standardów kształcenia.

Według założeń obecnie planowanej reformy jednym z podstawowych celów projektowanych zmian jest zapewnienie polskim studentom wyższej jakości kształcenia. Wśród środków prowadzących do realizacji tego celu wymienia się stworzenie mechanizmów racjonalnego finansowania szkolnictwa wyższego, które opierać się będą na efektach pracy naukowej i dydaktycznej. W projektowanych zmianach rola PKA ma więc ulec poszerzeniu. Przede wszystkim PKA co 2 lata wyłaniać ma nie więcej niż 25 kierunków studiów o najwyższej jakości kształcenia spośród jednostek organizacyjnych uczelni publicznych i niepublicznych. Podmioty te będą mogły liczyć na dodatkowe, znaczące finansowanie. Ponadto, w ramach założeń do reformy, postuluje się, by PKA zwiększyła swą aktywność i skupiała się bardziej niż dotychczas na:

- efektach kształcenia i ich zgodności z misją uczelni;
- powiązaniu działalności dydaktycznej z działalnością naukową;
- ocenie efektów działalności dydaktycznej przez pracodawców (*Założenia do nowelizacji ustawy... 2010:32*).

Koncentracja na efektach kształcenia ma być ułatwiona po wprowadzeniu w życie Krajowych Ram Kwalifikacji. Ramy te zastąpią wprowadzone wcześniej standardy kształcenia.

Czy wiemy, czym jest jakość w szkolnictwie wyższym?

Choć nie ulega wątpliwości, że jakość od wielu lat jest przedmiotem troski decydentów, to równocześnie podkreślić należy, że jest to pojęcie, zarówno na gruncie szkolnictwa wyższego, jak i w odniesieniu do wielu innych dziedzin, trudne czy wręcz, zdaniem niektórych, niemożliwe do zdefiniowania (por. Bugdol 2008).

Na jakość i jej niebagatelne znaczenie zwraca się uwagę w wielu kontekstach. Trudno natomiast znaleźć publikacje, które w sposób precyzyjny i niebudzący wątpliwości definiowałyby, czym owa jakość jest. W świecie biznesu, jak piszą Hamrol i Mantura (2005:10), *Sukces lub porażka przedsiębiorstwa zależą od jakości jego funkcjonowania, a w szczególności od jakości oferowanych produktów. Poziom jakości produktów i obsługi nabywców w pierwszym rzędzie determinuje pozycję rynkową i sukces każdego wytwórcy*. Choć zacytowana opinia podkreśla znaczenie jakości w kontekście walki konkurencyjnej, to nie odnosi się do istoty jakości, tj. tego, czym miałyby ona być. Pojęcie „jakości” jest więc terminem kłopotliwym. Pomimo intuicyjnego rozumienia terminu, z trudem przychodzi wyrażenie jego znaczenia poprzez słowa. Jest to również konstrukt silnie obciążony wartościami, tzn. wiąże się z subiektywnym postrzeganiem tego, co dobre i co należy cenić. „Jakość” jest też pojęciem wielowymiarowym i niejednorodnym, a w związku z tym nie ma zgody co do tego, jak powinna być definiowana. Przykładowo, Vroenijensstijn (por. Wójcicka 2001), podkreślając ogromną ilość prac odnoszących się do istoty jakości w szkolnictwie wyższym, bezradnie uznaje, że jest to pojęcie wymykające się ramom definicyjnym: *jakość jest jak miłość. Każdy o niej mówi i wie, o czym mówi. Każdy wie i czuje, kiedy przychodzi. Jeśli jednak spróbuje zdefiniować to pojęcie, pozostaje z pustymi rękami*.

W praktyce w literaturze funkcjonuje szereg podstawowych ujęć „jakości”. Według Green (1994) najważniejsze koncepcje to:

- **Jakość jako doskonałość** – jakość ujmowana jest w związku z koniecznością dostarczenia produktu lub usługi, które stanowią swego rodzaju luksus, są wyróżniające i podwyższają status właściciela lub osoby korzystającej z usługi. Wysoki standard może być osiągnięty tylko przez wysokie nakłady, co powoduje, że produkty lub usługi dostępne są dla niewielkiej części populacji.
- **Jakość jako niewystępowanie usterek** – takie rozumienie jakości ma swe źródła w przemyśle, gdzie kontrola jakości odbywa się przez sprawdzanie zgodności procesu produkcyjnego z ustalonymi standardami.
- **Jakość jako przygotowanie do osiągnięcia celów instytucji** – w tym ujęciu jakość ma znaczenie jedynie w odniesieniu do celów stawianych sobie przez organizację. Znaczenie jakości wyraża się w ocenie stopnia, w jakim instytucji udaje się realizować ustalone wcześniej cele.
- **Jakość jako spełnienie oczekiwań i potrzeb klienta** – w tym podejściu nacisk kładziony jest na identyfikację potrzeb i oczekiwań klientów jako punktu wyjścia do projektowania produktu lub usługi. Generowanie jakości polega więc na określeniu potrzeb klienta i przełożeniu ich na mierzalne wymiary, które, realizując się w produkcie lub usłudze, dadzą klientowi zadowolenie.
- **Jakość jako ciągły rozwój** – z tej perspektywy osiągnięcie jakości jest procesem ciągłym, w którym podkreśla się odpowiedzialność członków organizacji za to, by – realizując cele instytucji – najlepiej wykorzystywali posiadane zasoby.

Przedstawiony katalog warto uzupełnić o kolejne ujęcie, czyli jakość rozumianą poprzez kulturę organizacyjną. Jedną z najpopularniejszych definicji kultury organizacyjnej zaproponował Schein, który rozumiał ją jako zbiór dominujących wartości i norm postępowania, charakterystycznych dla danej organizacji, podbudowany założeniami co do natury rzeczywistości, przejawiający się poprzez artefakty i wytwory (por. Kostera 1996). Jak pisze Bugdol (2008:19): *Dzisiaj już wiemy, że jakość jest w nas samych, w działaniach, które podejmujemy. Jakość jest bardzo mocno zakorzeniona w kulturze.*

Każde z przedstawionych wyżej ujęć ma swoje wady i zalety, z mniejszym lub większym powodzeniem daje się zaadaptować na potrzeby określania jakości w poszczególnych dziedzinach życia. Przykładowo, w sektorze prywatnym często przyjmowanym rozwiązaniem jest ujęcie jakości w odniesieniu do oczekiwań i potrzeb klientów (w tym kierunku zmierza

m.in. definicja zapisana w normie ISO 9000:2005 [por. Wawak 2007]). Również instytucje publiczne zachęca się do adaptowania metod i technik zarządzania z sektora prywatnego. Okazuje się jednak, że prosta analogia nie jest możliwa, przez co definicje jakości, użyteczne z perspektywy np. przemysłu, nie będą prowadzić do satysfakcjonujących rozwiązań w obszarze szkolnictwa wyższego. System szkolnictwa wyższego cechuje bowiem m.in. duża liczba interesariuszy. Za Frazerem (1994) można wśród nich wymienić: władze państwowe czy samorządowe, płacących podatki obywateli, pracodawców, studentów oraz środowisko akademickie. Oznacza to, że nie istnieje proste rozumienie chociażby terminu „klient”.

Idąc dalej, jednym z istotnych założeń konkurencyjnego modelu rynku jest dostęp do informacji na temat ceny, jakości czy innych charakterystyk nabywanego dobra lub usługi (por. Amaral 2007). W przypadku szkolnictwa wyższego i usług oferowanych studentom okazuje się, że:

- Wiedza na temat dobra nabywana jest w momencie jego konsumowania. Student dopiero rozpoczynając naukę i uczęszczając na zajęcia, wyrabia sobie pogląd dotyczący specyfiki produktu, z jakim ma do czynienia na uczelni.
- Student z reguły posiada ograniczone doświadczenie tj. pobiera naukę w jednej uczelni. W efekcie nie ma możliwości porównania usług oferowanych przez wiele uczelni.
- Koszty zmiany kierunku lub uczelni przez studenta są na tyle wysokie, że rzadko dochodzi do zejścia z obranej wcześniej ścieżki.

Przytoczone powyżej przykłady pokazują, że proste przełożenie mechanizmów rynkowych, terminologii i narzędzi funkcjonujących w sektorze prywatnym napotykać będzie na szereg ograniczeń. Z tego względu definicja jakości w szkolnictwie wyższym zaproponowana przez UNESCO jest szczególnie złożona i łączy w sobie szereg przedstawionych wcześniej ujęć. Prezentowane podejście podkreśla przede wszystkim kwestie związane z zakładanymi przez instytucje celami. Jakość w szkolnictwie wyższym rozumiana jest jako wielowymiarowe, wielopoziomowe i dynamiczne pojęcie, które odnosi się do misji i celów danej instytucji, jak również określonych standardów w ramach danego systemu, instytucji, programu czy dyscypliny naukowej. W rozwinięciu powyższej definicji jej autorzy podkreślają, że jakość przyjmować może różne znaczenia w zależności od:

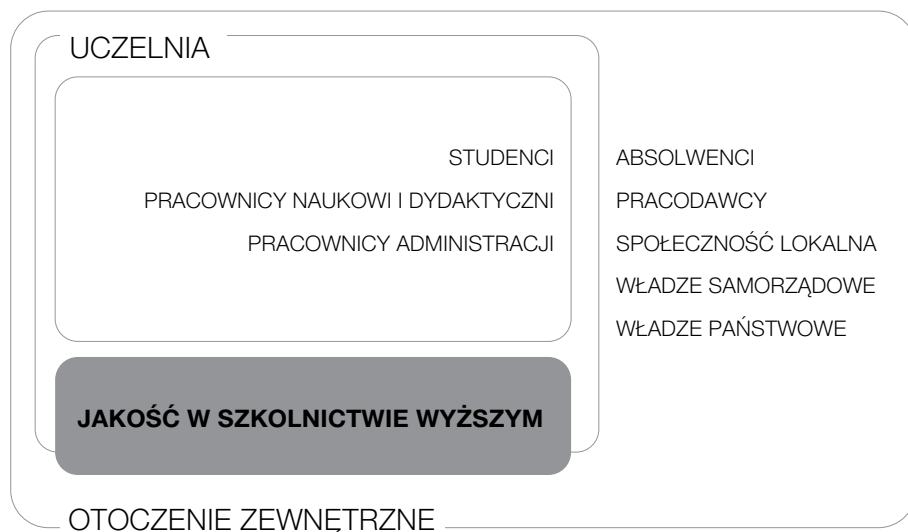
- interesariuszy szkolnictwa wyższego;

- przyjęcia perspektywy nakładów, procesu, misji, celów uczelni itp.;
- charakterystyki wybranego systemu szkolnictwa wyższego;
- stopnia rozwoju wybranego systemu szkolnictwa wyższego.

Warto również nadmienić, że w ostatnim czasie w dyskusji nad jakością w szkolnictwie wyższym mocno uwydatnianym aspektem jest wspomniana wcześniej kultura jakości (por. *Embedding Quality Culture in Higher Education 2007*; *Quality Culture in European Universities: a Bottom-up Approach 2006*). Cechuje się ona ciągłym dążeniem do poprawy jakości. Należy ją również rozpatrywać w odniesieniu do dwóch aspektów: po pierwsze, do wspólnie podzielanych przez członków organizacji wartości i przekonań względem jakości, po drugie, do systemu zarządzania organizacją, który, tworząc struktury projakościowe, sprzyjać będzie wykorzystaniu potencjału tkwiącego w członkach organizacji (*Embedding Quality Culture in Higher Education 2007*).

Podsumowując i odpowiadając na pytanie zadane w tytule podrozdziału, możemy stwierdzić, że nie ma jednego, uniwersalnego rozumienia jakości w szkolnictwie wyższym. Tym bardziej nie ma jednoznacznej definicji szczególnie interesującej z naszego punktu widzenia jakości kształcenia. Jakość w dużej mierze zależy od kontekstu, w jakim ją umieszczamy. Jakość nie jest też czymś stałym, lecz tworzy się w przebiegu interakcji pomiędzy różnymi uczestnikami systemu szkolnictwa wyższego. Podkreślić należy, że pojęcie jakości ma różne znaczenie dla różnych grup interesariuszy (por. rysunek poniżej) – studentów, pracowników akademickich, absolwentów i innych. Dbając o jakość, nie można abstrahować od ich potrzeb i oczekiwań, stąd konieczne jest prowadzenie ciągłego dialogu z reprezentantami społecznego otoczenia szkół wyższych i uwzględnianie różnorodnych punktów widzenia.

Rysunek 1. Jakość w szkolnictwie wyższym w odniesieniu do różnych interesariuszy



Źródło: opracowanie własne

Jakość – skuteczność – efektywność

Najczęściej spotykane definicje jakości kształcenia odnoszą się do jednego z trzech komponentów: stopnia spełnienia założonego standardu (mierzonego w sposób ilościowy lub jakościowy), stopnia zadowolenia klientów/studentów z usługi, jaką oferuje szkoła (podejście rynkowe), stopnia spełnienia założonych przez uczelnię celów, związanych z procesem kształcenia (skuteczność w osiągnięciu tych celów). Dodatkowo jakość może być oceniana w sposób zewnętrzny, w oparciu o obiektywne kryteria, może też być oceniana subiektywnie jako interpretowana użyteczność jednostkowa lub społeczna. Wspomniana wielowymiarowość pojęcia implikuje konieczność definiowania jakości kształcenia w odniesieniu do innych terminów, takich jak: skuteczność, efektywność czy nawet produktywność szkoły wyższej.

Niektórzy autorzy starają się oddzielić wymienione pojęcia, utożsamiając efektywność i produktywność z komponentami ilościowymi, obiektywnymi, natomiast jakość kształcenia

z podejściem jakościowym, trudnym do zmierzenia subiektywnym. W niniejszym artykule przeciwstawiamy się takiemu sposobowi ujmowania jakości, traktując ją jako pojęcie silnie związane ze skutecznością i efektywnością działań konkretnej szkoły. Jakość wiążemy więc nie tylko ze stopniem spełnienia założonych standardów formalnych procesu kształcenia, ale przede wszystkim z efektem, do którego doprowadzić ma właściwie zaplanowany i przeprowadzony proces kształcenia (skuteczność w osiągnięciu założonych celów). Można więc powiedzieć, że indywidualny i społeczny efekt netto³ kształcenia powinien stać się wskaźnikiem wysokiej jakości nauczania, a stosowanie jednolitych standardów (często tzw. standardów minimum formalnego) nie gwarantuje kształcenia na wysokim poziomie⁴.

Nie jest więc zaskakujące, że w wielu opracowaniach zamiennie używane są pojęcia „skuteczność” (*effectiveness*)⁵ oraz „jakość” (*quality*) – z założeniem, że termin „skuteczność” (zdolność osiągnięcia celu) w szerokim zakresie pokrywa się z terminem „jakość” (por. Gates i inni 2002). Oba pojęcia często łączone są z kwestią stawianych przed organizacją zadań i ewaluacją stopnia ich osiągnięcia. W tym kontekście jakość rozumie się jako: rodzaj doskonałości, perfekcję i konsekwencję działania, działanie dopasowane do celu bądź najlepszy stosunek wartości towaru do jego ceny, tzw. *value for money* (Cullen 1987, za: Cave i inni 2006). Coraz częstsze łączenie pojęcia „jakość” z terminem „skuteczność” związane jest z tendencją do indywidualizowania funkcji, jakie mają pełnić poszczególne szkoły wyższe (np. kształcenie elit, upowszechnienie wykształcenia na poziomie wyższym, kształcenie dla potrzeb rynku lokalnego, kształcenie dla potrzeb rynku europejskiego itp.). Nie jest zatem możliwe ustalenie jednego, powszechnie obowiązującego standardu czy sposobu kształcenia (chyba, że jest to standard minimum), zgodnie z którym funkcjonować powinny szkoły wyższe. Standard ustalony powinien zostać wewnątrz szkoły⁶ i być interpretowany w kontekście celów (najczęściej zapisanych w strategii), jakie planuje ona osiągnąć. Cele te

3 Efekt netto jest różnicą pomiędzy pewnym stanem „na wejściu” (np. kompetencjami i wiedzą studentów) a stanem „na wyjściu” z systemu kształcenia na poziomie wyższym.

4 Przedstawiony powyżej sposób myślenia zgodny jest z opisanymi już współczesnymi tendencjami w zakresie zmiany planowania i ewaluowania procesu kształcenia. Niepraktyczne standardy kształcenia powoli zastępowane są bardziej elastyczną strukturą kwalifikacji.

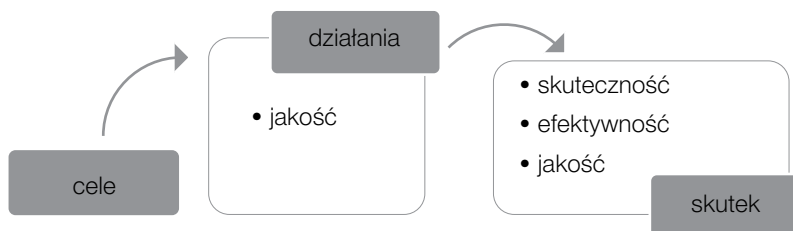
5 Zgodnie z prakseologiczną definicją „skuteczność” (*effectiveness*) charakteryzuje ogólną zdolność do uzyskania określonego celu (por. Kotarbiński 2003).

6 Nie oznacza to oczywiście, że szkoła nie powinna być zewnętrznie kontrolowana, a jakość edukacji monitorowana przez instytucje zewnętrzne. Powinny one jednak skupiać się na ocenie efektów kształcenia (np. porównując je pomiędzy szkołami), a nie tylko na weryfikacji formalnie ustalonych kryteriów.

określają, do jakich efektów społecznych (np. upowszechnienie wykształcenia wyższego, obniżenie wykluczenia edukacyjnego pewnych grup społecznych, kreowanie elit politycznych, zawodowych, rozwój regionalnego rynku pracy) i indywidualnych (wzrost poziomu wiedzy studentów, wzrost ich kompetencji społecznych, wzrost zaangażowania społecznego) powinno doprowadzić kształcenie w jednostce akademickiej. Podkreślimy jeszcze raz, że efekty te nie muszą być tak samo sprofilowane dla każdej ze szkół wyższych, a raczej każda z uczelni powinna sama określać cele społeczne i indywidualne, jakie stawia przed sobą. Funkcje te oczywiście muszą być zgodne z kierunkami zapisanymi w strategiach krajowych i w obowiązujących dokumentach europejskich, nie oczekujemy jednak, że będą pokrywać całe spektrum zawartych w tych dokumentach zadań.

Kolejnym konceptem, który w opracowaniach teoretycznych często pojawia się w opozycji do pojęcia „jakość”, jest termin „efektywność”, w ekonomii rozumiany jako stosunek uzyskanych efektów do poniesionych nakładów (efektywność wzrasta, gdy wkład potrzebny do wytworzenia danego produktu maleje w stosunku do osiągniętej jego wartości). Nie jest to oczywiście pojęcie tożsame z pojęciem jakości, jednak coraz częściej zwraca się uwagę na konieczność dwubiegunowej analizy procesu kształcenia, z uwzględnieniem zarówno jego komponentu jakościowego, jak i efektywnościowego.

Rysunek 2. Ocena jakości, skuteczności i efektywności w procesie kształcenia



Źródło: Opracowanie własne

Można zatem powiedzieć, że w cyklu projektowania procesu kształcenia, który rozpoczyna się od ustalenia konkretnych celów kształcenia⁷, poprzez zaprojektowanie i wdrożenie konkretnych inicjatyw, aż do uzyskania (bądź nie) określonego efektu, pomiar skuteczności i efektywności działań odbywa się „post factum” (po zakończeniu procesu). Weryfikacja jakości kształcenia

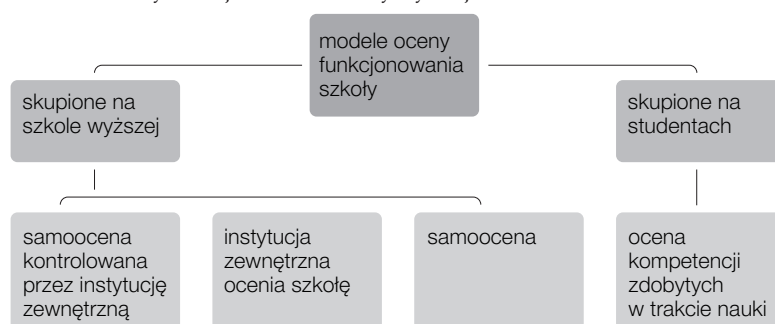
⁷ Cele te powinny być określone w strategii rozwoju jednostki akademickiej.

może odbywać się zarówno na poziomie projektowania samego działania (poprzez ustalenie standardów), jak i po jego zakończeniu. Pamiętać jednak należy, że jakość mierzona „post factum” i analizowana w kontekście zdolności instytucji do osiągnięcia danych celów edukacyjnych jest w zasadzie tożsama ze skutecznością tego działania.

Modele oceny funkcjonowania szkoły wyższej⁸

W obszarze szkolnictwa wyższego zostały ukształtowane cztery klasyczne modele stosowane współcześnie w procesie oceny funkcjonowania uczelni. Warto zwrócić uwagę na fakt, że mechanizmy te mogą być wykorzystywane zarówno do oceny kształcenia, jak i aktywności badawczej jednostki akademickiej, w niniejszym opracowaniu charakteryzowane będą jednak z punktu widzenia oceny jakości kształcenia. Modele zostały zaprezentowane na poniższym rysunku.

Rysunek 3. Modele oceny funkcjonowania szkoły wyższej



Źródło: Gates i inni 2002, s. 24

Modele oceny funkcjonowania szkoły wyższej możemy podzielić na dwie klasy – ze względu na podmiot, który w sposób bezpośredni uczestniczy w ocenie. Pierwszy typ to grupa

⁸ Niniejsza część powstała w oparciu o fragmenty pracy doktorskiej dr Magdaleny Jelonek, pt. Społeczna produktywność szkół wyższych. Nowa koncepcja produktywności nauczania (badania do pracy finansowane były ze środków MNiSW w ramach grantu promotorskiego nr N N116 1110340.

metod skupionych na samej szkole wyższej i oceniających różne aspekty jej działania, drugi natomiast ogniskuje uwagę na studentach, w sposób pośredni oceniając pracę instytucji.

Pierwsza z metod nazywana jest samooceną kontrolowaną przez instytucję zewnętrzną i powstała na gruncie rozwiązań biznesowych. Polega ona na wewnętrznym ustalaniu kryteriów oceny przez poddawaną ewaluacji instytucję, a następnie na weryfikacji spełnienia tych kryteriów przez jednostkę zewnętrzną. Jako przykład samooceny kontrolowanej może posłużyć coraz powszechniejszy w środowisku uniwersyteckim audyt akademicki, implementowany na różnych uczelniach: w Hong Kongu, Skandynawii, Wielkiej Brytanii, Australii czy Nowej Zelandii. Metodę audytu akademickiego możemy zaklasyfikować do grupy technik zorientowanych na procesie. Mimo iż wywodzi się ona bezpośrednio z metod oceny jakości stosowanych w biznesie, takich jak TQM czy ISO, jest powszechnie akceptowanym sposobem ewaluacji pracy szkół wyższych wewnątrz środowiska naukowego (por. Gates i inni 2002). Audyt akademicki jest praktyką europejską i po raz pierwszy został zaimplementowany w Wielkiej Brytanii w latach 90. XX wieku z inicjatywy rządu brytyjskiego. Ocena w ramach audytu akademickiego najczęściej ogranicza się do analizy kluczowych procesów, w tym może dotyczyć weryfikacji systemu zapewniania jakości i efektywności pracy instytucji. Metoda ta jest krytykowana właśnie za zbytne zaangażowanie w badanie procesów, a nie analizowanie produktywności i efektywności instytucji. Kolejną wadą audytu jest brak porównywalności jednostkowych wyników oceny, co jest efektem zbytniego zindywidualizowania kryteriów ewaluacji.

Instytucje zewnętrzne oceniające szkołę wyższą zazwyczaj skupiają się na celach systemowych, a tym samym nie indywidualizują procesu oceny na poszczególne szkoły. Podmioty ewaluujące to najczęściej organizacje państwowe bądź niezależne od szkół wyższych instytucje publiczne. W Polsce za przykład mogą posłużyć wszelkie komisje akredytacyjne, w tym Państwowa Komisja Akredytacyjna, natomiast w Stanach Zjednoczonych najsłynniejszą tego typu inicjatywą jest nagroda Baldrige'a (*The Malcolm Baldrige National Quality Award*) czy *U.S. News & World Report*.

Samoocena uważana jest za jedną z najefektywniejszych i najpopularniejszych metod ewaluacji pracy szkoły wyższej. Może ona przybierać najróżniejsze formy, począwszy od

najprostszyc, rocznych zestawień wskaźników, a skończywszy na coraz powszechniejszych portfolioch elektronicznych, takich jakie prezentują m.in. instytucje biorące udział w *The Urban Universities Portfolio Project*⁹.

Ostatni model oceny wyróżniony przez autorów raportu *Ensuring Quality and Productivity in Higher Education. An Analysis of Assessment Practices*, w przeciwieństwie do poprzednich, nie koncentruje się na samej instytucji, ale raczej na efektach jej działania w postaci zdobytych przez studentów kompetencji. Nie analizuje samego procesu przekazywania wiedzy i umiejętności (metody skupione na procesie), ale raczej ogniskuje się na wynikach (metody skoncentrowane na produkcie bądź rezultacie), nie wnikając w proces dochodzenia do takiej a nie innej ich formy. Jako klasyczny przykład takiego podejścia możemy wymienić badania, prowadzone w Stanach Zjednoczonych na szeroką skalę, mające na celu oszacowanie wartości dodanej działań edukacyjnych (por. Klein i inni 2008).

Wymienione powyżej modele oceny pracy szkoły wyższej zwiększają swoją skuteczność w określonych warunkach. Specyficzne cechy systemu szkolnictwa wyższego wyznaczają rozwiązania, które mogą się okazać bardziej lub mniej przydatne w danej sytuacji.

W poniższej tabeli skrzyżowane zostały dwa wymiary¹⁰. W pierwszym z nich ulokowano modele oceny pracy szkół wyższych, w drugim specyficzne cechy systemu szkolnictwa wyższego. W polu powstałym po skrzyżowaniu tych dwóch wymiarów umieszczono informację dotyczącą przydatności zastosowania danego modelu w określonych warunkach systemowych. Kolor ciemniejszy pola wyznaczonego przez przecięcie się wymiarów oznacza, że istnieje duże prawdopodobieństwo skuteczności takiego rozwiązania w określonych warunkach. Kolor jaśniejszy sugeruje, że model może, ale nie musi, odnieść sukcesu. Brak koloru mówi nam, że istnieje nikła szansa, iż model będzie w określonych okolicznościach miał sens.

Cechy systemu, które wyznaczają przesłanki do zastosowania określonego modelu to:

- ukierunkowanie na wymierność;
- ukierunkowanie na poprawę;

⁹ Jest to inicjatywa sześciu uniwersytetów, która ma na celu skonstruowanie wzoru uniwersyteckiego portfolio.

¹⁰ Opisywana przeze mnie koncepcja Gatesa może okazać się użyteczna w wyborze odpowiednich modeli oceny pracy polskich instytucji akademickich (Gates i inni 2002).

- określone zasoby;
- określony poziom centralizacji systemu;
- homo- bądź heterogeniczność;
- kompleksowość bądź prostota układu.

Ukierunkowanie działań jednostki na wymierność wiąże się z postulatem zwiększenia możliwości rozliczania instytucji akademickich z efektów ich pracy. Ukierunkowanie na poprawę wskazuje, że jednostki są mobilizowane w sposób zewnętrzny do ciągłego reformowania własnych działań. Decentralizacja polega na przeniesieniu uprawnień z centralnych ośrodków kierowniczych na indywidualne uczelnie. Heterogeniczność systemu świadczy natomiast o jego różnorodności i dużym skomplikowaniu.

Tabela 1. Ocena przydatności zastosowania modeli ewaluacji pracy szkoły wyższej w różnych okolicznościach

	SAMOOCENA KONTROLOWANA	OCENA PRZEZ ZEWNĘTRZNĄ INSTYTUCJĘ	SAMOOCENA	OCENA KOMPETENCJI STUDENTÓW
Ukierunkowanie na wymierność				
Ukierunkowanie na poprawę				
Ograniczone zasoby				
Decentralizacja				
Wysoka heterogeniczność systemu				

Jak widać w powyższej tabeli, najbardziej elastycznym modelem, który może być wdrożony w celu oceny funkcjonowania jednostki akademickiej, jest model oparty na samoocenie. Słabym punktem, wynikającym z subiektywizmu w wyznaczaniu kryteriów oceny, jest jego mała użyteczność w zakresie spełniania postulatu wymierności działań instytucji. Model ten przynosi zatem większe korzyści pojedynczym podmiotom podlegającym ocenie niż jednostkom zarządzającym sektorem czy instytucjom alokującym środki finansowe (por. Gates i inni 2002).

Kolejnym elastycznym rozwiązaniem może okazać się wdrożenie modelu opartego na samoocenie kontrolowanej. Wybór ten ma szansę odnieść sukces bez względu na warunki systemowe, najlepiej jednak sprawdza się w systemach ukierunkowanych na poprawę, zdecentralizowanych, o wysokim stopniu heterogeniczności. Nie jest to jednak model, który gwarantuje wymierność pracy szkoły wyższej. Ten warunek spełnia natomiast ocena przeprowadzana przez zewnętrzną instytucję, która jednak nie sprawdza się w strukturach zdecentralizowanych, niskofinansowanych. Wybór taki może okazać się ponadto mało efektywnym narzędziem o poprawy funkcjonowania jednostek akademickich.

Metody oparte na ocenie zdobytych w toku nauczania kompetencji studentów dobrze sprawdzają się w typowych zdecentralizowanych systemach szkolnictwa wyższego. Są to jednak zazwyczaj modele „kosztochłonne” – związane z szacowaniem wartości dodanej procesu dydaktycznego poprzez porównanie osiągnięć „na wejściu” i „na wyjściu” z systemu. „Kosztochłonność” powyższego rozwiązania wiąże się przede wszystkim z koniecznością zastosowania standaryzowanych testów dwukrotnie: przed rozpoczęciem nauki w danej szkole oraz po jej zakończeniu.

Zakończenie

W dobie gospodarki opartej na wiedzy przed szkołami wyższymi postawiono niezwykle ambitne wyzwania. Ze względu na swoją wagę debata na temat jakości w szkolnictwie wyższym wyszła znacznie poza mury samych uczelni, a zagadnienie to jest przedmiotem szczególnej troski zarówno na forum międzynarodowym, jak i krajowym. Podejmowane dyskusje nie stanowią jedynie teoretycznej refleksji, ale za sprawą m.in. Procesu Bolońskiego przekładają się na projektowanie konkretnych rozwiązań, które w znaczący sposób modyfikują i reformują istniejące systemy szkolnictwa wyższego.

Kompleksowość problematyki jakości kształcenia uniemożliwia dokonanie szczegółowej analizy zagadnienia w jednym artykule. Z tego względu koniecznym było przyjęcie skromniejszego celu, który w naszym przypadku określić należy jako zarysowanie ogólnego

kontekstu i uporządkowanie podstawowych kwestii związanych z definiowaniem jakości, jej mierzaniem oraz sposobami oceny.

Z niejednoznaczności oraz niepewności co do znaczenia pojęcia „jakość” wynikać może szereg nieporozumień. Jak starałyśmy się wykazać, jakość kształcenia w szkolnictwie wyższym, choć powiązana z pojęciami efektywności i skuteczności, nie jest z tymi terminami tożsama i daje się definicyjnie wyodrębnić. Współcześnie jednak w związku z ukierunkowaniem pracy instytucji na cele systemowe i indywidualne pojęcie jakości staje się coraz bliższe pojęciu skuteczności.

Mimo że jakości kształcenia nie można sprowadzić do jednej uniwersalnej definicji, to współczesne rozumienie tego terminu zwraca uwagę m.in. na jego wielowymiarowość, uzależnienie od perspektywy poszczególnych grup interesariuszy oraz osadzenie w misjach i celach przyjmowanych przez konkretne instytucje. W konsekwencji w ocenie jakości koniecznym staje się czerpanie inspiracji z wielu istniejących podejść i metod, a także przemyślane tworzenie nowych, jak najbardziej użytecznych w danym kontekście narzędzi.

Potrzeba zapewnienia możliwie najlepszego poziomu kształcenia w szkołach wyższych stała się podstawą do ukształtowania czterech podstawowych modeli oceny funkcjonowania uczelni (samoocena kontrolowana, ocena przez zewnętrzną instytucję, samoocena, ocena kompetencji studentów). Omówione w tekście metody oceny pracy szkół wyższych wykazują różną skuteczność w zależności od warunków, w jakich zostałyby zaimplementowane.

Biorąc pod uwagę znaczenie problematyki jakości kształcenia i jej oceny oraz istnienie obszernej literatury na ten temat, należy wyrazić nadzieję, że przyjmowane rozwiązania (zarówno na poziomie systemowym, jak i w odniesieniu do pojedynczych uczelni) w sposób optymalny sprzyjać będą poprawie jakości kształcenia.

Bibliografia:

Amaral, Alberto (2007), *Higher Education and Quality Assessment. The Many Rationales for Quality*, [w:] *Embedding Quality Culture in Higher Education*, Brussels: European University Association, s. 6–10.

Bauman, Teresa (2003), *Zagrożona tożsamość uniwersytetu*, [w:] Ładyżyński Andrzej, Raińczuk Jacek (red.), *Uniwersytet – między tradycją a wyzwaniami współczesności*, Kraków: Oficyna Wydawnicza „Impuls”, s. 53–68.

Brzeziński, Jerzy, Elias, Andrzej (red.) (2003), *Ocena wewnętrzna i zewnętrzna jakości kształcenia w szkołach wyższych*, Warszawa: Wydawnictwo SWPS ACADEMICA.

Bugdol, Marek (2008), *Zarządzanie jakością w urzędach administracji publicznej*, Warszawa: Difin.

Cave, Martin, Hanney, Stephen, Henkel, Mary, Kogan, Maurice (2006), *The Use of Performance Indicators in Higher Education – The Challenge of the Quality Movement*, Higher Education Policy Series 34, London, Bristol, Pennsylvania: Jessica Kingsley Publishers.

Chmielecka, Ewa (2003), *Korzyści i koszty związane z przystąpieniem Polski do Unii Europejskiej w sferze szkolnictwa wyższego*, Warszawa: Urząd Komitetu Integracji Europejskiej.

Frazer, Malcolm (1994), *Quality in Higher Education: an International Perspective*, [w:] Green Diana (red.), *What is Quality in Higher Education?*, London: Society for Research in Higher Education, s. 101–111.

Gates, Susan, Augustine, Catherine, Benjamin, Roger, Bikson, Tora, Kaganoff, Tessa, Levy, Dina, Moini, Joy, Zimmer Ron (2002), *Ensuring Quality and Productivity in Higher Education. An Analysis of Assessment Practices*, „ASHE-ERIC Higher Education Report”, 29 (1).

Gieysztor, Aleksander (1998), *Systemy wartości w tradycji uniwersyteckiej*, [w:] Samsonowicz Henryk (red.), *Idea uniwersytetu u schyłku tysiąclecia*, Warszawa: Wydawnictwo Naukowe SCHOLAR, s. 3–11.

Green, Diana (1994), *What is Quality in Higher Education? Concepts, Policy and Practice*, [w:] Green Diana (red.), *What is Quality in Higher Education?*, London Society for Research in Higher Education, s. 3–20.

Hamrol, Adam, Mantura, Władysław (2005), *Zarządzanie jakością: teoria i praktyka*, Warszawa: PWN.

Jakość kształcenia w szkolnictwie wyższym. Podręcznik metod zapewnienia jakości (2000), Opole: European Training Foundation.

Klein, Sheri, Freedman, David, Shavelson, Richard, Bolus, Roger (2008), *Assessing School Effectiveness*, „Evaluation Review”, 32(6), s. 511–525.

Kostera, Monika (1996), *Postmodernizm w zarządzaniu*, Warszawa: PWE.

Kotarbiński, Tadeusz (2003), *Dzieła wszystkie. Prakseologia 2*, Wrocław: Wydawnictwo Ossolineum.

Szulc, Tadeusz (2005), *Prawo o jakości szkolnictwa wyższego*, „Forum Akademickie” nr 12/2005, http://www.forumakad.pl/archiwum/2005/12/13-za-prawo_o_jakosci_szkolnictwa_wyzszego.htm [05.05.2010].

Vlasceanu, Lazar, Grünberg, Laura, and Parlea, Dan (2004), *Quality Assurance and Accreditation: A Glossary of Basic Terms and Definitions*, Bucharest: Papers on Higher Education, <http://www.cepes.ro/publications/pdf/QA&A%20Glossary.pdf> [05.05.2010].

Wawak, Sławomir (2007), *Podręcznik wdrażania ISO 9001:2000*, Gliwice: Helion.

Winławski, Włodzimierz (red.) (1994), *Tożsamość uniwersytetu: antologia tekstów Profesorów Uniwersytetu Mikołaja Kopernika*, Toruń: Uniwersytet Mikołaja Kopernika.

Wnuk-Lipińska, Elżbieta, Wójcicka, Maria (1995), *Przesłanki budowy system zewnętrznego recenzowania jakości kształcenia w polskich uczelniach*, „Nauka i Szkolnictwo Wyższe”, nr 6, s. 60–85.

Wójcicka, Maria (red.) (2001), *Jakość kształcenia w szkolnictwie wyższym. Słownik tematyczny*, Warszawa: Centrum Badań Polityki Naukowej i Szkolnictwa Wyższego Uniwersytetu Warszawskiego.

Raporty:

Diagnoza stanu szkolnictwa wyższego w Polsce (2009), Ernst&Young Business Advisory, Instytut Badań nad Gospodarką Rynkową, <http://www.uczelnie2020.pl/zobacz/diagnoza-stanu-szkolnictwa-wyz-2> [11.05.2010].

Embedding Quality Culture in Higher Education (2007), Brussels: European University Association, http://www.eua.be/fileadmin/user_upload/files/Publications/EUA_QA_Forum_publication.pdf [05.05.2010].

Fulton, Oliver, Santiago, Paul, Edquist, Charles, El-Khawas, Elaine, Hackl, Elsa (2007), *OECD Reviews of Tertiary Education „Poland”*, OECD, http://www.oecd.org/document/13/0,3343,en_2649_39263238_35585357_1_1_1_1,00.html [10.05.2010].

Górniak, Jarosław, Jelonek, Magdalena, Kwinta-Odrzywołek, Joanna, Skrzyńska, Joanna, Uhl, Hanna (2009), *Ocena wpływu projektów zrealizowanych dzięki funduszom europejskim przez uczelnie wyższe z terenu województwa małopolskiego na wzrost konkurencyjności gospodarczej regionu oraz wzmocnienie potencjału sektora badawczo-naukowego w województwie*, Kraków, http://www.politykarozwoju.obserwatoria.malopolska.pl/Files.mvc/134/Raport-SZK_WYZ.pdf [01.05.2010].

Polska 2030. Wyzwania rozwojowe (2009), Kancelaria Prezesa Rady Ministrów.

Szkolnictwo wyższe: jakość, równość i skuteczność (2006), OECD, Ateny.

Strategia rozwoju szkolnictwa wyższego 2010–2020. Projekt środowiskowy (2009), KRASP, FRP, UW, Warszawa, <http://www.frp.org.pl/?page=aktualnosci&cid=83> [08.05.2010].

Tuning. Harmonizacja struktur kształcenia w Europie. Wkład uczelni w Proces Boloński (2008), Fundacja Rozwoju Systemu Edukacji.

Quality Culture in European Universities: a Bottom-up Approach (2006), Brussels: European University Association.

Założenia do nowelizacji ustawy – Prawo o szkolnictwie wyższym oraz ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (2010), Ministerstwo Nauki i Szkolnictwa Wyższego.

Seweryn Rudnicki

Anna Szwed

Ocenianie oceniających. Trzy pytania o ewaluację zajęć dydaktycznych¹

Artykuł prezentuje problematykę ewaluacji nauczania w szkolnictwie wyższym poprzez próbę odpowiedzi na konkretne pytanie: czym tak naprawdę jest (lub powinna być) ewaluacja zajęć, jak przeprowadza się ewaluację w uznanych światowych ośrodkach akademickich oraz czy opinie wyrażane przez studentów w ankietach ewaluacyjnych są wiarygodnym źródłem danych.

W pierwszej części autorzy wyjaśniają pojęcie ewaluacji, zwracając uwagę na fakt, że obejmuje ona nie tylko surowe dane, ale także ich niezbędną interpretację oraz podkreślając jej podwójny, formatywny i sumatywny, charakter. W części drugiej przedstawiono „dobre praktyki” funkcjonujące w ramach systemów ewaluacji na wybranych uniwersytetach z kręgu anglosaskiego. Autorzy wspominają tutaj o problemach, takich jak: transparentność systemu, triangulacja zbieranych danych, procedury ewaluacyjne, zaangażowanie studentów itp. Trzecia część dotyczy zagadnień rzetelności i trafności ocen studenckich oraz czynników, które mogą potencjalnie wpływać na wyniki dokonywanego przy ich użyciu pomiaru jakości sposobu prowadzenia zajęć.

