

Czynniki wpływające na ogólną ocenę zajęć przez studentów

Autor: Katarzyna Sokołowska

Abstrakt

Głównym celem artykułu jest zbadanie, jakie czynniki wpływają na ogólną ocenę zajęć przez studentów. W badaniu wykorzystane zostały dane pochodzące z badania ankietowego „Ocena przedmiotu w Wyższych Szkołach Bankowych”, przeprowadzonego w semestrze zimowym roku akademickiego 2012/2013, w którym ocenie poddano 1802 zajęcia. Do zidentyfikowania czynników wpływających na ogólną ocenę zajęć wykorzystano model logitowy. Praca zawiera również przegląd angielskojęzycznych artykułów dotyczących ewaluacji zajęć i uzyskanych w nich wyników.

Słowa kluczowe: ewaluacja zajęć, model logitowy

JEL: C25

Wprowadzenie

Oceny studenckie są początkiem drogi wykładowcy w kierunku doskonalenia się, a nie jej końcem. William Cashin (1995)

Ewaluacja zajęć przez studentów jest jednym z najczęściej stosowanych sposobów oceny jakości nauczania (Wright, 2006). Ocena dokonywana w ten sposób nie jest zjawiskiem nowym: pierwsza skala oceny nauczyciela została opublikowana w 1915 roku (Dalton, Denson, 2009). Głównym celem ewaluacji zajęć przez studentów jest uzyskanie informacji zwrotnej dla nauczyciela dotyczącej jakości jego zajęć i nauczania, tak aby umożliwić mu podjęcie działań prowadzących do poprawy wyników nauczania i zapewnić odpowiedni poziom szkoleń. Drugim celem jest zmierzenie efektywności nauczania, a następnie wykorzystanie wyników pomiaru do podejmowania decyzji kadrowych, czyli dotyczących awansu, wysokości wynagrodzenia czy warunków zatrudnienia (Beran i in., 2005). Kolejnym celem, w przypadku gdy wyniki badań są udostępniane studentom, jest ułatwienie im wyboru zajęć oraz osób prowadzących, jeśli na uczelni istnieje taka możliwość. Powstaje

podstawowe pytanie: czy można na podstawie przeprowadzonych badań ankietowych podejmować właściwe decyzje, mogące mieć wpływ na przebieg kariery akademickiej nauczyciela i ocenę wydziału? Niektórzy krytycy oceniają jakość nauczania nauczycieli akademickich przez studentów uważają, że studenci nie powinni brać udziału w ocenie swoich nauczycieli. Jednak z większości badań wynika, że studenci mogą dostarczyć cennych informacji na temat wielu aspektów nauczania, które są przydatne w realizacji celów związanych z doskonaleniem nauczycieli. Niewłaściwe jest natomiast nadużywanie danych z ocen studenckich, na przykład w interpretowaniu różnic, które są wyrażane w wielkościach dziesiętnych (Abrami, 2001), czy wykorzystaniem otrzymanych na ich podstawie ratingów jako jedyne źródła dowodów skuteczności nauczania. Uważa się, że wiedza wykładowcy, jakość programu nauczania, poziom nauczania czy ilość przekazywanych treści należą do tych obszarów, które nie powinny być oceniane przez studentów (Coren, 2001; Theall i Franklin, 2001) ze względu na ich ograniczone doświadczenie i wiedzę w danej dziedzinie. Jednak na pytania dotyczące tego, jak wiele

można się nauczyć na danych zajęciach lub na temat umiejętności przekazywania przez prowadzącego wiedzy czy wymaganych nakładów pracy, nikt nie jest w stanie odpowiedzieć równie dobrze jak studenci, ponieważ to oni są najbliższym procesowi nauczania. Powstaje jednak pytanie, czy odpowiedzi te nie są zniekształcone przez nastawienie do danego przedmiotu lub inne czynniki, takie jak wielkość grupy, poziom studentów czy forma zajęć. W literaturze przedmiotu można znaleźć wiele badań, których wyniki zostały przedstawione poniżej.