¹ Prace badawcze, których jednym z rezultatów jest niniejszy tekst, były prowadzone w ramach realizowanego przez Wyższą Szkołę Europejską projektu „Kompetentnie ku przyszłości”, finansowanego z Europejskiego Funduszu Społecznego w ramach Programu Operacyjnego „Kapitał Ludzki” Działanie 4.1.1. Tekst ukazał się po raz pierwszy w czasopiśmie naukowym „Kultura i Polityka” (nr 5/2009), wydawanym przez Wyższą Szkołę Europejską.

Wprowadzenie

Liczba publikacji naukowych poświęconych ewaluacji dydaktyki po koniec ubiegłego wieku sięgała 2 tysięcy pozycji na całym świecie (McKeachie, Kaplan 1996), co zdaje się świadczyć zarówno o stopniu skomplikowania problemu, jak i o jego znaczeniu. O ile przekonywanie o konieczności zapewnienia wysokiej jakości kształcenia w szkolnictwie wyższym jest truizmem, o tyle rzeczowa refleksja nad sposobem prowadzenia ewaluacji dydaktyki wydaje się zasadna. Potoczna obserwacja sposobu funkcjonowania uczelni wyższych w naszym kraju wskazuje bowiem, że mimo wprowadzania ewaluacji studenckich na coraz szerszą skalę, ciągle nie jest powszechne traktowanie studenta jako klienta instytucji edukacyjnej (którego zadowolenie z wysokiej jakości produktu stanowi jedną z podstawowych racji bytu uczelni) czy jako jednostki współtworzącej unikalną *communitas*, jaką ma być uniwersytet.

Niniejsze opracowanie ma na celu udzielenie odpowiedzi na zestaw podstawowych pytań związanych z prowadzeniem ewaluacji zajęć dydaktycznych w oparciu przede wszystkim o literaturę przedmiotu, rozwiązania funkcjonujące w krajach anglosaskich, a także doświadczenie autorów w przeprowadzaniu ewaluacji dydaktyki.

Czy dane mówią same za siebie?

Pojęcie ewaluacji dydaktyki wymaga szczególnej uwagi w związku z pojawiającymi się często nieścisłościami w jego rozumieniu. Na gruncie polskiego szkolnictwa wyższego, gdzie ewaluacja nauczania i problematyka zapewniania odpowiedniej jakości kształcenia są wciąż tematami stosunkowo nowymi choćby w porównaniu z krajami anglosaskimi, w których należą one do akademickiej rutyny, ewaluację najczęściej utożsamia się z semestralnymi ocenami studencki-

mi dokonywanymi za pomocą ankiet. W rezultacie uznaje się, że jeśli uczelnia posługuje się, w sposób bardziej lub mniej regularny, ankietami (najczęściej mającymi dodatkowo charakter wyłącznie lub głównie ilościowy), w których studenci oceniają przedmiot i wykładowcę, to możemy mówić, że taka jednostka posiada system ewaluacji. Nic bardziej mylnego. Oceny studentów zbierane za pomocą ankiet są tylko jednym ze źródeł danych wykorzystywanych w procesie ewaluacji, który nie sprowadza się do zaprezentowania surowych danych uzyskanych z jednego źródła, lecz zakłada interpretację ocen będących, w wersji optymalnej, wynikiem tzw. triangulacji (wykorzystania różnych narzędzi zbierania danych, takich jak: ankiety studenckie, hospitacje, *peer observations*, autoewaluacja). Interpretacja, a więc nadanie znaczenia (także w oparciu o kontekst oceniania²), jest zatem równie istotnym elementem ewaluacji, co same dane.

Drugie ważne zastrzeżenie terminologiczne związane jest z celem ewaluacji. W literaturze przedmiotu najczęściej stosuje się niezwykle praktyczne rozróżnienie na tzw. ewaluację formatywną i ewaluację sumatywną (Felder, Brent 2004). W przypadku pierwszej z nich nacisk kładziony jest na rozwijanie kompetencji dydaktycznych prowadzącego. Na podstawie zgromadzonych danych (m.in. z ankiet studenckich, obserwacji przez zaprzyjaźnionych wykładowców, hospitacji) prowadzący, wspomagany przez doradcę edukacyjnego lub inną kompetentną osobę, podejmuje refleksję dotyczącą mocnych i słabych stron swojego stylu nauczania, możliwych obszarów poprawy jego jakości i kierunku rozwoju. W ewaluacji formatywnej punktem wyjścia i celem, na którym się koncentrujemy, jest zatem sam proces dydaktyczny. Ewaluacja sumatywna ma natomiast na celu dostarczenie danych, które mogą zostać wykorzystane w rozstrzygnięciach o charakterze personalnym, takich jak: decyzja o dalszym prowadzeniu określonego przedmiotu lub jego zmianie, rozstrzygnięcia dotyczące zatrudnienia i jego formy, awansu, wynagrodzenia, nagród itp.

Choć w obu typach ewaluacji możliwe jest wykorzystanie podobnych narzędzi zbierania danych, można jednak z dużym prawdopodobieństwem zakładać, że osoby dokonujące ewaluacji dla celów podsumowujących będą raczej koncentrowały się na stwarzających możliwość porównywania danych ilościowych. W przypadku ewaluacji formatywnej szczególnego znaczenia nabierać będą natomiast wszelkie dane o charakterze jakościowym (np. uzyskane za pomocą pytań o charakterze otwartym).

2 Tym kontekstem może być na przykład typ przedmiotu, którego uczy prowadzący/a (kurs obligatoryjny, kurs fakultatywny), doświadczenie zawodowe prowadzącego/ej (czy jest to jego/jej pierwszy semestr nauczania).

Na czym polega system ewaluacji zajęć dydaktycznych?

W odpowiedzi na to pytanie proponujemy przegląd „dobrych praktyk”³ stosowanych w wybranych uniwersytetach z kręgu anglosaskiego⁴. Tym, co łączy wszystkie brane pod uwagę uczelnie, jest z pewnością duży stopień transparentności systemu ewaluacji (nie oznacza on oczywiście upublicznienia danych o charakterze poufnym, osobowym itp.). Świadczyć o tym może sam fakt zamieszczenia szczegółowych informacji o funkcjonującym na uczelni systemie ewaluacji zajęć dydaktycznych na ogólnodostępnej stronie internetowej. Z pewnością oprócz roli informacyjnej zabieg ten pełni także funkcję uwiarygodniającą, wskazującą na dbałość o jakość nauczania, a pośrednio również i promocyjną.

W analizowanych dokumentach opisujących „polityki ewaluacyjne” uczelni ważnym elementem jest określenie celów, jakim ma służyć ewaluacja nauczania. Przykładowo, Stanford University zwraca uwagę na trzy zasadnicze zadania: dostarczenie informacji zwrotnej wykładowcom, dostarczenie studentom informacji pomocnej w wyborze kursów, dostarczenie ogólnej informacji o jakości kursów dla celów administracyjnych (polityka kadrowa, wynagrodzenia itp.). W podobnym duchu ewaluacja rozumiana jest w polityce University of British Columbia, gdzie za jej podstawowe cele uznaje się: dostarczenie danych, które pozwolą na polepszanie nauczania studentów, dostarczenie danych o ogólnej jakości nauczania na danym kierunku, wydziale itp., dostarczenie danych wykładowcom (także po to, by mogli dalej rozwijać swoje umiejętności), dostarczenie danych uniwersytetowi dla celów operacyjnych (ocena pracownika, wynagrodzenie, nagrody itp.). Zarówno w tych, jak i w innych analizowanych przypadkach, zwraca uwagę silny nacisk na podwójną funkcję ewaluacji – formatywną i sumatywną.

3 Wykorzystane tu informacje zostały zaczerpnięte z ogólnodostępnych źródeł w postaci materiałów zamieszczonych na stronach internetowych uczelni (ich lista znajduje się w bibliografii). Jest to zatem raczej analiza typów idealnych rozwiązań ewaluacyjnych, niż ich realnego funkcjonowania (choć można w sposób uzasadniony zakładać, że wypracowane dokumenty są wynikiem nabytych przez konkretną uczelnię doświadczeń).

4 Uczelnie, do których rozwiązań odwołujemy się, to (w nawiasach podano skróty stosowane w dalszej części artykułu): University of British Columbia (BC), Ohio State University (OSU), Stanford University (SU), University of Auckland (UA). W ograniczonym zakresie korzystamy także z propozycji Oxford University (OxU) oraz Washington State University (WSU).

Wszystkie analizowane uczelnie zwracają także szczególną uwagę na konieczność triangulacji metodologicznej w zbieraniu danych, czyli na wykorzystanie różnych sposobów uzyskiwania informacji na temat prowadzonych zajęć dydaktycznych. Najbardziej popularną, podstawową i obecną we wszystkich jednostkach formą są oczywiście ankiety studenckie (*student ratings*), wśród których można wyróżnić zarówno takie, które oceniają prowadzenie pojedynczego przedmiotu, jak i dokonujące oceny bardziej ogólnej – całości programu studiów, warunków studiowania itp. (np. Univeristy of Auckland – „University Review”, „Graduating Year Review”, „Student Experience and Satisfaction Survey” ; „Oxford University – Student Course Experience Questionnaire”). Ankiety studenckie uznawane są za najlepszą pojedynczą metodę zbierania danych wykorzystywanych w procesie ewaluacji dydaktyki; są też metodą najczęściej stosowaną i mającą największe znaczenie (zob. Paswan, Young 2002). Oprócz ankiet opracowywanych przez ekspertów na potrzeby konkretnych uczeni, wiele placówek wykorzystuje także narzędzia standaryzowane o zweryfikowanej trafności i rzetelności, które dodatkowo dają możliwość porównywania wyników ewaluacji między uczelniami⁵.

Większość analizowanych systemów ewaluacji uwzględnia także jakąś formę recenzji koleżeńskiej, najczęściej w postaci tzw. *peer observations*, w trakcie których wykładowcy obserwują i oceniają swoje zajęcia (często określana jest wymagana częstotliwość tej formy oceny dla każdego z uniwersyteckich stanowisk, np. w przypadku asystentów – raz w roku). Innym proponowanym narzędziem ewaluacji jest samoocena prowadzącego zajęcia. W wybranych przypadkach stanowić ona może część tzw. *teaching portfolio*. Obszerny i szczegółowy opis tej formy zbierania danych zamieszcza na swoich stronach Washington State University. Przygotowywane przez samego prowadzącego portfolio, choć oryginalne i twórcze, powinno zgodnie z zaleceniami uczelni zawierać między innymi następujące elementy: cele stawiane sobie przez prowadzącego w zakresie dydaktyki; obowiązki – w tym: prowadzone kursy, indywidualna praca ze studentem, praca ze specjalnymi grupami studentów, doradztwo, doskonalenie dydaktyczne prowadzącego wprowadzane innowacje; ewaluacje – i m.in. wyniki ocen studenckich, *peer review*, uzyskane nagrody itp.

5 Przykładowo Instructional Development and Effectiveness Assesment (IDEA), Student Instructional Report, Student Evaluation of Educational Quality (SEQ), Instructor and Course Evaluation System, Student Instructional Rating System, Instructional Assessment System (Hobson, Talbot 2001).

Ostatnia część portfolio nosi nazwę „rezultaty” i obejmuje m.in. informacje dotyczące sukcesów studentów, w których osiągnięciu wykładowca miał swój udział, opublikowane artykuły na temat nauczania oraz przygotowane przez prowadzącego materiały dydaktyczne (podręczniki, tzw. readers czy wizualne pomoce naukowe). Innym źródłem cennych opinii na temat dydaktyki (choć w mniejszym stopniu dotyczących poszczególnych zajęć), a także innych aspektów funkcjonowania uczelni, mogą być wyniki badania absolwentów.

W publikowanych dokumentach uczelnie na różnych stopniach szczegółowości określają procedury związane z ewaluacją. Ze względu na duże rozmiary instytucji, często decydują się na jakąś formę decentralizacji. Dokonują jej, delegując opracowanie procedury na poziom wydziału czy departamentu lub dopuszczając do dywersyfikacji narzędzi ze względu na specyfikę jednostek uczelnianych, potrzeby poszczególnych wykładowców czy organizacji studenckich. Ważnym elementem każdego z analizowanych systemów ewaluacji są kwestie poufności w procesie zbierania danych i publikowania informacji. Pierwsza z nich wiąże się z zapewnieniem anonimowości studentów oceniających kurs (np. Stanford University deklaruje, że studenci wypełniają elektroniczne ankiety na zewnętrznym, niezależnym serwerze). W drugim przypadku uczelnie często zastrzegają sobie prawo do nieewaluowania lub niebrania pod uwagę wyników ewaluacji kursów prowadzonych po raz pierwszy (lub będących pierwszymi kursami prowadzącymi), kursów ze zbyt małą liczbą oceniających studentów itd. Niemal wszystkie z analizowanych uczelni organizują oceny studenckie w końcowej części semestru (lub innej jednostki rozliczeniowej), przed egzaminami.

Te z uczelni, które w sposób szczegółowy określają procedurę ewaluacji dydaktyki, często umieszczają w niej także obowiązujący w ramach instytucji obieg informacji związanych z oceną prowadzących i zajęć, w tym kto i do jakich informacji ma dostęp. Warto podkreślić także uwzględnienie roli studentów w procesie ewaluacji, zawarte w większości analizowanych dokumentów. W najbardziej oczywistym wymiarze wiąże się ono z oparciem ewaluacji dydaktyki na ocenach studentów. Inne wybrane aspekty tego zaangażowania to w zależności od uczelni: rola organizacji studenckich jako podmiotu zainteresowanego wynikami prowadzonej ewaluacji, możliwość wprowadzenia przez taką organizację własnych pytań do formularza ankiety, a także przeprowadzenia oceny zajęć na wniosek studentów (UA).

Ostatni element, o którym warto wspomnieć w kontekście systemów ewaluacji funkcjonujących w analizowanych uczelniach, związany jest z formatywnym wymiarem ewaluacji. Mowa tutaj o rozmaitych mechanizmach wsparcia, jakie udzielane jest wykładowcom zarówno w samym procesie dydaktycznym, jak i w kontekście prowadzonej ewaluacji i jej wyników. Niemal wszystkie uczelnie dysponują specjalnie wydzielonymi jednostkami zajmującymi się jakością nauczania, które stanowią nie tylko podmiot w mniejszym lub większym stopniu czuwający nad realizacją polityki ewaluacyjnej, ale także udzielający porad i wsparcia dydaktykom. Godne uwagi są także proste w swej zasadzie, a niezwykle użyteczne przedsięwzięcia w postaci przygotowania zbioru najczęściej zadawanych przez kadre i studentów pytań dotyczących ewaluacji – FAQ (SU), czy też dokumentu wyjaśniającego przyjęte kryteria oceny oraz wspomagającego proces interpretacji uzyskanych danych (OSU).

Czy oceny studentów są wiarygodne?

W podtekście tego pytania kryje się brak zaufania do poprawności ocen przyznawanych prowadzącym przez studentów i przekonanie o ich możliwym zniekształceniu przez różne czynniki. Czy studenci mają wystarczające kompetencje, by prawidłowo oceniać sposób prowadzenia zajęć? Czy ich ocena nie jest swoistym „konkurem” na najbardziej lubianego prowadzącego/ą? Czy najlepszych ocen nie dostają zwłaszcza te osoby, u których łatwo dostać dobrą ocenę? Czy nie lepiej byłoby, gdyby prowadzących oceniali inni przedstawiciele kadry?

Warto przypomnieć, że ocena zajęć dokonywana przez studentów jest w istocie specyficznym rodzajem pomiaru, który zgodnie z podstawami metodologii powinien być *rzetelny* (a więc m.in. powtórne zmierzenie tego samego obiektu tym samym narzędziem powinno przynieść ten sam wynik) i *trafny* (a więc narzędzie powinno mierzyć to, do czego zostało zaprojektowane, np. ocena sposobu prowadzenia zajęć może dotyczyć umiejętności wyjaśniania trudnych zagadnień przez wykładowcę, ale jej wynik nie powinien zależeć od motywacji studentów do nauki, trudno bowiem karać prowadzącego za gorsze oceny uzy-

skiwane w grupach wyraźnie mniej chętnych do pracy). Idealne spełnienie tych warunków pozostaje oczywiście poza zasięgiem jakiegokolwiek pomiaru, jednak ankiety studenckie są uznawane za rzetelne i dość trafne narzędzia.

Rzetelność ankiet potwierdzono w wielu badaniach, w których zazwyczaj stosowano wystandaryzowane i używane powszechnie na uczelniach narzędzia (ze względu na mnogość prac w dziedzinie ewaluacji podawane poniżej wyniki zazwyczaj opierają się na metaanalizach wielu badań). Większość poprawnie skonstruowanych narzędzi legitymuje się wewnętrzną spójnością na co najmniej akceptowalnym poziomie (współczynniki spójności na poziomie między 0,7 a 0,9 [Cashin 1995; Ali, Sell 1998]). Wykazano również, że podstawowym czynnikiem decydującym o ocenie jest osoba prowadzącego, a nie wykładany przedmiot, co dodatkowo wpływa na stabilność ocen (Cashin 1995). O wysokiej rzetelności świadczy także duża zgodność oceny danych zajęć wśród uczestniczących w nich studentów, a także podobieństwo oceny tych samych zajęć przez obecnych studentów i absolwentów minimum rok po ukończeniu studiów (mediana korelacji na poziomie 0,83 [Overall, Marsh 1980 za: Overall, Marsh 1982]). Nie do utrzymania byłby zatem argument, że studenci nie są w stanie adekwatnie ocenić zajęć lub że mając większe doświadczenie zawodowe i życiowe, będą oceniać je inaczej. Wreszcie, okazuje się, że studenci są średnio znacznie bardziej zgodni w ocenie danych zajęć niż przedstawiciele kadry dydaktycznej, którzy również je obserwowali (wewnętrzna korelacja ocen wśród przedstawicieli kadry była na poziomie 0,26 [Hobson, Talbot 2001]). Co więcej, oceny innych dydaktyków nie korelowały wystarczająco wysoko z innymi kryteriami jakości dydaktyki, chyba że oceniający zostali wcześniej przeszkoleni pod kątem zasad dokonywania oceny.

Warto zaznaczyć, że warunkiem poprawności ewaluacji jest wykorzystanie danych zebranych w odpowiedni sposób. Szczególnie istotne jest tutaj traktowanie z ostrożnością wyników ankietyzacji w małych grupach zajęciowych (poniżej 15 osób [Cashin 1995]), ponieważ mała liczba badanych jest zagrożeniem dla rzetelności wyników. Warunkiem koniecznym jest również wysoka stopa zwrotów, bowiem osoby, które uczestniczyły w zajęciach, a nie oceniły ich, mogą różnić się znacząco od tych, które wyraziły swoje opinie (przykładowo, może się okazać, że ankiety są najchętniej wypełniane przez najbardziej niezadowolonych lub najbardziej zadowolonych studentów, co w oczywisty sposób zniekształca wyniki). Czynnikiem stopy zwrotów jest bardzo istotny zwłaszcza w wypadku ankiet internetowych

(Johnson 2002), często wybieranych z racji niskich kosztów zbierania i obróbki danych. Popularne sposoby „zachęcenia” do wypełniania ankiet (np. zablokowanie możliwości podglądu ocen egzaminacyjnych, brak możliwości zaliczenia semestru itp.) mogą jednak budzić wątpliwości co do jakości danych uzyskiwanych w ten sposób (ankiety mogą być wypełniane pośpiesznie lub nieuważnie).

Większe wątpliwości wiążą się z problemem trafności ocen studenckich. Zakłada się, że jeżeli metody oceny są trafne, prowadzący, na których zajęciach studenci uczą się najwięcej i najefektywniej, powinni być oceniani najwyżej. Sprzeczny z tym przekonaniem jest tzw. efekt doktora Foxa – udający prowadzącego aktor, który wygłosił dla studentów efektowny, ale pozbawiony merytorycznej treści wykład, został bardzo wysoko oceniony (Naftalin, Ware, Donnelly 1973 za: Overall, Marsh 1982, s. 11). Jak się okazało, eksperyment miał istotne metodologiczne słabości, ale pewne wątpliwości pozostały, tym bardziej że inne badania pokazały zależność ocen od ekspresyjności, pasji i dynamiki prowadzącego.

Najczęściej stosowanym sposobem wykazania trafności ocen studenckich jest porównanie ich wyników z ocenami przeprowadzonymi w inny sposób. Metaanalizy pokazują, że istnieje umiarkowanie silny związek oceny zajęć przez studentów ze stopniem wynoszonej przez nich wiedzy, potwierdzonej przez zewnętrzne egzaminy (d’Apollonia, Abrami 1997, za: Ali, Sell 1998) i wspólne egzaminy grup, które uczyli różni prowadzący (Cashin 1995; Felder 1992). Choć związek ten nie jest bardzo silny, typowo uzyskiwane wartości współczynników korelacji (ok. 0,3–0,5) są zwyczajowo uznawane w naukach społecznych czy w psychologii za pokazujące ważne związki. Dodatkowo warto zauważyć, że ostateczny wynik procesu kształcenia nie zależy przecież wyłącznie od umiejętności prowadzącego, ale także od takich czynników, jak: motywacja studentów, czas poświęcony na naukę czy kapitał kulturowy wyniesiony z domu rodzinnego. Większość badań potwierdza również trafność ocen studenckich, udokumentowaną korelacjami z wynikami innych metod oceny prowadzących. Stwierdzono podobieństwo na średnim poziomie między wynikami ankiet studenckich a samooceną prowadzących, średnie lub wysokie podobieństwo ocen studentów i absolwentów oraz różny poziom podobieństwa opinii studentów i innych prowadzących, którzy z kolei są zazwyczaj mniej zgodni ze sobą niż studenci (Ali, Sell 1998; Cashin 1995; Felder 1992; Gursoy, Umbreit 2005; Hobson, Talbot 2001).

Jak się wydaje, wokół ankiet studenckich narosło także wiele błędnych przekonań dotyczących czynników, od których ich wyniki są uzależnione i które mogą zniekształcać rezultaty oceny, czyniąc ją nietrafną (Simpson, Siguaw 2000). Warto zatem zdawać sobie sprawę z tego, jakie czynniki rzeczywiście wpływają na wysokość ocen studenckich – oto ich lista oparta na metaanalizach wielu badań (Cashin 1995; Felder 1992; Stapleton, Murkison 2001)⁶:

- motywacja studentów – wyższe oceny przyznają studenci, którzy są zainteresowani przedmiotem lub mocno zmotywowani do nauki;
- dobrowolność udziału w kursie – wyżej oceniane są zajęcia wybierane niż obowiązkowe;
- spodziewana ocena – ten czynnik był szczególnie starannie badany (zob. Centra 2003; Eiszler 2002; Greenwald 1996; Gump 2007), a uzyskiwane wyniki nie są zupełnie jednoznaczne, jednak istotna część badań pokazuje niestety niewielki (współczynniki korelacji 0,1–0,3) dodatni związek między oceną za kurs, jakiej spodziewają się studenci, a ich opinią na temat prowadzącego, istnieje zatem pewna możliwość „kupowania” przychylności studentów za podwyższanie ocen (a właściwie za dawanie do zrozumienia, że oceny będą wysokie).
- ekspresyjny styl prezentacji – pokazuje to wspomniany wyżej „efekt doktora Foxa”, ale akurat ta cecha wydaje się wspierać proces uczenia się, zatem nie przekreśla trafności ocen;
- poziom zaawansowania kursu – wyżej oceniane są kursy zaawansowane niż wprowadzające, jednak związek ten jest słaby;
- dyscyplina akademicka – zgodnie z oczekiwaniami wyżej oceniane są kursy z dziedzin humanistycznych, nieco niżej z nauk społecznych, a najniżej przedmioty matematyczno-przyrodnicze;
- wymagany nakład pracy – przeciwnie do oczekiwań wyżej (choć związek ten jest słaby) oceniane są zajęcia wymagające od studentów znacznego nakładu pracy, jakkolwiek część badań pokazuje związek krzywoliniowy, tzn. przekroczenie pewnego poziomu wymagań skutkuje pogorszeniem wyników ankiet;

⁶ Warto jednak zaznaczyć, że zdecydowaną większość z nich przeprowadzono w Stanach Zjednoczonych, a więc w innym kontekście kulturowym. Co więcej, z reguły mają one charakter korelacyjny, potwierdzają zatem pewne ogólne i niejako uśrednione prawidłowości – w poszczególnych przypadkach mamy jednak zapewne do czynienia z sytuacjami, w których np. wymagający prowadzący są oceniani nisko albo pleć osoby prowadzącej obniża lub zawyża oceny.

- brak zapewnionej anonimowości oceniających – opinie wyrażane nieanonimowo są zazwyczaj wyższe;
- obecność osoby prowadzącej zajęcia w czasie oceny – ten pozornie nieważny szczegół w istocie ma znaczenie, ponieważ wypełniane w obecności wykładowców ankiety prowadzą do wyższej oceny ich pracy;
- ujawnienie celu ocen – jeżeli studenci zdają sobie sprawę, że wyniki ewaluacji będą używane do decyzji personalnych, oceniają wyżej, niż gdy opinie mają służyć jedynie do poprawy warsztatu osoby prowadzącej;
- niektóre cechy osobowości osoby prowadzącej – szczególnie entuzjazm, energiczność i pozytywna samoocena; związek ten jest o tyle istotny, że osobowość nie ulega zasadniczym zmianom w ciągu życia, trudno więc w tym zakresie doskonalić swój warsztat⁷.

Zaskakujące jest jednak, jak wiele czynników w istocie nie ma wpływu na wysokość ocen studenckich. Większość badań nie pokazuje wpływu (Clayson 1999): wieku, płci i rasy osoby prowadzącej, doświadczenia w nauczaniu (oceny prowadzących są stabilne w czasie z wyjątkiem pierwszego okresu pracy, dokładniejsze badania pokazują jednak, że nie zmienia się komponent związany z oceną osobowości, a poprawić się może wiedza czy warsztat osoby prowadzącej), dorobku naukowego, wieku, płci i osobowości studentów, momentu dokonywania ewaluacji (pod warunkiem, że ewaluacja jest przeprowadzana w 2. połowie semestru), rozmiaru grupy zajęciowej (choć część badań pokazuje, że mniejsze grupy oceniają prowadzących wyżej).

Warto przypomnieć, że ankiety badają opinie studentów, są więc w procesie ewaluacji *danymi wejściowymi*, a nie jej *wynikiem*. Oceny studentów pozostają najbardziej rzetelnym i trafnym sposobem zbierania danych ewaluacyjnych, jakkolwiek ostateczna ewaluacja nie jest zadaniem studentów, lecz osób zarządzających instytucją edukacyjną. Podatność na pewne zniekształcenia nie przekreśla ich wartości (tym bardziej że trudno je zastąpić), lecz wymaga świadomości i sposobów przeciwdziałania, z których najbardziej polecaną jest wykorzystanie danych z dodatkowych źródeł. Prosty sposób jest wykorzystanie w ewaluacji nie tylko odpowiedzi na pytania ilościowe (np. mierzone na skali 1–5), ale także

7 Istnieją również badania pokazujące, że postrzegana przez studentów osobowość prowadzącego jest czynnikiem, którym można wyjaśnić na 73% wariacji ocen zajęć (Clayson, Sheffet 2006). Co więcej, istotny wpływ tego czynnika da się stwierdzić już po 5 minutach od rozpoczęcia pierwszych zajęć.

na jakościowe pytania otwarte (np. o słabe i mocne strony kursu czy osoby prowadzącej). Naturalnie konieczne jest w tym wypadku uniemożliwienie identyfikacji osoby udzielającej odpowiedzi. Informacje uzyskiwane w ten sposób są znacznie bogatsze i zawierają o wiele więcej konkretnych wskazówek dla prowadzących niż niewiele mówiące wartości liczbowe. Ich wartość formacyjna jest zatem o wiele większa.

* * *

Na koniec kilka słów refleksji natury ogólnej. Kultury ewaluacyjnej, którą zdefiniować można jako formalne i nieformalne zakorzenienie (również na poziomie zwyczajów, praktyk i sposobów myślenia) zasad wzajemnego i opartego na jasnych kryteriach oceniania oraz nie mniej ważnej otwartości na wyniki tej oceny, nie buduje się ani łatwo, ani szybko. Mimo że ocena jest stałym elementem życia uniwersyteckiego, zwykle dokonywana jest ona hierarchicznie (np. ocena studenta przez egzaminatora czy wnioskodawcy grantu przez komisję składającą się z uznanych ekspertów w dziedzinie), a jej kryteria nie zawsze są czytelne i intersubiektywne. Braków w tej sferze polskiego szkolnictwa wyższego, o których może najlepiej świadczy niewielka liczba krajowych publikacji i konferencji na ten temat, nie usuną formalne wymagania Państwowej Komisji Akredytacyjnej. Praktycznym czynnikiem wymuszającym rozwój może być natomiast presja demograficzna i rosnąca konkurencja, także na globalnym rynku edukacyjnym. W tej sytuacji wykorzystanie ewaluacji jako narzędzia podwyższania jakości kształcenia może być zarówno czynnikiem wprowadzania zmian związanych z przeorientowaniem uczelni w kierunku nastawienia na studenta, jak i budowania przewagi konkurencyjnej względem innych placówek.

Bibliografia:

Ali, Donna L., Sell, Yvonne (1998), *Issues Regarding the Reliability, Validity and Utility of Student Ratings of Instruction: A Survey of Research Findings*, Calgary: University of Calgary APC Implementation Task Force on Student Ratings of Instruction [15.02.2009].

Cashin, William 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.

Centra, John A. (2003), *Will Teachers Receive Higher Student Evaluations by Giving Higher Grades and Less Course Work?*, „Research in Higher Education”, 44(5), s. 495–518.

Clayson, Dennis E. (1999), *Students' Evaluation of Teaching Effectiveness: Some Implications of Stability*, „Journal of Marketing Education”, 21(1), s. 68–75.

Clayson, Dennis E., Sheffet, Mary Jane (2006), *Personality and the Student Evaluation of Teaching*, „Journal of Marketing Education”, 28(2), s. 149–160.

Eizsler, Charles F. (2002), *College Students' Evaluations of Teaching and Grade Inflation*, „Research in Higher Education”, 43(4), s. 483–501.

Felder, Richard M. (1992), *What Do They Know, Anyway?*, „Chemical Engineering Education”, 26(3), s. 134–135.

48 Felder, Richard, Brent, Rebecca (2004), *How to evaluate teaching?*, „Chemical Engineering Education”, 38(3), s. 200–202.

Greenwald, Anthony G. (1996), *Applying Social Psychology to Reveal a Major (But Correctable) Flaw in Student Evaluations of Teaching*, Papers presented at the Annual Meeting of the American Psychological Association, New York.

Gump, Steven E. (2007), *Student Evaluations of Teaching Effectiveness and the Leniency Hypothesis: A Literature Review*, „Educational Research Quarterly”, 30(3), s. 56–69.

Gursoy, Dogan, Umbreit, Terry W. (2005), *Exploring Students' Evaluations of Teaching Effectiveness: What Factors Are Important?*, „Journal of Hospitality and Tourism Research”, 29(1), s. 91–109.

Hobson, Suzanne M., Talbot, Donna M. (2001), *Understanding Student Evaluations: What All Faculty Should Know*, „College Teaching”, 49(1), s. 26–31.

Johnson, Trav (2002), *Online Students Rating: Will Students Respond?*, Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Overall, Jesse U. IV, Marsh, Herbert W. (1982), *Students' Evaluations of Teaching: An Update*, „AAHE Bulletin” 1982, s. 9–12.

Paswan, Audhesh K., Young, Joyce A. (2002), *Student Evaluation of Instructor: A Nomological Investigation Using Structural Equation Modeling*, „Journal of Marketing Education”, 24(3), s. 193–202.

Simpson, Penny M., Siguaw, Judy A. (2000), *Student Evaluations of Teaching: An Exploratory Study of the Faculty Response*, „Journal of Marketing Education”, 22(3), s. 199–213.

Stapleton, Richard John, Murkison, Gene (2001), *Optimizing the Fairness of Student Evaluations: A Study of Correlations Between Instructor Excellence, Study Production, Learning Production, and Expected Grades*, „Journal of Management Education”, 25(3), s. 269–291.

Strony internetowe uczelni:

Ohio State University, http://oaa.osu.edu/eval_teaching/index.html, <http://www.ureg.ohio-state.edu/ourweb/scansurvey/index.html> [13.07.2009].

Oxford University, <http://ceq.oucs.ox.ac.uk/intro.cfm>, <http://www.learning.ox.ac.uk/>, <http://ceq.oucs.ox.ac.uk/interpretOSCEQ.cfm> [13.07.2009].

Stanford University, http://registrar.stanford.edu/faculty/course_evaluations/index.htm [13.07.2009].

University of Auckland, http://cad.auckland.ac.nz/content/files/apg/evaluation_of_teaching.pdf, <http://www.education.auckland.ac.nz/uoa/fms/default/education/about/faculty/committee/docs/Review%20and%20Evaluation%20of%20Teaching%20TLQ%20Policy.pdf> [13.07.2009].

Washington State University, <http://www.wsu.edu/provost/teaching.htm> [13.07.2009].

University of British Columbia, http://www.vpacademic.ubc.ca/policy_on_student_evaluation_of_teaching.pdf [13.07.2009].

Grzegorz Bryda

Magdalena Jelonek

Barbara Worek

Jak skonstruować dobre narzędzie do ewaluacji jakości zajęć dydaktycznych? Refleksje praktyczne

Niniejszy artykuł poświęcony został praktyce ewaluacji jakości zajęć dydaktycznych, w tym możliwościom w zakresie wykorzystania ilościowych i jakościowych metod oceny pracy nauczycieli akademickich. Zaprezentowane zostały w nim takie kwestie, jak:

- etapy konstruowania narzędzia do oceny jakości kształcenia w szkole wyższej;
- podstawowe zasady konstrukcji pytań oceniających pracę nauczycieli akademickich;
- czynniki, które mogą zaburzyć rzetelną ocenę jakości dydaktyki oraz możliwości w zakresie zapobiegania wpływowi tych czynników na wyniki badań;
- najczęściej popełniane błędy w procesie konstruowania narzędzi ewaluacyjnych, zasady analizy oraz prezentacji materiału empirycznego.

Artykuł oprócz opracowania teoretycznego zawiera przykłady praktyczne oparte zarówno na praktykach międzynarodowych, jak i na rozwiązaniach, które wdrożył Uniwersytet Jagielloński.

Wprowadzenie

Wydaje się, że ocenianie zajęć dydaktycznych przez studentów przestaje być powoli w polskim systemie szkolnictwa wyższego czymś unikatowym, świadczącym o tym, że uczelnia przywiązuje wagę do jakości kształcenia oraz że traktuje studentów w sposób partnerski, doceniając znaczenie ich opinii. Ponieważ posiadanie wewnętrznego systemu zapewniania jakości kształcenia, obejmującego też badania opinii studenckich, jest warunkiem uzyskania pozytywnej oceny w procesie akredytacji, każda uczelnia starająca się o nią musi wykazać się posiadaniem takiego systemu. Formalnie można więc przyjąć, że wszystkie uczelnie posiadające akredytację do prowadzenia studiów danego stopnia na konkretnym kierunku, podejmują działania zmierzające do oceny jakości kształcenia i uwzględniają w tych działaniach opinie studenckie jako źródło danych na temat przebiegu procesu dydaktycznego¹.

Doceniając zatem fakt, że prowadzenie badań oceniających jakość dydaktyki staje się powoli obowiązującym standardem, chcemy zauważyć, że badania te będą wartościowym i użytecznym źródłem informacji tylko wtedy, gdy będą prowadzone w sposób gwarantujący trafność i rzetelność wyników oraz gdy ich realizacji będzie towarzyszyła refleksja dotycząca celu i sposobu wykorzystania zebranych danych. W tym przypadku, podobnie jak i w wielu innych, kluczowy jest bowiem nie tyle sam fakt badania opinii studenckich, ile to, jak owe badania są prowadzone oraz jak wykorzystywane są ich wyniki.

1 Warto tutaj zaznaczyć, że samo pojęcie „jakość kształcenia” jest różnie definiowane i rozumiane. Jakość może być rozumiana jako ceniony zbiór cech procesu dydaktycznego wobec stanu rzeczywistego lub możliwego do osiągnięcia czy, też jako stabilność założonych parametrów przyjętego modelu kształcenia. Jakość kształcenia jest postrzegana z dwóch perspektyw: wewnętrznej i zewnętrznej (Payne 1996:264). Pierwsza, wskazuje na zgodność procesu kształcenia z określonymi normami (np. standardy akredytacyjne), druga, na subiektywną ocenę doświadczeń jednostek uczestniczących w procesie kształcenia (np. ocena zajęć dydaktycznych przez studentów). Definicje jakości kształcenia można sprowadzić do dualnego podziału, zgodnie z którym jakość to zgodność z przyjętymi standardami (kryteria obiektywne) lub jakość to interpretowana użyteczność jednostkowa (kryteria subiektywne). W praktyce trudno utrzymać ten podział. Obecnie łączy się te dwa aspekty jakości z wyraźnym naciskiem na użyteczność – kształcenie umiejętności. Świadczą o tym np. dyskusje dotyczące zwiększenia autonomii programowej szkół wyższych i wypracowania sposobów ewaluacji efektów procesu kształcenia.

Na trafność² i rzetelność³ badań wpływa wiele czynników, które nie sposób omówić w tak krótkim opracowaniu. Upraszczając nieco problem, można przyjąć, że wpływa na nie poprawna konceptualizacja badań, ich właściwa operacjonalizacja, przeprowadzenie zgodnie z przyjętymi dla danej techniki standardami, poprawna analiza danych, uwzględniająca wpływ potencjalnych źródeł obciążeń wyników⁴. Omówienie wszystkich tych kwestii wymagałoby obszernego opracowania, dlatego w niniejszym artykule skoncentrujemy się przede wszystkim na narzędziu badawczym i możliwych błędach, jakie mogą się pojawiać przy jego konstrukcji. Budowę narzędzia badawczego będziemy jednak rozpatrywać w powiązaniu z innymi etapami realizacji badań: odniesiemy się do niezbędnych działań, które powinny poprzedzać konstrukcję kwestionariusza do badań opinii studenckich na temat jakości zajęć, omówimy czynniki, które mogą obciążać wyniki takich badań, przedstawimy też podstawowe zasady analizy ich wyników i sposoby prezentacji danych. Na koniec odniesiemy się do różnego rodzaju obaw związanych z oceną zajęć, występujących zarówno wśród nauczycieli akademickich, jak i wśród studentów.

Etapy konstruowania narzędzia do oceny jakości kształcenia w szkole wyższej

Posiadanie trafnego i rzetelnego narzędzia badawczego, które pozwoli monitorować jakość usług edukacyjnych, jest kluczową potrzebą każdego systemu oceny jakości kształcenia w szkole wyższej. Najczęściej narzędziem tym jest kwestionariusz ankiety ewaluacyjnej, dystrybuowany w formie papierowej lub elektronicznej (online), np. w systemie USOS⁵. Kwestionariusz ankiety ewaluacyjnej musi wynikać z logiki zastosowanego modelu ewa-

2 Cecha narzędzia, która wskazuje na to, że mierzy ono to, co w intencji twórców powinno mierzyć.

3 Cecha narzędzia, która określa dokładność pomiaru.

4 Wymienione czynniki odpowiadają przede wszystkim etapom i działaniom związanym z realizacją badań ilościowych, prowadzonych technikami sondażowymi. Pomijamy tutaj kwestie związane z trafnością i wiarygodnością danych zbieranych za pomocą innych technik oceny zajęć dydaktycznych, choć techniki te mogą, a nawet powinny być wykorzystywane jako komplementarne w stosunku do technik sondażowych. Ponieważ badania ilościowe są jednak wykorzystywane najpowszechniej, zdecydowaliśmy się do nich zawęzić analizę.

5 Niektóre jednostki akademickie stosują uproszczone, wewnętrzne systemy prowadzenia badań ewaluacyjnych, oparte na innych niż USOS rozwiązaniach. Najczęściej sytuacja ta dotyczy szkół prywatnych.

luacji i strategii metodologicznej przyjętej do oceny procesu jakości kształcenia w szkole wyższej. Przykładowo, na Uniwersytecie Jagiellońskim zastosowanie do oceny zajęć dydaktycznych systemu USOS (ankiety w wersji online) spowodowało systematyczny wzrost liczby zajęć i pracowników objętych oceną⁶. Zważywszy na to, że system jest względnie nowy, przyjęta strategia ma na celu przede wszystkim „pokrycie oceną” w formie ankiety elektronicznej wszystkich zajęć i pracowników uczelni, zaś dopiero w drugiej kolejności zwiększanie frekwencji oceniających studentów.

W praktyce problem konstruowania narzędzia ewaluacji zajęć sprowadza się do wyboru jednej z dwóch strategii działania. Pierwsza to zbudowanie od podstaw modelu jakości kształcenia poprzez wyróżnienie kluczowych wymiarów pojęcia, a następnie skonstruowanie modelu pomiarowego (systemu wskaźników) do każdego z jego podwymiarów. Druga opiera się na rekonstrukcji modelu jakości kształcenia w oparciu o istniejące już rozwiązania cząstkowe (poprawa systemu, który już funkcjonuje w uczelni wyższej). Jeszcze inne rozwiązanie to kopiowanie wzorców – narzędzi ewaluacji – europejskich lub polskich szkół wyższych (tzw. izomorfizm mimetyczny).

Oczywiście zanim przystąpimy do budowy własnego narzędzia – kwestionariusza – należy sobie zadać pytanie, czy kwestionariusz ankiety jest rzeczywiście najlepszym narzędziem umożliwiającym odpowiedź na pytania ewaluacyjne? Czy inne metody nie są bardziej adekwatne do oceny jakości kształcenia? Niezależnie od tego, modelowa ścieżka konstrukcji narzędzia ewaluacji w formie kwestionariusza ankiety wymaga wykorzystania w pierwszej fazie metod jakościowych, najczęściej w postaci serii grup fokusowych (ewentualnie wywiadów pogłębionych) wśród studentów i pracowników. Pomagają one w rekonstrukcji wymiarów oceny jakości kształcenia, poprzez zidentyfikowanie elementów procesu kształcenia, które zdaniem badanych są najważniejsze. W niektórych przypadkach wywiady grupowe mogą w operacjonalizacji (m.in. uzupełnienie kafeterii odpowiedzi, pierwszy test narzędzia). Niezwykle ważnym, a jednak często pomijanym etapem w procesie konstrukcji kwestionariusza oceny studenckiej, jest pilotaż narzędzia. Pilotaż, będący drugim po preteście testem kwestionariusza, dostarcza racjonalnych podstaw ku temu, żeby wyselekcjonować

⁶ W roku akademickim 2006/2007 ocenionych w formie elektronicznej było 67% wszystkich zajęć/86% pracowników, w roku 2007/2008 – 78% zajęć/93% pracowników, zaś w roku 2008/2009 – 86% zajęć/96% pracowników.

te pozycje, które najlepiej wskazują na ukryte wymiary jakości kształcenia, a także te, które w największym stopniu różnicują badane obiekty. Tym samym umożliwia maksymalną redukcję liczby pytań w ankiecie, przy minimalnej stracie wyjaśnionej wariancji ukrytego konstruktów. Najczęściej wykorzystywanymi w tym zakresie technikami statystycznymi są: eksploracyjna analiza czynnikowa (EFA) oraz analiza głównych składowych (PCA).

Kwestionariusz ankiety ewaluacyjnej musi być krótki, lecz treściwy, co nie jest łatwe, szczególnie wtedy, gdy w ocenie elektronicznej student dostaje przeciętnie 6–10 kwestionariuszy odnoszących się do zajęć, na które uczęszcza. W praktyce – dzięki temu, że system jest otwarty przez określony czas – studenci wracają, aby wypełnić ankiety, które im pozostały z wcześniejszej oceny.

Jak pokazują doświadczenia konstruowania kwestionariusza ankiety ewaluacyjnej na Uniwersytecie Jagiellońskim, aby możliwe były późniejsze porównania między jednostkami, zajęciami czy pracownikami uczelni, niezbędna jest standaryzacja narzędzia. Jej dobrym uzupełnieniem jest „kontekstualizacja” narzędzia, uzyskiwana m.in. przez zastosowanie skal z opisanymi krańcami (tzw. dyferencjał semantyczny). W przypadku Uniwersytetu Jagiellońskiego wykorzystanie dyferencjału semantycznego dało klarowny układ odniesienia dla każdego pytania ankiety oceny zajęć dydaktycznych (podejście to jednak nadal jest w fazie testów).

Na podstawie wcześniejszych doświadczeń z oceną zajęć dydaktycznych na każdym wydziale uznano, że najlepszym rozwiązaniem z punktu widzenia budowy uniwersyteckiego systemu doskonalenia jakości kształcenia jest elektroniczna ocena zajęć dydaktycznych. Standaryzacja elektronicznej ankiety ewaluacyjnej została przeprowadzona na poziomie całego Uniwersytetu, z uwzględnieniem typów zajęć, różnic między jednostkami i przedmiotami w danej jednostce. Przy standaryzacji ankiety ewaluacyjnej kierowano się zasadą zachowania autonomii i uwzględniania specyfiki poszczególnych jednostek Uniwersytetu Jagiellońskiego (wydziałów, instytutów lub katedr). Istotnym i cenionym, szczególnie przez wykładowców, elementem oceny zajęć stała się możliwość uzupełnienia oceny ilościowej poprzez studenckie komentarze, uwagi i sugestie.

Ocena pracy nauczyciela akademickiego – zasady konstrukcji pytań kwestionariuszowych

Skonstruowanie dobrego narzędzia oceniającego jakość prowadzonych zajęć wymaga podjęcia szeregu decyzji. W pierwszej kolejności odpowiedzieć musimy na pytania koncepcyjne i operacyjne. Jakich informacji potrzebujemy?⁷ W jaki sposób te informacje będą wykorzystywane wewnątrz uczelni? W jaki sposób, w jakiej formie (wartości zagregowane/indywidualne) oraz komu prezentowane będą wyniki? W drugiej kolejności podjąć musimy decyzje metodologiczne. Jakiej metody użyć? Jaką formę pytania wybrać (pytania otwarte/zamknięte)? Jak sformułować pytania i kafeterie odpowiedzi? Czy z punktu widzenia założonej metodologii liczba pytań jest wystarczająca, a może za duża? Jak uniknąć zniechęcenia badanych? W jakiej kolejności zadać pytania?

W większości kwestionariuszy oceniających pracę nauczycieli akademickich znajdują się zarówno pytania o fakty (np. dostępność sylabusu, literatury przedmiotu), jak i o opinie czy subiektywne doświadczenia (ocena prowadzącego zajęcia, por. Brzezińska, Brzeziński 2004; Oppenheim 1992). W przypadku obu typów pytań kwestionariuszowych wyróżnić można obowiązujące schematy ich konstrukcji oraz najczęściej spotykane przy ich tworzeniu błędy. Przykładowo, pytania o fakty, które mają na celu uzyskanie obiektywnych informacji, zawsze powinny umożliwiać respondentowi wybór opcji „nie wiem”/„nie potrafię sobie przypomnieć”. W procesie oceny pracy nauczycieli akademickich zdecydowanie częściej wykorzystuje się klasyczne skale ocen szacunkowych, w których to kategorie odpowiedzi stanowią kwantyfikatory wskazujące na stopień intensywności danej oceny (por. Brzezińska, Brzeziński 2004). Przykładowo, skale

⁷ W związku z wielowymiarowym charakterem jakości kształcenia i nieraz odmiennym definiowaniem przez autorów składających się na to pojęcie wymiarów trudno jest wyróżnić wszystkie elementy budujące wysoką jakość nauczania. Za Williamem E. Cashinem (1995) możemy wymienić najczęściej oceniane w ankietach ewaluacyjnych obszary jakości, którymi są: sposób organizacji zajęć, jasność prezentacji i umiejętności komunikacyjne nauczyciela akademickiego, stosunek nauczyciela do studentów, stopień trudności kursu, sposób oceny pracy studenta, samoocena studenta w zakresie efektów uczenia się. Ponadto kategorie te uzupełnić można o: ocenę koncepcji i formy zajęć, atrakcyjność zajęć, przygotowanie nauczyciela do zajęć, zaangażowanie prowadzącego, umiejętność zainteresowania przedmiotem nauczania, umiejętność wykorzystania czasu czy gotowość do udzielania dodatkowej pomocy. W badaniach często pojawia się dodatkowe pytanie o skłonność do rekomendowania danego kursu innym studentom, które traktowane jest jako jeden z kluczowych wskaźników stopnia zadowolenia z uczestnictwa w kursie.

takie przybrać mogą postać klasycznego pytania typu Likerta z kategoriami odpowiedzi od „zdecydowanie się zgadzam” do „zdecydowanie się nie zgadzam”.

W związku z ograniczoną objętością niniejszego artykułu poniżej, w sposób skrótowy, przedstawiono zasady konstrukcji pytań kwestionariuszowych służących do oceny pracy nauczyciela akademickiego. Zasady te ograniczone zostały do najważniejszych reguł, a jednym z kryteriów ich wyróżnienia była częstość „łamania” danego założenia w praktyce badawczej. W związku z tym, że najłatwiej uczyć się na własnych bądź cudzych błędach, do każdej z reguł dopasowane zostały przykłady błędnie sformułowanych pytań, które łamią daną regułę. Pytania te pochodzą z kwestionariuszy ewaluacyjnych, stosowanych w praktyce, które zgromadzone zostały dla potrzeb niniejszego artykułu⁸.

Reguła 1. Zadawaj pytania w sposób jasny, jednoznaczny, zrozumiały i dopasowany do typu zajęć (np. wykład / konwersatorium / laboratorium).

W wielu analizowanych kwestionariuszach ewaluacyjnych pojawiają się pytania, które mogą być **wieloznacznie rozumiane**, jak np. prośba o ocenę stwierdzenia: „zachowane są proporcje między zajęciami i przerwami” czy o ocenę kategorii nazwanej: „poziom organizacji struktury treści wykładu”. Pytania takie mogą niejednokrotnie generować odpowiedzi różnorodnie interpretowane.

W jednym z kwestionariuszy znajdujemy tak sformułowane pytanie: „Czy zajęcia odbywały się regularnie?” O czym będzie świadczyła przewaga odpowiedzi, że nie? O tym, że prowadzący nagminnie przekładał zajęcia czy może o tym, że zajęcia odbywały się w systemie blokowym (a więc nieregularnie)? Samo sformułowanie „regularnie” budzi wiele wątpliwości, podobnie zresztą jak kategorie odpowiedzi typu: „często”, „czasami”, „rzadko”, które mogą posiadać odmienne znaczenia dla każdego z respondentów. Błąd taki wkraść się do niżej zaprezentowanego pytania: „proszę podać częstotliwość obecności na zajęciach z ocenianego przedmiotu” (kafeteria odpowiedzi: „zawsze”, „w większości”, „sporadycznie”, „nigdy”)⁹.

⁸ Nie identyfikujemy w tekście szkół, z których pochodzą przykładowe pytania.

⁹ Alternatywne, wyżej oceniane podejście zastosowane zostało w narzędziu ewaluacyjnym innej szkoły wyższej, w którym umieszczono pytanie: „jaka była Pani/Pana frekwencja na zajęciach z tego przedmiotu?” (kafeteria odpowiedzi: 1=0–20%; 2=21–40%; 3=41–60%; 4=61–80%; 5=81–100%).

Podobny problem z niejednoznacznością może wystąpić, kiedy prosimy respondenta o **ocenę bardzo ogólnej kategorii**, jaką jest przykładowo organizacja zajęć. W związku z tym, że organizacja ta składa się z wielu elementów, takich jak punktualność lub dostęp do sylabusu, należy zadawać pytania raczej o elementy szczegółowe, a nie o ogólne.

Reguła 2. Unikaj pytań długich, skomplikowanych i napisanych „językiem branżowym”, technicznym czy żargonem.

Jako „język branżowy” należy rozumieć np. sformułowania takie jak „jakość kształcenia”. Jeżeli zatem chcemy o nią zapytać, to albo musimy ją skrótowo zdefiniować, co może okazać się jednak zbyt dużym wyzwaniem, lub potraktować jako pojęcie latentne (nieobserwowane w sposób bezpośredni) i zanalizować w oparciu o baterię pytań kwestionariuszowych (innymi słowy, zbudować skalę do jej pomiaru).

Reguła 3. Nie zadawaj pytań, na które odpowiedź jest oczywista.

Zadając pytanie: „Czy zawsze byłeś/aś skoncentrowany/a w trakcie zajęć?”, nie spodziewajmy się, że odpowiedzi studentów różnicować będą w sposób istotny zajęcia, w których uczestniczyli. Nie ma chyba osoby, która nawet na najbardziej interesujących zajęciach byłaby bez przerwy skoncentrowana.

Reguła 4. Unikaj zadawania pytań, na które nie jest możliwa jednoznaczna odpowiedź.

Częstą praktyką łamiącą powyższą regułę są **pytania o dwie rzeczy na raz**, które pojawiają się w większości kwestionariuszy ewaluujących pracę nauczycieli akademickich i wprowadzają badanego w stan zakłopotania, w momencie gdy dwa elementy zawarte w pytaniu powinny zostać ocenione odmiennie. „Czy prowadzący zajęcia był życzliwy, otwarty i taktowny w stosunku do studentów?” (jaką odpowiedź powinien zaznaczyć respondent, jeżeli prowadzący był życzliwy, ale nie otwarty?); „Czy materiały dydaktyczne do zajęć były dostępne i przydatne?” (a jeśli były dostępne, ale nieprzydatne?); „Czy ustalone terminy zajęć i konsultacji są przez prowadzącego przestrzegane?” (a jeśli przestrzegane są terminy zajęć, a konsultacji już nie?); „Czy prowadzący podał przejrzyste zasady zaliczenia zajęć na początku semestru i zastosował się do nich?” (a jeśli podał, ale się do nich nie zastosował?).

Ponadto nie można wymagać od respondenta dokonywania wyborów niemożliwych. Jeśli w jednym z pytań: „Podaj orientacyjną średnią ocen ze wszystkich przedmiotów, uzyskaną

w ostatnim semestrze” autor zaproponował następujące odpowiedzi: 2,0–3,0; 3,0–4,0; 4,0–5,0, to w związku z **nierozłącznością przedziałów** respondent, którego średnia ocen wynosi 3.0 lub 4.0 nie jest w stanie dokonać jednoznacznego wyboru.

Reguła 5. Skoncentruj się na pytaniach zamkniętych, ale nie unikaj pytań otwartych.

Mimo że główna zasada konstruowania narzędzi do badań kwestionariuszowych mówi, iż należy unikać pytań otwartych, w przypadku ewaluacji jakości kształcenia pytania te mogą dostarczyć wielu cennych informacji, takich jak: rekomendacje odnośnie zmiany formy, struktury czy treści zajęć, opinie na temat słabych i mocnych stron zajęć lub wykaz elementów wymagających poprawy itp.

Reguła 6. Unikaj błędnych założeń, w tym błędnego założenia znawstwa.

Błędne założenia ukryte mogą być zarówno w samym pytaniu kwestionariuszowym, jak i w kafeterii odpowiedzi. Przykładem drugiej możliwości jest przypuszczenie, że prowadzący zajęcia nie mógł być do zajęć całkowicie nieprzygotowany. Założenie takie zawarte zostało w kafeterii odpowiedzi, która pozwala na ocenę stopnia przygotowania prowadzącego do zajęć: „wzorowo”, „bardzo dobrze”, „dobrze” lub „zadowolająco” – nie pozwala jednak na ocenę negatywną. Błędne założenie znawstwa ukryte jest w pytaniu z przykładowego narzędzia do oceny pracy nauczycieli akademickich: „Czy na zajęciach omawiany był najnowszy (z ostatnich lat) dorobek naukowy w tej dziedzinie?” Przypuszczenie, że przeciętny student jest w stanie ocenić stopień zgodności przekazywanego materiału z najnowszym dorobkiem naukowym, jest oczywiście zbyt optymistyczne, w szczególności jeśli weźmiemy pod uwagę stopień specjalizacji poszczególnych dziedzin naukowych. Decydując się na tak sformułowane pytanie, bezwzględnie zawsze powinniśmy umożliwić badanemu wybór kategorii „nie wiem”.

Reguła 7. Zwracaj uwagę na to, żeby kafeteria odpowiedzi była dopasowana (tematycznie i gramatycznie) do zadawanego pytania.

Częstym błędem jest brak gramatycznego dopasowania kafeterii odpowiedzi do pytania kwestionariuszowego. Przykładem takiego „metodologicznego potknięcia” jest pytanie: „Czy prowadzący zajęcia inspirował studentów do samodzielnego myślenia?” połączone z kafeterią odpowiedzi: „wzorowo”, „bardzo dobrze”, „dobrze”, „zadowolająco”.

Reguła 8. Jeśli nie posiadasz merytorycznego uzasadnienia, żeby postępować inaczej, stosuj skale zrównoważone.

Niezrównoważone skale stosuje się najczęściej w momencie wyjątkowo wysokiej skłonności badanych do wybierania odpowiedzi z jednego końca skali (np. w punkcie skrajnie pozytywnym lub negatywnym). Zastosowanie kafeterii: „bardzo dobrze”, „dobrze”, „zadowolająco”, „negatywnie” do pytania kwestionariuszowego: „Jak ocenia Pan/Pani całość zajęć?” związane jest zapewne z założeniem, że większość badanych oceniać będzie pozytywnie zajęcia, w związku z czym skoncentrować należy się na zróżnicowaniu opinii wewnątrz ocen pozytywnych. Równocześnie niezadowolenie z zajęć wyrażone jest w jednej, dodatkowo skrajnej kategorii „negatywnie”. Taki układ proponowanych odpowiedzi może prowadzić do zawyżenia deklarowanych ocen.

Reguła 9. Nie umieszczaj w środku skali szacunkowej kategorii „nie wiem”/ „nie mam zdania”/ „trudno powiedzieć”/ „nie potrafię odpowiedzieć”.

W pierwszej kolejności zastanowić należy się, czy kategoria ta jest w ogóle potrzebna (zazwyczaj generuje kłopotliwe dla analityka braki danych), a jeśli jest konieczna, to umieścić należy ją na końcu skali. Pozostawiając ją w środku skali, narażamy się na zaburzenie struktury i porządku ocen.

Zasady analizy oraz prezentacji materiału empirycznego

Konstruując narzędzie ewaluacji jakości kształcenia w formie kwestionariusza ankiety, musimy zadać sobie pytanie, czego oczekujemy po analizie danych, a tym samym, jaka będzie forma pytań kwestionariuszowych i kafeterii odpowiedzi, czy analizować będziemy pojedyncze wskaźniki, czy raczej będziemy bazować na wskaźnikach syntetycznych w formie skal lub indeksów. Wszystko to prowadzi do przyjęcia już na wstępie co najmniej zarysu planu analizy danych i sposobu prezentacji wyników ewaluacji zajęć dydaktycznych. Sposób analizy i prezentacji rezultatów oceny dydaktyki może wpłynąć na poprawne osza-

cowanie poziomu jakości kształcenia. Z tego też powodu należy zwrócić uwagę na kilka zasad i „pułapek” związanych z analizą oceny zajęć dydaktycznych.

Pierwszym krokiem, który musimy wykonać jeszcze przed rozpoczęciem analizy, jest walidacja danych, czyli sprawdzenie kompletności i poprawności danych ankietowych. Szczególną uwagę należy na to zwrócić wówczas, gdy budujemy indeks sumujący oceny na poszczególnych wskaźnikach w jedną wartość syntetyczną, bowiem w takim przypadku występowanie braków danych w zmiennych wchodzących w skład indeksu może prowadzić do błędnych ocen. Problem ten pojawia się, gdy osoby oceniające unikają odpowiadania na niektóre pytania ankiety ewaluacyjnej, a my wykorzystujemy te pytania w algorytmie obliczeniowym. W takiej sytuacji należy rozważyć, czy w algorytmie obliczeniowym wprowadzić warunek wykluczający braki danych, czy też zastosować jedną z metod imputacji¹⁰. W praktyce istnieje wiele różnych metod uzupełniania braków danych (ang. *missing data imputation*, *missing value imputation*). Najprostszą jest zastąpienie braków średnią ze wszystkich ocen pracowników danej jednostki (wydziału, instytutu) w odniesieniu do danego pytania ankiety ewaluacyjnej. Inne popularne metody imputacji wykorzystują średnią z sąsiednich obserwacji, medianę¹¹ lub interpolację liniową¹². Czasami w celu imputacji braków w ocenach zajęć (pracowników) można zastosować regresję liniową.

Jeśli w ocenie zajęć dydaktycznych wykorzystujemy skale pomiarowe (np. skale Likerta 5- czy 7-stopniowe), a tak w większości przypadków konstruowane są ankiety ewaluacyjne, to schemat analizy danych zakłada porównywanie zajęć lub pracowników względem siebie. W tym celu wykorzystujemy zwykle klasyczne miary położenia (tendencji centralnej), zmienności i asymetrii. Miary tendencji centralnej, tzw. miary przeciętne (średnia arytmetyczna, mediana, modalna), charakteryzują typowy poziom (przeciętne zajęcia, przeciętny pracownik), wokół którego skupiają się wszystkie pozostałe analizowane obserwacje (inne zajęcia, inni pracownicy). Średnia arytmetyczna jest powszechnie stosowana

10 Imputacja to w statystyce sztuczne wstawienie pewnych wartości do tabeli danych. Najczęściej stosuje się ją wtedy, gdy braki danych nie mają losowego charakteru.

11 Mediana zdecydowanie lepiej sprawdza się przy skośności rozkładów danej zmiennej.

12 Interpolacja – metoda numeryczna polegająca na wyznaczeniu w danym przedziale tzw. funkcji interpolacyjnej, która przyjmuje z góry zadane wartości w ustalonych punktach. Stosowana jest wtedy, gdy dysponuje się zazwyczaj skończoną liczbą ocen do określenia zależności między danymi oraz w celu uproszczenia skomplikowanych funkcji arytmetycznych. Interpolacja jest szczególnym przypadkiem metod numerycznych typu aproksymacja. Interpolacja liniowa to przypadek interpolacji za pomocą funkcji liniowej.

w analizie ankiet ewaluacyjnych, bo przez wielu decydentów jest uznawana za intuicyjną miarę, pozwalającą na pozycjonowanie zajęć lub pracowników. Jednakże obliczanie średniej arytmetycznej, o czym nie każdy pamięta, nie jest wskazane, gdy w zbiorze danych występują obserwacje odstające, co bezpośrednio wiąże się z nieproporcjonalnym udziałem obserwacji w tworzeniu sumy. Podobnie, gdy rozkład empiryczny jest silnie asymetryczny, nie należy liczyć średniej arytmetycznej¹³.

Aby skutecznie monitorować i zarządzać jakością kształcenia w szkole wyższej, niezbędnym staje się uwzględnianie w analizie danych rozwiązań z zakresu *benchmarkingu*. W najprostszy sposób można go scharakteryzować jako podejście nakierowane na poprawę własnych działań w oparciu o monitorowanie i porównanie osiągnięć instytucji z osiągnięciami konkurencji bądź innych jednostek, będących elementem szkoły wyższej (por. Woźnicki 2008; Hamalainen 2002). W praktyce wyróżniamy trzy rodzaje *benchmarkingu*: wewnętrzny (szukanie najlepszego wzorca, np. zajęć lub pracownika, i porównywanie się z nim), zewnętrzny – zwany konkurencyjnym (np. porównywanie się z innymi uczelniami) i funkcjonalny (dostosowanie się do wzorca realizującego tę samą lub podobną funkcję, por. Zimmewicz 2009). Przykładowo, na Uniwersytecie Jagiellońskim po zakończeniu oceny zajęć dydaktycznych każdy pracownik może porównać swoje wyniki z wynikami jednostki. W sprawozdaniach analitycznych z oceny zajęć na Uniwersytecie, jakie otrzymuje rektor, uwzględniane są porównania międzywydziałowe, wewnątrzwydziałowe, między pracownikami instytutu lub wydziału (choćby w kontekście typowania najlepszych pracowników na wydziałach do nagród rektorskich), między typami zajęć itp.¹⁴ Wynik takiego benchmarkingu analitycznego może służyć jako podstawa do doskonalenia procesu jakości kształcenia w szkole wyższej na różnych poziomach struktury uczelni¹⁵.

13 Średnia arytmetyczna jest estymatorem zgodnym i nieobciążonym wartości oczekiwanej rozkładu. Jeśli jest to rozkład normalny, to średnia jest również estymatorem efektywnym. Średnia arytmetyczna jest podatna na obserwacje odstające (czyli w tym przypadku wartości zmiennej losowane spoza rozkładu, którego wartość oczekiwaną chcemy estymować, np. pomyłki w danych). W przypadku, gdy jest ich dużo, inne miary tendencji centralnej, takie jak mediana czy średnia obciążona (5% średnia arytmetyczna), mogą dawać lepsze wyniki.

14 Otrzymujemy szereg porównań ocen pracownika w porównaniu ze średnią oceną jednostki, jednostki z innymi jednostkami, ocen jednostki z oceną wydziału, ocen wydziału z oceną globalną uczelni itd.

15 Warto dodać, że systemy *benchmarkingowe* współcześnie coraz częściej skupiają się nie tyle na porównywaniu wartości określonych mierników, co raczej na wymianie informacji o procesie dochodzenia do wysokiej jakości dydaktyki (*benchmarking* procesu).

Aby sensownie porównywać ze sobą zajęcia i pracowników z różnych jednostek (wydziałów lub instytutów) tej samej uczelni (czy też uczelnie między sobą), musimy otrzymywane dane poddać odpowiednim przekształceniom algebraicznym. Wobec danych ilościowych stosowanych w analizach statystycznych formułowane są określone wymogi formalne: warunek symetrii rozkładu, addytywność cech (zwłaszcza gdy są różnoimienne), stałość rozstępu cech (wartości ekstremalnych), jednokierunkowość oddziaływania na zjawisko (Malarska 2005: 39). Spełnienie tych postulatów – jeśli dana cecha ich nie posiada – wymaga zastosowania przekształceń algebraicznych określanych jako normalizacja danych. Normalizacja odnosi się do wstępnej obróbki danych, ujednoczenia cech w celu umożliwienia ich wzajemnego porównywania i dalszej analizy. Powszechnymi rodzajami ujednoczania, tj. normalizacji, które można zastosować do analizy oceny zajęć, są: przekształcenia ilorazowe, standaryzacja, unitaryzacja lub usunięcie ocen odstających. W praktyce wykorzystuje się standaryzację i przekształcenia ilorazowe. Standaryzacja jest rodzajem normalizacji zmiennej losowej, w wyniku której zmienna uzyskuje średnią wartość oczekiwaną zero i wariancję jeden. Najczęściej spotykanym sposobem standaryzacji jest tzw. standaryzacja Z. W opisywanym powyżej przykładzie Uniwersytetu Jagiellońskiego jako formułę normalizacyjną w algorytmie obliczania ocen pracowników uczelni wykorzystuje się przekształcenie ilorazowe oparte na stosunku średniej ważonej z wszystkich zajęć danego pracownika do zakresu ocen (rozstępu skali) dla danego pytania lub baterii pytań. Normalizacja algorytmu analitycznego pozwala uzyskać identyczny dla wszystkich jednostek przedział porównań od 0 do 100 punktów dla zajęć i pracowników, w odniesieniu do semestru, roku akademickiego, jednostki macierzystej, typów zajęć itp. Dzięki temu wynik końcowy jest łatwy w interpretacji dla osób nieobytych ze statystyczną analizą danych.

Na koniec warto dodać, że opisane powyżej metody są tymi, które najczęściej spotyka się w raportach analitycznych z oceny zajęć dydaktycznych, niemniej jednak należy zwrócić uwagę na szereg dodatkowych możliwości w zakresie eksploracji i prezentacji danych z kwestionariuszy ewaluacyjnych. Mamy tu na myśli wspomniane już techniki redukcji przestrzeni danych, takie jak analiza głównych składowych, które umożliwiają identyfikację wymiarowości pojęcia „jakość kształcenia”, a także pozwalają na graficzną prezentację pozycji poszczególnych grup pracowników (np. doktorów, doktorów habilitowanych i profesorów) w wielowymiarowej przestrzeni jakości. W zależności od poziomu pomiaru

zmiennych, którymi dysponujemy, zastosować możemy ponadto szereg technik segmentacyjnych, takich jak choćby analiza skupień.

Rzetelność i trafność studenckich ocen jakości nauczania

Niezależnie od technik badawczych stosowanych do oceny jakości nauczania, podstawowym warunkiem możliwości wykorzystania jej wyników jest pewność, że stanowią one dobre przybliżenie rzeczywistości, czyli faktycznej jakości prowadzenia zajęć, oraz że są spójne i stabilne. W metodologii badań kwestie te określa się mianem trafności i rzetelności, poświęcając im wiele uwagi, bowiem to one decydują o jakości uzyskiwanych wyników, a co za tym idzie, o jakości i słuszności podejmowanych w oparciu o nie działań i decyzji. Ponieważ najczęstszą techniką zdobywania informacji na temat jakości nauczania są badania ankietowe, odniesiemy się tutaj do czynników, które mogą mieć wpływ na wyniki uzyskiwane za pomocą tych technik.

Trafność

Najprościej rzecz ujmując, trafność odnosi się do tego, czy określone narzędzie mierzy to, co powinno mierzyć (Babbie 2003:166; Frankfort-Nachmias i Nachmias 2001:180). Narzędzie służące ocenie jakości nauczania powinno zatem mierzyć faktyczną jakość nauczania, być czułe na różnice w jej poziomie i wiernie je oddawać. Jak ujmuje to Wiliam E. Cashin (1995), trafność w przypadku ankiet studenckich sprowadza się do pytania, do jakiego stopnia oceny studenckie mierzą jakość i skuteczność nauczania oraz różne ich wymiary.

Ocena trafności pomiarów jakości nauczania była przedmiotem wielu badań i eksperymentów metodologicznych, cytowanych w artykułach i opracowaniach poświęconych tej problematyce (por. m.in. Cashin 1995; Erdle, Murray i Rushton 1985; Kulik 2001). W badaniach tych testowano wiele hipotez odnoszących się do różnych wymiarów trafności. Sprawdzano m.in. czy można zaobserwować związek pomiędzy wynikami uzyskanymi

przez studenta na egzaminie, a tym, jak ocenił on dany kurs (Cohen 1981; Feldman 1989b). Przyjmowano tutaj założenie, że jednym z elementów jakości nauczania jest jego skuteczność, tę zaś można mierzyć poprzez wyniki uzyskane przez studentów na egzaminie¹⁶. W konsekwencji można oczekiwać, że studenci, którzy więcej się nauczyli podczas danego kursu (w praktyce – otrzymali lepsze oceny), będą też lepiej oceniać prowadzącego. Cytowane badania potwierdziły istnienie takiego związku, jednak ich wyniki należy interpretować ostrożnie, ponieważ oceny końcowe studenta zależą od wielu innych czynników, w dużej mierze niezależnych od prowadzącego.

Trafność studenckich ocen kursów potwierdzają też analizy zależności pomiędzy wynikami tych ocen a samooceną prowadzącego oraz ocenami wystawionymi prowadzącemu przez inne osoby: przełożonych, innych nauczycieli akademickich, absolwentów, przeszkolonych obserwatorów zewnętrznych (Feldman 1989a; Overall i Marsh 1980; Braskamp i Ory 1994).

Rzetelność

Jak piszą Chava Frankfor-Nachmias i David Nachmias, „rzetelność dotyczy wielkości błędu związanego z danym narzędziem pomiarowym, tj. błędu, który powstaje w sposób losowy, w kolejnych pomiarach dokonywanych za pomocą tego samego narzędzia pomiarowego” (2001:185). W literaturze poświęconej mierzeniu jakości kształcenia przez rzetelność rozumie się zazwyczaj spójność i stabilność ocen (por. Cashin 1995), a w ramach analizy ankiet studenckich rzetelność odnosi się najczęściej do spójności ocen pomiędzy poszczególnymi oceniającymi (np. czy wśród uczestników jednego kursu, prowadzonego przez jednego uczącego istnieje zbieżność ocen dla poszczególnych pozycji ankiety).

Na rzetelność studenckich ocen zajęć wpływa wiele czynników. Kluczowe znaczenie odgrywa sama technika badawcza oraz wykorzystywane narzędzie badawcze, w tym przede wszystkim jasność, jednoznaczność i zrozumiałość pytań, o czym pisaliśmy w poprzedniej części tego opracowania. Niezmiernie istotną kwestią jest też liczba i sposób doboru studentów oceniających dane zajęcia: generalnie im większy odsetek oceniających, tym bardziej rzetelne wyniki. W literaturze poświęconej ocenie jakości nauczania spotkać można

16 W eksperymentach wykorzystywano egzaminy przygotowywane nie przez samego prowadzącego, lecz przez osoby zewnętrzne, aby zminimalizować wpływ różnicowania stopnia trudności egzaminów.

konkretne zalecenia dotyczące tego, jak duży musi być poziom realizacji próby dla danej wielkości grupy studentów, aby oceny mogły zostać uznane za reprezentatywne i rzetelne (por. Thall i Franklin 1991). Zalecenia prezentujemy w poniższej tabeli:

Tabela 1. Zalecany poziom realizacji próby w studenckich ocenach zajęć

WIELKOŚĆ GRUPY	ZALECANY POZIOM REALIZACJI PRÓBY
5–20	80%
21–30	75%
31–50	66%–75%
51–100	60%–75%
Ponad 100	50%–75%

Źródło: opracowanie własne na podstawie M. Theall, J.I. Franklin (1991)

Istotny jest jednak nie tylko sam poziom realizacji próby, czyli odsetek uczestników danych zajęć, którzy wypełnili ankietę, ale i reprezentatywność próby, którą możemy tutaj rozumieć przede wszystkim jako brak występowania autoselekcji podczas wypełniania kwestionariusza. Autoselekcja może bowiem w tym przypadku stanowić źródło systematycznych obciążań uzyskanych ocen. Bardziej chętni do wypełnienia ankiety mogą być studenci silnie zmotywowani do wyrażenia swojej opinii, czyli na przykład bardzo z zajęć zadowoleni bądź też bardzo niezadowoleni.

Problem ten w dużym stopniu dotyczy wspomnianych przez nas wcześniej ankiet internetowych (Computer Aided Web Interview, w skrócie CAWI), stosowanych coraz powszechniej w badaniach opinii studentów na temat zajęć. O wzrastającej popularności tej techniki decyduje przede wszystkim łatwość jej zastosowania (zwłaszcza w sytuacji, gdy uczelnia dysponuje elektronicznym systemem obsługi studentów), niski koszt zbierania danych, szybkość zbierania danych, redukcja błędów przy wprowadzaniu danych i ogólna łatwość zarządzania takimi badaniami. Atuty te doskonale streszcza Jadwiga Przewłocka, pisząc, że technika „opierając się na nowych technologiach i umożliwiając atrakcyjną z punktu widzenia respondenta formę kontaktu, pozwala na szybkie uzyskiwanie dużych prób stosunkowo niskim kosztem” (2009:99). Jak w każdym badaniach internetowych, problemem pozostają jednak wspomniane wyżej: niski poziom realizacji próby oraz autoselekcja badanych.

W badaniach studenckich ocen zajęć realizowanych za pomocą ankiety CAWI mamy zazwyczaj do czynienia z próbami wyczerpującymi, czy też inaczej, z badaniami realizowanymi na całych populacjach: zaproszenie do wzięcia udziału w badaniu kierowane jest do wszystkich uczestników danego kursu. Uczestnicy ci, zachęceni na różne sposoby, m.in. przez bannery pojawiające się na stronie obsługi studentów, plakaty rozwieszane na uczelniach, zachęty prowadzących zajęcia, artykuły i apele w prasie uczelnianej, mogą wypełnić przypisaną do danych zajęć ankietę. Zazwyczaj takich ankiet do wypełniania mają kilka: liczba ankiet odpowiada liczbie kursów, w jakich wzięli udział i formom tych kursów – zazwyczaj oddzielnie wypełnia się ankietę dla wykładów, ćwiczeń, laboratoriów czy praktyk¹⁷. Tak założony schemat realizacji badania sprawia, że wszystkie osoby mają jednakowe szanse na wzięcie udziału w badaniach, zakłada się bowiem, że każdy student uczęszczający na zajęcia ma dostęp do Internetu, mając tym samym możliwość wypełnienia ankiety. Jak jednakże pisze Przewłocka:

odpowiedni dobór próby nie zapewni jednak wysokiej jakości próby, jeśli wylosowane do próby jednostki nie wezmą udziału w badaniu. Jest to tzw. problem braku odpowiedzi (*non-response error*), którego skala zależy z jednej strony od stopnia realizacji próby (*response rate*), z drugiej – od różnic między osobami, które wypełniły ankietę i tymi, które jej nie wypełniły. (2009:104)

Niski poziom realizacji próby, będący jednym z najpoważniejszych problemów badań CAWI może też silnie wpływać na reprezentatywność wyników uzyskanych dzięki ankietom studenckim. Niechęć do wypełniania ankiet w przypadku oceny zajęć może być dodatkowo wzmacniana przez swoisty brak kultury ewaluacji na polskich uczelniach. Ów brak kultury ewaluacji, związany z nowością działań z tego zakresu i często z brakiem przekonania o ich znaczeniu, widoczny jest zarówno wśród studentów, jak i wśród prowadzących. Jak można przypuszczać, istnienie tego zjawiska zdecydowanie nie sprzyja gotowości do uczestnictwa w badaniach oceniających jakość nauczania.

Naturalnie, podejmowanie działań zachęcających do wzięcia udziału w ocenie zajęć, w tym podkreślanie anonimowości tych badań, często niepokojącej studentów i kwestionowanej przez nich, może wpłynąć na poziom realizacji próby, zwiększając tym samym prawomoc-

17 Sytuacja ta może się jednak w praktyce komplikować, gdy jedno zajęcia, np. wykład z danego przedmiotu, prowadzone są przez kilku wykładowców.

ność generalizacji. Problemem pozostaje jednak wciąż możliwość występowania różnic pomiędzy tymi, którzy ankietę wypełnili i tymi, którzy jej nie wypełnili.

Stabilność z kolei jest związana ze spójnością ocen danego kursu, ewaluowanego przez różnych oceniających w różnych momentach czasu. Jak pokazują badania (por. np. Braskamp i Ory 1994; Centra 1993; Overall i Marsh 1980), oceny tego samego prowadzącego są raczej podobne w różnych punktach czasowych. **Możliwość uogólniania** określa związek pomiędzy pojedynczymi ocenami danego prowadzącego a jego rzeczywistymi umiejętnościami jako wykładowcy. Przeprowadzono wiele badań dotyczących tej problematyki (por. np. Marsh 1982), a wyniki wskazują, że to osoba prowadzącego i sposób prowadzenia zajęć w większym stopniu wpływa na oceny studenckie niż sam kurs (jego zakres tematyczny, przedmiot itp.).

Możliwe źródła obciążenia danych

Obciążać wyniki ocen danego prowadzącego, czy oceny danego kursu mogą te wszystkie czynniki, które nie są zależne od prowadzącego czy też nie wynikają ze specyfiki kursu, a wpływają na efektywność uczenia lub zadowolenie studentów z uczestnictwa w zajęciach. Do takich czynników należą np. wielkość grup ćwiczeniowych, warunki, w jakich odbywają się zajęcia, ale też zainteresowanie studentów tematyką zajęć.

Wskazuje się, że podczas analizy wyników ankiet studenckich, a zwłaszcza przy wykorzystaniu ich w polityce kadrowej, warto kontrolować te czynniki, które mogą obciążać pomiar. Prowadzone badania pozwoliły wyróżnić wiele takich czynników. Najważniejsze z nich przedstawia poniższa tablica:

Tablica 2. Czynniki wpływające na studenckie oceny zajęć

- **Motywacja studentów do uczestnictwa w kursie**, zainteresowanie problematyką. Wyżej oceniane są zazwyczaj kursy, które studenci wybierają dobrowolnie lub gdy uczestniczą w kursach związanych z ich zainteresowaniami.
- **Oczekiwana ocena**: wyżej oceniane przez studentów są kursy, z których otrzymują lepsze oceny. Wyniki badań nie są pod tym względem jednoznaczne, podobnie jak wyjaśnienia przyczyn tego zjawiska. W jednym z przytaczanych wyjaśnień wskazuje się, że ci prowadzący, którzy wystawiają studentom lepsze oceny od tych, na jakie zasłużyli, są także przez studentów lepiej oceniani, niż na to zasłużyli.
- **Poziom kursu, miejsce w cyklu studiowania**: różnie oceniane są kursy podstawowe i bardziej zaawansowane.
- **Dyscyplina naukowa**: badania wskazują, że oceny kursów związane są z kierunkiem studiów. Najwyżej oceniane są kursy z dziedziny nauk humanistycznych, następnie z dziedziny nauk społecznych, najniżej z dziedziny nauk ścisłych. Oznacza to, że ocen tych kursów nie można porównywać.
- **Stopień trudności kursu**: wbrew potocznym opiniom, wiele badań wskazuje jednak, że wyżej oceniane są te kursy, które wymagają większego nakładu pracy, są trudniejsze.
- **Brak anonimowości**: wpływa na wyższe oceny kursu/prowadzącego.
- **Obecność prowadzącego zajęcia** podczas oceniania: również wpływa na wyższe oceny kursu.

Źródło: opracowanie własne na podstawie Cashin (1995)

Obawy związane z oceną zajęć dydaktycznych

Ostatnią rzeczą, na którą należy zwrócić uwagę, są obawy nauczycieli akademickich wobec wyników oceny zajęć dydaktycznych, sposobu ich wykorzystania oraz prezentacji i upubliczniania w społeczności akademickiej danej szkoły wyższej. Do najczęstszych obaw związanych z ewaluacją zajęć dydaktycznych można zaliczyć: obawę przed byciem

ocenianym oraz obawy związane z gloryfikowaniem ewaluacji jako źródła pozyskiwania informacji zwrotnej, szczególnie, gdy jest to jedyna forma gromadzenia informacji na temat jakości prowadzonych zajęć w szkole wyższej. Obawy budzi także stosunek studentów do wypełniania ankiet ewaluacyjnych, co rodzi pytanie o rzetelną informację zwrotną i obiektywność ocen studenckich. Oceniani mają poczucie, że studenci traktują ankietę ewaluacyjną jako sposób karania i nagradzania pracowników naukowo-dydaktycznych. Dochodzi do tego przekonanie o tym, że ankietę wypełniają albo studenci zadowoleni, albo niezadowoleni (co może prowadzić do obciążenia wyników ocen). W takiej sytuacji wykorzystywana w pozycjonowaniu zajęć lub pracowników średnia arytmetyczna jest wypadkową ocen skrajnych.