Wyróżnia się szereg czynników, które mogą mieć wpływ na ogólną ocenę przedmiotu przez studenta, jednak wpływ ten, nawet jeśli jest statystycznie istotny, zwykle jest bardzo niewielki (Gravestock, Gregor-Greenleaf, 2008). Czynniki te można podzielić według następujących kryteriów (podano za Gravestock, Gregor-Greenleaf, (2008) i na podstawie własnego przeglądu literatury):

Warunki administracyjne:

1. Czas przeprowadzania oceny – jeśli ocena jest dokonywana po egzaminach czy kolokwium, wyniki są niższe, w związku z tym zaleca się, aby ewaluację realizować przed ostatnimi egzaminami i wystawieniem końcowych ocen.
2. Instrukcje przekazane studentom – zauważono, że wskaźniki mogą być lepsze, gdy prosi się uczniów o bardziej krytyczne opinie lub gdy informuje się, że uzyskane oceny będą wykorzystywane do podejmowania decyzji kadrowych.
3. Anonimowość – w przypadku prośby o podpisanie ankiety studenci obawiają się konsekwencji, dlatego wystawiają wówczas wyższe oceny.
4. Obecność wykładowcy – niektóre badania podają, że w przypadku obecności wykładowcy mogą być wystawiane wyższe oceny, inne, że jego obecność wpływa na ocenę tylko wtedy, gdy ankieta nie jest anonimowa (Algozzine i in., 2004).

Cechy zajęć, które nie mogą być kontrolowane przez wykładowcę:

1. Wielkość grupy – niektóre badania wskazują, że w mniej licznych grupach są wystawiane wyższe oceny (Aleamoni, Hexner, 1980; Koh, Tan, 1997), jednak w innych badaniach reprezentowany jest pogląd, że różnica ta jest nieistotna. Arreola (2007) i Feldman (1984) zauważają jednak, że większość zajęć

prowadzonych w dużych grupach jest prowadzona dla studentów I lub II roku i są to przedmioty obowiązkowe, a nie opcyjne, i to może mieć większy wpływ na niższe oceny niż wielkość grupy.

2. Kurs kierunkowy (do wyboru) a kurs obowiązkowy – przedmiotom ogólnym (obowiązkowym) studenci wystawiają nieco niższe oceny niż kierunkowym, lecz nie jest to statystycznie istotna różnica (Cashin, 1995; March, Dunkin, 1992). Wachtel (1998) twierdzi, że istotny wpływ na ocenę ma to, czy studenci są zainteresowani zajęciami, czyli czy są one związane z ich kierunkiem studiów i zainteresowaniami.
3. Trudność kursu (nakład pracy) – uzyskano niespójne wyniki: część badaczy uważa, że trudniejsze zajęcia uzyskują niższe oceny, inni tego nie potwierdzają (Cashin, 1988).
4. Poziom kursu – kursy bardziej zaawansowane uzyskują wyższe oceny, ale wynik jest niejednoznaczny (Algozzine i in. 2004; Cashin, 1995).
5. Dziedzina – niektóre badania wykazały, że najbardziej pozytywne wyniki otrzymują nauki humanistyczne, następnie społeczne, a na końcu przyrodnicze, ale uważa się również, że porównania między różnymi dziedzinami mogą nie być dokładne (Neumann, 2001; Cashin, 1988, 1990, 1995). Przypuszcza się, że nauki humanistyczne mogą otrzymywać wyższe oceny niż nauki ścisłe, ale przeprowadzono do tej pory zbyt mało badań, aby potwierdzić tę zależność.

Cechy wykładowcy

1. Wiek, płeć – wpływ na ocenę jest nieistotny (Algozzine i in., 2004).
2. Osobowość – jeśli osobowość mierzy pewien sposób zachowania wykładowcy w trakcie zajęć, na przykład entuzjazm, wyrazistość, ekspresyjność, to może mieć to wpływ na skuteczność nauczania (Ory, 2001). Do podobnych wniosków doszli Williams i Ceci (1997), którzy w swoich badaniach wykazali, że instruktor, którego sposób nauczania był bardziej entuzjastyczny, osiągnął wyższe oceny, nawet jeśli inne aspekty pozostały bez zmian (na podstawie eksperymentu z wykładowcą, który zmienił swój styl nauczania, nie zmieniając przekazywanej wiedzy i materiałów). Natomiast Marsh i Roche (1997) uważają, że jeżeli ktoś jest zabawny i pełen entuzjazmu, a nie jest biegły w przekazywaniu wiedzy, to studenci w swoich ocenach będą potrafili to