Często artykułowane są także obawy związane z niedoskonałością narzędzi ewaluacyjnych m.in. wielością pytań (szczególnie w przypadku ankiet elektronicznych, gdy student ma do wypełnienia przeciętnie po kilka ankiet w semestrze) i niejednoznacznością pytań (gdy student ma problem z udzieleniem odpowiedzi zgodnej z własnymi przekonaniem i/lub stanem faktycznym). Należy jednak zdać sobie sprawę, iż kwestionariusz ankiety ewaluacyjnej wymaga wielu testów i czasu zanim stanie się rzeczywiście trafnym i rzetelnym narzędziem oceny jakości kształcenia.

Ostatnią kategorią są obawy dotyczące sposobu prezentacji i wykorzystania wyników ewaluacji. Szczególnie w przypadku ankiet elektronicznych podkreślane są dwie kwestie: wiarygodność oceny (jaki powinien być odsetek oceniających/ile ankiet trzeba wypełnić, aby wyniki uznać za miarodajne) oraz późniejsze konsekwencje dla pracownika (problem transmisji oceny z ewaluacji na ocenę przełożonego, szczególnie w sytuacji, gdy nie jest ona adekwatna do ewaluowanej rzeczywistości). Niezależnie od tego występuje obawa, że wynik ewaluacji może stanowić podstawę do rozwiązania stosunku pracy z nauczycielem akademickim lub być przyczyną braku propozycji prowadzenia zajęć dydaktycznych.

W przypadku studentów obawy w zakresie oceny zajęć dydaktycznych są głównie związane z brakiem lub pozorami zachowania anonimowości (co może przekładać się na negatywne konsekwencje wobec studentów), zachowaniem poufności danych, niedoskonałością narzędzia ewaluacyjnego i sposobem wykorzystania wyników ewaluacji. Do tych obaw dochodzi także problem upubliczniania wyników ewaluacji wśród studentów i pracowników danej uczelni. Należy bowiem pamiętać, że istnieje wyraźny związek między powszechnością

wyników a frekwencją oceniających. Raporty z oceny winny być publikowane na stronie internetowej uczelni oraz na stronach internetowych jednostek dydaktycznych, tak aby oceniający mieli wgląd w wyniki. A to dla wielu decydentów stanowi poważny problem.

Jak widać, zarówno studenci jak i pracownicy, wśród swoich obaw odnośnie oceny zajęć dydaktycznych podkreślają dwie kluczowe kwestie: posiadanie dobrego, tzn. trafnego i rzetelnego, narzędzia ewaluacyjnego oraz sposób wykorzystania, prezentacji i upubliczniania wyników. Dlatego wszelka ocena – zwłaszcza jeśli ma ona mieć praktyczne konsekwencje (a ocena jakości kształcenia ma je z pewnością) – powinna spełniać cztery postulaty: **prawdziwości, dokładności, wszechstronności i uczciwości**. Warto więc zadbać o te postulaty, aby nie tylko doskonalić proces zbierania i analizy oceny zajęć dydaktycznych w każdej szkole wyższej, ale mieć realny wpływ na zarządzanie jakością kształcenia.

Zakończenie

Niniejszy tekst należy traktować jako zestawienie refleksji na temat metodologii konstruowania narzędzia i analizy wyników ankiet oceniających pracę nauczycieli akademickich. Refleksje te dotyczyły kwestii związanych z trafnością i rzetelnością pomiaru, standardami konstrukcji pytań kwestionariuszowych czy sposobem prezentacji materiału empirycznego. Świadomi jesteśmy, że tematy poruszone przez nas w artykule nie wyczerpują problematyki, a raczej mają na celu pobudzenie do wspólnej refleksji nad metodologicznym aspektem ewaluacji praktyki dydaktycznej. Tak jak świadomi jesteśmy, na co zwracaliśmy już uwagę, że ewaluacja jakości kształcenia w szkole wyższej nie sprowadza się do eksploracji ankiet ewaluacyjnych, a ankiety te są jedynie drobnym, dodatkowo narażonym na duży subiektywizm oceniających, elementem pomiaru szeroko rozumianej jakości kształcenia. Ponadto ewaluacja jakości kształcenia nie jest ewaluacją samego prowadzącego kurs, ale raczej powinna skupiać się na procesie, w którym uczestniczą nauczyciel akademicki i student. Proces ten dodatkowo odbywa się w określonej przestrzeni (sale wykładowe, laboratoria, wyposażenie), która może mieć wpływ na samą jakość kształcenia. Każdy z tych elementów poddany powinien zostać dogłębnej ocenie.

Współczesne definicje jakości kształcenia raczej rozszerzają, a nie zawężają pojęcie, sytuując je zazwyczaj na tle terminów, takich jak: skuteczność i efektywność kształcenia, osiągnięcia szkolne czy produktywność nauczania. Rodzi to wymóg zastosowania wielu metod, podejść, zespolenia koncepcji i opracowania nowych narzędzi pomiarowych. Implikuje to ponadto głębsze spojrzenie w sam proces funkcjonowania szkoły wyższej, który powinien być analizowany z uwzględnieniem różnorodnych źródeł i typów danych.

Przykładowo, pamiętać należy, że jakość kształcenia to nie tylko zadowolenie ze studiowania, ale też zdobyta przez studentów wiedza/umiejętności i ich użyteczność, ich sukces zawodowy i społeczny, a także wizerunek tej jakości ukształtowany w świadomości kandydatów na studia. Tym samym informacje o jakości kształcenia ukryte są nie tylko w ankietach ewaluacyjnych, ale też: w wynikach badania losów absolwentów, w danych sprawozdawczych szkół wyższych (np. sprawozdania składane do GUS, informacje o osiągnięciach nauczycieli akademickich raportowane w ramach ankiety oceny jednostki naukowej przygotowywanej dla OPI, dane z działów zajmujących się rekrutacją studentów), a nawet w ocenach i wyborach kandydatów na studia.

Sprawne wykorzystanie narzędzi badawczych wymaga określenia wspólnych standardów ich konstrukcji, czy to na poziomie uczelnianym, czy nawet ogólnopolskim. Nie oznacza to oczywiście, że szkoły wyższe zmuszone są do stosowania tego samego „szablonu” podczas ewaluacji jakości kształcenia, a wszelka innowacyjność powinna być w nich ograniczana. Sugeruje się raczej wspólne wypracowanie kilku wskaźników, które pozwolą na porównywanie ocen pomiędzy uczelniami.

Na koniec warto dodać, że wzrost świadomości znaczenia badań prowokuje do dyskusji na temat poprawności wnioskowania w procesie oceny pracy nauczycieli akademickich. Miejmy nadzieję, że wzrost ten doprowadzi do rozwoju metodologii ewaluacji zajęć dydaktycznych i do opracowania wspólnych reguł konstruowania narzędzi i analizy materiału empirycznego.

Bibliografia:

Babbie, Earl (2003), *Badania społeczne w praktyce*, Warszawa: PWN.

Braskamp, Larry A., Ory, John C. (1994), *Assessing Faculty Work: Enhancing Individual and Institutional Performance*, San Francisco: Jossey-Bass.

Brzezińska, Anna, Brzeziński, Jerzy (2004), *Skale szacunkowe w badaniach diagnostycznych*, [w:] Jerzy Brzeziński (red.), *Metodologia badań psychologicznych. Wybór tekstów*, Warszawa: PWN, s. 232–307.

Cashin, William 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.

Cave, Martin, Hanney, Stephen, Henkel, Mary, Kogan, Maurice (2006), *The Use of Performance Indicators in Higher Education. The Challenge of the Quality Movement*, „Higher Education Policy Series”, 34.

Centra, John A. (1993), *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*, San Francisco: Jossey-Bass.

Cohen, Peter A. (1981), *Student Ratings of Instruction and Student Achievement: A Meta-analysis of Multisection Validity Studies*, „Review of Educational Research”, 51, s. 281–309.

Erdle, Steve, Murray, Harry G., and Rushton, Philippe J. (1985), *Personality, Classroom, Behavior, and College Teaching Effectiveness: A Path Analysis*, „Journal of Educational Psychology”, 77, s. 394–407.

Feldman, Kenneth A. (1989), *The Association Between Student Ratings of Specific Instructional Dimensions and Student Achievement: Refining and Extending the Synthesis of Data from Multisection Validity Studies*, „Research in Higher Education”, 30, s. 583–645.

Gaither, Gerald, Nedwek, Brian, Neal, John (1994), *Measuring Up. The Promises and Pitfalls of Performance Indicators in Higher Education*, „ASHE-ERIC Higher Education Report”, 23(5).

Gates, Susan, Augustine, Catherine, Benjamin, Roger, Bikson, Tora, Kaganoff, Tessa, Levy, Dina, Moini, Joy, Zimmer, ROn (2002), *Ensuring Quality and Productivity in Higher Education. An Analysis of Assessment Practices*, „ASHE-ERIC Higher Education Report”, 29(1).

Hamalainen, Kauko (2002), *Benchmarking in the Improvement of Higher Education*, ENQA Workshop Reports 2, Helsinki.

Kulik, J.A. (2001), *Student Ratings: Validity, Utility, and Controversy*, „New Directions for Institutional Research”, 109, s. 9–25.

Marsh, Herbert W. (1982), *The Use of Path Analysis to Estimate Teacher and Course Effects in Student Rating's of Instructional Effectiveness*, „Applied Psychological Measurement”, 6, s. 47–59.

Nachmias-Frankfort, Chava, Nachmias, David (2001), *Metody badawcze w naukach społecznych*, Poznań: Zysk i S-ka.

Oppenheim, Abraham (1992), *Kwestionariusze, wywiady, pomiary postaw*, Poznań: Zysk i S-ka.

Overall, Jesse U., Marsh Herbert W. (1980), *Students' Evaluations of Instruction: A Longitudinal Study of Their Stability*, „Journal of Educational Psychology”, 72, s. 321–325.

Przewłocka, Jadwiga (2009), *CAWI – specyfika, wykorzystanie, perspektywy rozwoju*, [w:] Agnieszka Haber, Maciej Szałaj (red.), *Ewaluacja wobec wyzwań stojących przed sektorem finansów publicznych*, Warszawa: PARP.

Sztabiński, Paweł, Sawiński, Zbigniew, Sztabiński, Franciszek (2005), *Fieldwork jest sztuką*, Warszawa: Wydawnictwo IFiS PAN.

Theall, Michael, Franklin, Jennifer (2001), *Looking for Bias in all the Wrong Places: a Search of Truth or a Witch Hunt in Student Ratings of Instruction?*, „New Directions for Institutional Research”, 109, s. 45–56.

Theall, Michael, Franklin, Jennifer (red.)(1990), *Student Ratings of Instruction: Issues for Improving Practice. New Directions for Teaching and Learning*, 43, San Francisco: Jossey-Bass.

Theall, Michael, Franklin, Jennifer (1991), *Using Student Ratings for Teaching Improvement*, [w:] Michael Theall, Jennifer I. Franklin (red.), *Effective Practices for Improving Teaching. New Directions for Teaching and Learning*, 48, San Francisco: Jossey Bass.

Woźnicki, Jerzy (red.) (2008), *Benchmarking w systemie szkolnictwa wyższego*, Warszawa: FNP.

Yorke, Mantz (1996), *Indicators of Programme Quality: a Project Report Prepared for HEQC*, London: Higher Education Quality Council.

Mikołaj Jasiński

System Ewaluacji Jakości Kształcenia na Uniwersytecie Warszawskim jako przykład rozwoju kultury ewaluacyjnej na dużej uczelni wyższej

Na Uniwersytecie Warszawskim System Ewaluacji Jakości Kształcenia istnieje od 2007 roku. Opiera się on na trzech zasadniczych założeniach: funkcjonowaniu stałych mechanizmów pozwalających na wykorzystywanie wyników badań ewaluacyjnych przez władze uczelni przy podejmowaniu decyzji, jasnym oddzieleniu działań badawczych od bieżącej „polityki” uczelni, anonimowości i dobrowolności badań. Spełnienie tych warunków na UW umożliwiło rozwój systematycznych badań ewaluacyjnych, wykorzystujących zróżnicowane, często nowatorskie rozwiązania metodologiczne i obejmujących różne grupy społeczności akademickiej, a w konsekwencji pozwoliło na rozwój kultury ewaluacyjnej (ewaluacja nie jest utożsamiana z procedurami kontrolnymi).

Integralnym elementem opisywanego w artykule systemu jest Pracownia Ewaluacji Jakości Kształcenia (PEJK). W wyniku jej działań na UW powstał zintegrowany, koordynowany centralnie, system studenckich badań ankietowych oceniających zajęcia dydaktyczne. Wyniki dotyczące poszczególnych zajęć przekazywane są bezpośrednio do prowadzących dane zajęcia i władz jednostek dydaktycznych. Co roku PEJK prowadzi ogólnouniwersyteckie badania ankietowe adresowane, osobno, do studentów i nauczycieli akademickich, dotyczące szeroko rozumianych warunków kształcenia. Pracownia analizuje także liczne zasoby informacyjne uczelni. Na szczególną uwagę zasługuje program „Trajektoria studiowania” – badania ścieżki edukacyjnej studenta: od kandydata do uczestnika rynku pracy.

„Uniwersytet ma nie tylko utrzymywać w narodzie nauki i umiejętności w takim stopniu, na jakim już w świecie uczonym stanęły, ale nadto doskonalić je rozkrzewiać i teorię ich do użytku społeczności zastosowywać”.

(*Misja UW, [w:] Strategia UW, 2008*)

Jakość kształcenia wyższego to pojęcie niezwykle szerokie. Świadczy o niej nie tylko nabywanie przez studentów podczas zajęć umiejętności oraz wiedzy. Wskazywać na nią powinien również przekazywany podczas studiów całokształt postaw i osobistych doświadczeń, cennych w dalszym życiu absolwenta uczelni wyższej.

Potrzeba dbania o jakość kształcenia wyższego nie budzi z pewnością w nikim wątpliwości. Wynika nie tylko z państwowych aktów prawnych (ustawa „Prawo o szkolnictwie wyższym”) czy wewnętrznych regulacji uczelni wyższych (statuty, uchwały i zarządzenia władz uczelnianych). Winna być podyktowana zdrowym rozsądkiem ludzi dobrej woli, pragnących dobrze kształcić studentów.

Zasadniczy problem, zwłaszcza w kwestii tak oczywistej jak konieczność dbałości o jakość kształcenia, tkwi zwykle w szczegółach realizacji:

- Jak pogodzić procedury doskonalenia jakości dydaktyki z potrzebą zapewnienia swobody intelektualnej nauczycieli akademickich?
- W jaki sposób zaangażować społeczność akademicką w ideę doskonalenia jakości kształcenia?
- Jaki wpływ na ocenę jakości zajęć dydaktycznych prowadzonych przez „mistrzów” – nauczycieli akademickich winny mieć opinie ich „uczniów” – studentów?

To tylko niektóre pytania, na które należy odpowiedzieć, rozważając zagadnienie jakości kształcenia na uczelni wyższej. Wydaje się, że Uniwersytet Warszawski z powodzeniem podjął te i wiele innych wyzwań związanych z jakością kształcenia. Dodatkowym czynnikiem,

który trzeba było uwzględnić, budując System Ewaluacji Jakości Kształcenia, była skala przedsięwzięcia. Na Uniwersytecie Warszawskim – największej uczelni wyższej w Polsce – –55 tysięcy studentów kształci ponad 3 tysiące nauczycieli akademickich, obsługiwanych przez ponad 2,5 tysiąca pracowników administracyjnych (por. *Sprawozdanie Rektora...*, 2009).

Celem niniejszego tekstu jest prezentacja wdrożonego na UW Systemu Ewaluacji Jakości Kształcenia oraz odpowiedź na przedstawione powyżej pytania. Wprowadzony na UW system ciągle się rozwija, jest doskonalony – modyfikując go, staramy się na bieżąco odpowiadać na potrzeby dydaktyki na naszej uczelni. Niemniej zasady, które zostały przyjęte jako podstawy tego przedsięwzięcia, pozostają stałe. Mam nadzieję, że ich przedstawienie służyć będzie promocji dobrych praktyk.

Podstawy Systemu Ewaluacji Jakości Kształcenia na UW

Historia studenckich ankiet oceny zajęć dydaktycznych na UW sięga początku lat 90. Za pomocą uchwał Senatu oraz zarządzeń Rektora Uniwersytet Warszawski zobowiązał się jednak do stworzenia całego systemu badań (por. akty prawne UW)¹. System Ewaluacji Jakości Kształcenia na UW przewiduje ich systematyczne prowadzenie w całej społeczności akademickiej – również wśród pracowników i doktorantów, a także podejmowanie badań wśród kandydatów na studia i naszych absolwentów. Ważnym elementem systemu badań jakości kształcenia jest również analiza danych zastanych, w tym dostępnych na UW bogatych baz danych oraz analiza danych urzędowych. Opisywany system opiera się na **trzech zasadniczych założeniach**:

1. jasnego oddzielenia działań badawczych od bieżącej „polityki” uczelni;
2. funkcjonowania stałych mechanizmów pozwalających na wykorzystywanie wyników badań ewaluacyjnych przez władze uczelni przy podejmowaniu decyzji;
3. anonimowości i dobrowolności badań.

¹ Statut UW, par. 115, 2006; Uchwała Senatu UW nr 167, 1998; Uchwała Senatu UW nr 240, 2007; Uchwała Senatu UW nr 241, 2007; Zarządzenie Rektora UW nr 12, 2008; Zarządzenie Rektora UW nr 16, 2010.

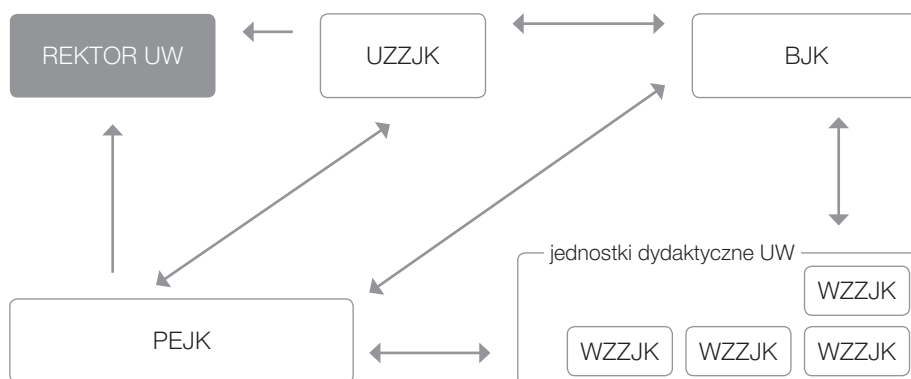
Powodzenie całego przedsięwzięcia zapewniło Uniwersytetowi Warszawskiemu spełnienie jeszcze jednego, zasadniczego warunku – zaangażowania w ewaluację jakości kształcenia **całej społeczności akademickiej UW**.

Omówienie sposobu realizacji na UW wymienionych założeń musi zostać poprzedzone choć krótką charakterystyką struktur projakościowych istniejących na naszej uczelni.

Struktura zapewniania jakości kształcenia na UW

Funkcjonująca dziś na UW struktura zapewniania jakości kształcenia krystalizowała się w latach 2007–2008. 20 czerwca 2007 roku Senat rozpoczął przedsięwzięcie, którego zadaniem jest „dbanie o zapewnienie i doskonalenie jakości kształcenia na UW” (*Uchwała Senatu UW nr 240, par. 2, 2007*). Powstały dwie równoległe struktury: na poziomie ogólnouczelnianym – Uczelniany Zespół Zapewniania Jakości Kształcenia (UZZJK) oraz na poziomie jednostek organizacyjnych – Zespoły Zapewniania Jakości Kształcenia jednostek (tzw. Wydziałowe Zespoły Zapewniania Jakości Kształcenia – WZZJK). Poniższy diagram przedstawia sposób funkcjonowania struktury zapewniania jakości kształcenia na Uniwersytecie Warszawskim.

Rysunek 1. Struktura zapewniania jakości kształcenia na Uniwersytecie Warszawskim



Wymienione w uchwale zespoły spełniają na UW funkcje doradcze i opiniodawcze w kwestiach związanych z jakością kształcenia. **Uczelniany Zespół Zapewniania Jakości Kształcenia** (UZZJK) pełni przede wszystkim funkcję doradczą i rekomendacyjną dla Rektora. W skład UZZJK wchodzi nauczyciele akademicy z jednostek dydaktycznych UW, przedstawiciele studentów i doktorantów. Przewodniczącym UZZJK jest Pełnomocnik Rektora ds. Jakości Kształcenia.

Wydziałowe Zespoły Zapewniania Jakości Kształcenia spełniają w jednostkach w znacznej mierze funkcję analogiczną do funkcji UZZJK na UW – na podstawie informacji o dydaktyce (z dostępnych dokumentów oraz bazując na wynikach badań) przedstawiają władzom jednostek rekomendacje dobrych praktyk oraz proponują zasady monitorowania jakości kształcenia w poszczególnych jednostkach dydaktycznych.

Obsługa organizacyjna Uczelnianego Zespołu Zapewniania Jakości Kształcenia na UW oraz wspomaganie podejmowanych przez Pełnomocnika Rektora ds. Jakości Kształcenia działań o charakterze koncepcyjnym i organizacyjnym, a także koordynacja przedsięwzięć projakościowych na uczelni to, zasadnicze zadania **Biura ds. Jakości Kształcenia** (BJK).

Przedstawiona powyżej konstrukcja uwzględnia złożoność struktury Uniwersytetu Warszawskiego. Obok kwestii ogólnouniwersyteckich, którymi zajmuje się UZZJK, znaczna część problemów związanych z kształceniem pojawia się wyłącznie na poziomie jednostek – są dla nich specyficzne i często niepowtarzalne. Niektóre Wydziałowe Zespoły Zapewniania Jakości Kształcenia wyodrębniają wewnątrz siebie mniejsze Zespoły Instytutowe, aby umożliwić faktyczne odniesienie się do specyfiki kształcenia w jednostkach. Poszczególne zespoły – zarówno UZZJK, jak i zespoły jednostek – są miejscem cennej współpracy przedstawicieli różnych części środowiska akademickiego: studentów, doktorantów, młodszych oraz starszych pracowników naukowych. Mimo wielkiej różnorodności charakteru kształcenia w poszczególnych jednostkach, wiele problemów jest wspólnych. Zespoły jednostek pozostają ze sobą w stałym kontakcie – kilka razy w ciągu roku akademickiego odbywają się spotkania przedstawicieli wszystkich zespołów jednostek, podczas których istnieje możliwość skoordynowania działań, wymiany doświadczeń, dzielenia się dobrymi praktykami. Kontakty między zespołami jednostek ułatwia Biuro ds. Jakości Kształcenia „spinające” tę strukturę w całość.

Integralnym elementem systemu zapewniania i doskonalenia jakości kształcenia na Uniwersytecie Warszawskim jest **Pracownia Ewaluacji Jakości Kształcenia** (PEJK). Tworzą ją socjologowie, naukowcy – nauczyciele akademicy i doktoranci wspierani przez liczne grono studentów – praktykantów, wolontariuszy oraz osób uczestniczących w zajęciach poświęconych tematyce ewaluacyjnej. We współpracy z Pracownią, przy wykorzystaniu danych opracowywanych przez PEJK, przygotowywane są obecnie dwie prace magisterskie i jedna doktorska.

Głównym zadaniem Pracowni jest przygotowanie koncepcji ewaluacji jakości kształcenia na UW i procedur metodologicznych oraz realizacja i opracowanie wyników badań dotyczących różnych aspektów kształcenia na uczelni. Wyniki tych badań mają dostarczyć kluczowych informacji jej władzom na temat odpowiednich kierunków działań i w efekcie przyczynić się do doskonalenia jakości kształcenia na UW oraz promowania tzw. dobrych praktyk dydaktyki uniwersyteckiej, poprawy standardów i warunków kształcenia, modernizacji programów i sposobów nauczania, zwiększania konkurencyjności absolwentów UW na rynku pracy itd.

Jasne oddzielenie działań badawczych od bieżącej „polityki” uczelni

Ewaluacja prowadzona przez Pracownię jest – z formalnego punktu widzenia – ewaluacją wewnętrzną. Ten typ ewaluacji narażony jest więc na uwikłanie w konflikt z różnymi wewnętrznymi „politykami” na różnych poziomach tak dużej instytucji jak Uniwersytet. Typ pytań badawczych stawianych przez nas w związku z ewaluacją sprawia, że Pracownia odgrywa niewdzięczną rolę „zwiastuna złych wieści”. Dlatego istotnym zamysłem przy tworzeniu Systemu Ewaluacji Jakości Kształcenia na UW było zapewnienie swobody badań ewaluacyjnych. Pracownia Ewaluacji Jakości Kształcenia podlega bezpośrednio JM Rektorowi UW i Rektor jest pierwszym, często krytycznym, odbiorcą raportów PEJK. Stanowi z jednej strony gwarancję niezależności badań, z drugiej – co nie mniej ważne – kontakt z odbiorcą oczekującym informacji przydatnych w zarządzaniu uczelnią ułatwia zachowanie utylitarnego charakteru naszych opracowań.

Mechanizmy pozwalające na wykorzystywanie wyników badań ewaluacyjnych przez władze uczelni przy podejmowaniu decyzji

Wyniki badań Pracowni Ewaluacji Jakości Kształcenia są wykorzystywane przez Uczelniany Zespół Zapewniania Jakości Kształcenia i Biuro ds. Jakości Kształcenia przy tworzeniu planu działań na rzecz jakości kształcenia dla władz rektorskich UW oraz dla efektywnej współpracy z jednostkami dydaktycznymi. Rezultaty naszej pracy są także rozpowszechniane wśród całej społeczności akademickiej².

Stworzony na UW System Ewaluacji Jakości Kształcenia zakłada, że poprawa jakości kształcenia musi odbywać się z udziałem wszystkich zainteresowanych, nie tylko rektora i dziekanów. Naszym celem jest wypracowanie sprzężenia zwrotnego pomiędzy potrzebami całej społeczności akademickiej a podejmowanymi działaniami organizacyjnymi. Istotne jest więc dla nas zaangażowanie wszystkich członków tej społeczności. W swoich badaniach staraliśmy się poznać potrzeby i opinie zarówno studentów, jak i doktorantów, pracowników UW, a także kandydatów na studia oraz absolwentów³. Realizacja badań opiera się na stałej interakcji z władzami uczelni, Samorządem Studentów i Doktorantów, UZZJK, Zespołami Zapewniania Jakości Kształcenia w jednostkach i BJK. Bez wytworzenia stałego mechanizmu oddziaływania wyników badań na działania władz uczelni nasze inicjatywy stałyby się „kolejną ankietką”, czy może nawet „kolejną jałową kontrolą”, która szybko umarłaby śmiercią naturalną, jak wszystkie oderwane od rzeczywistości przedsięwzięcia.

2 Znaczna część raportów PEJK dostępna jest na stronie Pracowni. Pozostałe, skierowane do poszczególnych jednostek dydaktycznych UW oraz innych jednostek (władze Biblioteki Uniwersyteckiej w Warszawie, kierownicy biur ogólnouniwersyteckich, Kanclerz UW itd.), przekazywane są bezpośrednio do adresatów i stanowią podstawę do dalszych studiów i konsultacji z udziałem PEJK oraz do podejmowania decyzji projakościowych.

3 Sposoby realizacji badań bardziej szczegółowo przedstawiam w części poświęconej przedsięwzięciom badawczym PEJK.

Anonimowość i dobrowolność badań

Zasadniczym wymogiem skuteczności działań Pracowni – wewnętrznego ewaluatora – jest zaufanie. Bez niego przedsięwzięcia projakościowe na uczelni stałyby się wyłącznie formalno-administracyjnymi procedurami kontroli. Prowadziłyby nieuchronnie do fikcyjnych działań, niemających nic wspólnego z poprawą jakości kształcenia. Konieczne jest więc zapewnienie **anonimowości** badanych, a w przypadku niemożności jej zachowania – np. podczas prowadzenia badań jakościowych – **poufności**. Środowisko akademickie jest szczególnie wyczulone na najmniejsze nawet znamiona przymusu, tym bardziej należy dbać o to, by zapewnić pełną **dobrowolność** uczestnictwa w badaniu. Zbudowanie jasnych zasad wykorzystywania wyników badań ewaluacyjnych przy podejmowaniu decyzji oraz zapewnienie anonimowości i dobrowolności badań pozwoliło na UW pogodzić procedury doskonalenia jakości dydaktyki z potrzebą zapewnienia swobody intelektualnej nauczycieli akademickich, a więc rozwiązać jeden z problemów wymienionych na początku tego tekstu.

Zaangażowanie w ewaluację jakości kształcenia całej społeczności akademickiej

Rozpoczynając przed trzema laty budowę Systemu Ewaluacji Jakości Kształcenia, zdawaliśmy sobie sprawę z podwójnego celu ewaluacji – poza ściśle badawczym, a w konsekwencji utylitarnym jej wymiarem (umożliwienie władzom uczelni oceny procesów kształcenia i w ten sposób ułatwienie podejmowania właściwych decyzji), ewaluacja jakości kształcenia na wyższej uczelni ma w istocie również charakter badawczej interwencji. Jak można wyczytać w dowolnym podręczniku metodologii badań społecznych, każde badanie społeczne oddziałuje na badaną rzeczywistość. W tym wypadku wszelako nie traktujemy tej zależności jako problemu, lecz jako szansę. Nieprzypadkowo hasło II edycji Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia brzmiało: „Zmieniamy Uniwersytet!?”. Zaangażowanie wszystkich członków społeczności akademickiej w proces przekazywania informacji (np. opinii na temat własnej uczelni) warunkuje nie tylko skuteczność samych

działań badawczych (zwiększenie liczby respondentów, bardziej przemyślane odpowiedzi na pytania w ankietach⁴), ale staje się podstawą wzrostu tzw. **kultury ewaluacyjnej**. Głównym problemem, z którym przychodzi się zmierzyć ewaluatorom, jest bowiem przekonanie przedstawicieli ocenianej instytucji, że **ewaluacja nie jest w żadnej mierze formą kontroli**, zaś **pierwszym odbiorcą wniosków ewaluacyjnych** – z zasady zawierających również informacje o słabych stronach przedsięwzięcia – **jest osoba/instytucja poddana ewaluacji**.

Spełnienie przedstawionych warunków na Uniwersytecie Warszawskim umożliwiło rozwój systematycznych badań ewaluacyjnych, wykorzystujących zróżnicowane, często nowatorskie rozwiązania metodologiczne, obejmujących różne grupy społeczności akademickiej. Zapewniło też masowy udział studentów w naszych badaniach – w pierwszych dwóch edycjach wzięło udział prawie 25 tysięcy osób. Rośnie również liczba pracowników naukowo-dydaktycznych oraz administracyjnych zaangażowanych w realizację naszych badań i wykorzystujących ich wyniki. Coraz rzadziej spotyka się również obawy przed ewaluacją jako formą kontroli. Wydaje się, że przedstawione warunki pokazują właściwą odpowiedź na drugie z pytań postawionych na początku tekstu (w jaki sposób zaangażować społeczność akademicką w ideę doskonalenia jakości kształcenia?).

System Ewaluacji Jakości Kształcenia na Uniwersytecie Warszawskim

86

Pracownia Ewaluacji Jakości Kształcenia realizuje swoje zadania, skupiając się na dwóch typach działań badawczych:

- badanie opinii na temat jakości i warunków kształcenia na UW wśród:
 - członków społeczności akademickiej: studentów, nauczycieli akademickich, doktorantów;

4 Obszerniej pisze o tym zagadnieniu Albert Izdebski w tekście zaprezentowanym w tym samym zbiorze: *Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia na Uniwersytecie Warszawskim – przykład dobrych praktyk ewaluacji jakości kształcenia wyższego. Przegląd kluczowych aspektów badania*, s. 123.

- kandydatów na studia i absolwentów;
- analiza warunków kształcenia oraz „Trajektorii studiowania” na UW, uwzględniająca dynamikę zjawisk na podstawie dostępnych na UW zasobów informacyjnych dotyczących kształcenia – danych zastanych:
 - dokumentów, rankingów;
 - zasobów Internetowej Rejestracji Kandydatów (IRK);
 - zasobów Uniwersyteckiego Systemu Obsługi Studiów (USOS).

W działaniach badawczych PEJK wykorzystuje dwa podejścia metodologiczne: ilościowe i jakościowe. Zagadnienie integracji tych dwóch podejść jest przedmiotem ciągłych dyskusji metodologicznych. Naszym zdaniem **integracja ilościowej i jakościowej metodologii badań** jest warunkiem koniecznym pełnego przedstawienia problematyki społecznej (Kowalski, Jasiński 2007). W swoich działaniach badawczych traktujemy oba podejścia jako komplementarne. Szczegółowe omówienie tego zagadnienia wykracza poza ramy niniejszego tekstu. Prezentacja zagadnienia **triangulacji metodologicznej** jest dostępna w licznych podręcznikach metodologii nauk społecznych, odniesienie zaś tej kwestii do problematyki ewaluacji znajduje się m.in. w publikacjach Polskiej Agencji Rozwoju Przedsiębiorczości (por. Kowalski, Jasiński 2007).

W skrócie przedstawię poniżej przegląd przedsięwzięć PEJK. Pozwoli to przybliżyć sposób funkcjonowania Systemu Ewaluacji Jakości Kształcenia na Uniwersytecie Warszawskim.

Studenckie ankiety oceny zajęć dydaktycznych

Ankiety oceny zajęć dydaktycznych realizowane są w wybranych jednostkach Uniwersytetu Warszawskiego od początku lat 90., jednak poziom zaawansowania tych badań jest na tak dużej uczelni bardzo zróżnicowany. Od lat trwa debata na temat ich funkcji, sposobu realizacji i interpretacji. Sprawa była przedmiotem obrad Senatu, debat w komisjach eksperckich, dyskusji w jednostkach dydaktycznych. Wydaje się, że kwestia ta doczekała się uporządkowania i właściwych konkluzji. Tym ważniejsze, że jest to ten obszar, w którym

szczególnie jaskrawo ujawnia się potrzeba pogodzenia specyfiki jednostek z koniecznością wsparcia ich działań przez specjalistów. Równocześnie tu właśnie pojawia się opór nauczycieli akademickich („mistrzów”) przed poddaniem się ocenie studentów („uczniów”). W jaki sposób Uniwersytet Warszawski poradził sobie z tymi wyzwaniami?

Paradoksalnie, właśnie studenckie ankiety oceny zajęć dydaktycznych stały się okazją do wyrażenia jedności społeczności akademickiej, wspólnie, uwzględniając partnerskie relacje „mistrz – uczeń”, dbającej o jakość kształcenia. Kluczem do rozwiązania okazało się właściwe określenie założeń i oczekiwań wobec tzw. studenckich ankiet wydziałowych. Postulaty te zostały sformułowane w 2006 roku podczas prac komisji senackiej ds. studenckich i procesu kształcenia, na czele której stała wówczas obecna Prorektor UW ds. studenckich prof. dr hab. Marta Kicińska-Habior:

1. Należy dążyć do tego, aby na UW istniał jednolity, powszechnie obowiązujący system ewaluacji zajęć dydaktycznych.
2. Ankiety studenckie powinny być tylko jedną z form ewaluacji zajęć.
3. Poszczególne zagadnienia mogą być ewaluowane przez osoby kompetentne w danej dziedzinie, a więc:
 - a. ocena merytoryczna zajęć może być przeprowadzana tylko przez osoby kompetentne w dziedzinie, która była przedmiotem zajęć dydaktycznych, a więc przez pracowników naukowych wskazanych przez kierownika jednostki;
 - b. ankiety studenckie mają służyć przede wszystkim ocenie postawy prowadzącego podczas zajęć, a także jakości ujawnionych w czasie zajęć umiejętności dydaktycznych prowadzącego. **Studenci posiadają kompetencje do oceny postawy prowadzącego.**

Kluczem do rozwiązania problemu oceny „mistrzów” przez „uczniów” okazało się właściwe określenie zakresu ankiet oraz ich roli w procesie ewaluacji zajęć. Jest to odpowiedź na kolejne z trzech postawionych na początku artykułu pytań – o to, jaki wpływ na ocenę jakości zajęć dydaktycznych prowadzonych przez „mistrzów” – nauczycieli akademickich, winny mieć opinie ich „uczniów” – studentów.

Formalną drogę do utworzenia na UW **zintegrowanego systemu studenckich ankiet oceny zajęć dydaktycznych** otworzyła w roku 2008 Rektor UW, obligując wszystkie jednostki organizacyjne do corocznej realizacji badań oceniających zajęcia dydaktyczne

za pomocą ankiety składającej się z dwóch modułów: „ankieta zawiera treści wspólne dla wszystkich jednostek (...) oraz treści właściwe dla danej jednostki organizacyjnej, odzwierciedlające specyfikę jej dydaktyki” (por. *Zarządzenie Rektora UW* nr 12, par. 7 i 9, 2008).

We współpracy z przedstawicielami Zespołów Zapewniania Jakości Kształcenia poszczególnych jednostek zespół Pracowni przygotował w roku 2009 propozycję wspomnianego systemu studenckich ankiet oceny zajęć dydaktycznych. Po szerokich konsultacjach przygotowaliśmy projekt kwestionariusza, w którym staraliśmy się sformułować pytania w możliwie najbardziej uniwersalny sposób. W kwestionariuszu znalazły się, obok pytań zadawanych na wszystkich zajęciach, pytania uwzględniające specyfikę niektórych jednostek. Nie jest to łatwe zadanie, jeśli wziąć pod uwagę dodatkowy, ważny, metodologiczny wymóg – **kwestionariusz oceny zajęć musi być krótki**, w przeciwnym razie znaczna część studentów zrezygnuje z wypełnienia ankiety, a jeśli już odpowie na pytania, potraktuje je lekceważąco. Dobrowolny charakter badania jest tu natomiast kluczowy, właśnie ze względu na konieczność dbałości o rzetelność odpowiedzi. Poza tym udział w badaniu tylko małej części uczestników zajęć stawia pod znakiem wartości wyników.

Nasze konsultacje rozpoczęliśmy od przeglądu wszystkich kwestionariuszy wykorzystywanych dotychczas w jednostkach UW. Do prac nad nowym kwestionariuszem zaangażowaliśmy, poza zespołem PEJK, studentów uczestniczących w seminarium badawczym zatytułowanym „Ewaluacja jakości kształcenia na UW”, prowadzonym przez PEJK w Instytucie Socjologii UW. Przygotowanie nowego kwestionariusza pozwoliło przede wszystkim wyeliminować rażące błędy metodologiczne obecne we wcześniej przeprowadzanych ankietach jednostek dydaktycznych. Wśród różnych wniosków metodologicznych, które wypracowane zostały w trakcie studium kwestionariusza, na szczególną uwagę zasługują dwa:

1. **Kwestionariusz oceny zajęć musi zawierać wyłącznie proste, łatwo interpretowalne pytania**, nawet kosztem ograniczenia ich wrażliwości na specyfikę zajęć, czy też studiów prowadzonych w jednostce. Badanie to nie ma bowiem charakteru eksploracyjnego studium socjologicznego, lecz ma za zadanie dostarczyć nauczycielowi akademickiemu jasnej informacji na temat percepcji jego zajęć i jego samego jako wykładowcy przez studentów. Na ogół też nie ma możliwości prowadzenia pogłębionych, wielowątkowych analiz zajęć. **Badanie ma w istocie odpowiedzieć na pytanie, czy na zajęciach pojawiły się jakieś rażące problemy, czy też nie?**

Z tego powodu m.in. zrezygnowaliśmy z zadawania pytań otwartych. Tę możliwość pozostawiliśmy nauczycielom akademickim, którym zaproponowaliśmy, aby rozdali studentom kartki i zachęcili ich do wypisania, przykładowo, słabych i mocnych stron zajęć. W innym przypadku trudno wyobrazić sobie systematyczną i rzetelną analizę odpowiedzi na pytania otwarte z tak wielkiej liczby ankiet (oszacowaliśmy, że docelowo rocznie na całym Uniwersytecie Warszawskim będzie przeprowadzanych około 300 tysięcy ankiet oceniających zajęcia).

2. **Studenckie ankiety oceny zajęć dydaktycznych powinny być realizowane metodą ankiety audytoryjnej** podczas jednego z ostatnich zajęć dydaktycznych. Świadomie **rezygnowaliśmy więc z realizacji badania drogą ankiety internetowej** przez system USOS-web. Oto zasadnicze przyczyny takiej decyzji:

- a. Wypełniając ankietę podczas zajęć, w grupie ćwiczeniowej, student odpowiada na pytania, pozostając w realiach tych właśnie zajęć – **na jego odpowiedź mają mniejszy wpływ czynniki przypadkowe.**
- b. Badanie jest realizowane na ogół przed wystawieniem ocen przez wykładowców. Mamy większą szansę na to, że **oceny studenckie odnoszą się do przebiegu zajęć, a nie opinii na temat uzyskanych zaliczeń** (sprawiedliwość oceniania przez wykładowcę, to również interesująca kwestia, jednak realizacja ankiet po wystawieniu ocen spowodowałaby, że opinie na temat zaliczeń zdominowałyby inne).
- c. Zasadniczą **wadą ankiety internetowej jest znaczny spadek udziału studentów w badaniu.** W zasadzie oznacza to rezygnację z jakichkolwiek dróg wpływania na wzrost udziału studentów w badaniu⁵. W przypadku wypełniania ankiety podczas zajęć, studenta zachęca do tego fakt, że w badaniu uczestniczy cała grupa.
- d. **Społeczny, kulturotwórczy wymiar ankiety** realizowanej podczas zajęć polega na tym, że studenci i wykładowcy mają szansę nabrać słusznego przekonania o udziale we wspólnym przedsięwzięciu „poprawiania” dydaktyki na UW. Nauczyciel, zgadzając się na realizację ankiety podczas zajęć (poświęcając około 10–15 minut), staje „po stronie studentów”, którzy mogą bez skrępowania, anonimowo wyrazić swoje opinie na temat zajęć.

5 Stosowanie przez niektóre uczelnie zachęt w postaci losowania nagród wśród osób, które wypełniły ankietę, jest przykładem całkowitego niezrozumienia zasad i sensu prowadzenia badań ewaluacyjnych. Narusza nie tylko poczucie anonimowości badanych, ale przede wszystkim skłania do masowego wypełniania ankiet bez dbałości o jakość wypowiedzi. Pomysł administracyjnego obligowania studentów do udziału w badaniu pozostawiam bez komentarza.

Ankieta audytoryjna oznacza wyższe koszty niż ankieta internetowa (koszty druku formularzy). Aby ograniczyć kosztowne i stwarzające duże ryzyko pomyłek ręczne kodowanie odpowiedzi, zdecydowaliśmy się na wprowadzenie **ankiet skanowanych**. Formularze drukowane są centralnie, co znacznie ogranicza koszty druku. Szczegółowe instrukcje, precyzujące działania konieczne do realizacji badania, administracja jednostek, nauczyciele akademicy i studenci otrzymują wraz z formularzami. Następnie jednostki dostarczają wypełnione formularze do PEJK. Pozostaje zeskanowanie formularzy i przygotowanie raportów. To już zadanie administracji centralnej UW (skanowanie) i PEJK (raporty). W tym celu został przygotowany specjalny program, który po wczytaniu obrazów ankiet, automatycznie przygotowuje raporty i rozsyła je do wszystkich zainteresowanych – poszczególnych nauczycieli akademickich oraz ich przełożonych. Docelowo każdy prowadzący będzie mógł odnaleźć wyniki oceny własnych zajęć po zalogowaniu się do systemu USOS-web. To nowatorskie rozwiązanie przyjęte na UW bardzo usprawniło realizację studenckich ankiet oceny zajęć dydaktycznych, a w niektórych jednostkach dydaktycznych w istocie umożliwiło systematyczną realizację tych badań.

Kierownicy jednostek mają możliwość wykorzystania otrzymanych wyników do poprawy jakości kształcenia w swoich jednostkach. W każdej z nich poszczególne rady naukowe uchwałyły „Zasady i procedury zapewnienia i doskonalenia jakości procesu dydaktycznego”, które zawierają m.in. szczegóły wykorzystania wyników badań ewaluacyjnych.

W czerwcu 2009 roku, w pierwszym etapie wdrażania systemu ankiet oceny zajęć dydaktycznych, traktowanym przez PEJK jako etap pilotażowy, wzięło udział 28 jednostek. Wypełniono i przeanalizowano przeszło 40 tysięcy ankiet. Włączenie się do systemu nie jest obowiązkowe⁶, jednak jak wynika ze zgłoszeń, obecnie w przedsięwzięcie zaangażowała się większość jednostek dydaktycznych. Spodziewamy się, że wkrótce system obejmie całość dydaktyki na UW.

Poza raportami przekazanymi poszczególnym odbiorcom, PEJK uzyskuje co semestr w wyniku tego badania informacje pozwalające na konstruowanie bezcennego materiału empirycznego, dzięki któremu będzie można dokładniej monitorować opinie na temat ja-

6 Zarządzenie Rektora stanowi, że „w jednostkach organizacyjnych o formie, treści, sposobie realizacji badań ankietowych oraz sposobie opracowania danych zebranych w wyniku badań decyduje kierownik jednostki organizacyjnej” (por. Zarządzenie Rektora UW nr 16, par. 9, 2010).

kości kształcenia w jednostkach UW. Umożliwia on prowadzenie analiz dynamiki opinii na temat zajęć dydaktycznych. Adresatami tych wyników są władze UW oraz władze jednostek.

Omówione powyżej tzw. ankiety wydziałowe to w istocie ważny, ale niewielki wycinek Systemu Ewaluacji Jakości Kształcenia na UW. PEJK przygotowuje i przeprowadza systematycznie ogólnouniwersyteckie badania ankietowe, adresowane do różnych grup społeczności akademickiej. Szczegóły organizacji tych badań reguluje osobne *Zarządzenie Rektora UW*⁷. Analizujemy również liczne zasoby informacyjne uczelni.

Badania kandydatów na studia oraz absolwentów

Obie te grupy nie zaliczają się w ścisłym tego słowa znaczeniu do społeczności akademickiej – spośród pierwszych tylko nieliczni do niej trafiają, ci drudzy zaś mają ze swoją macierzystą uczelnią coraz słabszy związek. Badania obu tych grup stanowią jednakże niezwykle ważny element Systemu Ewaluacji Jakości Kształcenia na UW.

Do tej pory przeprowadziliśmy we współpracy z Biurem ds. Rekrutacji dwie edycje **badania wśród kandydatów na studia na UW**. Badanie było realizowane za pomocą programu ankietującego zintegrowanego z systemem Internetowej Rejestracji Kandydatów (IRK). W badaniu pytaliśmy kandydatów o ich opinie na temat funkcjonowania Biura ds. Rekrutacji oraz o uwarunkowania dokonanego przez nich wyboru studiów. Można już mówić o systematyczności tych badań. Ich efektem jest coroczny raport rekrutacyjny, prezentowany na listopadowym posiedzeniu Senatu UW. Integralnym elementem raportu jest analiza wyników ogólnouniwersyteckiej ankiety dla kandydatów.

Na temat kandydatów i procesu rekrutacji można wiele powiedzieć także na podstawie istniejących baz danych. Dzięki połączeniu informacji charakteryzujących kandydatów na studia (baza IRK) z ich późniejszymi wynikami jako studentów (baza USOS) można wypowiedzieć się nie tylko na temat jakości kandydatów, ale również znaczenia wyników maturalnych dla ich kształcenia na uczelni. Analiza tych danych (prowadzimy ją od trzech

7 Zob. *Zarządzenie Rektora UW* nr 12, 2008; *Zarządzenie Rektora UW* nr 16, 2010.

lat) stanowi drugą część wspomnianych raportów rekrutacyjnych. Służy ona poprawie procedur rekrutacyjnych, stanowi podstawę namysłu nad programami studiów w jednostkach UW. Bardziej szczegółowo tę kwestię podjął w swoim tekście zamieszczonym w niniejszym zbiorze jeden z naszych analityków, Tomasz Zajac⁸.

Wykorzystanie wspomnianych baz danych jest częścią szerszego programu badawczego, którego realizację Pracownia rozpoczęła w tym roku akademickim. Program ten nazwaliśmy „Trajektorią studiowania”, gdyż jego zasadniczym celem jest prześledzenie ścieżek edukacyjnych naszych studentów – od kandydata do uczestnika rynku pracy. Program zakłada połączenie w przyszłości informacji rekrutacyjnych oraz danych ze studiów z informacjami z rynku pracy. Analiza możliwości predykcyjnych różnych cech studenta – wyników maturalnych oraz wyników uzyskiwanych podczas studiów – dla jego sukcesu na rynku pracy wzbogaci możliwości systematycznego i wielostronnego monitorowania efektów kształcenia na różnych kierunkach studiów na UW.

We współpracy z Biurem Karier UW przygotowujemy ponadto program badania opinii absolwentów na temat programu studiów i możliwości na rynku pracy po zakończeniu studiów. Program zakłada przeprowadzanie panelu z dwukrotnym pomiarem: po raz pierwszy w momencie kończenia studiów i drugi – około pół roku później.

Badania wśród studentów, doktorantów i nauczycieli akademickich

93

Zgodnie z uchwałami Senatu UW PEJK co roku prowadzi **ogólnouniwersyteckie badania ankietowe**, adresowane, osobno, do studentów, doktorantów i nauczycieli akademickich. Badania te dotyczą szeroko rozumianych warunków kształcenia. Realizowane są przy pomocy nowoczesnego, informatycznego systemu ankietującego, zintegrowanego z USOS-web. Do tej pory zrealizowaliśmy dwie edycje **studenckich ankiet ogólnouniwersyteckich**.

⁸ Zob. Zajac, Wykorzystanie danych z Uniwersyteckiego Systemu Obsługi Studiów oraz systemu Internetowej Rejestracji Kandydatów w procesie ewaluacji jakości kształcenia, s. 143.

Na podstawie badań przygotowaliśmy raporty badawcze (Jasiński i inni 2008; Jasiński i in. 2009), na podstawie których UZZJK przygotował rekomendacje działań władz UW. 31 maja 2010 roku rozpoczęliśmy trzecią edycję tego badania. Szczegółowo problematykę studenckich badań ogólnouniwersyteckich omówił, koordynujący w PEJK badania ilościowe, Albert Izdebski w tekście zamieszczonym w niniejszym zbiorze.

W tym roku rozpoczęliśmy realizację szeroko zakrojonych badań wśród nauczycieli akademickich. Obejmują one następujące obszary tematyczne:

- wartość Uniwersytetu Warszawskiego dla nauczycieli akademickich;
- funkcjonowanie społeczności uniwersyteckiej i stosunków instytucjonalnych;
- postrzeganie Uniwersytetu jako instytucji otwartej na dialog wewnętrzny i współpracę międzynarodową;
- zaangażowanie nauczycieli akademickich na rzecz Uniwersytetu, wydziału czy instytutu;
- warunki pracy nauczycieli akademickich, organizacja pracy przez jednostkę dydaktyczną;
- świadome współtworzenie kultury jakości przez nauczycieli akademickich;
- program studiów.

Obecnie realizujemy pierwszy etap – badania jakościowe. Zespół PEJK prowadzi wywiady pogłębione (IDI) z pracownikami naukowymi naszej uczelni. Ich wyniki pozwolą lepiej poznać kondycję naszej kadry naukowo-dydaktycznej oraz jej sposób widzenia problemów Uniwersytetu Warszawskiego. Pozwolą również lepiej przygotować się do etapu ilościowego, który rozpoczniemy w listopadzie 2010 roku. Zebrane dotychczas sygnały wskazują na niezwykle dużą potrzebę wyrażenia przez nauczycieli akademickich ich opinii i doświadczeń oraz zapotrzebowanie na przełożenie tych opinii na konkretne działania na rzecz Uniwersytetu.

W kolejnym roku planowane jest **badanie adresowane do doktorantów**. Będzie to najtrudniejsze wyzwanie. Grupa ta w niektórych sytuacjach ma problemy podobne do problemów studentów, w innych zbliża się do problemów nauczycieli akademickich. Przygotowanie badania ułatwią doświadczenia zdobyte podczas wywiadów i ankiet przeprowadzanych wśród studentów i nauczycieli akademickich. Będzie ono realizowane, podobnie jak w przypadku studentów, przy pomocy systemu ankietującego „Ankieter”, zintegrowanego z systemem USOS-web.

Zakończenie

Oto w skrócie prezentacja Systemu Ewaluacji Jakości Kształcenia na Uniwersytecie Warszawskim – systemu, który stale się rozwija. W roku 2007, kiedy przygotowywane były założenia i plany działań ewaluacyjnych na UW, termin wdrożenia systemu został przewidziany na 2010 rok. Rzeczywiście, udało się wdrożyć wszystkie planowane działania – na Uniwersytecie Warszawskim są prowadzone systematyczne badania odwołujące się do różnych dostępnych źródeł danych (badania opinii, analiza danych z oficjalnych rejestrów – IRK i USOS, analiza danych ze statystyk publicznych), pozwalające na wykrywanie i diagnozowanie problemów związanych z kształceniem. Działają mechanizmy decyzyjne pozwalające na wprowadzanie w życie rekomendacji formułowanych pod wpływem badań Pracowni. Wreszcie, co nie mniej ważne, zmienia się sposób postrzegania ewaluacji – pojęcia tak niedawno jeszcze słabo znanego, a dziś traktowanego coraz powszechniej jako szansa na poprawę rzeczywistości, a nie jako forma procedur kontrolnych.

Przed kilkoma laty wydawało nam się, że widzimy horyzont, za którym będziemy mogli powiedzieć, że rozwijanie systemu jest zakończone. Kolejne doświadczenia przekonują nas, że wraz z rozwojem kultury ewaluacyjnej na Uniwersytecie Warszawskim pojawiają się nowe wyzwania i nowe oczekiwania. Zmieniająca się rzeczywistość kształcenia wyższego wymaga ciągłego doskonalenia narzędzi ewaluacji. Wymaga również wymiany doświadczeń między ewaluatorami – propagowania dobrych praktyk.

Bibliografia:

Jasiński, Mikołaj, Izdebski, Albert, Kowalski, Michał, Zając, Tomasz, Pękala, Justyna (2008), *Raport z pierwszej edycji ogólnouniwersyteckich badań jakości kształcenia na UW*, Warszawa: Uniwersytet Warszawski.

Jasiński, Mikołaj, Izdebski, Albert, Zając, Tomasz, Kowalski, Michał wraz ze współpracownikami (2009), *Raport z drugiej edycji ogólnouniwersyteckich badań jakości kształcenia na UW*, Warszawa: Uniwersytet Warszawski.

Jasiński, Mikołaj, Kowalski, Michał (2007), *Falszywa sprzeczność: metodologia jakościowa czy ilościowa?*, [w:] Agnieszka Haber (red.), *Ewaluacja ex post. Teoria i praktyka badawcza*, Warszawa: PARP, s. 97–114.

Akty prawne UW dostępne na stronie: <http://www.monitor.uw.edu.pl>.

Prawo o szkolnictwie wyższym, Ustawa z dnia 27 lipca 2005 roku: [http://orka.sejm.gov.pl/opinie.nsf/nazwa/2720_u/\\$file/2720_u.pdf](http://orka.sejm.gov.pl/opinie.nsf/nazwa/2720_u/$file/2720_u.pdf)

Sprawozdanie Rektora Uniwersytetu Warszawskiego za rok 2008 (2009), Warszawa: Uniwersytet Warszawski.

Statut Uniwersytetu Warszawskiego, załącznik do Uchwały nr 115 Senatu Uniwersytetu Warszawskiego z dnia 21 czerwca 2006 roku w sprawie nadania Statutu Uniwersytetowi Warszawskiemu (Monitor UW z 2006 r. nr 7, poz. 94), <http://www.monitor.uw.edu.pl/get.php?id=278>.

Uchwała Senatu Uniwersytetu Warszawskiego nr 167 z dnia 18 lutego 1998 r. w sprawie zapewnienia jakości zajęć dydaktycznych.

Uchwała Senatu Uniwersytetu Warszawskiego nr 240 z dnia 20 czerwca 2007 r. w sprawie systemu zapewnienia i doskonalenia jakości kształcenia na Uniwersytecie Warszawskim (Monitor UW z 2007 r. nr 6, poz. 280).

Uchwała Senatu Uniwersytetu Warszawskiego nr 241 z dnia 20 czerwca 2007 r. w sprawie badań ankietowych na Uniwersytecie Warszawskim (Monitor UW z 2007 r. nr 6, poz. 281).

Strategia Uniwersytetu Warszawskiego, załącznik do Uchwały Senatu Uniwersytetu Warszawskiego, nr 34 z dnia 17 grudnia 2008 r. w sprawie Strategii Uniwersytetu Warszawskiego (Monitor UW z 2008 r nr 10, poz. 246), <http://www.monitor.uw.edu.pl/get.php?id=194>.

Zarządzenie Rektora Uniwersytetu Warszawskiego nr 12 z dnia 7 maja 2008 r. w sprawie szczegółowych zasad i trybu przeprowadzania ankiet (Monitor UW z 2008 r. nr 5, poz. 118).

Zarządzenie Rektora Uniwersytetu Warszawskiego nr 16 z dnia 2 marca 2010 r. w sprawie zasad i trybu przeprowadzania ogólnouniwersyteckich badań ankietowych oraz badań ankietowych na wydziałach i w innych jednostkach organizacyjnych Uniwersytetu Warszawskiego kształcących studentów (Monitor UW z 2010 r. nr 3, poz. 63)⁹.

⁹ Aktualna, znowelizowana postać *Zarządzenia Rektora UW nr 12* z 7 maja 2008 r.

Wojciech Przybylski

Seweryn Rudnicki

Anna Szwed

System ewaluacji jakości zajęć dydaktycznych w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera

Artykuł opisuje proces konstruowania i ostateczny kształt systemu ewaluacji zajęć dydaktycznych w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera. WSE jest niewielką prywatną uczelnią, kładącą wyjątkowo duży nacisk na jakość nauczania, dlatego – mimo że do ewaluacji zajęć przywiązywano bardzo dużą wagę od początku istnienia uczelni – potrzebne było opracowanie nowoczesnego i kompleksowego systemu ewaluacji, dostosowanego do specyficznych potrzeb placówki. Tekst opisuje poszczególne etapy opracowywania systemu: analizę potrzeb i doświadczeń WSE, zapoznanie się z bogatą literaturą przedmiotu i rozwiązaniami stosowanymi przez uczelnie anglosaskie; a następnie prezentuje przygotowany i wdrożony system ewaluacji oraz stosowane narzędzia. Przyjęta w tekście perspektywa – pokazywanie praktycznych problemów i sposobów ich rozwiązywania – może być ciekawym źródłem wiedzy i inspiracji dla innych zespołów przygotowujących rozwiązania w zakresie ewaluacji dydaktyki. W tekście omawiane są także, kluczowe w kontekście ewaluacji zagadnienia, takie jak: triangulacja źródeł danych, formatywny i sumatywny charakter ewaluacji oraz budowanie kultury ewaluacyjnej w środowisku akademickim.

Wstęp: po co nam w ogóle ewaluacja jakości dydaktyki?

„Ewaluacja” jest jednym z najpopularniejszych słów w języku, który opanował niemal wszystkie sfery życia publicznego. W węższym kontekście nauki i edukacji, ewaluację można postrzegać jako swoistą modę, napędzaną mechanizmem wzrastającej konkurencji między uczelniami, niezależnie od ich statusu. Mimo że wszystkie polskie uczelnie są prawnie zobowiązane do ewaluacji jakości kształcenia, rzeczywiste (głębokie) kompetencje w tym zakresie dopiero się tworzą, i to w wybranych szkołach wyższych. Optymalne byłoby jednak podejście, w którym działania monitorujące jakość dydaktyki, rozumianej jako istotny element zapewniania jakości nauczania, wypływają z autentycznego przekonania co do ich słuszności. Oczywiście motywacja konkurencyjna (wolnorynkowa) nie jest sama w sobie złą, może jednak nie wystarczyć do wdrożenia systemu, który będzie rzetelny i użyteczny.

Sytuacja Wyższej Szkoły Europejskiej jest o tyle nietypowa, że uczelnia od początku postawiła na wysoką jakość kształcenia, zaś ewaluacja była wdrażana od pierwszego roku jej istnienia. Nie oznaczało to jednak automatycznie spełniania wyśrubowanych standardów w tym zakresie, znanych z najlepszych zachodnich szkół wyższych. Dlatego ujęcie ewaluacji w ramach jednego z zadań projektu rozwojowego uczelni, realizowanego w ramach priorytetu 4.1 POKL z Europejskiego Funduszu Społecznego, należy traktować jako świadomy, strategiczny wybór wzmocnienia takiego obszaru działalności, który decyduje o przewadze konkurencyjnej na rynku edukacyjnym.

Realność dwóch podstawowych funkcji ewaluacji: formatywnej i sumatywnej (będą one odmieniane w niniejszym artykule przez wszystkie przypadki) wymaga dwóch dojrzałości. Pierwsza to dojrzałość prowadzących do bycia ocenianymi, a także otwartość na zmiany własnego warsztatu dydaktycznego. Tylko przy spełnieniu tego warunku ewaluacja może być narzędziem poprawy jakości dydaktyki. Druga dojrzałość jest jeszcze bardziej problematyczna i dotyczy systemowej gotowości uczelni do wyciągania wniosków o cha-

rakterze personalnym, zarówno negatywnych, jak i pozytywnych. Sprowadzając dylemat do brutalnego banału, pytamy tu o zdolność do zwolnienia lub odsunięcia od dydaktyki wykładowcy, który źle uczy i nie rokuje w tym względzie poprawy. Nie chodzi jednak o spłylenie sensu ewaluacji – opisana sytuacja jest skrajną z jej możliwych konsekwencji – ale określenie warunków brzegowych całego procesu. Uczelnia o zakonserwowanej strukturze zatrudnienia może mieć podstawowy problem z przekonaniem studentów do sensu dzielenia się opinią o jakości zajęć dydaktycznych. O ile pierwsza dojrzałość występuje zapewne w jakimś stopniu w każdej placówce edukacyjnej, to druga dzieli polskie uczelnie na dwa wyraźne zbiory.

Ewaluacja jest zatem procesem osadzonym w konkretnej, specyficznej rzeczywistości kulturowej. Kultura ewaluacyjna wywodzi się z krajów anglosaskich o konkurencyjnym modelu wyższej edukacji – przeszczepienie tej tradycji na grunt etatystycznej kultury akademickiej musi stawiać jej funkcję sumatywną pod znakiem zapytania. W tradycyjnie hierarchicznej polskiej rzeczywistości akademickiej problemem może być także bardzo podmiotowe traktowanie studentów. Ustawowo silna pozycja Samorządu Studenckiego w strukturze uczelni wyższej może tu być istotnym sprzymierzeńcem, wiele jednak zależy od samych studentów. W przeciwnym razie może się skończyć na kolejnym złudzeniu, tym razem złudzeniu podmiotowości. Uwaga ta dotyczy zarówno uczelni publicznych, jak i niepublicznych.

Niniejszy artykuł ma w krótki i przystępny sposób przedstawić proces reformy systemu ewaluacji w Wyższej Szkole Europejskiej. Pierwszy podrozdział opisuje fazę koncepcyjną: analizę literatury, a także doświadczeń własnych i wybranych uczelni zagranicznych. Drugi podrozdział przedstawia praktyczne aspekty nowego i kompleksowego systemu ewaluacji jakości dydaktyki, wdrażanego w WSE. Proces ten nie przebiegał bezproblemowo, choć wszystkie istotne trudności – jak się wydaje – udało się nam pokonać. Podstawowe wnioski i rekomendacje zostaną w formie skróconej zebrane w zakończeniu, co w zamierzeniu powinno pomóc w praktycznym zastosowaniu zebranych przez nas doświadczeń przez inne uczelnie.

Od potrzeby do analizy: w poszukiwaniu wyższego standardu ewaluacji

W trakcie projektowania i realizowania działań zmierzających do wypracowania systemu ewaluacji dydaktyki w WSE towarzyszyło nam przekonanie, że wymagają one przyjęcia dwojakiej optyki. Z jednej strony – spojrzenia do wewnątrz organizacji, jaką jest WSE, analizy dotychczasowych mechanizmów oceny dydaktyki, dostrzeżenia zarówno słabych, jak i mocnych punktów funkcjonujących rozwiązań. Z drugiej strony – spojrzenia na zewnątrz, poszukiwania i oceny rozwiązań stosowanych przez bliższe i dalsze otoczenie edukacyjne uczelni. **Tę swoistą filozofię projektu można ująć skrótowo jako chęć uczenia się zarówno na własnym, jak i na cudzym doświadczeniu.**

Proces wypracowywania systemu ewaluacji dydaktyki składał się z kilku bloków działań. Warto w tym miejscu zaznaczyć, że choć poniżej są one opisywane rozłącznie, w rzeczywistości często nakładały się na siebie w czasie i funkcjonalnie zazębiały. Dość często powracaliśmy także do wcześniejszych etapów, krytycznie konfrontując dane, które uzyskaliśmy w późniejszych fazach projektu.

Przegląd literatury

Uczenie się na doświadczeniu innych, o którym wspomniano powyżej, rozpoczęliśmy od przeglądu literatury¹ poświęconej ewaluacji dydaktyki w szkolnictwie wyższym. Niestety **w Polsce wciąż niewiele jest publikacji dotyczących tej tematyki**, stąd koncentrowaliśmy się głównie na artykułach anglojęzycznych. Zdecydowaną większość z nich stanowiły artykuły poświęcone zagadnieniom metodologicznym związanym z tym procesem, wśród nich **szczególnie użyteczne okazały się zaś publikacje o charakterze metaanaliz** (np. Ali, Sell 1998; Baldwin, Blattner 2003; Cashin 1995; Felder 1992; Felder, Brent 2004; Overall, Marsh 1982). Problematyka ewaluacji dydaktyki jest bowiem niezwykle rozwiniętym

¹ Wybrane przykłady artykułów stanowią bibliografię do innego artykułu w tym tomie – S. Rudnicki, A. Szwed, *Ocenianie ocenających. Trzy pytania o ewaluację zajęć dydaktycznych*, s. 35.

i zróżnicowanym obszarem badań naukowych. Wśród poruszanych w artykułach tematów znalazły się m.in. zagadnienia takie jak trafność i rzetelność narzędzi ewaluacyjnych (w tym głównie ankiet studenckich), problem triangulacji (wykorzystywania różnych sposobów gromadzenia danych) w ewaluowaniu dydaktyki, rozróżnienie na ewaluację sumatywną i formatywną oraz jego konsekwencje dla procesu oceny nauczania itp. Opisy tych wielokrotnie i gruntownie przebadanych zagadnień stały się dla nas ważnym teoretycznym odnośnikiem w dalszych zorientowanych praktycznie poszukiwaniach modelu ewaluacji dydaktyki, który najlepiej odpowiadałby potrzebom WSE.

Dobre praktyki

Kolejnym istotnym etapem realizowanego projektu była analiza systemów ewaluacji dydaktyki funkcjonujących na innych uczelniach. Ze względu na ograniczone możliwości czasowe i organizacyjne postanowiliśmy poddać ocenie przede wszystkim polityki ewaluacyjne lub opisy systemów ewaluacyjnych, zamieszczone na stronach internetowych uczelni. Poszukując danych, początkowo swoją uwagę koncentrowaliśmy zarówno na uczelniach zagranicznych, jak i krajowych. Należy jednak uczciwie zaznaczyć, że **dostęp do informacji na temat ewaluacji dydaktyki w przypadku polskich uczelni jest dość trudny**. Instytucje te albo nie posiadają odnośnych systemów², albo nie upubliczniają danych na ich temat³. Wobec zaistniałej sytuacji podstawowymi punktami odniesienia stały się dla nas uczelnie zagraniczne, konkretnie zaś wybrane uniwersytety z obszaru anglojęzycznego: University of British Columbia – znajdujący się w czołówce najlepszych uniwersytetów kanadyjskich; Ohio State University – uniwersytet stanowy znajdujący się w pierwszej dwudziestce najlepszych uniwersytetów amerykańskich, najlepszy uniwersytet publiczny w stanie; Stanford University – prywatny uniwersytet w ścisłej czołówce najlepszych uniwersytetów amerykańskich (4. miejsce); University of Auckland – najstarszy i jeden z najlepszych uniwersytetów w Nowej Zelandii. W ograniczonym zakresie korzystaliśmy także z propozycji Oxford University oraz Washington State University.

2 Nie oznacza to oczywiście, że nie prowadzą w ogóle oceny jakości dydaktyki (do tego zobowiązuje je Państwowa Komisja Akredytacyjna).

3 Jednym z pozytywnych przykładów jest tutaj Uniwersytet Warszawski i działająca w jego ramach Pracownia Ewaluacji Jakości Kształcenia.

Jak już o tym wspominaliśmy, przedmiotem zainteresowania były dla nas dokumenty publikowane na stronach internetowych uczelni. Należy zatem uznać, że była to raczej analiza modeli rozwiązań ewaluacyjnych niż ich realnego funkcjonowania (choć można w sposób uzasadniony zakładać, że wypracowane dokumenty są wynikiem nabytych przez konkretną uczelnię doświadczeń). W przypadku większości uczelni analizie poddawane były dokumenty ogólne, mające charakter polityk ewaluacyjnych, niekiedy zaś także bardziej szczegółowe instrukcje, np. określające procedurę ewaluacji, zasady interpretacji danych itp. Taki sposób potraktowania zebranego materiału pozwolił na dostrzeżenie interesujących i ważnych rozwiązań, niekoniecznie występujących „w pakiecie” w ramach jednej instytucji. Od początku było też dla nas jasne, że bezpośrednie przejęcie systemów działających na innych uczelniach jest niemożliwe chociażby ze względu na odmienny charakter tych instytucji (najczęściej duże uniwersytety funkcjonujące w innej kulturze akademickiej niż polska). *Benchmarking*, którego pewne elementy chcieliśmy zastosować, nie zakłada zresztą prostego naśladownictwa, a raczej twórczą adaptację wybranych rozwiązań czy sposobów działania.

Szczegółowy opis wniosków wynikających z analizy systemów ewaluacji funkcjonujących na wspomnianych wyżej uczelniach przekracza możliwości tego artykułu⁴. Warto jednak przywołać kilka najważniejszych punktów, które stały się dla nas ważnym punktem odniesienia w ocenie dotychczasowych procesów oceny jakości dydaktyki w WSE. **Podstawowy wniosek dotyczy transparentności systemów ewaluacji i nacisku, jaki kładziony jest na poinformowanie wszystkich zainteresowanych grup (studentów, kadry, władz) o jego mechanizmach.** Ważnym elementem jest także jasne sformułowanie celów, jakim służy ewaluacja. Z jednej strony ma więc ona charakter sumatywny (dostarcza danych, które mogą zostać wykorzystane w rozstrzygnięciach o charakterze personalnym), z drugiej – formatywny, związany z rozwojem kompetencji dydaktycznych prowadzącego. **Wszystkie analizowane przez nas uczelnie uwzględniają potrzebę wykorzystania różnych sposobów uzyskiwania informacji na temat prowadzonych zajęć dydaktycznych (triangulacja).** Najbardziej popularnym narzędziem są tutaj oczywiście ankiety studenckie, którym towarzyszą formy gromadzenia danych, takie jak: *peer observations*,

4 Nieco szerzej problem ten został potraktowany w innym artykule w tym tomie: S. Rudnicki, A. Szwed, *Ocenianie oceniających. Trzy pytania o ewaluację zajęć dydaktycznych*, s. 35.

samoocena czy *teaching portfolios*⁵ itp. Warte podkreślenia jest także akcentowanie roli studentów w procesie ewaluacji, zawarte w większości analizowanych dokumentów oraz istnienie mechanizmów wsparcia dla ocenianych dydaktyków w zakresie doskonalenia warsztatu pracy.

Analiza doświadczeń WSE

Niemal równoległe z opisaną powyżej analizą polityk i systemów innych uczelni dokonywaliśmy krytycznej oceny rozwiązań dotychczas stosowanych w WSE. **Ewaluacja zajęć dydaktycznych w postaci cosemestralnej oceny zajęć za pomocą ankiety była prowadzona od samego początku istnienia uczelni.** Warte podkreślenia jest, że w tym samym czasie na innych uczelniach krakowskich proces ten odbywał się najczęściej w sposób dużo mniej uporządkowany i konsekwentny (np. tylko w wybranych jednostkach lub jako działanie inicjowane przez konkretnych prowadzących). Nie oznacza to jednocześnie, że w WSE udało się od razu wprowadzić spójny i kompletny system ewaluacji zajęć dydaktycznych. Wypadałoby raczej mówić o szeregu działań, przynoszących bardziej lub mniej wymierne rezultaty, które miały na celu usprawnienie procesu oceny.

W trakcie dotychczasowych sześciu lat działalności WSE zmieniała się między innymi forma studenckich ocen zajęć dydaktycznych – stosowano ankiety papierowe, następnie elektroniczne i na powrót papierowe. Jak wyjaśnić taki kierunek zmian? **Pomimo znacznie szybszego opracowania danych, jakie zapewniały ankiety elektroniczne, ich główną wadą była bardzo niska, w porównaniu z papierowymi, stopa zwrotu.** Warto w tym miejscu zaznaczyć, że nie wprowadzono ani żadnych negatywnych form motywowania studentów do wypełniania ankiet, ani stosowanego często motywowania finansowego (np. losowanie nagrody). Odwoływano się jedynie do poczucia obowiązku, chęci wpływu na jakość nauczania itp. Niska stopa zwrotu ankiet elektronicznych i przypuszczenie, że skłonność do ich wypełniania wykazują głównie dwie grupy studentów: ci, którzy byli wybitnie zadowoleni oraz ci, którzy byli zdecydowanie nieusatysfakcjonowani prowadzi-

5 Portfolio przygotowuje najczęściej przez samego prowadzącego; obejmuje m.in. dokumentację prowadzonych kursów, informacje o indywidualnej pracy ze studentami, doskonaleniu dydaktycznym prowadzącego, wprowadzanych innowacjach, przygotowanych materiałach i pomocach dydaktycznych, a także zawiera wyniki ocen studenckich, *peer observations* itp.

niem zajęć (oba czynniki wpływałyby na reprezentatywność ocen), sprawiły, że w roku akademickim 2008/2009 podjęto decyzję o powrocie do papierowej formy ankietyzacji. Warto w tym miejscu nadmienić, że oprócz formy przeprowadzania ankiety zmieniała się także jej zawartość jako narzędzia – od prostego kilkupytniowego kwestionariusza po rozbudowany formularz z kilkunastoma zagadnieniami.

Opisany powyżej problem narzędzia zbierania danych pociągał za sobą także pytanie o inne zagadnienia, takie jak wspomniana już stopa zwrotu ankiet, problem ich dystrybucji, a także kwestię opracowania danych (głównie tempo pracy i sposób przygotowania raportów). Innymi ważnym problemem był także sposób wykorzystania danych uzyskanych w ramach ocen studenckich.

Opisując doświadczenia WSE w zakresie ewaluacji dydaktyki, warto również zwrócić uwagę na częściową dywersyfikację form oceny, jaka występowała pomiędzy poszczególnymi kierunkami. Mowa tu przede wszystkim o działającym w ramach filologii angielskiej – lingwistyki stosowanej – a nieobecnym na innych kierunkach, systemie hospitacji zajęć. Miały one charakter zapowiedzianych wizyt, podczas których osoba wizytująca (przełożony) wypełniała wystandaryzowaną ankietę oceniającą obserwowane zajęcia. Po zakończeniu hospitacji (najczęściej w przeciągu kilku dni) wyniki hospitacji były szczegółowo analizowane wspólnie z prowadzącym. Wyraźnie dostrzegalny jest tutaj formatywny charakter takiej formy ewaluacji. W ramach kierunku prowadzone były także tzw. *peer observations* – wzajemne odwiedziny wykładowców na swoich zajęciach (oparte na zasadzie większego doświadczenia lub pokrewieństwa przedmiotów). W trakcie obserwacji odwiedzający wypełniał odrębną ankietę, która zawierała nie tylko opis sposobu prowadzenia zajęć, ale także element autorefleksji (pytania dotyczące przydatności tych zajęć z punktu widzenia praktyki dydaktycznej odwiedzającego).

Obszary zmian

Analiza porównawcza rozwiązań istniejących w WSE i na wymienionych wcześniej uczelniach zagranicznych pozwoliła nam na zarysowanie obszarów potencjalnych zmian. Można wśród nich wymienić problemy takie, jak: konieczność wypracowania formalnego, przejrzystego dokumentu opisującego ewaluację dydaktyki w WSE i jego masowe rozpowszechnienie

w ramach organizacji, powszechniejsze stosowanie innych niż ankiety studenckie narzędzi pozyskiwania danych o jakości dydaktyki, większy nacisk na formatywny aspekt ewaluacji (przy istniejącym silnym wymiarze sumatywnym). **W wymiarze organizacyjnym łatwo dostrzegalny stał się także brak w WSE wydzielonej jednostki zajmującej się całościowo procesem ewaluacji dydaktyki (zadania podzielone między różne działy uczelni).**

Konsultacje z grupami interesariuszy

Jednym z najważniejszych etapów realizowanego przez nas projektu były z pewnością wywiady indywidualne i grupowe, jakie przeprowadziliśmy z władzami uczelni (rektor, dziekani), kadrami dydaktyczną i studentami. Ta część projektu była szczególnie istotna ze względu na towarzyszące nam od początku przekonanie, że **planowaną zmianę w organizacji należy budować nie tyle przez odgórne rozstrzygnięcia (choć w którymś momencie stają się one niezbędne), ale przede wszystkim w procesie oddolnych konsultacji**. Wspomniane badania jakościowe pełniły więc podwójną funkcję. Z jednej strony ich celem było zebranie informacji o tym, jak jest (lub raczej, jak zainteresowane grupy postrzegają obecną sytuację związaną z ewaluacją dydaktyki w WSE), oraz jak być powinno (próba zdefiniowania systemu idealnego). Z drugiej strony wywiady te angażowały poszczególne grupy (władze, kadrami dydaktyczną i studentów) w budowanie systemu, dając poczucie wpływu i sprawstwa. Takie nastawienie stwarzało też duże szanse, że wypracowany ostatecznie system ewaluacji dydaktyki w WSE będzie dość dobrze odpowiadać zróżnicowanym potrzebom poszczególnych grup.

Ze względu na ograniczenia niniejszej publikacji nie sposób w tym miejscu przedstawić wszystkich wniosków z przeprowadzonych przez nas badań, warto jednak przywołać choć kilka przykładów zebranych opinii. Wśród przedstawicieli władz uczelni idea zastosowania ocen studenckich do ewaluacji dydaktyki nie budziła wątpliwości – były one postrzegane jako podstawowe i najważniejsze w praktyce kryterium oceny pracy wykładowców oraz cenne źródło informacji, ułatwiające prowadzenie polityki kadrowej. Mimo przekonania o wiarygodnej i zasadniczo sprawiedliwej ocenie dydaktyki przez studentów, większość osób wyrażała jednak przekonanie, że ilościowe wyniki ocen studenckich nie powinny być kryterium jedynym czy bezwzględnie rozstrzygającym.

W przypadku wywiadów z kadrami dydaktyczną szczególnie uderzający był niewielki poziom wiedzy na temat szczegółowego przebiegu ewaluacji w WSE. Wykładowcy, także ci, którzy współpracują z WSE od kilku lat, kojarzyli w zasadzie jedynie dwa jej etapy: wypełnianie ankiet przez studentów oraz dystrybucję raportów podsumowujących ich wyniki. U rozmówców można było jednocześnie wyczuć bardzo silną potrzebę bycia poinformowanym. Mechanizmy ewaluacji nie były dla nich przejrzyste. Prowadzący nie wiedzieli na przykład, jaki jest związek między wynikami ocen studenckich a podejmowaniem decyzji personalnych lub postrzegali go w sposób uproszczony (złe oceny = zwolnienie).

W trakcie dyskusji z kadrami dydaktyczną prosiłiśmy także o opinie związane z innymi formami zbierania danych dotyczących prowadzonych zajęć, takimi jak *hospitacje* czy *peer observations*. W trakcie wywiadu z wykładowcami lingwistyki dyskusja nad zasadnością tych form w ogóle się nie rozwinęła (funkcjonują one od samego początku istnienia kierunku i są uznawane za normę). W przypadku kadry pozostałych kierunków pomysł wprowadzenia hospitacji zajęć spotkał się ze stosunkowo dużą rezerwą, choć niekoniecznie z bezpośrednią dezaprobatą. Wielu prowadzącym hospitacje kojarzyły się negatywnie – ze szkołą, wizytacją kuratora, lekcją pokazową, jednorazową mobilizacją prowadzącego i ocenami za aktywność dla uczestników. Problematyczne wydawało im się także rozwiązanie organizacyjne hospitacji: kto miałby jej dokonywać? kto byłby jej poddawany? Wskazywano również na brak odpowiedniej kultury oceniania w polskim świecie akademickim. W nieco bardziej pozytywny sposób oceniano pomysł wprowadzenia obserwacji koleżeńskich. W trakcie rozmowy ze strony jednego z prowadzących padł pomysł wprowadzenia tzw. hospitacji wspomagającej (w przypadku słabych ocen studentów), której warunkiem powodzenia musiałaby być jednak atmosfera wzajemnego zaufania.

W opiniach studentów na temat dotychczasowego funkcjonowania ewaluacji dominowały wyraźnie dwa wątki: przekonanie o sensowności samej idei oceny prowadzących oraz brak wiedzy na temat rzeczywistego wykorzystania zebranych w ankietach ewaluacyjnych opinii. Znaczna część dyskusji dotyczyła m.in. udzielania studentom informacji o wynikach ewaluacji, informowania o zmianach wprowadzonych na jej podstawie oraz rozpropagowania idei i zasad ewaluacji wśród studentów. Nasi rozmówcy przedstawili też pewne konkretne propozycje rozwiązań, które ich zdaniem mogłyby stanowić część systemu ewaluacji dydaktyki w WSE, takie jak: ankieta śródsesemestralna czy mailowa skrzynka kontaktowa, które pozwalałyby na ewentualną interwencję w trakcie trwania semestru.

luacji miał być nie tylko technologią *sensu stricto*, ale także sposobem budowania kultury ewaluacyjnej w ramach organizacji.

Przyjęliśmy, że podstawowym celem ewaluacji zajęć dydaktycznych w WSE będzie zapewnienie wysokiej jakości ich prowadzenia, która stanowi jeden z podstawowych czynników wpływających na jakość kształcenia w ogóle (a przy tym bodaj najważniejszy spośród pozostających pod kontrolą uczelni) oraz ściśle wpisuje się w misję WSE, którą jest kształcenie studentów na najwyższym poziomie i przygotowanie ich do konkurencji na europejskim rynku pracy. Przyjęte założenia dotyczące podwójnej (formatywnej i sumatywnej) funkcji ewaluacji wpłynęły na dobór i konstrukcję narzędzi.