zróżnicować, dając wyższe oceny związane z entuzjazmem, a niższe z umiejętnością przekazywania wiedzy. McKeachie (1997), opierając się na pracy Scrivena (1981), twierdzi, że styl nauczania (entuzjazm, sposób organizacji zajęć, okazywane ciepło) nie powinien być oceniany, gdyż wysoka efektywność nauczania może być osiągnięta na wiele sposobów. Korzystanie z cech, które na ogół pozytywnie korelują z efektywnością nauczania, karze nauczycieli, którzy są skuteczni, pomimo iż osiągają w którymś z tych aspektów gorsze wyniki.

3. stopień naukowy i doświadczenie – wpływ na ocenę jest nieistotny, zwykle profesorem nie uzyskują wyższych ocen (Arreola, 2007).

Cechy studenta

1. Wiek – im starszy uczeń, tym wyższe wystawia oceny (Nasser i Hagtvet, 2006).
2. Płeć – istnieje tendencja do oceniania wykładowców tej samej płci nieco wyżej (Ory, 2001).
3. Rok studiów – Arreola (2007) i Marsh (1987) stwierdzili, że istotny wpływ na ocenę wykładowcy ma to, na którym roku studiują osoby oceniające – im wyższy rok, tym mniej surowe są oceny. Jest to prawdopodobnie odzwierciedleniem faktu, że uczniowie po kilku latach nauki mają bardziej realistyczne oczekiwania w zakresie swoich doświadczeń akademickich.
4. Motywacja – im wyższą motywację do nauki danego przedmiotu ma student, tym wyższe wystawia oceny (Cashin, 1988, 1995).
5. Frekwencja – jest pozytywnie skorelowana z wyższymi ocenami, ale nie jest postrzegana jako coś, co może zniekształcić wyniki ewaluacji, gdyż wyższa frekwencja na zajęciach prowadzi do poprawy wyników nauczania, a one mają wpływ na wyższą ocenę wykładowcy (Davidovitch & Soen, 2006).
6. Uzyskiwane stopnie – zauważono, że jeśli studenci spodziewają się, że wykładowca będzie pobłażliwy w trakcie oceniania, to mogą wystawiać mu wyższe oceny (Wachtel, 1998). McKeachie (1997) twierdzi natomiast, że studenci najprawdopodobniej nie będą wyżej oceniać nieskutecznego nauczyciela, który stara się „kupić” ich dobre oceny łagodnym traktowaniem, jednak część badaczy uważa, że studenci mają skłonność do karania nauczycieli za wystawianie im niskich ocen (Crumbley, Henry i Kratchman, 2001). Heckert i in. (2006) na podstawie

swoich badań doszli natomiast do wniosku, że bardziej wymagający wykładowcy uzyskiwali wyższe oceny. Podobnie Marsh (1987) stwierdził, że wyższy poziom trudności zajęć i zajęcia wymagające więcej wysiłku oraz czasu są oceniane przychylniej niż pozostałe zajęcia. Nie potwierdziły tego inne badania, np. Pounder (2007) wykazał, że trudniejsze kursy uzyskują zwykle niższe oceny. Inni badacze uważają, że studenci porównują swój wkład, wysiłek, obciążenie pracą na danych zajęciach w stosunku do innych zajęć. Jeśli są one podobne, a uzyskana ocena jest wyższa, to wykładowca zostaje wyżej oceniony. Można również znaleźć wnioski mówiące o tym, że ci, dla których zajęcia były trudniejsze niż przewidywali, oceniali je mniej korzystnie, natomiast studenci, dla których były łatwiejsze, niż oczekiwali, oceniali je bardziej pozytywnie. Katiliūtė (2010) doszedł do wniosku, że studenci uzyskujący lepsze oceny z danego przedmiotu wyżej oceniali wykładowcę, dlatego zaproponował prowadzenie ewaluacji w formie elektronicznej, aby pomóc wychwycić tego rodzaju wpływ przez połączenie ocen uzyskiwanych przez studentów z danego przedmiotu z ocenami wystawianymi przez nich wykładowcom. Zaproponował, aby w analogiczny sposób wychwycić wpływ uzyskanych przez studentów ocen na ocenę kursu jako trudnego/**łatwego**.