Ankiety ewaluacyjne

Podstawowym narzędziem zbierania danych ewaluacyjnych w WSE były (i są) ankiety ewaluacyjne wypełniane przez studentów. W poprzednich latach testowano na uczelni różne kwestionariusze (najdłuższe składały się z 15 pytań zamkniętych i kilku otwartych), w których oceniano różne wymiary dydaktyki – zarówno związane wprost z osobą prowadzącą zajęcia, jak i z treściami merytorycznymi czy organizacją przedmiotu. W ramach opracowywania systemu ewaluacji zdecydowaliśmy się na kwestionariusz, który miał zaledwie 6 pytań zamkniętych w przypadku wykładów, a 7 w przypadku ćwiczeń i konwersatoriów (odpowiedzi udzielane były na skali od 1 do 5), poświęconych przede wszystkim ocenie pracy osoby prowadzącej zajęcia. Ta decyzja opierała się na założeniu, że ewaluacji (zarówno w sensie sumatywnym, jak i formatywnym) powinien podlegać przed wszystkim sposób prowadzenia zajęć. Konkretnie pytania oceniały: kontakt prowadzącego ze studentami, umiejętność zainteresowania studentów przedmiotem, umiejętność wyjaśniania, sposób traktowania studentów, umiejętność prowadzenia dyskusji (w przypadku zajęć ćwiczeniowych), zadowolenie z zajęć z daną osobą oraz ogólną opinię na temat zajęć. Przy doborze pytań oddających poszczególne kryteria ewaluacji brano pod uwagę wyrażone we wcześniej przeprowadzonych wywiadach indywidualnych i grupowych zdanie władz uczelni, pracowników dydaktycznych i przedstawicieli studentów. Kwestionariusz zawierał także pytania jakościowe, w których studenci mogli wyrażać opinie na temat słabych i mocnych stron zajęć lub zgłaszać dodatkowe uwagi. Zakładano, że odpowiadając na pytania jakościowe, studenci odniosą się nie tylko do

osoby prowadzącej, ale także do strony merytorycznej i organizacyjnej zajęć – analiza danych w pełni potwierdziła to założenie.

Czynnikiem wyróżniającym skonstruowane narzędzie (także spośród rozwiązań stosowanych na wielu uczelniach zachodnich, gdzie dominują ankiety składające się z kilkunastu lub kilkudziesięciu pytań) była niewątpliwie jego zwięzłość. Dzięki temu spodziewano się, że narzędzie nie będzie powodować zmęczenia studentów i zniechęcenia do ewaluacji. Zakładano, że krótkie narzędzie będzie nie tylko wystarczająco rzetelne, ale że wszystkie istotne aspekty zajęć, o które nie zapytano w części ilościowej, pojawią się w części jakościowej, na którą przeznaczono w formularzu dużo miejsca. Miała ona także stanowić dodatkowe źródło danych i być jednym ze sposobów zapewniania triangulacji wyników. Chyba zbyt rzadko mówi się w kontekście ewaluacji o tym, że pytania jakościowe wnoszą często znacznie więcej konkretnych informacji niż pytania ilościowe, co może być szczególnie przydatne w wymiarze formatywnym ewaluacji. Przykładowo, dzięki pytaniu ilościowemu prowadzący może dowiedzieć się, że potrafił zainteresować studentów omawianymi zagadnieniami na poziomie 3,6 w skali od 1 do 5, podczas gdy odpowiedzi na pytania jakościowe powiedzą mu, przykładowo, że prezentacje, które przygotował były zbyt długie, że podawał zbyt mało konkretnych przykładów lub że zbyt rzadko wchodził w dyskusje ze studentami. W kontekście doskonalenia warsztatu takie uwagi mogą być bezcenne, choć dla przełożonych prawdopodobnie będą miały nieco mniejsze znaczenie – trudno bowiem wymagać od zwierzchników dokładnej lektury wszystkich tego typu uwag (jakkolwiek przedstawiciele władz WSE deklarowali duże zainteresowanie także tymi informacjami). Na pewno jednak pogłębienie analizy dzięki pytaniom jakościowym może mieć miejsce w przypadku szczególnie wysokich lub szczególnie niskich ocen w części ilościowej. Istotnym problemem (zwłaszcza dla dużych jednostek akademickich) może się jednak okazać żmudny proces przepisywania odpowiedzi na pytania otwarte, niezbędny przy dystrybucji ankiet w formie papierowej (szerzej ten problem opisano poniżej).

Trafność⁶ przygotowanego narzędzia sprawdzono, przeprowadzając pilotaż na grupie studentów i sprawdzając sposób rozumienia pytań (niektóre pytania nieznacznie zmieniono

6 Trafność narzędzia związana jest z pytaniem, czy narzędzie mierzy to, co ma mierzyć (zob. Frankfort-Nachmias, Nachmias 2001:180). Dobre narzędzie powinno być trafne, czyli mierzyć ten (w naukach społecznych zwykle niedostępny bezpośrednio) przedmiot, który ma mierzyć. W odniesieniu do ankiet ewaluacyjnych można przyjąć, że narzędzie trafne powinno mierzyć jakość sposobu prowadzenia zajęć, a nie np. sympatyczność osoby prowadzącej czy łatwość egzaminu.

po pilotażu). Dodatkowym sposobem potwierdzenia trafności części ilościowej narzędzi było porównanie wyników z odpowiedziami na pytania jakościowe. Analiza danych z pierwszej tury ewaluacji przy zastosowaniu tego narzędzia (ewaluacja zajęć w semestrze letnim roku akademickiego 2008/2009) potwierdziła także zadowalającą rzetelność⁷ narzędzia: obliczone niezależnie dla każdej z wylosowanych próby grup studenckich wartości współczynników alfa Cronbacha⁸ najczęściej zawierały się w przedziale 0,8–0,9. Trafność czynnikową⁹ sprawdzono, przeprowadzając analizę czynnikową¹⁰ na wylosowanej próbie grup studenckich. Jej wyniki wskazywały dość wysoką interkorelację pozycji (współczynniki korelacji najczęściej zawierały się w przedziale 0,45–0,65), jednak nie na tyle wysoką, by można było powiedzieć, że poszczególne pytania nie wnoszą istotnej dodatkowej wiedzy. Zauważono także, że pytanie o sposób traktowania studentów prawdopodobnie mierzy nieco inny czynnik niż pozostałe (zdecydowano się jednak pozostawić je w niezminionej formie, uznając, że ten wymiar oceny jest niezbędny). Z kolei pytanie o ogólną ocenę zajęć, które w naszej intencji miało być oceną całokształtu zajęć (a więc zarówno sposobu prowadzenia, jak i treści merytorycznej czy organizacji), było wysoko skorelowane z pozostałymi, co potwierdzało prawidłowość, że kluczowym czynnikiem w ocenie zajęć jest ocena prowadzącego, ale oznaczało także, że pytanie o zajęcia, które ocenia *de facto* osobę prowadzącą, jest niepoprawnie sformułowane.

Ankieta ewaluacyjna w ostatecznej formie, zaproponowanej przed kolejną rundą ewaluacji (semestr zimowy roku akademickiego 2009/2010), zawierała pytania dotyczące opinii studentów na temat: umiejętności zainteresowania ich przedmiotem, umiejętności wyjaśniania, postrzeganego przygotowania prowadzącego do zajęć, sposobu traktowania studentów, ogólnej oceny sposobu prowadzenia zajęć i ogólnego zadowolenia z zajęć

7 Rzetelność narzędzia związana jest z wielkością błędu losowego, powstającego przy kolejnych pomiarach tego samego obiektu dokonanych przy użyciu tego samego narzędzia (zob. Frankfort-Nachmias, Nachmias 2001:185–186). Dobre narzędzie powinno być rzetelne, a więc dawać te same lub zbliżone wyniki przy kolejnych pomiarach tego samego obiektu.

8 Alfa Cronbacha jest jedną z podstawowych miar rzetelności narzędzia rozumianej jako jego jednorodność. Wartości współczynnika zmieniają się od 0 do 1 – w naukach społecznych za rzetelne przyjmuje się zwykle skale o wartości współczynnika powyżej 0,7; rzetelność na poziomie 0,8–0,9 jest uważana za wysoką (więcej o współczynniku alfa Cronbacha zob. Brzeziński 2002).

9 Trafność narzędzia wyznaczana przy użyciu analizy czynnikowej. Ankieta ewaluacyjna spełniająca to kryterium powinna mieć założoną strukturę czynników (np. być jednorodna, jeśli wszystkie pytania mają mierzyć jakąś jednorodną własność). Więcej o trafności czynnikowej zob. Brzeziński 2002.

10 Analiza czynnikowa jest techniką statystyczną redukcji danych, tj. pozwala zidentyfikować ukryte czynniki leżące pod obserwowanym zróżnicowaniem odpowiedzi na poszczególne pytania.

oraz – w przypadku ćwiczeń i konwersatoriów – umiejętności prowadzenia dyskusji i zapewniania odpowiedniej możliwości zadawania pytań. Poprawność rozumienia pytań sprawdzono w pilotażu. Rzetelność narzędzia mierzona po badaniu współczynnikiem alfa Cronbacha była wysoka (średnia wartość 0,84)¹¹. Trafność czynnikową sprawdzono, analizując wyniki analizy czynnikowej wykonanej na wylosowanej próbie grup – tak jak w poprzedniej wersji narzędzia pytanie o sposób traktowania studentów dość często generowało dodatkowy wymiar.

Ankiety dystrybuowane są w formie papierowej, odrębnie dla przedmiotów, form zajęć i grup (wykład i ćwiczenia z tego samego przedmiotu są oceniane oddzielnie; oddzielnie zbierane i raportowane są także wyniki dla różnych grup). Ankietyzacja jest prowadzona zasadniczo na przedostatnich zajęciach w semestrze, tak by w możliwie największym stopniu uniezależnić opinie studentów od przebiegu i wyniku zaliczenia zajęć, który mógłby być oczywistym czynnikiem zniekształcającym oceny. Studenci wypełniają ankiety w trakcie zajęć, a osoba prowadząca jest w tym czasie proszona o opuszczenie sali (jak wynika z wielu badań, obecność prowadzących w sali w trakcie wypełniania ankiet podwyższa oceny, jakie wystawiają im studenci). Wypełnione ankiety są następnie skanowane (część ilościowa) i przepisywane (część jakościowa). Przepisywanie odpowiedzi na pytania otwarte okazało się, zgodnie z przewidywaniami, zadaniem bardzo pracochłonnym (a więc i kosztownym), było jednak konieczne dla zapewnienia pełnej anonimowości studentów, co jest z kolei niezbędnym warunkiem budowania zaufania do procesu ewaluacji. Zabieg taki może być jednak efektywnie prowadzony prawdopodobnie jedynie w niewielkich uczelniach lub na poziomie mniejszych jednostek organizacyjnych, co może stać w sprzeczności z obecną tendencją do centralizacji procesów ewaluacji na poziomie całego uniwersytetu czy szkoły wyższej.

Wiele uwagi w projektowaniu i wdrażaniu systemu poświęcono sposobowi raportowania wyników ankiet ewaluacyjnych. Zdecydowano, że raportowane będą zarówno średnie, jak

11 Wartość współczynnika wyznaczono w następujący sposób: spośród 322 grup, które oceniały zajęcia w semestrze zimowym 2009/2010, wylosowano 64 grupy i dla każdej z nich obliczono niezależnie współczynnik alfa Cronbacha, o ile wariancja wszystkich zmiennych (pytań) była niezerowa. W praktyce ten ostatni warunek dość często nie był spełniony – zwłaszcza w niewielkich i wysoko oceniających zajęcia grupach zdarzało się, że przynajmniej na jedno pytanie wszyscy odpowiadali tak samo. Błąd standardowy średniej wartości współczynnika wyniósł 0,019, zatem z dużym prawdopodobieństwem można przyjąć, że spójność wewnętrzna narzędzia istotnie jest wysoka.

i częstości odpowiedzi na poszczególne pytania, zakładając, że informacja, ile osób wybrało poszczególne odpowiedzi na kolejne pytania, jest cennym źródłem danych (pozwala m.in. zidentyfikować sytuację, w której jedna lub dwie osoby konsekwentnie wybierają odpowiedzi najniższe, co może znacznie zaniżyć średnią w niewielkich grupach; bądź sytuację, gdy grupa jest w swoich opiniach wyraźnie podzielona). Zrezygnowano natomiast z wyliczania ogólnej średniej z wszystkich pytań, nie tylko dlatego, że zestawienie pytań w jedną skalę błędnie zakładałoby ich jednakową wagę (bądź wymagałoby wyznaczenia wag odgórnie), lecz także po to, by uniknąć pokusy oceny osoby prowadzącej za pomocą jednej wartości liczbowej zamiast pełnego profilu. W niedługiej perspektywie planowane jest uruchomienie systemu informatycznego, który będzie przechowywał dane z ewaluacji oraz umożliwiał ich wygodne zestawianie i raportowanie, np. porównywanie ocen osoby prowadzącej w kolejnych latach bądź na tle średnich dla uczelni lub kierunku.

Opracowane wyniki ankiet są przekazywane władzom uczelni i osobom odpowiedzialnym za kształcenie na poszczególnych kierunkach oraz prowadzącym. W toku opracowywania systemu oraz w czasie spotkań z władzami uczelni wykorzystującymi dane ewaluacyjne w polityce kadrowej, zwracano szczególną uwagę na możliwe zniekształcenia wyników ewaluacji prowadzonych metodami ankietowymi¹², wskazując na konieczność sięgnięcia po dodatkowe dane i uwzględnienia szerszego kontekstu oceny (np. wyników osoby prowadzącej w poprzednich latach), szczególnie przy podejmowaniu ważnych decyzji kadrowych.

Tak przygotowany sposób zbierania danych ewaluacyjnych za pomocą ankiet okazał się wartościowy (zarówno w opiniach władz uczelni, jak i studentów, z którymi naturalnie konsultowano proponowane rozwiązania), lecz jednocześnie pracochłonny i kosztowny. W ciągu jednego semestru konieczna była dystrybucja ok. 4–5 tysięcy formularzy ankiet, a następnie zebranie, raportowanie i analiza wyników, co było niemałym przedsięwzięciem badawczym. Sama obróbka czytanych przez skaner danych i formatowanie raportów okazała się dość pracochłonna, bo niemożliwa do pełnej automatyzacji informatycznej. Czasochłonność tego procesu powodowała opóźnienie w raportowaniu wyników, co było niekorzystne, biorąc pod uwagę zasadę jak najszybszej informacji zwrotnej dla prowadzących. Pozornie najłatwiejszym rozwiązaniem mogłaby być zmiana formy ankiet na elektroniczną, jednak w istocie – jak zaznaczyliśmy już wcześniej – oznaczałoby to

12 Zob. wspomniany wcześniej artykuł w tym tomie – S. Rudnicki, A. Szwed, *Ocenianie ocenających...*

pojawienie się nowych problemów: autoselekcji badanych, konieczności wprowadzenia nagród lub przymusu (obydwa sposoby mają wpływ bądź na zakłócenie losowości doboru próby, bądź na nastawienie i staranność udzielania odpowiedzi przez studentów) oraz niskie stopy zwrotu ankiet. Każdy z tych czynników może być istotnym źródłem błędu pomiaru na etapie zbierania danych – podczas gdy jakość danych „na wejściu” jest czynnikiem zasadniczo wpływającym na jakość uzyskanych wyników. Rozwiązaniem, na które się ostatecznie zdecydowano, było wprowadzenie reguły, że osoby prowadzące w WSE zajęcia z danego przedmiotu od co najmniej dwóch lat i uzyskujące wystarczająco wysokie wyniki, będą mogły być oceniane rzadziej niż co rok, lecz nie rzadziej niż raz na dwa lata. Taką decyzję podyktowała obserwacja, że wyniki stałych pracowników i współpracowników uczelni są stabilne, zatem zbieranie danych co rok powoduje niepotrzebną redundancję, a jednocześnie bardzo duże obciążenia organizacyjne. Każdorazowo listę osób ocenianych na dany semestr ustala dziekan, zatem osoby prowadzące (poza zatrudnionymi w pierwszym lub drugim roku pracy) do ostatnich zajęć nie wiedzą, czy będą w danym semestrze oceniane, czy nie. Ryzyko obniżenia przez nich jakości sposobu prowadzenia zajęć jest tym samym zminimalizowane. Jednocześnie zarówno studenci, jak i prowadzący, mogą z góry poprosić o ocenę konkretnych zajęć metodą ankietową. Aby zapewnić studentom jeszcze większe możliwości wyrażenia opinii i uruchomić dodatkowy „mechanizm bezpieczeństwa”, wprowadzono także internetową skrzynkę kontaktową dla studentów, na którą mogą nadsyłać uwagi w ciągu całego semestru (zob. opis poniżej).

Inne formy zbierania danych ewaluacyjnych

116

Pozostałe formy zbierania danych ewaluacyjnych mają służyć z jednej strony triangulacji wyników zebranych metodą ankietową, a z drugiej wspierać funkcję formacyjną ewaluacji. Podstawową metodą wprowadzoną wzorem zasady *peer review* są **obserwacje koleżeńskie**. Osoba, której zajęcia są obserwowane, sama wybiera obserwatora spośród innych prowadzących, ponieważ celem tej formy ewaluacji nie jest ocena, ale wzajemne doskonalenie warsztatu przez obydwie strony. Zdecydowano, że co najmniej raz na 2 lata tej formie obserwacji powinni poddać się wszyscy pracownicy, a obowiązkowo pracownicy w pierwszym roku swojej pracy. Mimo pewnego dystansu części kadry do tej metody, sygnały napływające w pierwszym semestrze jej stosowania były zdecydowanie pozytywne.

Narzędziem wspomagającym przede wszystkim ewaluację sumatywną są z kolei **obserwacje przeprowadzane przez osoby odpowiedzialne za dydaktykę** (zwłaszcza dziekanów lub wyznaczone przez nich osoby), z której korzysta się rzadko – w celu weryfikacji danych z ankiet ewaluacyjnych, jeżeli spadają one poniżej wymaganego poziomu. Przy obu formach obserwacji obowiązuje zasada jak najszybszej informacji zwrotnej – zakłada się bowiem, że sens formatywny mają przede wszystkim nie „suche” wyniki tej czy innej formy ewaluacji, ale ich odpowiednie omówienie i wyciągnięcie wniosków.

Jako opcjonalne i mające cel formatywny narzędzie zbierania danych wprowadzono **ankiety śródsemestralne**, przeprowadzane z własnej woli i samodzielnie przez osoby prowadzące. Przygotowana propozycja formularza zawiera wyłącznie pytania otwarte, które mają pomóc wykładowcom zebrać swobodne opinie studentów w trakcie semestru, a więc w momencie, kiedy możliwa jest jeszcze zmiana niektórych aspektów prowadzenia zajęć (wykładowcy mogą także stosować formularze przygotowane samodzielnie). Internetowa „**skrzynka kontaktowa**” dla studentów funkcjonuje przez cały rok, pozwala zatem na nadsyłanie opinii także już po przeprowadzeniu ankiet ewaluacyjnych (np. po egzaminie). Podstawową rolę pełni tutaj samorząd studencki, który posiada wyłączne prawo do administrowania skrzynką, a jednocześnie obowiązek porządkowania, anonimizowania i przekazywania dziekanowi opinii nadsyłanych przez studentów (obowiązuje go naturalnie także poufność odnośnie prowadzących, o których opinie są wyrażane). Ostatnią formą ewaluacji jest **autoewaluacja** przeprowadzana formalnie w ramach regularnej oceny pracowniczej, a mniej formalnie w ramach obserwacji koleżeńskich. W obecnej chwili nie ma obowiązku tworzenia *teaching portfolios*, głównie z uwagi na pracochłonność ich przygotowania i zastosowania w ewaluacji.

Opisany powyżej system został spisany w formie wewnętrznego dokumentu WSE oraz zaakceptowany zarówno przez władze uczelni, jak i samorząd studentów. Przedstawiono go także kadrze dydaktycznej, która została poinformowana z odpowiednim wyprzedzeniem o planowanych zmianach. Dla wszystkich studentów i szerszego otoczenia uczelni przygotowano opis „polityki ewaluacyjnej WSE” przedstawiający główne założenia i rozwiązania systemu. Dokument rozesłano wszystkim studentom wraz z wyjaśniającym

wprowadzone zmiany listem dziekana; jest on także dostępny na stronie internetowej uczelni dla wszystkich zainteresowanych¹³ (zob. także Aneks).

Warto dodać, że w tym artykule opisano rozwiązania dotyczące tylko ewaluacji zajęć dydaktycznych na studiach stacjonarnych i niestacjonarnych I i II stopnia. W zmodyfikowanej formie ewaluacja jest przeprowadzana na studiach podyplomowych. Odpowiednimi narzędziami badani są także kandydaci na studia w WSE, studenci kończący studia licencjackie oraz absolwenci uczelni. W badaniach dwóch ostatnich grup uwzględnia się przede wszystkim inne niż jakość sposobu prowadzenia zajęć czynniki, mające wpływ na jakość kształcenia, w tym w szczególności: związane z organizacją zajęć, działalnością biura karier i wymianą międzynarodową. Ze względu na ograniczenia objętościowe artykułu tej formy ewaluacji nie mogły być tutaj omówione.

Wnioski

Sensowna i sprawnie działająca ewaluacja zajęć dydaktycznych nie jest prosta do osiągnięcia i wymaga spełnienia określonych warunków. Po pierwsze, musi mieć charakter systemowy. Dla przykładu, jakościowe dane dotyczące słabych stron prowadzonych zajęć mogą nie wystarczyć do poprawienia warsztatu prowadzącego, jeśli nie zostaną wsparte interpretacją oraz radą bardziej doświadczonego dydaktyka. Po drugie, jeśli system ma być stabilny i wpisać się w kulturę organizacji, ewaluacja powinna znaleźć zrozumienie i wsparcie wszystkich interesariuszy: zarządu uczelni, kadry dydaktycznej i studentów. Po trzecie ewaluacja, która nie ma przełożenia na politykę kadrową uczelni, na dłuższą metę nie ma racji bytu. Nie uda się zmobilizować studentów, jeśli nie będą oni widzieć realnych konsekwencji swoich ocen. Po czwarte należy pamiętać, że ankiety studenckie są podstawowym, lecz nie jedynym źródłem danych, zaś do pełnej oceny sytuacji niezbędna jest triangulacja wyników. Po piąte, struktura organizacyjna uczelni powinna zapewniać płynny przebieg złożonego procesu ewaluacji. Optymalne wydaje się powołanie dedykowanej jednostki czuwającej nad jakością kształcenia, pełniącej również rolę archiwum

¹³ Zob. http://www.wse.krakow.pl/pl/Jakosc_zajec_na_WSE/do_pobrania.

danych. Po szóste wreszcie, niezbędna jest stała analiza kosztów funkcjonowania ewaluacji i monitoring jej użyteczności w oczach wszystkich zainteresowanych. Dobra ewaluacja sama wymaga ewaluacji.

Przy tym wszystkim należy mieć świadomość, że budowa trwałej kultury ewaluacyjnej to zadanie na lata, zarówno dla każdej uczelni z osobna, jak i dla całego polskiego systemu wyższej edukacji.

Bibliografia:

Ali, Donna L., Sell, Yvonne (1998), *Issues Regarding the Reliability, Validity and Utility of Student Ratings of Instruction: A Survey of Research Findings*, Calgary: University of Calgary APC Implementation Task Force on Student Ratings of Instruction [15.02.2009].

Baldwin, Tamara, Blattner, Nancy (2003), *Guarding Against Potential Bias in Student Evaluations. What Every Faculty Member Needs To Know*, „College Teaching”, 51(1), s. 27–32.

Brzeziński, Jerzy (2002), *Metodologia badań psychologicznych*, Warszawa: Wydawnictwo Naukowe PWN.

Cashin, William 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.

Felder, Richard M. (1992), *What Do They Know, Anyway?*, „Chemical Engineering Education”, 26(3), s. 134–135.

Felder, Richard M., Brent, Rebecca (2004), *How To Evaluate Teaching?*, „Chemical Engineering Education”, 38(3), s. 200–202.

Frankfort-Nachmias, Chava, Nachmias, David (2001), *Metody badawcze w naukach społecznych*, Poznań: Zysk i S-ka.

Overall, Jesse U. IV, Marsh, Herbert W. (1982), *Students' Evaluations of Teaching: An Update*, „AAHE Bulletin”, s. 9–12.

ANEKS

Ewaluacja zajęć dydaktycznych w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera

Misja

Misją WSE jest kształcenie studentów na najwyższym poziomie. Ostateczna jakość kształcenia jest uwarunkowana wieloma czynnikami (m.in. programem studiów, standardami oceniania studentów, zasobami uczelni wspierającymi kształcenie itd.), wśród których jednym z najistotniejszych jest jakość zajęć dydaktycznych. Unikalny poziom dydaktyki jest czynnikiem wyróżniającym WSE na tle innych uczelni i wynika z założenia, że na współczesnym rynku pracy kluczowe są rzeczywiste kompetencje, których zdobycie zależy w głównej mierze od jakości zajęć oferowanych studentom.

Na WSE obowiązuje specjalny system ewaluacji zajęć dydaktycznych, opracowany w oparciu o doświadczenia renomowanych uniwersytetów z całego świata i dostosowany do specyficznych potrzeb Uczelni.

Cele

Celem ewaluacji zajęć prowadzonej w WSE jest **doskonalenie jakości dydaktyki** oraz **warsztatu prowadzących**. Ewaluacja dydaktyki jest także ważnym elementem **regularnej oceny wszystkich prowadzących** zajęcia, będącej częścią polityki personalnej Uczelni. Dzięki temu, że podstawowym źródłem danych ewaluacyjnych są opinie studentów na temat zajęć, nie tylko władze WSE i prowadzący, ale także studenci, włączeni są w proces tworzenia uczelni jako wspólnoty ludzi uczących się.

Narzędzia

Podstawową metodą zbierania danych ewaluacyjnych są **ankiety** wypełniane przez studentów na zakończenie każdego kursu. Jak pokazuje bogata literatura przedmiotu oraz

doświadczenia WSE i innych uczelni, opinie studenckie są najlepszym pojedynczym wskaźnikiem jakości dydaktyki. Wyniki ankiet ewaluacyjnych są zbierane osobno dla każdej grupy studentów i formy zajęć oraz raportowane do osoby prowadzącej i władz Uczelni. Opinie studentów nie są jednak traktowane automatycznie jako ostateczne oceny prowadzących, lecz poddawane analizie w celu wyjaśnienia przyczyn ocen i uwzględnienia szerszego kontekstu czynników wpływających na ich kształt.

Cennym źródłem danych w ewaluacji dydaktyki w WSE są także **obserwacje zajęć** przeprowadzane w dwóch formach: jako regularne obserwacje koleżeńskie (*peer observations*) oraz obserwacje przeprowadzane z polecenia osób odpowiedzialnych za dydaktykę. Celem obserwacji jest doskonalenie warsztatu zarówno osoby obserwującej, jak i obserwowanej (zwłaszcza w przypadku obserwacji koleżeńskich), oraz dodatkowa weryfikacja danych z ankiet studenckich.

Dodatkowymi źródłami danych ewaluacyjnych są także: **ankiety śródsesemestralne** (rekomendowane zwłaszcza osobom o mniejszym doświadczeniu lub w sytuacji nowych przedmiotów); „**skrzynka kontaktowa**”, dzięki której studenci mogą przekazywać swoje opinie za pośrednictwem Samorządu Studenckiego, oraz **samoocena** prowadzących przeprowadzana w ramach oceny pracowniczej i obserwacji zajęć.

Osoby odpowiadające za dydaktykę w WSE mogą także w zależności od potrzeb stosować bardziej rozbudowane formy i narzędzia ewaluacji. Przy każdej z form ewaluacji zasadą jest, by prowadzący otrzymywali możliwie szybko informację zwrotną na temat jej wyników.

Podsumowanie

Ewaluacja zajęć dydaktycznych w WSE nie jest jedynie formalną procedurą wynikającą z regulacji obowiązujących w polskim szkolnictwie wyższym, ale realnym standardem, który zapewnia wysoką jakość dydaktyki oraz upowszechnia zasady kultury ewaluacyjnej – wzajemnego i opartego na przejrzystych kryteriach oceniania.

Albert Izdebski

Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia na Uniwersytecie Warszawskim – przykład dobrych praktyk ewaluacji jakości kształcenia wyższego. Przegląd kluczowych aspektów badania

Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia na UW ma już trzyletnią historię. Projekt tego badania został przygotowany jako integralna część budowy Systemu Ewaluacji Jakości Kształcenia. Dzięki wykorzystaniu specyficznych rozwiązań metodologicznych i nowoczesnych technologii na etapie realizacji przedsięwzięcia badawczego możliwe stało się zebranie opinii na temat warunków kształcenia wszystkich zainteresowanych studentów Uniwersytetu Warszawskiego. Zebrane w trakcie tego badania opinie studentów stanowią ważne, choć jedno z wielu, źródło informacji o procesach dydaktycznych na uczelni. Zebrane dane stanowią z kolei podstawę do diagnozy sytuacji i podejmowania strategicznych działań nakierowanych na poprawę jakości kształcenia.

- Badanie opinii studentów na temat jakości kształcenia na uczelni techniką ankiety wymagało rozstrzygnięcia szeregu istotnych zagadnień. Wśród nich należy wymienić:
- zakres przedmiotowy badania – a zatem to, jakie opinie studentów są kluczowe z punktu widzenia procesu poprawy jakości kształcenia;
- dobór respondentów – czyli to, w jaki sposób należy dobierać studentów do badania;
- sposób realizacji – tj. jak dostarczyć kwestionariusz ankiety do badanych.

Powyższe kwestie będą przedmiotem niniejszego opracowania, w którym zostaną pokazane najważniejsze dylematy pojawiające się przy projektowaniu ankietowych badań studentów oraz praktyczne rozwiązania tych problemów przyjęte w procesie budowy Systemu Ewaluacji Jakości Kształcenia na Uniwersytecie Warszawskim.

Wprowadzenie

Od 2007 roku na Uniwersytecie Warszawskim istnieje System Ewaluacji Jakości Kształcenia. Zasadniczą rolę w tym systemie odgrywa działalność Pracowni Ewaluacji Jakości Kształcenia UW (PEJK UW). Jej głównym zadaniem jest prowadzenie badań ewaluacyjnych obejmujących najważniejsze obszary życia akademickiego. Przekonanie, że poprawa jakości kształcenia jest możliwa jedynie w wyniku współpracy wszystkich zainteresowanych uczestników życia akademickiego stanowiło fundamentalne założenie w procesie projektowania, a następnie wprowadzania badań ewaluacyjnych na naszej uczelni. Poza gromadzeniem i analizą danych na temat procesów kształcenia i ich efektów, co jest immanentną cechą przedsięwzięć ewaluacyjnych w szkolnictwie wyższym (por. Brzezińska 2000), jednym z głównych celów wdrażanego systemu było więc wypracowanie – a obecnie jest utrwalenie – mechanizmu wspólnych działań na rzecz jakości kształcenia. Rdzeniem tej instytucji są relacje zwrotne między potrzebami społeczności akademickiej a działaniami władz, podejmowanymi na rzecz poprawy jakości kształcenia.

Badania ewaluacyjne na Uniwersytecie Warszawskim z jednej strony umożliwiają diagnozę sytuacji i zdefiniowanie strategicznych kierunków działań – a następnie ocenę ich efektów; zaś z drugiej strony, tworzą nowe, szersze kanały artykulacji potrzeb dla poszczególnych grup środowiska akademickiego.

Badanie opinii wśród studentów

Biorąc pod uwagę powyższe założenie oraz fakt, że na Uniwersytecie Warszawskim kształci się przeszło 55 tysięcy osób (*Sprawozdanie Rektora Uniwersytetu Warszawskiego 2009:36*),

czyli 10% wszystkich studentów polskich uniwersytetów (por. GUS 2009), konieczne staje się poznanie ich zdania na temat szeroko rozumianych warunków kształcenia. Opinie studentów w procesie ewaluacji jakości kształcenia na wyższych uczelniach są ważne z wielu powodów. Choć wystarczy wymienić tylko jeden, mianowicie: studenci są odbiorcami zaprojektowanych procesów dydaktycznych na uczelni, a więc bezpośrednio doświadczają wszelkich pozytywnych i negatywnych skutków wynikających z praktyki funkcjonowania uczelni. To z kolei oznacza, że analiza tych opinii umożliwia „(...) diagnozę pożądaných i niepożądaných zjawisk w systemie kształcenia na UW (...)” (Jasiński i inni, 2008:9).

Na wyższej uczelni istnieje szereg danych zastanych opisujących różne aspekty życia akademickiego. Podobnie jak inne dane urzędowe informują one o warunkach, w jakich funkcjonuje organizacja, o przebiegu działań czy w końcu o ich efektach (por. Sułek 1986). Niestety te dane – relatywnie łatwo dostępne – przede wszystkim mówią o strukturalnych aspektach życia uczelni, selektywnie i dość rzadko ukazując opinie uczestników życia akademickiego. Dla zespołu PEJK UW oznaczało to, że należy zaprojektować, przygotować i w końcu wdrożyć system zbierania opinii studentów.

Wybór wystandaryzowanego wywiadu jako narzędzia badawczego na wstępnym etapie prac konceptualizacyjnych nie wzbudzał kontrowersji, bowiem, jak to się często podkreśla, dzięki ich zastosowaniu: „(...) stosunkowo łatwo i we względnie krótkim czasie można uzyskać dużą ilość zestandaryzowanych informacji odnoszących się do różnych zjawisk, a wybranych dowolnie przez badacza” (Lutyński 2000:153). Badanie opinii studentów wymaga jednak przyjęcia szeregu założeń metodologicznych oraz wielu rozwiązań praktycznych dotyczących sposobu realizacji badania. Dzięki pracy zespołu PEJK UW oraz zaangażowaniu środowiska akademickiego aktualnie na Uniwersytecie Warszawskim funkcjonuje przedsięwzięcie pod nazwą Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia.

Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia – krótka historia

Omówienie podstawowych założeń – a także problemów i sposobów ich rozwiązania – projektu należy poprzedzić krótkim wprowadzeniem, gdyż Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia ma już trzyletnią historię, zaś poniższe rozważania odwołują się wprost do doświadczeń z tego okresu.

Zamysł badań ankietowych powstał wiosną 2007 roku, wówczas określono ogólne ramy przedmiotowe i schemat badania.

Cele badań sformułowano następująco:

- Zebranie informacji o opiniach społeczności studentów UW na temat warunków kształcenia.
- Zdiagnozowanie kluczowych problemów kształcenia na UW.
- Dostarczenie władzom centralnym UW informacji na temat ogólnouniwersyteckich zagadnień dotyczących warunków kształcenia.
- Dostarczenie władzom jednostek dydaktycznych informacji na temat zagadnień dotyczących warunków kształcenia na poziomie jednostek.

W czerwcu 2007 roku zostało przeprowadzone badanie pilotażowe wśród celowo dobranych studentów 7 kierunków. Łącznie pilotażem objęto 450 osób. Wyniki tych badań stały się bardzo użytecznym materiałem na kolejnym etapie tj. w trakcie przygotowań do badań o zasięgu ogólnouniwersyteckim. Pierwszą edycję Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia na UW przeprowadzono w okresie od 30 maja do 30 września 2008 roku, wzięło w niej udział ponad 13 tysięcy studentów. Druga edycja trwała od 30 maja do 31 lipca 2009 roku i objęła ponad 10 tysięcy osób. Biorąc pod uwagę skalę badania, jak i jego tematykę, należy Ogólnouniwersytecką Ankietę Oceniającą Jakość Kształcenia na UW uznać za jedno z największych tego rodzaju przedsięwzięć w kraju. Aktualnie trwają prace nad trzecią edycją badań wśród studentów, która rozpocznie się 30 maja 2010 roku.

Założenia i wybrane zagadnienia metodologiczne

Przyjęta w 2007 roku koncepcja badań ankietowych wśród studentów Uniwersytetu Warszawskiego zakładała:

- **systematyczny charakter obserwacji** – dynamiczne badania w przeciwieństwie do przekrojowych umożliwiają nie tylko diagnozę sytuacji, ale także monitorowanie zmian;
- **rotację tematyki badań w kolejnych edycjach** – w taki sposób, aby zakres przedmiotowy każdej kolejnej ankiety zawierał zarówno problematykę z poprzedniej edycji, jak i tematykę całkowicie nową;
- **objęcie badaniem wszystkich studentów uczelni** – ściśle biorąc kształcących się na I i II stopniu studiów oraz na studiach jednolitych magisterskich; świadomie wyłączono studentów III stopnia (doktorantów), gdyż stanowią oni wyraźnie odrębną zbiorowość w społeczności akademickiej¹;
- **ścisłą współpracę pomiędzy PEJK UW a innymi jednostkami i przedstawicielami uczelni** – w celu definiowania zakresów przedmiotowych badania;
- **wykorzystanie dwóch typów pytań** – wzajemnie uzupełniających się wystandaryzowanych pytań zamkniętych, które pozwalają określić częstości występowania określonych poglądów, oraz pytań otwartych, dających możliwość rekonstrukcji uwarunkowań opinii studentów;
- **zgodność badań z zasadami etyki badawczej** – naruszenie etycznych zasad prowadzenia badań społecznych zawsze stanowi zagrożenie dla wiarygodności materiału empirycznego i podważa sens działań ewaluacyjnych;
- **powszechny dostęp do wyników badań w ramach społeczności akademickiej** – jako jeden z warunków potrzebnych do ożywienia dyskusji na temat jakości kształcenia oraz integracji środowiska uczelni wokół tej problematyki.

1 Kształcenie na III stopniu studiów jest dość specyficzne przede wszystkim dlatego, że są to studia bardziej zindywidualizowane i ukierunkowane na pracę naukową. Większość doktorantów co prawda uczestniczy w kursach dydaktycznych, lecz także znaczna część z nich prowadzi zajęcia. A więc są oni zarówno odbiorcami, jak i twórcami procesów dydaktycznych na uczelni. Ta dualność ról przesądza o specyfice doświadczeń, a w rezultacie o opiniach tej zbiorowości. Z tego powodu uznaliśmy, że badanie doktorantów powinno być odrębnym projektem badawczym.

Powtarzalny charakter badania

Założenie o systematycznej obserwacji przesądziło o dynamicznym charakterze badań, jednak pozostały problemy: które z podejść badawczych zastosować i jak często je aplikować. Ogólnie biorąc, wybór między typami badań sprowadzał się do odpowiedzi na pytanie, czy wykorzystać podejście trackingowe czy panelowe?. W przypadku pierwszych prowadzi się powtarzalne pomiary wśród tak samo zdefiniowanej zbiorowości, wykorzystując stały schemat doboru próby, identyczne lub bardzo podobne narzędzie badawcze oraz schemat realizacji badania (por. Dyjas-Pokorska, Poławska 2004). Wyniki badań trackingowych dają możliwość oceny kierunku i tempa zmian. Z kolei w badaniach panelowych prowadzonych w zbiorowości dokładnie tych samych respondentów, poza ogólną oceną trendu pojawia się także możliwość analizy zmian opinii pojedynczych respondentów, a więc dają one bardziej precyzyjny obraz. Niestety w przypadku badań panelowych występuje szereg trudności, które skutecznie zniechęciły nas do wykorzystania tego podejścia. Po pierwsze: następuje zjawisko „wymierania” próby, czyli w warunkach uczelni z próby w kolejnych okresach obserwacji wypadają studenci kończący naukę, jednocześnie w badanej zbiorowości nie pojawiają się nowi studenci z I roku. Po drugie: następuje efekt uczenia się respondentów polegający na tym, że respondent zapamiętuje pytania oraz swoje wcześniejsze odpowiedzi. Efekt ten może w sposób istotny zniekształcać wyniki. Po trzecie: w badaniach panelowych używa się identycznego (lub nieznacznie zmodyfikowanego) kwestionariusza, a to stoi w sprzeczności z założeniem, że tematyka badań w kolejnych okresach będzie ulegała zmianie. I w końcu: ponieważ badanie jest powtarzane dokładnie w tej samej zbiorowości, to badany student może nie mieć poczucia anonimowości, co także może wpływać na wiarygodność jego odpowiedzi. Powyższe argumenty przesądziły o tym, że **Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia na UW jest realizowana jako specyficzny typ badania trackingowego.**

W badaniach powtarzalnych pojawia się problem związany z częstością dokonywania pomiarów. Biorąc pod uwagę możliwości organizacyjne, planowaną tematykę badań oraz to, że zbyt częste badania ankietowe wśród studentów mogą ich zniechęcić do uczestnictwa w przedsięwzięciu, zdecydowaliśmy, iż **badanie ankietowe będzie odbywało się raz w roku akademickim.** Dodatkowo w zgodzie z założeniem, że w kolejnych edycjach badań

zakres tematyczny będzie ulegał zmianie, określiliśmy zasady rotacji pytań. Pierwszą grupę pytań stanowią te powtarzane co roku, drugą – zadawane co kilka lat, i trzecią pytania *ad hoc* – niepowtarzane w kolejnych edycjach. Przyjęcie takiego rozwiązania wynikało ze specyfiki funkcjonowania wyższej uczelni: są tutaj ważne i/lub dynamicznie ulegające przeobrażeniom przestrzenie – należy je zatem obserwować w sposób możliwie dokładny. Istnieją też takie aspekty, które ulegają zmianie w dłuższych odcinkach czasu i w tym wypadku coroczne śledzenie przeobrażeń nie jest już konieczne. Są też pojedyncze, ważne wydarzenia, które także należy objąć obserwacją.