Metodyka badań. Dobór zmiennych

Celem artykułu jest sprawdzenie, jakie czynniki mają wpływ na ocenę zajęć dokonywaną przez studentów Wyższych Szkół Bankowych. Podmiotem badawczym byli studenci, którzy korzystali w tych szkołach z usługi edukacyjnej w semestrze zimowym roku akademickiego 2012/2013. Do analizy wykorzystano dane pochodzące z trzech szkół. Badanie przeprowadzono na ostatnich zajęciach studiów stacjonarnych i niestacjonarnych, w każdej z grup, wręczając każdemu studentowi obecnemu na zajęciach do wypełnienia kwestionariusz. W rezultacie otrzymano 60 421 wypełnionych ankiet. Ankieta składała się z następujących 7 pytań dotyczących oceny prowadzonych zajęć:

Przyjęta skala ocen: 1 – *zdecydowanie nie*, 2 – *nie*, 3 – *ani tak, ani nie*, 4 – *tak*, 5 – *zdecydowanie tak*.

Pytanie 1: Zajęcia były prowadzone zgodnie z programem.

Pytanie 2: Wymagania dotyczące zajęć i zaliczenia były jasno sformułowane.

Pytanie 3: Zajęcia prowadzone były w zrozumiałym sposobie.

Pytanie 4: Zajęcia prowadzone były w interesujący sposób.

Pytanie 5: Na tych zajęciach dużo się dowiedziałem.

Pytanie 6: Prowadzący traktował studentów z szacunkiem.

Pytanie 7: Ogólnie jestem zadowolony z zajęć.

Następnie dla każdego ocenianego zajęcia obliczana była średnia arytmetyczna odpowiedzi wszystkich studentów obecnych na danych zajęciach na każde pytanie. Średnia uzyskana ze wszystkich pytań przez wszystkich nauczycieli akademickich wyniosła w ocenianym okresie 4,27. Dla autorki artykułu dostępne były tylko dane w postaci wyliczonych średnich. Mimo, że

zastosowana w ankiecie skala Likerta zalicza się do skal porządkowych, to w podejściu liberalnym zakłada się, że jeśli skala jest przynajmniej pięciopunktowa, to można już zastosować techniki analityczne przewidziane dla skali przedziałowej. W pracy Elliota i Woodwarda (2007) można znaleźć informację, że skala Likerta może być traktowana jako ilościowa, jeżeli można racjonalnie założyć, że różnice między każdymi kolejnymi kategoriami są równe. Strukturę ocenianych zajęć ze względu na przydatne w badaniu cechy przedstawia tabela 1.

W badaniu jako zmienną objaśnianą wykorzystano zmienną opartą na ocenie stwierdzenia: *Ogólnie jestem zadowolony z zajęć*. Ze względu na to, że początkowo była ona zmienną mierzoną na skali porządkowej (od 1 – *zdecydowanie nie* do 5 – *zdecydowanie tak*), natomiast po obliczeniu średniej stała się zmienną ciągłą przyjmującą wartości z przedziału $<1,5;5>$, postanowiono powrócić do skali porządkowej przez przypisanie wartościom tej zmiennej następujących kategorii:

Tabela 1. Struktura ocenianych zajęć według wybranych cech w semestrze zimowym roku akademickiego 2012/2013 (N=1802)

Podział	Podział	Podział	Podział
Forma zajęć	% odpowiedzi	Liczba przedmiotów, dla których liczba udzielonych odpowiedzi mieściła się w poniższym przedziale	Liczba przedmiotów w %
ćwiczenia	28,86	(0,20>	47,38
konwersatorium	3,86	(20,40>	22,99
laboratorium	5,56	(40,60>	12,35
lektorat	13,12	(60,90>	9,26
projekt	0,46	(90,120>	4,01
seminarium	7,25	(120,150>	1,54
wykład	40,90	(150,180>	1,39
Specyfika przedmiotu I	Liczba przedmiotów w %	(180,210>	0,77
przedmiot ścisły	23,37	(210,240>	0,30
przedmiot humanistyczny	76,63		
Specyfika przedmiotu II	Liczba przedmiotów w %		
kierunkowy	56,57		
ogólny	43,43		

Źródło: opracowanie własne na podstawie danych WSB

Tabela 2. Wartości zmiennej *Ogólnie jestem zadowolony z zajęć* i przypisane im kategorie

Ocena z przedziału	Przypisana kategoria	Ocena z przedziału	Przypisana kategoria
$<1,5;2>$	1	$(3,5;4>$	5
$(2;2,5>$	2	$(4;4,5>$	6
$(2,5;3>$	3	$(4,5;5>$	7
$(3;3,5>$	4		

Źródło: opracowanie własne na podstawie danych WSB

Stworzenie większej liczby kategorii było podyktowane tym, aby zachować jak najwięcej informacji o zmiennej objaśnianej, która powstała, gdy dla każdego zajęcia obliczono średnią ocen.

W przypadku, gdy zmienna objaśniana jest zmienną porządkową, najwłaściwszym sposobem wyjaśniania przyczyn jej zmienności jest zastosowanie modeli uporządkowanej zmiennej objaśnianej. Najczęściej przyjmuje się, że modele takie mają postać modelu probitowego (wtedy zakłada się, że dystrybuanta składnika losowego jest dystrybuantą standaryzowanego rozkładu normalnego) lub modelu logitowego (wówczas jest to dystrybuanta rozkładu logistycznego). Konstrukcję i estymację tego typu modeli opisano m.in. w pracach Cramera (2001, 2003), Greene'a (2003, 2009) i Gruszczyńskiego (2001, 2010).

W niniejszej pracy wykorzystano model logitowy dla kategorii uporządkowanych. Model ten opiera się na założeniu, że zmienna objaśniana y^* jest pewną nieobserwowalną zmienną ciągłą, która jest liniową funkcją zmiennych objaśniających X zapisanych w wektorze x oraz nieznanymi parametrami, zapisanych w wektorze β :

$$y_i^* = x_i' \beta + u_i'$$

gdzie: $i=1,2,\dots,1802$ oznaczają kolejne ankietowane zajęcia,

$$x_i' \beta = (1 \ X_{1i} \ X_{2i} \ \dots \ X_{ki}) (\beta_0 \ \beta_1 \ \beta_2 \ \dots \ \beta_k)'$$

a u_i jest składnikiem losowym o rozkładzie logistycznym (Gruszczyński 2010, s.108).

Wartość uporządkowanej zmiennej objaśnianej y_i zależy od wartości zmiennej nieobserwowalnej y_i^* zgodnie z regułą:

$$y_i = j \Leftrightarrow k_{j-1} < y_i^* \leq k_j \quad (j = 1, \dots, 7),$$

gdzie k_j , zwane progami, są zwykle nieznanymi i estymuje się je razem z parametrami β . Ponieważ y^* jako zmienna nieobserwowalna może przyjmować dowolne wartości rzeczywiste, można założyć, że $k_0 = -\infty, k_7 = +\infty$, dzięki czemu liczba progów wymagających oszacowania ogranicza się do 6 (Gruszczyński, 2010). Prawdopodobieństwo p_j , że dla i -tego ocenianego zajęcia zmienna y przyjmie wartość j , można zapisać jako:

$$\begin{aligned} p_j &= P(y_i = j) = \\ &= P(k_{j-1} - x_i' \beta < u_i \leq k_j - x_i' \beta) = \\ &= P(u_i \leq k_j - x_i' \beta) - P(u_i < k_{j-1} - x_i' \beta) = \\ &= F(k_j - x_i' \beta) - F(k_{j-1} - x_