Badana zbiorowość

Kolejnym założeniem przy projektowaniu przedsięwzięcia był pogląd mówiący, że aby badanie było użyteczne z punktu widzenia procesu ewaluacji jakości kształcenia, to objąć nim należy wszystkich studentów uczelni. Uwzględnienie różnicowań studentów zwłaszcza wg takich cech, jak: kierunek, rodzaj, tryb, rok studiów – w naszym przekonaniu – odgrywa doniosłą rolę z punktu widzenia poziomu ogólności formułowanych wniosków oraz pozwala uchwycić uwarunkowania opinii studentów. Rozwiązania wymagał zatem dylemat: jaki zastosować typ doboru respondentów do badania? Może on mieć charakter reprezentatywny albo wyczerpujący. Aby zastosować dobór losowy, cechujący podejście reprezentatywne, musi zostać spełniony warunek konieczny: „Zbiorowość generalna musi być dostępna w sensie fizycznym lub symbolicznym” (Mayntz i inni, 1985:90). Na uczelni istnieje operat losowania – jest nim lista studentów. Z podejściem reprezentacyjnym wiąże się jednak szereg problemów: selektywność badania, czyli wyłączenia z badanej zbiorowości części studentów chcących podzielić się opiniami, konieczność wypracowania złożonego schematu doboru próby, tak aby możliwe było wnioskowanie na poziomie jednostki dydaktycznej, zagrożenie poczucia anonimowości badanych, „(...) odmowy odpowiedzi spowodowałyby nieprzewidywalny wzrost błędów systematycznych popełnianych podczas wnioskowania (...)” (Jasiński i inni, 2008:9). Ponadto procedury przygotowania operatu, schematu losowania i dotarcie do ściśle określonego studenta przekładają się na znaczny wzrost kosztów badania. Z kolei w przypadku badań wyczerpujących pojawia się innego

rodzaju problem. Zwykle badania społeczne prowadzone na próbach wyczerpujących nie są badaniami dobrowolnymi. Zaś wprowadzenie obowiązku wypełnienia ankiety kłóciło się z założeniem zachowania standardów etycznych badań społecznych (por. Babbie 2004). Przymuszalnie wywołałoby to niepożądane skutki zarówno dla samej idei badań ewaluacyjnych, jak i wpłynęłoby na wiarygodność uzyskiwanych odpowiedzi (por. Jasiński i inni, 2008).

W naszych badaniach **zastosowaliśmy zmodyfikowane podejście prowadzenia obserwacji na próbie wyczerpującej** – nazwane „głos każdego się liczy”. Jako kryteria definiujące populację studentów określono: aktywne konto w Uczelnianym Systemie Obsługi Studiów (USOS) oraz chęć uczestnictwa w procesach poprawy jakości kształcenia na uczelni. Przyjęliśmy więc, że interesują nas w tym wypadku opinie tych studentów „(...) którzy chcą zaangażować się w działania projakościowe”. (por. Jasiński i inni, 2008:9). To ustalenie bezpośrednio przesądziło o sposobie realizacji badania, konieczne było stworzenie możliwości udziału w badaniu każdemu studentowi.

Technika realizacji

Z przyjętych celów i założeń przedsięwzięcia badawczego wynikała bezpośrednio technika zbierania danych. Biorąc pod uwagę planowany zakres przedmiotowy, cykliczny przebieg obserwacji, badaną zbiorowość oraz koszty przedsięwzięcia przewaga badania ankietowego nad pozyskiwaniem danych przy pomocy wywiadu kwestionariuszowego jest bezsporna. Niemniej pojawił się inny trudny wybór dotyczący tego, jaki rodzaj ankiety należy zastosować. Poważnie rozważaliśmy dwa sposoby aplikacji: ankietę audytoryjną oraz ankietę internetową. Inne sposoby takie jak tradycyjna ankieta pocztowa lub ankieta prasowa zostały odrzucone, ponieważ charakteryzowały się one wyraźnymi wadami. W przypadku pierwszej bardzo wysokimi kosztami realizacji, w przypadku drugiej brakiem odpowiedniego medium dystrybucji spełniającego wymóg inkluzywności badania.

Ankieta audytoryjna, a więc taka, która może być przeprowadzana w trakcie zajęć, ma jednak szereg wad. W 2007 roku przeprowadziliśmy pilotażowe badanie, wykorzystując

ankietę audytoryjną. Wyniki tych badań ukazały wiele trudności wynikających z przyjętej w pilotażu techniki realizacji². Do najważniejszych ograniczeń tego podejścia należą bardzo wysokie koszty związane z drukiem, dystrybucją i wprowadzeniem danych. Poważne trudności wiążą się także z logistyką badania. Po pierwsze: należy dysponować rozkładem zajęć wszystkich studentów uczelni i dostarczyć ankiety na właściwe zajęcia. Problem w tym, iż może zdarzyć się, że zajęcia, z różnych powodów, nie odbywają się w terminie – potwierdziły to wyniki badania pilotażowego. Po drugie: pojawia się problem zwany błędem pokrycia, czyli niektórym studentom wielokrotnie zostanie dostarczona ankieta, a z kolei inni mogą zostać wyłączeni z badania. Po trzecie: realizacja ankiet w trakcie zajęć dydaktycznych może nieprzychylnie nastawiać nauczycieli akademickich do przedsięwzięcia.

W badaniach można także wykorzystać ankietę internetową. Zaletą tego podejścia jest to, że spadają znacząco koszty realizacji oraz w dużej mierze zostają wyeliminowane problemy związane z logistyką badania. **W procesie zbierania danych w I i II edycji Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia na UW zastosowaliśmy ankietę internetową.** Problem sprowadza się w tym wypadku do dwóch zasadniczych kwestii: dotarcia do studentów oraz wypracowania odpowiedniego narzędzia informatycznego.

W celu dotarcia do studentów wykorzystaliśmy potencjał systemu USOS. Dzięki temu powszechnie używanemu na UW systemowi informatycznemu, możliwe było dotarcie z ankietą do wszystkich studentów. Niemniej wymagało to szeregu działań. Przede wszystkim konieczne było nawiązanie ścisłej współpracy z osobami zarządzającymi USOS na uczelni. Dzięki temu zostały precyzyjnie określone: możliwości aplikacji ankiety internetowej i wymagania wobec informatycznego systemu ankietującego „Ankieter”. System informatyczny został wypracowany pod merytorycznym nadzorem PEJK UW jako aplikacja służąca do prowadzenia ankiet internetowych. Warto przy tym podkreślić, że może być ona wykorzystywana przez wszystkie jednostki UW. „Ankieter” został zaprojektowany w taki sposób, aby charakteryzował się wysoką funkcjonalnością, czyli aby można było w jednym czasie prowadzić kilka rodzajów badań, zadawać szereg typów pytań, stosować w jednym przedsięwzięciu wiele wersji ankiet, powracać do niedokończonych wywiadów w celu

2 Skala problemów była tak znaczna, że w ogólnouniwersyteckich badaniach studentów w latach 2008 i 2009 zastosowaliśmy ankietę internetową.

ich uzupełnienia, automatycznie zapisywać odpowiedzi respondentów w elektronicznych zbiorach danych. Ważną cechą tej aplikacji jest jej otwarta struktura, która umożliwia modyfikację typów pytań oraz innych elementów ankiety np. komunikatów i dyspozycji.

„Ankieter” jest niezależną aplikacją współpracującą z systemem USOS w związku z tym wymaga odpowiedniego administrowania. W trakcie badań system ankietujący wykorzystuje niektóre – zdefiniowane uprzednio – informacje o osobach zarejestrowanych w USOS, po zakończeniu badania zebrane odpowiedzi na pytania są zapisywane w odrębnych plikach, w efekcie organizatorzy badania nie mają dostępu do informacji zawartych w Uniwersyteckim Systemie Obsługi Studiów. Jest to szczególnie ważne z punktu widzenia anonimowości badanych oraz bezpieczeństwa informacyjnego uczelni.

Realizacja badania przy pomocy ankiety internetowej – poza wypracowaniem procedur wykorzystywania USOS i samej aplikacji ankietującej – wiąże się z innymi, ważnymi zagadnieniami. Wśród nich, poza specyficznym charakterem ankiety, o czym nieco dalej, wymienić należy:

- właściwą koordynację działań związanych z uruchomieniem aplikacji na różnych serwerach USOS tak, aby badanie rozpoczęło się dokładnie w planowanym terminie;
- konieczność każdorazowych testów narzędzia przed badaniem pod względem: zawartości merytorycznej (co odpowiada badaniom pilotażowym) i układu formalnego (co wynika ze specyfiki badań internetowych);
- ustalenie procedur kontroli badania, a w tym systematyczne monitorowanie, czy zakładane funkcjonowanie aplikacji nie uległo zakłóceniu w skutek działania czynników zewnętrznych np. awarii serwera.

Zakres tematyczny badania

Jednym z ważniejszych problemów był zakres przedmiotowy ankiety, a zatem odpowiedź na pytanie: o co należy pytać studentów w kontekście problematyki jakości kształcenia na uczelni? Powszechnie wiadomo, że wielu ważnych wymiarów jakości kształcenia – np. standardów kształcenia – student nie jest w stanie prawidłowo ocenić. Może jednak oce-

niać warunki, w jakich się kształci – w tym procesy organizacyjne czy wyposażenie uczelni w pomoce dydaktyczne – a te stanowią ramy dla urzeczywistnienia określonych standardów edukacyjnych. Toteż w naszych badaniach koncentrowaliśmy się przede wszystkim na różnych aspektach warunków kształcenia.

Warto zwrócić uwagę na niektóre bardziej ogólne problemy związane z pomiarem opinii studentów. Nieco upraszczając, opinie o warunkach kształcenia wśród studentów są kształtowane przez dwa podstawowe mechanizmy: po pierwsze przez indywidualne doświadczenia, i po drugie poprzez utrwalone w przestrzeni społecznej poglądy innych. Oznacza to, że udzielając odpowiedzi na pytania, studenci niekoniecznie przedstawiają własne, ugruntowane opinie, mogą odtwarzać dominujące stereotypy z macierzystych jednostek. Na poziomie analitycznym nie zawsze jest możliwe wychwycenie tych subtelnych różnic. Jednak z punktu widzenia diagnozy sytuacji nie jest to poważna wada, bowiem nawet powielone opinie innych studentów wskazują na wagę problemu wewnątrz uczelni. Kolejna kwestia polega na tym, że warunki kształcenia na dużej uczelni można rozpatrywać na dwóch poziomach: ogólnouczelnianym i jednostkowym. Przykładowo student, oceniając obsługę procesu dydaktyki, czyni to z perspektywy własnej jednostki, zaś wyrażając opinie na temat systemu zapisów na WF odnosi się do rzeczywistości ogólnouniwersyteckiej. Ma to dwie ważne konsekwencje praktyczne. Przede wszystkim student musi jasno określić swój kierunek studiów. Zaś ustalenie kierunku studiów badanego w praktyce jest zadaniem trudnym, gdyż studenci często kształcą się jednocześnie na kilku kierunkach. W naszym badaniu przyjęliśmy następujące rozwiązanie: na początku respondent określał jednoznacznie kierunek swoich studiów na podstawie listy jednostek dydaktycznych. Dalsze pytania zadawane były w kontekście tak określonej jednostki dydaktycznej. Dodatkowo dwa poziomy organizacyjne na dużej uczelni także wymagają namysłu i staranności w prowadzeniu analiz, trzeba bowiem bardzo wyraźnie rozdzielać te dwie przestrzenie, aby uniknąć błędnych wnioskowań.

W ramach prac koncepcyjnych, a następnie szerokich konsultacji, zostało określonych kilkanaście ważnych aspektów warunków kształcenia. Zaś badania pilotażowe wskazywały, które z pytań mogą sprawiać studentom ewentualne problemy. Wiedza ta pozwoliła na wypracowanie pytań oraz rozmieszczenie ich w blokach tematycznych.

W poniższym zestawieniu przedstawiłem zakres przedmiotowy dwóch edycji badań ankietowych. Jak z niego wynika, część tematów – zgodnie z założeniem – pojawiła się w dwóch okresach pomiaru. Większość z pytań w powtórzonych blokach tematycznych zostało ponownie zadanych, choć nie wszystkie.

Tabela 1. Zakres przedmiotowy oraz liczba pytań w blokach badawczych w dwóch edycjach Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia na UW*

I EDYCJA		II EDYCJA	
BLOK PYTAŃ	LICZBA PYTAŃ	BLOK PYTAŃ	LICZBA PYTAŃ
Organizacja dydaktyki w jednostce dydaktycznej	28	Organizacja dydaktyki w jednostce dydaktycznej	28
Program studiów	11	Program studiów	14
Przestrzeganie praw studenta	6	Przestrzeganie praw studenta	10
Metryczka	27	Metryczka	20
Dydaktyka między jednostkami dydaktycznymi	12	Warunki rozwoju naukowego	24
Infrastruktura i wyposażenie jednostki dydaktycznej	16	Biblioteka	32
Organizacja ogólnouniwersyteckich zajęć WF i lektorów	10	Warunki socjalno-bytowe	35
Infrastruktura ogólnouniwersytecka	8		

* dane mają charakter orientacyjny – w badaniu stosowano kilka wersji kwestionariuszy o różnej liczbie pytań; dodatkowo niektóre z pytań były zadawane tylko wybranym respondentom

Blok pytań na temat organizacji dydaktyki w II edycji został zadany m.in. po to, by sprawdzić efekty wprowadzanych zmian inspirowanych wnioskami z I edycji badania. Z kolei pytania o program studiów zostały powtórzone, ponieważ jest to kwestia niezwykle ważna z punktu widzenia jakości kształcenia, i dodatkowo bardzo aktualna – programy obecnie są

dynamicznie przebudowywane. W wyniku doświadczeń z wcześniejszej edycji modyfikacji uległ sposób pytania o prawa studenta. Pomimo tego, że odsetek respondentów wyrażających przekonanie, że ich prawa są naruszane, był znikomy – uznaliśmy tę kwestię za ważną i wymagającą pogłębienia. Pozostałe bloki pytań pojawiły się tylko w jednej edycji badania, lecz zakładamy, że w kolejnych latach znowu powrócą do ankiety.

Z powyższego zestawienia wynika również to, że ankiety zawierały sto kilkadziesiąt pytań. Obszerność narzędzia nasuwała pewne obawy o to, czy badanie nie będzie zbyt obciążające dla respondenta i czy tym samym długość wywiadu nie będzie wpływała na wiarygodność zebranych danych, a także na przerywanie udzielania odpowiedzi w trakcie wypełniania ankiety. W związku z tym szczególną uwagę przykładaliśmy do sposobu formułowania pytań i układu kwestionariusza. W testach pilotażowych sprawdzaliśmy czas wypełniania ankiety. Okazało się, że zawiera się on w granicach kilkunastu minut. Analiza zebranych danych to potwierdziła. Ważnym ustaleniem było również i to, że kolejność pytań w ankiecie nie wpływała na zmiany odsetka odmów odpowiedzi, czyli długość ankiety nie wywoływała reakcji ucieczki wśród badanych. W innych analizach sprawdzaliśmy spójność i rzetelność pomiaru. Przykładowo w pilotażu w dwóch różnych miejscach ankiety pojawiło się pytanie o ocenę warunków sanitarnych. Wyniki pokazały, że 45% studentów dokonało identycznej oceny. Gdyby wskaźnik zgodności odpowiedzi konstruować w sposób mniej rygorystyczny, czyli dopuścić różnice w ocenach o jeden punkt na siedmiopunktowej skali, to procent zgodnych odpowiedzi wyniósł 80%. Te wyniki dowodzą, że studenci bardzo poważnie potraktowali nasze przedsięwzięcie.

Charakterystyka kwestionariusza ankiety

Formuła narzędzia badawczego w dużej mierze została określona poprzez specyfikę badanej zbiorowości oraz metodę zbierania danych. W badaniach zastosowano dwie wyraźnie odmienne wersje ankiety. Pierwsza była przeznaczona dla studentów studiów międzywydziałowych, druga dla pozostałych. To rozwiązanie było podyktowane tym, że z perspektywy procesu dydaktycznego te dwie grupy bardzo się różnią. W przypadku stu-

dentów międzywydziałowych kształcenie przebiega w wielu jednostkach dydaktycznych. A więc mogą oni oceniać zarówno warunki w jednostce macierzystej, jak również w innych. W związku z tym studentom studiów międzywydziałowych zadawaliśmy pytania na temat organizacji dydaktyki we własnej, międzywydziałowej jednostce. A w dalszej części ankiety spośród jednostek dydaktycznych, w których się kształcą, wybierali tę, którą będą oceniali. Ponadto nie byli oni pytani o oceny pewnych aspektów, w szczególności dotyczących programu studiów, bowiem studia międzywydziałowe zakładają indywidualnie ustalany tok kształcenia.

Zastosowanie systemu ankietującego wiąże się z wypracowaniem specyficznego modelu komunikacji z respondentem. Tutaj szczególnie ważnym aspektem była rekrutacja do badania. Zaproszenie do uczestnictwa w badaniu następowało w momencie, gdy student zalogował się do systemu USOS-web. Wówczas uaktywniał się system ankietujący i pojawiał ekran powitalny. Na nim były przedstawione informacje o badaniu oraz pytanie o chęć uczestnictwa. Naciskając pole „tak”, student przechodził do ankiety, a gdy zaznaczył pole „nie”, ekran zniknął, a student dalej pozostawał w USOS-webie. Ekran zapraszający rekrutował studentów w trakcie całego okresu badania. Wynikało to z przekonania, że badany, wchodząc do USOS-weba, może akurat nie mieć czasu na wypełnienie ankiety. Pojawiający się przy kolejnej wizycie ekran zapraszający miał na celu umożliwić studentowi udzielenie odpowiedzi na pytania zawarte w ankiecie. Ekran spełniał także swoje funkcje wówczas, gdy badany musiał przerwać udzielanie odpowiedzi na pytania. W takim przypadku student po ponownym zalogowaniu był informowany, że może dokończyć wypełnianie ankiety. To rozwiązanie mogło irytować – w przestrzeni wirtualnej wyskakujące strony budzą często negatywne emocje. Z naszych danych wynika jednak, że negatywne reakcje respondentów na pojawiający się ekran były relatywnie rzadkie. Zastosowaliśmy również szereg innych, niezbędnych komunikatów informujących np. o procedurze postępowania lub zmianie tematów w ankiecie.

Jak wcześniej wspominałem z uwagi na technikę realizacji i długość ankiety konieczne było zastosowanie przejrzystego schematu pytań. Większość pytań w *Ogólnouniwersyteckiej Ankiecie Oceniającej Jakość Kształcenia na UW* stanowiły pytania zamknięte, zwykle mające charakter pytań jednokrotnego wyboru. Badani byli proszeni o wyrażenie swojej oceny na siedmiopunktowej skali. W ankiecie znalazły się także bardziej skomplikowane

pytania mające charakter filtrujący, co oczywiście wymaga szczególnej uwagi na etapie projektowania i testowania ankiety internetowej. Poza pytaniami zamkniętymi w ankiecie znalazły się także pytania otwarte. Zazwyczaj występowały po każdym bloku tematycznym. Analiza jakościowa takich odpowiedzi daje możliwość pogłębionego wglądu w wyrażane opinie. Nasze doświadczenia wskazują, że umieszczenie pytań otwartych zdecydowanie poszerza możliwości interpretacji zjawisk, a ponadto spełnia też inne funkcje – pozwala studentom artykułować ważne dla nich problemy.

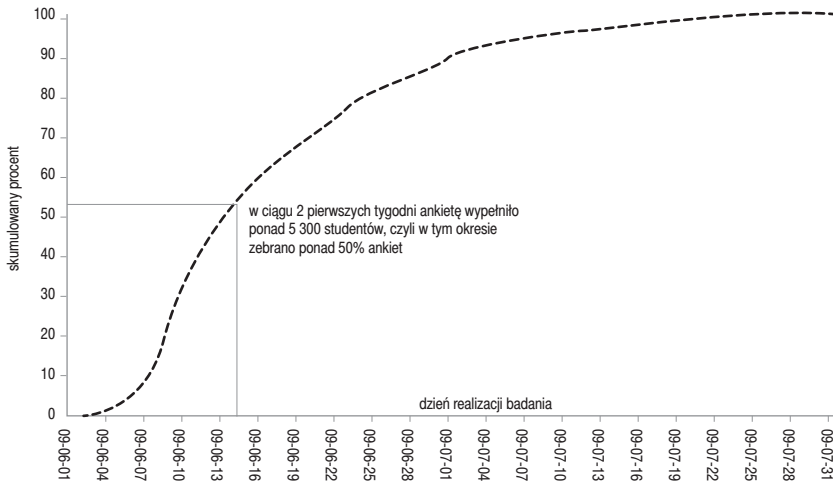
Realizacja badania

Wdrożenie Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia na UW poprzedzały szerokie konsultacje wśród przedstawicieli różnych środowisk akademickich. W przypadku przedstawicieli studentów wypełniały one nie tylko funkcje merytoryczne, ale także promocyjne. Jakkolwiek wyniki badania pilotażowego pokazały duże zainteresowanie studentów przedsięwzięciem, to do końca nie mogliśmy mieć pewności, jaki będzie poziom uczestnictwa. Chodziło więc o przekonanie studentów do masowego uczestnictwa w badaniu. Działania promocyjne prowadzone były zarówno przed rozpoczęciem badania, jak i w trakcie jego trwania. W celach promocyjnych poza wspomnianymi konsultacjami wykorzystywano także inne kanały komunikacji. I tak informacja o badaniach pojawiła się zarówno w prasie krajowej, jak i periodykach uczelni; jako medium wykorzystano także akademicką rozgłośnię Radio Kampus. Akcją informacyjną uzupełniona została poprzez plakaty, ulotki i internetowe bannery.

W I edycji badania jako termin realizacji wybrano okres od 29 maja do 1 października 2008 roku, czyli czas trwania sesji letniej, wakacji i sesji poprawkowej. Wybór ten opierał się na przekonaniu, że w tym okresie studenci najczęściej logują się do USOS-weba. Okazało się, że przeważająca część studentów odpowiedziała w trakcie trwania sesji letniej, w przeciągu pierwszych dwóch tygodni trwania badania zebrano ponad 50% wszystkich ankiet, zaś do końca sesji letniej uzyskano 75% spośród wszystkich wypowiedzi. Te doświadczenia wpłynęły na skrócenie okresu realizacji ankiety z czterech do dwóch miesięcy. W kolejnej

edycji system ankietujący działał w okresie od 1 czerwca do 31 lipca 2009 roku. Jak wynika z poniższego rysunku, faktycznie miało to nieznaczny wpływ na poziom realizacji badania, gdyż studenci przede wszystkim wypełniają ankietę w czerwcu, tj. w okresie sesji letniej.

Wykres 1. Skumulowany procent wypełnionych ankiet w kolejnych dniach realizacji badania II edycji Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia na UW



W I edycji badania wziął udział co czwarty student, zaś w II edycji badaniem objęto co piątego studenta UW. Jak pokazały szczegółowe analizy spadek poziomu realizacji był spowodowany skróceniem czasu trwania badania. Porównanie liczby wypełnionych ankiet w czerwcu i lipcu w dwóch edycjach pokazuje, że w II edycji w tym czasie zebrano o 5% więcej wypowiedzi niż w I edycji (por. Izdebski i inni, 2009). Świadczy to niewątpliwie o bardzo dużym zaangażowaniu studentów w działania na rzecz poprawy jakości kształcenia na naszej uczelni i rozwoju przedsięwzięcia.

Podstawowe znaczenie dla powodzenia przedsięwzięcia, jakim jest Ogólnouniwersytecka Ankieta Oceniająca Jakość Kształcenia na UW ma nie tylko zakres podejmowanych problemów badawczych czy proces podejmowania decyzji w oparciu o dane z badań, lecz także spójna strategia komunikacyjna informująca o wynikach badań i decyzjach podejmowanych na ich podstawie.

Zakończenie

Na zakończenie warto krótko scharakteryzować sposób wykorzystania wyników. Zebrane dane zostają opracowane analitycznie. Szerokie spektrum informacji można wielostronnie wykorzystać. Na ich podstawie powstaje wiele różnorodnych tematycznie raportów. Trafiają one do centralnych władz uczelni i umożliwiają podjęcie odpowiednich decyzji. Przykładowo, w badaniu ujawniliśmy problemy studentów z terminami zapisów na zajęcia ogólnouniwersyteckie – ta wiedza została wykorzystana w decyzji Senatu o zmianie terminów zapisów. Inny przykład: badani studenci wskazywali, że informacje zamieszczane na stronach internetowych jednostek dydaktycznych zawierają szereg luk. W oparciu o tak zdiagnozowany problem Rektor podjął decyzję obligującą jednostki do uzupełnienia informacji prezentowanych na stronach internetowych.

Wyniki badania są wykorzystywane nie tylko przez władze centralne. Raporty opisujące warunki studiowania na poziomie jednostek dydaktycznych są przekazywane dziekanom i Wydziałowym Zespołom Zapewnienia Jakości. Informacje w nich zawarte służą nie tylko do diagnozy sytuacji, ale są często wykorzystywane jako wskaźniki w procesie wewnętrznej ewaluacji jednostki dydaktycznej. Tematyczne raporty poświęcone wybranym zagadnieniom trafiają również do kierowników pozostałych jednostek organizacyjnych uczelni. Raporty są też przedkładane przedstawicielom Zarządu Samorządu Studentów. W wersji uproszczonej – pozbawione często trudnych i nużących dla czytelnika opisów analitycznych – dostępne są na stronie internetowej PEJK UW. Można więc powiedzieć, że badanie poza funkcjami poznawczymi, wypełnia i te związane z integracją środowiska uczelni.

Bibliografia:

- Babbie, Earl (2004), *Badania społeczne w praktyce*, Warszawa: Wydawnictwo Naukowe PWN.
- Brzezińska, Anna (2000), *Miejsce ewaluacji w procesie kształcenia*, [w:] Anna Brzezińska, Jerzy Brzeziński (red.), *Ewaluacja procesu kształcenia w szkole wyższej*, Poznań: Wydawnictwo Fundacji Humaniora.
- Dyjas-Pokorska, Anna, Poławska, Małgorzata (2004), *Trackingi jako specyficzny rodzaj pomiaru w badaniach marketingowych*, [w:] Paweł Sztabiński, Franciszek Sztabiński, Zbigniew Sawiński (red.), *Nowe metody, nowe podejścia badawcze w naukach społecznych*, Warszawa: Wydawnictwo IFiS PAN.
- GUS (2008), *Rocznik Statystyczny Rzeczypospolitej Polskiej 2009*, Warszawa: Wydawnictwo GUS.
- Izdebski, Albert, Jasiński, Mikołaj, Zając, Tomasz, Kowalski, Michał (2009), *Raport z drugiej edycji badań jakości kształcenia na UW*, Warszawa: Wydawnictwo UW.
- Jasiński, Mikołaj, Izdebski, Albert, Kowalski, Michał, Zając, Tomasz, Pękala, Justyna (2008), *Raport z pierwszej edycji badań jakości kształcenia na UW*, Warszawa: Wydawnictwo UW.
- Mayntz, Renate, Holm, Kurt, Hubner, Peter (1985), *Wprowadzenie do metod socjologii empirycznej*, Warszawa: PWN.
- Lutyński, Jan (2000), *Metody badań społecznych. Wybrane zagadnienia*, Łódź: Łódzkie Towarzystwo Naukowe.
- Sulek, Antoni (1986), *Socjolog wobec danych urzędowych; przyczynek do metodologii badań nieankietowych*, „Kultura i Społeczeństwo”, nr 1.
- Sprawozdanie Rektora Uniwersytetu Warszawskiego za rok 2008 (2009)*, Warszawa.

Tomasz Zając

Wykorzystanie danych z Uniwersyteckiego Systemu Obsługi Studiów oraz systemu Internetowej Rejestracji Kandydatów w procesie ewaluacji jakości kształcenia

Ważnym elementem procesu ewaluacji kształcenia jest analiza ścieżki edukacyjnej, którą przeszedł każdy student, czyli trajektorii studiowania. W przypadku studiów wyższych za jej początek można uznać proces rekrutacji, a za zakończenie – obronę pracy dyplomowej. Badając losy pojedynczych osób – od kandydata do absolwenta – można sprawdzić, jak wyniki uzyskane w procesie rekrutacji przekładają się na sukces w studiowaniu. To zaś daje lepszą (empiryczną) podstawę oceny jakości wykształcenia uzyskiwanego na uczelni.

W swoich badaniach koncentruję się na pierwszym etapie trajektorii studiowania: badam przebieg procesu rekrutacji na studia na Uniwersytecie Warszawskim oraz wartość predykcyjną wyników rekrutacyjnych dla sukcesu na studiach.

Celem badań jest nie tylko opis przebiegu procesów selekcji na Uniwersytecie Warszawskim oraz wzajemnych powiązań pomiędzy wynikami na różnych etapach studiów. Planowane jest także wypracowanie modelu analizy danych zbieranych przez administrację uczelnianą w ramach obsługi dydaktyki oraz procesu rekrutacji – danych z Uniwersyteckiego Systemu Obsługi Studiów (USOS) oraz systemu Internetowej Rejestracji Kandydatów (IRK).

Wprowadzenie

Dla procesu ewaluacji jakości kształcenia niezwykle istotnym elementem jest analiza ścieżki edukacyjnej, jaką przeszedł każdy student, czyli badanie trajektorii studiowania. W przypadku studiów wyższych za jej początek można uznać proces rekrutacji, a za zakończenie – obronę pracy dyplomowej. Badając losy pojedynczych osób, od kandydata do absolwenta, można sprawdzić, jak wyniki uzyskane w procesie rekrutacji przekładają się na sukces w studiowaniu oraz czy wyniki osiągnięte na poszczególnych etapach studiów są ze sobą powiązane. Daje to dodatkowe, empiryczne, kryterium oceny jakości wykształcenia uzyskiwanego na uczelni.

W swoich badaniach, będących elementem działalności Pracowni Ewaluacji Jakości Kształcenia UW, koncentruję się na pierwszym etapie trajektorii studiowania – badam przebieg procesu rekrutacji na studia na Uniwersytecie Warszawskim. Interesują mnie zarówno zachowania poszczególnych aktorów – władz jednostek dydaktycznych czy też komisji rekrutacyjnych oraz kandydatów – jak również możliwości przewidywania sukcesu w studiowaniu przy pomocy wyników rekrutacyjnych.

Edukacja na poziomie wyższym, szczególnie studia bezpłatne, jest cennym zasobem, który nie jest powszechnie dostępny, przynajmniej w przypadku najlepszych kierunków studiów. Ukończenie studiów wyższych, poza rozwojem intelektualnym oraz prestiżem, przynosi również bardzo wymierne korzyści – otwiera nowe możliwości pracy, umożliwia osiągnięcie wyższych dochodów. Bardzo ważne jest zatem, aby dobro, jakim jest edukacja na poziomie wyższym, w szczególności edukacja bezpłatna, było dystrybuowane właściwie. Problem jest niezwykle istotny z punktu widzenia całego społeczeństwa. W roku 2009 w Polsce na 131 uczelniach państwowych kształciło się około 1 270 000 studentów (w tym około 830 000 na studiach stacjonarnych; GUS, 2009). Zapewnienie edukacji tak licznej grupie osób wiąże się naturalnie z ogromnymi nakładami finansowymi. Wielkim marnotraw-

stwem zasobów byłoby przyjmowanie na studia nieodpowiednich osób, np. takich, które nie będą w stanie ich ukończyć. Równie poważnym błędem byłoby nieprzyjmowanie na studia osób, które na to zasługują. Dlatego tak ważne jest analizowanie procesów rekrutacji na studia – jakości przyjętych procedur oraz jakości egzaminów wstępnych. Nie chodzi tylko o zapewnienie poprawności formalnej (równego traktowania kandydatów, dążenia do obiektywizmu itp.). Należy dążyć do wyboru spośród kandydatów tych osób, które z największym prawdopodobieństwem odniosą sukces w studiowaniu. Jest to jeden z powodów, dla których warto prowadzić analizę procesów rekrutacji.

Chciałem w tym miejscu podkreślić rolę świadectwa maturalnego. We współczesnych społeczeństwach istnieje rozbudowany system certyfikatów, egzaminów potwierdzających zdobycie różnego rodzaju wiedzy i umiejętności. Posiadanie certyfikatów określa przynależność do grup, daje dostęp do zasobów. Jednym z tego typu certyfikatów jest świadectwo maturalne. Uczniowie nie mają wprawdzie obowiązku przystępowania do egzaminu, ale jego zdanie stanowi warunek konieczny do rozpoczęcia studiów wyższych w Polsce. W połowie lat 90. rozpoczęły się prace nad zmianą formy egzaminu maturalnego i zwiększenia jego roli w systemie edukacji. Zgodnie z deklaracjami Wojciecha Małeckiego, który stanął na czele grupy reformatorów, mającej przygotować nowy egzamin dojrzałości, matura miała przekształcić się „w certyfikat totalnej jakości absolwenta. Nowe świadectwo miało informować nie tylko o przeszłych osiągnięciach posiadacza, ale i o jego możliwościach na przyszłość, i dzięki temu być ‘przepustką’ do mniej lub bardziej atrakcyjnych pozycji społecznych” (Konarzewski 2004:164). W rzeczywistości zmiana roli matury nie okazała się aż tak znacząca. Nadal nie jest to egzamin obowiązkowy dla wszystkich kończących edukację na poziomie ponadgimnazjalnym. Najważniejsze zmiany dotyczyły sposobu przeprowadzania egzaminu oraz jego roli w procesie rekrutacji na studia. Po pierwsze, w 2005 roku wprowadzono ujednolicony dla całego kraju zewnętrzny egzamin maturalny. Po drugie, wyniki egzaminu maturalnego stały się podstawą przyjęcia na studia I stopnia lub jednolite magisterskie (por. art. 169 ustawy *Prawo o szkolnictwie wyższym*).

W poniższym artykule przedstawię schemat analizy przebiegu procesu rekrutacji oraz wartości predykcyjnej egzaminów maturalnych dla sukcesu w studiowaniu, opierającej się na danych pochodzących ze zbiorów tworzonych przez administrację uczelnianą w ramach obsługi dydaktyki oraz procesu rekrutacji – z Uniwersyteckiego Systemu Obsługi

Studiów (USOS) oraz systemu Internetowej Rejestracji Kandydatów (IRK). **Zaprezentuję ogólne założenia projektu, omówię podstawowe problemy i zaproponuję pewne ich rozwiązania.** W poniższym artykule czytelnik nie znajdzie natomiast opisu wszystkich kwestii metodologicznych oraz szczegółowych wyników analiz, które będą przedmiotem osobnej, znacznie obszerniejszej publikacji.

Źródła danych

Pomysł badania procesów selekcji kandydatów i analizowania właściwości egzaminów wstępnych sam w sobie nie jest nowy. W Polsce badania mające na celu określenie wartości predykcyjnej egzaminów wstępnych dla sukcesu w studiowaniu prowadzone były przynajmniej od kilkudziesięciu lat. W latach 60. ubiegłego stulecia prace na temat predyktorów powodzenia na studiach publikowali m.in. Jan Woskowski (1962), Jadwiga i Zygmunt Komorowscy (1963), Krystyna Kądzilska (1963), Barbara Łobodzińska (1964), Danuta Matuszewska-Mysior (1966) czy Wiesław Wiśniewski (1968). Również w późniejszych latach prowadzone były prace nad właściwościami systemów rekrutacyjnych. Wymienić w tym miejscu można prace Eugeniusza Plucińskiego, Ryszarda Wilczyńskiego (1974), Elżbiety Łuczak (1985), zespołu kierowanego przez Jana Osińskiego (1985) oraz Krystyny Ciekotowej, Ryszarda Jasińskiego (1990) i Bartłomieja Janusza, Józefa Kolonko (2001).

Wspomniane badania w przeważającej części odnoszą się do stosunkowo niewielkich zbiorowości. Z reguły też nie były to projekty prowadzone długofalowo. Przeważnie obejmowały do kilku roczników studentów przyjętych na wybrane wydziały jednej lub kilku uczelni. We wszystkich przypadkach poważnym utrudnieniem dla badaczy było zbieranie danych na temat poszczególnych studentów. Istotnym ograniczeniem, szczególnie dla badaczy podejmujących problem we wcześniejszych okresach, był brak możliwości wykorzystania komputerowych technik statystycznej analizy danych.

Obecnie badacze problemu znajdują się w znacznie bardziej komfortowej sytuacji. Postępująca od kilkunastu lat modernizacja systemów obsługi dydaktyki oraz procesu rekrutacji na studia, a przede wszystkim ich informatyzacja, spowodowała, że niemal do zera spadły

koszty zbierania na potrzeby badań danych na temat poszczególnych studentów. Eksport danych z dostępnych baz, aktualizowanych na bieżąco przez administrację uczelni, nie wymaga na ogół większych nakładów pracy. W przypadku Uniwersytetu Warszawskiego informacje pochodzące z Uniwersyteckiego Systemu Obsługi Studiów oraz systemu Internetowej Rejestracji Kandydatów dają potencjalną możliwość odtworzenia „trajektorii studiowania” całej populacji studentów¹. W systemie IRK znajdziemy bowiem następujące dane²:

- podstawowe zmienne socjodemograficzne, takie jak: płeć, wiek, obywatelstwo czy województwo, z którego kandydat pochodzi;
- informacje na temat szkoły, którą kandydat ukończył;
- informacje na temat kierunków studiów, na które dana osoba zdecydowała się kandydować;
- informacje o tym, na które spośród wybranych kierunków kandydat został zakwalifikowany, a na których nie udało mu się uzyskać odpowiednio wysokiej pozycji na liście rankingowej;
- informacje dotyczące wyników egzaminów zdawanych przez kandydata, na podstawie których przeprowadzana jest rekrutacja:
 - dla przeważającej większości kandydatów na studia pierwszego stopnia i jednolite magisterskie będą to wyniki:
 - egzaminów maturalnych (tzw. nowych matur) lub
 - Centralnych Egzaminów Wstępnych przeprowadzanych przez Uniwersytet Warszawski dla kandydatów, którzy zdawali egzamin dojrzałości sprzed reformy – tzw. starą maturę lub
 - wyniki matury międzynarodowej – The International Baccalaureate (IB);
 - dla kandydatów na studia II stopnia będą to wyniki z wcześniejszych etapów studiów (średnie ocen ze studiów, ocena pracy dyplomowej) lub wynik egzaminu wstępnego.

1 W 2009 roku na Uniwersytecie Warszawskim studiowało ponad 55 000 osób (Piotrowska 2010).

2 W systemie IRK znajdują się dane wszystkich kandydatów na studia, a nie tylko osób, które rozpoczęły naukę. W 2009 roku na studia na UW aplikowało blisko 34 000 kandydatów, w tym ponad 28 000 na studia jednolite magisterskie lub pierwszego stopnia (Zając 2009).

Uniwersytecki System Obsługi Studiów zawiera natomiast dane na temat przebiegu procesu kształcenia:

- informacje o tym, gdzie ostatecznie kandydat został przyjęty na studia;
- zapis ocen uzyskanych w trakcie studiów;
- informacje o ukończeniu studiów – o terminie oraz ocenie na dyplomie;
- informacje o ewentualnym skreśleniu ze studiów oraz jego przyczynie.

Do opisanych powyżej danych, podobnie zresztą jak w innych przypadkach, nie można podchodzić bezkrytycznie. Po pierwsze, proces komputeryzacji jest w różnym stopniu zaawansowania w poszczególnych jednostkach. Nie wszystkim udało się tak jak na przykład Wydziałowi Matematyki, Informatyki i Mechaniki czy Wydziałowi Fizyki przeprowadzić całkowitą informatyzację procesu obsługi dydaktyki³. Chociaż powszechne jest wykorzystanie systemu USOS-web np. przy zapisach na zajęcia, to nie we wszystkich jednostkach zapis przebiegu studiów został w pełni ucyfrowiony. W przypadku tych wydziałów, instytutów czy katedr prowadzenie analiz jest obecnie niemal niewykonalne⁴.

Poza dostępnością baz danych, tworzonych przez uczelnianą administrację, elementem, który czyni obecną sytuację badaczy wyjątkową, jest niespotykana dotąd forma egzaminu wstępnego na studia. Wprowadzone w 2005 roku reformy miały charakter rewolucyjny – zmieniły system egzaminów maturalnych i kwalifikacji na studia jednocześnie. Zewnętrzne egzaminy, jednakowe dla wszystkich maturzystów z danego rocznika, stwarzają zdecydowanie większe możliwości porównania jakości kandydatów ubiegających się o przyjęcie na różne kierunki studiów niż rezultaty egzaminów dojrzałości w poprzedniej formie, czy też pozostałe oceny uzyskane na zakończenie szkoły średniej, nie wspominając już o prowadzonych przez poszczególne jednostki egzaminach wstępnych.

3 Pierwsza wersja USOS powstała na Wydziale Matematyki Informatyki i Mechaniki UW w 2000 roku. Sprawność funkcjonowania systemu potwierdzają oceny studentów. W obu edycjach Ogólnouniwersyteckiej Ankiety Oceniającej Jakość Kształcenia na Uniwersytecie Warszawskim oceny poszczególnych aspektów funkcjonowania systemu USOS wśród studentów Wydziału Matematyki Informatyki i Mechaniki były zdecydowanie pozytywne (Jasiński i inni 2008, 2009).

4 Mimo braku danych na temat stopni na studiach, możliwe jest prowadzenie analizy zachowań kandydatów na etapie rekrutacji, opierające się o dane z systemu IRK, który obejmuje wszystkie jednostki.

Schemat analizy

Rozważania na temat wartości predykcyjnej egzaminów maturalnych dla przewidywania odniesienia sukcesu w studiowaniu wymagają uprzedniego zastanowienia się nad właściwościami nowej matury jako egzaminu zewnętrznego oraz nad przebiegiem samej rekrutacji. Należy odpowiedzieć na pytanie, co mówią wyniki egzaminu maturalnego – jakimi informacjami na temat kandydatów dysponują komisje rekrutacyjne, a także w jaki sposób wykorzystywana jest ta wiedza w procesie selekcji na studia. Nie sposób też nie uwzględnić wpływu preferencji osób ubiegających się o przyjęcie na studia, dotyczących kierunku i trybu studiów oraz wynikających z nich decyzji na ostateczne rezultaty procesu rekrutacji. Dlatego też należy prowadzić analizę na trzech, powiązanych ze sobą poziomach: na poziomie kraju, uczelni oraz wybranych jednostek dydaktycznych, np. wydziałów czy instytutów. Na poziomie kraju przebiega analiza właściwości matury jako egzaminu zewnętrznego, czyli źródło wiedzy na temat kandydatów, jakim dysponują komisje rekrutacyjne. Na poziomie uczelni badanie dotyczy przebiegu rekrutacji, czyli sposobu wykorzystania wyników matur w procesie selekcji kandydatów oraz strategię osób ubiegających się o przyjęcie na studia i jednostek dydaktycznych prowadzących rekrutację. Zaś na poziomie wybranych jednostek można zbadać, w jakim stopniu wyniki egzaminu maturalnego mogą posłużyć do przewidywania wyników w studiowaniu.

Analiza na poziomie kraju

W pierwszej kolejności należy zadać sobie pytanie o to, jakimi informacjami na temat kandydatów dysponują komisje rekrutacyjne, przyjrzeć się właściwościom egzaminów zewnętrznych, a w szczególności przeanalizować, jakim egzaminem jest matura. Analiza na poziomie kraju ma stanowić tło dla dalszych rozważań prowadzonych już na poziomie uczelni lub poszczególnych jednostek. Jej wykonanie nie wymaga prowadzenia własnych badań. Wystarczającą ilość informacji znaleźć można w dostępnych publikacjach.

Na temat zasad prawidłowej konstrukcji testów oraz czynników utrudniających właściwy pomiar umiejętności uczniów piszą między innymi Maciej Jakubowski i Artur Pokropek (2009:8–63). Natomiast informacje o różnych uwarunkowaniach wyników organizowanych w Polsce egzaminów zewnętrznych (sprawdzianu szóstoklasistów i egzaminu gimnazjalnego) można znaleźć w opracowaniach prezentowanych w „Biuletynie Badawczym Egzamin” (2004, 2007a, 2007b, 2007c).

Spośród wszystkich aspektów konstrukcji egzaminów kluczową kwestią z punktu widzenia wykorzystania egzaminów maturalnych jako podstawy rekrutacji jest zapewnienie porównywalności wyników różnych edycji egzaminów oraz rezultatów uzyskanych z różnych przedmiotów. Jest to problem o tyle istotny, że przy obliczaniu punktów rekrutacyjnych komisje rekrutacyjne na Uniwersytecie Warszawskim (zapewne nie tylko tam) traktują w ten sam sposób wyniki egzaminów maturalnych z różnych lat. Często także identyczne wagi przypisywane są punktom zdobytym z różnych przedmiotów.

Roman Dolata wskazuje na dwie podstawowe strategie budowania narzędzi egzaminacyjnych. Pierwsza strategia zakłada prowadzenie szeroko zakrojonych badań standaryzacyjnych, dzięki którym kolejne edycje egzaminu mają znane charakterystyki psychometryczne, jak w przypadku amerykańskiego SAT. W ramach drugiej strategii badania standaryzacyjne prowadzone są na znacznie mniejszą skalę i nie pozwalają na dokładne oszacowanie charakterystyk psychometrycznych wykorzystywanych narzędzi. „Ciężar działań przeniesiony jest na procedury post factum”. Po zebraniu danych na temat wyników egzaminu stosowane są ewentualne procedury korekcyjne (m.in. korekty trudności skal, harmonizacji wyników; Dolata i inni 2004:81–82). Niewiele wskazuje na stosowanie powyższych strategii w przypadku egzaminu maturalnego. Raporty dotyczące wyników matur, publikowane przez Centralną Komisję Egzaminacyjną (CKE 2009), pokazują, że rozkłady wyników egzaminów różnią się między sobą w istotny sposób, szczególnie gdy porównujemy wyniki matur z różnych przedmiotów (np. angielskiego i wiedzy o społeczeństwie). Trudno jest jednak orzec z całą pewnością, z czego wynikają obserwowane różnice – ze zmiennego poziomu trudności egzaminów, czy też może z różnic w stopniu opanowania materiału z poszczególnych przedmiotów przez kolejne roczniki maturzystów.

Ograniczona możliwość porównania wyników maturalnych z kolejnych lat oraz między przedmiotami poważnie komplikuje część analiz w kolejnych etapach. We wszystkich

momentach, w których brane są pod uwagę rezultaty matur, np. przy analizie jakości kandydatów lub badaniu wartości predykcyjnej egzaminów maturalnych, konieczne jest uwzględnienie nie tylko punktów procentowych, ale także zrelatywizowanych wyników. Miernikiem osiągnięć maturalnych mogłyby być tzw. punkty centylowe – względna pozycja kandydata wśród ogółu zdających dany przedmiot – lub wykorzystywane przez Centralną Komisję Egzaminacyjną skale staninowe (OKE w Krakowie 2005).

Analiza na poziomie uczelni

Drugi poziom analizy dotyczy wykorzystania wyników maturalnych w rekrutacji na studia pierwszego stopnia i jednolite magisterskie. W swoich badaniach zajmuję się naborem na studia na Uniwersytecie Warszawskim, prowadzonym od roku 2007. Wtedy to na niemal wszystkich kierunkach⁵ zaczął obowiązywać dwustopniowy podział studiów.

Na ostateczny rezultat rekrutacji ma wpływ wiele czynników. Po pierwsze, mamy do czynienia z działaniami uczelni, a w zasadzie z działaniami poszczególnych jednostek dydaktycznych i ich komisji rekrutacyjnych. Senat uczelni na wniosek władz poszczególnych jednostek dydaktycznych określa zasady rekrutacji – wybiera zestaw egzaminów maturalnych, które kandydat musi zdać, oraz ustala zasady przeliczania punktów uzyskanych na egzaminach maturalnych na punkty rekrutacyjne (czyli procedury dokonywania redukcji danych na temat kandydata) itp. Decyzje władz uczelni odnośnie do wymagań rekrutacyjnych są częściowo ograniczane przez działania Centralnej Komisji Egzaminacyjnej i Okręgowych Komisji Egzaminacyjnych, które decydują o kształcie egzaminu maturalnego – liczbie i poziomie przedmiotów, które maturzyści mogą zdawać, przedmiotach obowiązkowych itd. Do kompetencji uczelnianych komisji rekrutacyjnych należy także podejmowanie decyzji o ewentualnym otwarciu dodatkowej rekrutacji na kierunki, na których nie został wyczerpany limit miejsc.

Po drugie, mamy do czynienia ze zbiorowością osób ubiegających się o przyjęcie na studia. Preferencje i zachowania przede wszystkim tej grupy decydują o przebiegu i wynikach

⁵ Poza, częściowo, Wydziałem Prawa i Administracji oraz Wydziałem Psychologii.

rekrutacji. Kandydaci bowiem sami określają, na które kierunki chcą aplikować. Do nich też należy ostateczna decyzja, czy chcą rozpocząć studia na kierunku, na który zostali zakwalifikowani. Każda podjęta decyzja zmienia warunki, w jakich dokonują swoich wyborów pozostali uczestnicy gry, jaką jest nabór na studia. Każda dodatkowa osoba na liście chętnych na dany kierunek zwiększa konkurencję. Każda rezygnacja z podjęcia studiów tworzy wolne miejsce dla innego kandydata.

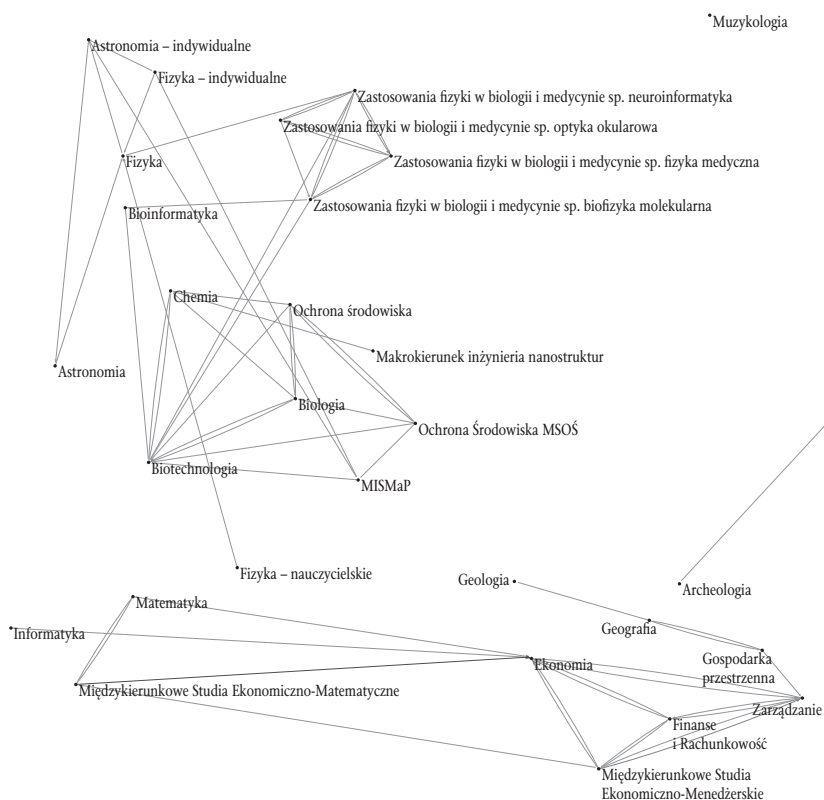
Należy pamiętać, że nie tylko kandydaci uczestniczą w pewnej „grze rekrutacyjnej”, w której konkurują o miejsca na najbardziej preferowanych kierunkach. Konkurują ze sobą również poszczególne kierunki studiów – starają się przyciągnąć jak najlepszych kandydatów. Część komisji rekrutacyjnych stara się możliwie przyspieszyć kolejne etapy procesu rekrutacji tak, aby jak najszybciej zachęcić kandydatów z wyższych pozycji na listach rankingowych do złożenia wszystkich wymaganych dokumentów.

Drugi poziom analizy ma kilka powiązanych ze sobą celów. Po pierwsze, ma odtworzyć pewne wzory zachowań kandydatów – przeanalizować wybory osób ubiegających się o przyjęcie na studia: zbadać zjawisko współkandydowania – jednoczesnego rejestrowania się na wiele kierunkach – oraz ostatecznych wyborów studiów już po etapie kwalifikowania kandydatów. Drugim problemem jest wykorzystanie wyników matur w ramach procesu naboru kandydatów – dobór przedmiotów egzaminacyjnych, sposoby obliczania punktów rekrutacyjnych itp. Kolejnym pytaniem jest stopień, w jakim wyniki matur determinują ostateczne rezultaty naboru na studia. Na koniec można przyrzeć się результатам selekcji – spróbować odpowiedzieć na pytanie, kto dostał się na studia, czyli ocenić jakość kandydatów i osób przyjętych na studia (na tym etapie korzystając wyłącznie z wyników egzaminów maturalnych). Kwestie te są częściowo podejmowane w corocznych raportach dotyczących rekrutacji na studia na UW (Batorski 2006; Abramczuk 2008; Zajac 2009). Dostępne opracowania dalekie są jednak od wyczerpania tematu.

Analiza wymagań rekrutacyjnych oraz zachowań kandydatów (rejestracje na studia oraz decyzje o podjęciu studiów) ma na celu ujawnienie pewnych grup kierunków, które są często jednocześnie wybierane przez osoby ubiegające się o przyjęcie na studia na Uniwersytecie Warszawskim oraz sprawdzenie, na ile wynika ono z zastosowania podobnych zestawów przedmiotów maturalnych jako kryteriów doboru kandydatów.

Przykładowym elementem tego typu analiz może być poniższa mapa współkandydowania na studia stacjonarne pierwszego stopnia i jednolite magisterskie w 2009 roku. Poszczególne koła to kierunki studiów, a ich wielkość jest proporcjonalna do liczby kandydatów ubiegających się o przyjęcie na nie. Połączenie pomiędzy dwoma kierunkami A i B oznacza, że przynajmniej 10% kandydatów na kierunek A ubiegało się jednocześnie o przyjęcie na kierunek B. Kolorami oznaczone zostały trzy grupy kierunków:

Rysunek 1. Powiązania pomiędzy kierunkami wynikające z jednoczesnych wyborów dokonywanych na etapie rejestracji kandydatów w 2009 roku

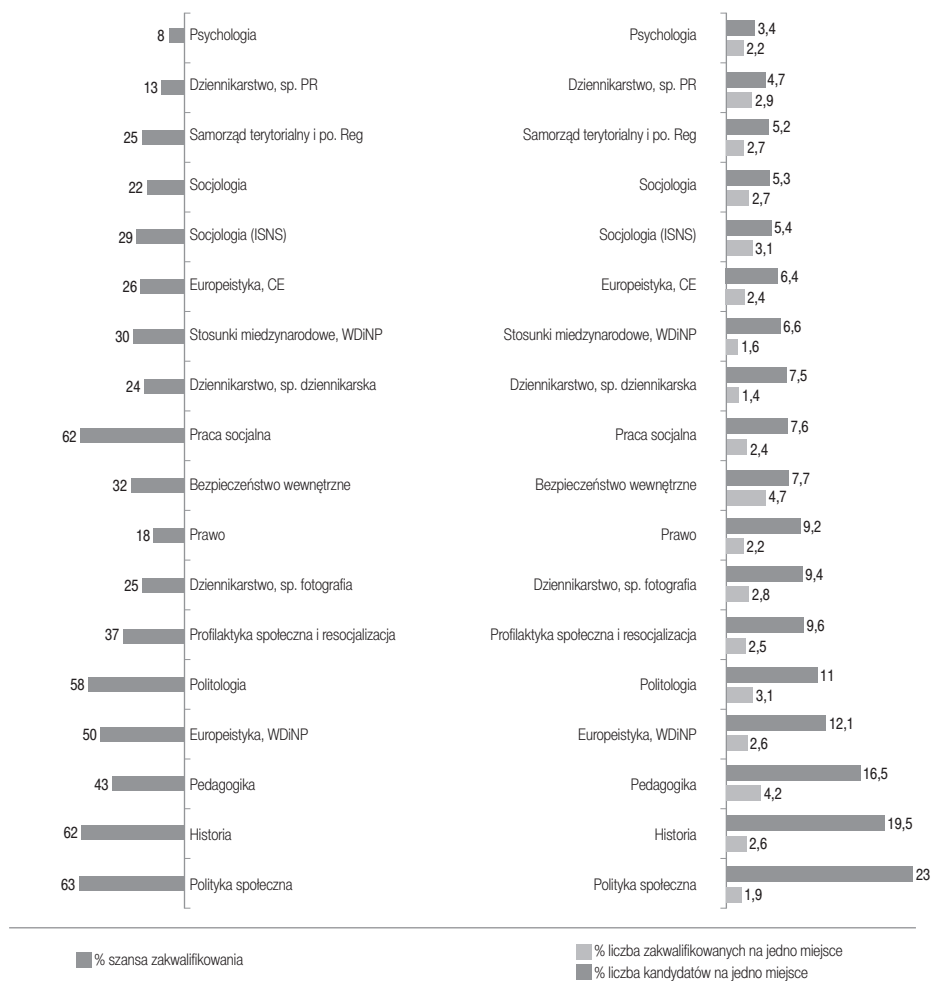


Źródło: opracowanie własne

Kolejny krok, to analiza popularności kierunków – badanie, w jakim stopniu poszczególne kierunki są „oblegane”. Do pewnego stopnia będzie to próba odtworzenia preferencji kandydatów – wskazania kierunków tzw. pierwszego wyboru.

Można wskazać kilka potencjalnych miar popularności kierunków. Najprostszą i najprawdopodobniej najczęściej stosowaną jest liczba kandydatów przypadających na jedno miejsce. Stosowana samodzielnie nie zdaje jednak w pełni sprawy z zainteresowania poszczególnymi kierunkami studiów. Nie mierzy ona bowiem stopnia, w jakim kandydaci na dany kierunek są faktycznie zdecydowani podjąć studia. Rzadko dochodzi do sytuacji, aby wszyscy zakwalifikowani w pierwszej turze rekrutacji kandydaci złożyli wymagane dokumenty i zostali przyjęci. Zazwyczaj, aby wyczerpać limit przyjęć, konieczne jest rozszerzenie listy zakwalifikowanych na studia. W skrajnych przypadkach dochodzi do sytuacji, gdy podjęcie studiów oferowane jest wszystkim kandydatom. Przy analizowaniu stopnia, w jakim „oblegane” są poszczególne kierunki studiów, wskazane jest więc wykorzystanie oprócz miernika, jakim jest liczba kandydatów przypadających na jedno miejsce, także informacji na temat tego, jak wielu kandydatów zostało zakwalifikowanych. Można więc stwierdzić, ilu kandydatom należało zaoferować studia, aby wypełnić wszystkie miejsca. Można także odnieść liczbę zakwalifikowanych do ogólnej liczby kandydatów, czyli obliczyć szansę zakwalifikowania się. Dzięki temu będzie można ocenić, na ile kandydaci byli rzeczywiście zainteresowani podjęciem studiów oraz na ile mamy faktycznie do czynienia z selekcją na poszczególnych kierunkach. Na kierunkach, na których liczba kandydatów nie przekracza lub przekracza tylko w niewielkim stopniu liczbę dostępnych miejsc, trudno mówić o jakiegokolwiek selekcji ze względu na egzaminy maturalne. O rezultatach rekrutacji decydują wtedy zupełnie inne czynniki.

Rysunek 2. Porównanie podstawowych mierników popularności kierunków w grupie kierunków społecznych i politologicznych w 2009 roku



Źródło: opracowanie własne

Odpowiedź na ostatnie postawione w tej części pytanie, dotyczące efektów rekrutacji będzie wymagała zdefiniowania w pierwszej kolejności pojęcia jakości kandydata. Z jednej strony możemy porównywać jakość osób ubiegających się o przyjęcie na studia, posługując się punktami uzyskanymi na egzaminie maturalnym. Z drugiej strony, ze względu na ograniczoną porównywalność wyników maturalnych między przedmiotami, warto skorzystać z miar zrelatywizowanych, np. badając pozycję w rankingu ogółu zdających maturę z da-

nego przedmiotu lub pozycję wśród osób ubiegających się o przyjęcie na studia. Warto porównać wyniki maturalne ogółu kandydatów, osób zakwalifikowanych oraz przyjętych. W połączeniu z informacjami na temat popularności kierunków da nam to pełen obraz selekcyjności poszczególnych kierunków.

Uzyskanie odpowiedzi na powyższe pytania może być niezwykle użyteczne z punktu widzenia władz uczelni i ewaluacji procesów kształcenia. Wyniki analiz mogą mówić o popycie na poszczególne kierunki studiów i pomóc w procesie budowania systemu rekrutacji wielokierunkowej. Identyfikacja kierunków, na których dochodzi do większej selekcji kandydatów, oraz zbadanie „jakości” przyjętych studentów może też rzucać nowe światło na wyniki innych badań ewaluacyjnych, np. ankiet przeprowadzanych wśród studentów, lub też być elementem badań poświęconych efektywności kształcenia. Bez znajomości stanu wyjściowego nie można prowadzić analizy efektów netto działań podejmowanych przez uczelnie.

Z punktu widzenia ewaluacji procesu rekrutacji ważne jest natomiast, aby przed przystąpieniem do oceny wartości predykcyjnej egzaminów maturalnych dla sukcesu na studiach mieć informacje na temat stopnia, w jakim determinują one wyniki naboru na studia w poszczególnych jednostkach oraz posiadać wiedzę na temat poziomu (i jego zróżnicowania) przyjętych kandydatów.

Analiza na poziomie jednostek dydaktycznych

Ostatnim, najważniejszym etapem analizy będzie badanie prowadzone na poziomie pojedynczych jednostek dydaktycznych. W ramach tej części pracy planuję sprawdzić, w jakim stopniu można posługiwać się wynikami maturalnymi do przewidywania sukcesu na studiach. Analiza odbywać się będzie dwutorowo. Z jednej strony będę badać wartość predykcyjną „surowych” wyników z poszczególnych egzaminów maturalnych. Z drugiej strony zamierzam przyjrzeć się wartości predykcyjnej pewnych funkcji wyników maturalnych, jakimi są punkty rekrutacyjne czy też pozycja na liście rankingowej. Jednoczesna analiza wartości predykcyjnej wyników egzaminów maturalnych oraz liczby punktów

rekrutacyjnych stworzy dodatkowo możliwość oceny jakości algorytmów przeliczania punktów w ramach procedury rekrutacyjnej.

Analiza tego typu będzie w pierwszym kroku wymagała zdefiniowania pojęcia sukcesu na studiach. Planuję odejść od stosowanego dotychczas w wielu badaniach kryterium sukcesu, jakim jest średnia arytmetyczna wyników na studiach i na jej miejsce zaproponować rozwiązanie, które będzie bardziej stosowne z punktu widzenia teorii pomiaru. Lepiej odpowiadające potrzebom analizy wydaje mi się zastosowanie do pomiaru sukcesu szeregu zmiennych dwuwartościowych, stanowiących zapis odpowiedzi na pytania, takie jak: czy student ukończył studia w terminie, czy zaliczył kolejny rok studiów lub czy zdał egzamin z określonego przedmiotu itp. Dzięki takiemu rozwiązaniu możliwe będzie stosowanie bardziej zaawansowanych metod statystycznych, takich jak chociażby regresja liniowa czy regresja logistyczna, których nie można wykorzystać, gdy posługujemy się porządkowymi skalami pomiarowymi, np. skalą ocen na studiach.

Niezwykle istotną kwestią na tym etapie analizy będzie właściwy dobór egzaminów, które mają świadczyć o odniesieniu sukcesu w studiowaniu. Należy wykorzystać w pierwszej kolejności informacje o wynikach z przedmiotów obowiązkowych dla wszystkich studentów, które są sprawdzane w sposób jednolity dla całego rocznika. Tylko wtedy można mówić o porównywalności ocen poszczególnych studentów.

Niestety, oceny na studiach zdają się być w jeszcze mniejszym stopniu porównywalne niż wyniki egzaminów maturalnych. Na każdym studiach można wskazać egzaminy łatwiejsze i egzaminy trudniejsze. Nie mamy także gwarancji, że poziom trudności egzaminu nie zmienia się z roku na rok. Dlatego przynajmniej w początkowej fazie badania konieczne jest rozpatrywanie osobno ocen z poszczególnych przedmiotów w kolejnych latach. Dopiero po zaobserwowaniu prawidłowości dotyczących ich związków z wynikami maturalnymi można przystąpić do budowy ogólniejszego modelu zależności.

Przed przystąpieniem do analizy możliwości wykorzystania wyników matury lub ich funkcji do przewidywania sukcesu na studiach warto także zbadać wzajemne związki pomiędzy różnymi wskaźnikami sukcesu na studiach.

W pierwszej fazie badań prowadzonych na poziomie jednostkowym warto ograniczyć analizę tylko do wybranych jednostek dydaktycznych. Faza ta powinna bowiem służyć

przede wszystkim testowaniu i doskonaleniu przyjętej metody badawczej. Konieczne mogą być na przykład modyfikacje wynikające ze specyfiki studiów w poszczególnych jednostkach. Dopiero wtedy można myśleć o stopniowym rozszerzaniu zasięgu badania tak, aby docelowo objęło ono wszystkie jednostki dydaktyczne.

Podsumowanie

W ostatnich latach znacząco zmieniły się możliwości prowadzenia badań ewaluacyjnych procesów rekrutacji. Z jednej strony, zrewolucjonizowany został system egzaminów maturalnych oraz proces naboru na studia. Dawne egzaminy maturalne sprawdzane w szkołach oraz egzaminy wstępne organizowane przez poszczególne jednostki uczelni zastąpiły jednolite dla całego rocznika, zewnętrznie sprawdzane egzaminy dojrzałości. Z drugiej strony, zmiana w ramach administracji uczelnianej – informatyzacja procesu obsługi dydaktyki – doprowadziła do wytworzenia baz danych pozwalających przy względnie małych nakładach odtworzyć trajektorię studiów każdego studenta: jego ścieżkę edukacyjną od rekrutacji aż do zakończenia studiów.

Opisałem schemat ewaluacji procesu rekrutacji, opierający się przede wszystkim na danych pochodzących z tworzonych przez administrację baz w ramach Uniwersyteckiego Systemu Obsługi Studiów oraz systemu Internetowej Rejestracji Kandydatów. Obecnie większość wysiłków koncentruje się na drugim poziomie analizy – rekrutacji na studia. Cały czas trwają prace nad doskonaleniem narzędzi, m.in. dyskusja na temat możliwości implementacji podejścia sieciowego do analizy wzorów zachowań kandydatów. Równocześnie trwają prace nad opracowaniem i sprawdzeniem baz danych oraz doбором właściwych metod statystycznych do prowadzenia analizy na poziomie poszczególnych jednostek, na razie tylko wybranych.

Wyniki analiz powinny umożliwić odpowiedź na pytanie, czy w procesie selekcji kandydatów opieramy się na informacjach o kandydacie, na podstawie których możemy trafnie prognozować sukces w studiowaniu? Ponadto wyniki analiz dotyczące przebiegu procesu rekrutacji oraz jego efektów mogą posłużyć jako istotne uzupełnienie innych badań ewaluacyjnych prowadzonych na uczelni.

Dodatkowo, wyniki prowadzonych głównie na potrzeby uczelni wyższych badań nad stosowanymi systemami rekrutacyjnymi mogą także posłużyć jako argumenty w dyskusji na temat kształtu samych egzaminów maturalnych.

Bibliografia:

Abramczuk, Katarzyna (2008), *Raport: Rekrutacja 2008*, Warszawa: Uniwersytet Warszawski.

Batorski, Dominik (2006), *Rekrutacja na Uniwersytet Warszawski na rok 2006/2007: Analiza danych rekrutacyjnych*, Warszawa: Uniwersytet Warszawski.

Centralna Komisja Egzaminacyjna (2004), *Uwarunkowania różnicowania wyników egzaminów zewnętrznych*, „Egzamin. Biuletyn Badawczy”, 2, http://www.cke.edu.pl/images/stories/badania/biuletyn_2.pdf.

Centralna Komisja Egzaminacyjna (2007a), *Ekonomiczne oraz geopolityczne uwarunkowania różnicowania wyników egzaminów*, „Egzamin. Biuletyn Badawczy”, 11, http://www.cke.edu.pl/images/stories/badania/biul_11.pdf.

Centralna Komisja Egzaminacyjna (2007b), *Społeczne uwarunkowania różnicowania wyników egzaminów*, „Egzamin. Biuletyn Badawczy”, 12, http://www.cke.edu.pl/images/stories/badania/biul_12.pdf.

Centralna Komisja Egzaminacyjna (2007c), *Psychologiczne i biologiczne uwarunkowania różnicowania wyników egzaminów*, „Egzamin. Biuletyn Badawczy”, 13, http://www.cke.edu.pl/images/stories/badania/biul_13.pdf.

Centralna Komisja Egzaminacyjna (2009), *Informacje o wynikach egzaminu maturalnego*, <http://www.cke.edu.pl/index.php?option=content&task=view&id=247&Itemid=147>.

Ciekotowa, Krystyna, Jasiński, Ryszard (1990), *Rekrutacja i Kwalifikowanie na studia a efektywność systemu edukacyjnego*, Warszawa-Łódź: Państwowe Wydawnictwo Naukowe.

Dolata, Roman, Putkiewicz, Elżbieta, Wilkomirska, Anna (2004), *Reforma egzaminu maturalnego – oceny i rekomendacje*, Warszawa: Instytut Spraw Publicznych.

GUS (2009), *Rocznik statystyczny Rzeczypospolitej Polskiej (2008)*, Warszawa: Wydawnictwo GUS.

Jakubowski, Maciej, Pokropek, Artur (2009), *Badając Egzamin. Podejście ilościowe w badaniach edukacyjnych*, Warszawa: Centralna Komisja Egzaminacyjna, s. 8–63.

Osiński, Jan (red.)(1985), *Dobór młodzieży na stacjonarne studia wyższe*, Warszawa: Państwowe Wydawnictwo Naukowe.

Janusz, Bartłomiej, Kolonko, Józef (2001), *Analysis of the Effectiveness of Candidate Selection Procedures Used in Student Enrolment*, [w:] Krystyna Lisiecka, Józef Kolonko (red.), *Quality Assurance Systems in Higher Education. Diagnosis. Improvement 1985*, Katowice: Wydawnictwo Akademii Ekonomicznej im. Karola Adamieckiego w Katowicach.

Jasiński, Mikołaj, Izdebski, Albert, Kowalski, Michał, Zając, Tomasz, Pękala, Justyna (2008), *Raport z pierwszej edycji ogólnouniwersyteckich badań jakości kształcenia na UW*, Warszawa: Uniwersytet Warszawski.

Jasiński, Mikołaj, Izdebski, Albert, Zając, Tomasz, Kowalski, Michał (2009), *Raport z drugiej edycji ogólnouniwersyteckich badań jakości kształcenia na UW*, Warszawa: Uniwersytet Warszawski.

Kądzielska, Krystyna (1963), *Egzamin wstępny jako punkt wyjścia do prognozy postępów w nauce na studiach wyższych*, Warszawa: Międzyuczelniany Zakład Badań nad Szkolnictwem Wyższym.

Komorowska, Jadwiga, Komorowski, Zygmunt (1963), *Przydatność egzaminów wstępnych dla prognozy postępów na studiach w zakresie architektury*, Warszawa: Międzyuczelniany Zakład Badań nad Szkolnictwem Wyższym.

Konarzewski, Krzysztof (2004), *Reforma oświaty, Podstawa programowa i warunki kształcenia*, Warszawa: Instytut Spraw Publicznych.

Łobodzińska, Barbara (1964), *Przydatność szkoły średniej dla studiów wyższych*, „Życie Szkoły Wyższej”, 11.

Łuczak, Elżbieta (1985), *Progностycznaść świadectwa maturalnego i egzaminu wstępnego*, „Życie szkoły wyższej”, 11.

Matuszewska-Mysior, Danuta (1966), *Rekrutacja a postępy w studiach*, „Życie szkoły wyższej”, 6, s. 12–21.

Okręgowa Komisja Egzaminacyjna w Krakowie (2005), *Pozycja wyniku maturalnego w roku 2005*, http://www.cke.edu.pl/images/stories/Wyniki/pozycja_wyniku_matury_2005_bis.pdf.

Piotrowska, Irmina (2010), *Uniwersytet w Liczbach*, <http://www.bip.uw.edu.pl/ViewBIP?exec=newy&pdm=1&tmt=1&dok=5533>, [8.02.2010].

Pluciński, Eugeniusz, Wilczyński, Ryszard (1974), *Ocena przygotowania kandydatów na studia*, „Życie szkoły wyższej”, 6, s. 83–89.

Wiśniewski, Wiesław (1968), *Wartość prognostyczna egzaminu wstępnego*, „Dydaktyka Szkoły Wyższej”, 1, s. 35–74.

Woskowski, Jan (1962), *Wartość prognostyczna świadectwa maturalnego i egzaminu wstępnego na studia w szkołach wyższych w świetle badań przeprowadzonych na wybranych kierunkach studiów Akademii Medycznej w Łodzi i Politechniki Łódzkiej*, Warszawa: Ministerstwo Szkolnictwa Wyższego Departament Planowania i Organizacji.

Ustawa z dnia 27 lipca 2005 r., *Prawo o szkolnictwie wyższym* (Dz.U. nr 164, poz. 1365 ze zm.).

Zajęc, Tomasz (2009), *Raport: Rekrutacja 2009*, Warszawa: Uniwersytet Warszawski.

Biogramy

dr Grzegorz Bryda

Socjolog i metodolog, adiunkt w Instytucie Socjologii UJ. Jego zainteresowania koncentrują się wokół zagadnień związanych ze współczesnymi teoriami socjologicznymi, metodologią nauki i badań socjologicznych, statystyką i *data mining*, komunikowaniem i modelowaniem zjawisk społecznych, wykorzystaniem nowoczesnych technologii w badaniach empirycznych i w analizie danych jakościowych/ilościowych. W latach 2003–2008 pracował jako starszy konsultant ds. analiz statystycznych i *data mining* oraz wykładowca w Centrum Szkoleniowym w SPSS Polska. Realizował szereg projektów naukowo-badawczych, marketingowych i analitycznych (m.in. dla GUS, Ministerstwa Finansów, Philip Morris, BRE Bank czy Ministerstwa Pracy i Polityki Społecznej). Przeprowadził ponad 400 zogniskowanych wywiadów grupowych i 250 wywiadów pogłębionych. Współpracuje z organizacjami pozarządowymi jako ekspert w dziedzinie metodologii badań społecznych i analizy danych w projektach z zakresu ewaluacji i polityki społecznej. Od 2004 roku zajmuje się ewaluacją jakości kształcenia w Instytucie Socjologii UJ. Od 2008 roku pełni funkcję pełnomocnika Rektora UJ ds. ewaluacji systemu jakości kształcenia oraz członka Stałej Komisji Rektorskiej UJ ds. Jakości Kształcenia. Jest odpowiedzialny za przygotowanie merytoryczne i analityczne Uczelnianego Systemu Doskonalenia Jakości Kształcenia na Uniwersytecie Jagiellońskim. Pełni funkcję sekretarza Oddziału Krakowskiego PTS, jest też członkiem Research Network – 20 European Sociological Association.

Albert Izdebski

Absolwent Instytutu Socjologii UW. Zainteresowania badawcze koncentruje na problematyce nierówności społecznych. Od trzech lat związany z budową systemu ewaluacji jakości kształcenia na Uniwersytecie Warszawskim. Koordynator badań ilościowych w Pracowni Ewaluacji Jakości Kształcenia Uniwersytetu Warszawskiego. W badaniach edukacyjnych ogniskuje swoją uwagę na społecznych uwarunkowaniach procesu osiągnięć edukacyjnych i roli, jaką odgrywa organizacja systemu edukacji w kształtowaniu procesów selekcji w systemie szkolnym.

dr Mikołaj Jasiński

Adiunkt w Zakładzie Statystyki, Demografii i Socjologii Matematycznej IS UW. Kierownik Pracowni Ewaluacji Jakości Kształcenia, inicjator i współtwórca Systemu Ewaluacji Jakości Kształcenia na UW. Wykładowca podczas licznych szkoleń dla instytucji publicznych oraz wykładowca na studiach podyplomowych na UW z dziedziny metodologii badań społecznych, w tym ewaluacji. Jego działalność dydaktyczna i naukowa obejmuje problematykę teorii wyboru społecznego, statystyki oraz metodologii badań społecznych, m.in. ewaluacji. Jest współpracownikiem Instytutu Badań Edukacyjnych, ekspertem m.in. w Centralnym Ośrodku Doskonalenia Nauczycieli (aktualnie: Ośrodek Rozwoju Edukacji) oraz Polskiej Agencji Rozwoju Przedsiębiorczości, a także autorem ekspertyz dla Departamentu Unii Europejskiej MSZ. Uczestniczy w ogólnopolskich i międzynarodowych programach badawczych.

dr Magdalena Jelonek

Wykładowczyni w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera i w Instytucie Socjologii Uniwersytetu Jagiellońskiego. Prowadzi zajęcia z szeroko rozumianego obszaru metodologii badań społecznych, włącznie z zaawansowaną analizą danych. Szczególną uwagę w pracy naukowej poświęca zagadnieniom związanym z socjologią edukacji, jakością kształcenia, w tym z analizą produktywności szkół wyższych. Autorka licznych badań diagnozujących sytuację, a także opracowań charakteryzujących sektor akademicki w Polsce (m.in. *Ocena wpływu projektów zrealizowanych dzięki funduszom europejskim przez uczelnie wyższe z terenu województwa małopolskiego na wzrost konkurencyjności gospodarczej regionu oraz wzmocnienie potencjału sektora badawczo-naukowego w województwie*). Brała udział w wielu projektach badawczych, m.in. w pilotażowej na skalę kraju ewaluacji śródo-kresowej Działania 1.2 i 1.3 Sektorowego PO Rozwoju Zasobów Ludzkich 2004–2006, w badaniach diagnozujących sytuację w zakresie kształcenia ustawicznego w województwie małopolskim. Jest także współautorką ekspertyzy „Warunki i zakres wprowadzenia audytu efektywnościowego (ewaluacji) w polskiej administracji”, powstałej na zamówienie Ministerstwa Finansów RP.

Wojciech Przybylski

Asystent w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera, absolwent Instytutu Socjologii UJ. Stypendysta Hertie Foundation oraz German National Academic Foundation. Naukowo interesuje się kapitałem społecznym (szczególnie teorią zaufania) i rozwojem regionalnym. W Wyższej Szkole Europejskiej członek zespołu reformującego system ewaluacji jakości dydaktyki w ramach projektu rozwojowego uczelni. W latach 2008–2009 był dyrektorem Centrum Analiz Regionalnych. Współtworzył system ewaluacji placówek oświatowych Powiatu Tatrzańskiego, koordynował projekt Edukacja dla Obywateli, ewaluujący system zarządzania oświatą w mieście Krakowie. Współtwórca cyklicznego badania opinii społecznej „Omnibus Krakowa i Małopolski”. Ekspert Klubu Jagiellońskiego, sekretarz Katedry Analiz Regionalnych KJ.

Seweryn Rudnicki

Socjolog i psycholog, asystent w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera, doktorant w Instytucie Socjologii UJ, wcześniej główny analityk w Centrum Analiz Regionalnych. Interesuje się socjologią kultury (zwłaszcza socjologią ciała, socjologią internetu i socjologią codzienności), a także badaniem rozwoju regionalnego, metodologią badań społecznych, psychometrią i ewaluacją dydaktyki. Uczestnik wielu projektów badawczych, analiz i raportów, współautor cyklicznego badania „Omnibus Krakowa i Małopolski”.

Joanna Skrzyńska

Doktorantka w Zakładzie Socjologii Gospodarki, Edukacji i Metod Badań Społecznych Uniwersytetu Jagiellońskiego, stypendystka programu Leonardo da Vinci. Prowadzi zajęcia w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera i w Instytucie Socjologii Uniwersytetu Jagiellońskiego. W pracy naukowej interesuje się metodami badań społecznych i marketingowych, zagadnieniami organizacji i dydaktyki szkoły wyższej. Współautorka publikacji na temat dialogu społecznego (*Diagnoza stanu rozwoju sektorowego dialogu społecznego*) i opracowań charakteryzujących sektor akademicki w Polsce (np. *Ocena wpływu projektów zrealizowanych dzięki funduszom europejskim przez uczelnie wyższe z terenu województwa małopolskiego na wzrost konkurencyjności gospodarczej regionu oraz wzmocnienie potencjału sektora badawczo-naukowego w województwie*).

Anna Szwed

Socjolożka, wykładowczyni w Wyższej Szkole Europejskiej im. ks. Józefa Tischnera w Krakowie (WSE) w Krakowie. Koordynatorka zadaniowa ds. jakości w projekcie rozwojowym uczelni. W latach 2003–2005 kierowniczką Działu Promocji WSE. Doktorantka na Wydziale Filozoficznym Uniwersytetu Jagiellońskiego. Współpracowniczką Katedry Porównawczych Studiów Cywilizacji UJ. Przygotowuje pracę doktorską na temat obrazu kobiety w nauczaniu Kościoła rzymskokatolickiego oraz w świadomości polskich księży. Zainteresowania naukowe: socjologia kultury, socjologia religii, *gender studies*.

dr Barbara Worek

Adiunkt w Zakładzie Socjologii Gospodarki, Edukacji i Metod Badań Społecznych Instytutu Socjologii UJ i w WSE im. Józefa Tischnera w Krakowie. Jej zainteresowania naukowe koncentrują się wokół metodologii badań jakościowych, zastosowania metod jakościowych w diagnozie i ewaluacji oraz wspomaganą komputerowo analizy danych jakościowych. Uczestniczyła w wielu polskich i międzynarodowych projektach badawczych dotyczących analizy polityk publicznych, wykluczenia społecznego, partycypacji obywatelskiej i kształcenia przez całe życie. Jest autorką i współautorką publikacji dotyczących zastosowania metod jakościowych w badaniach społecznych i marketingowych oraz licznych raportów badawczych z dziedziny polityk publicznych, ekonomii społecznej, kształcenia przez całe życie. Od 2008 roku jest członkinią Stałej Rektorskiej Komisji ds. Rozwoju Dydaktyki i Jakości Kształcenia UJ.

Publikacja została przygotowana w ramach Projektu Rozwojowego „Kompetentnie ku przyszłości” współfinansowanego przez Unię Europejską w ramach Europejskiego Funduszu Społecznego Działanie 4.1. Wzmocnienie i rozwój potencjału dydaktycznego uczelni oraz zwiększenie liczby absolwentów kierunków o kluczowym znaczeniu dla gospodarki opartej na wiedzy Poddziałanie 4.1.1 Wzmocnienie potencjału dydaktycznego uczelni.



KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

 **PROJEKT ROZWOJOWY**
KOMPETENTNIE
KU PRZYSZŁOŚCI



UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY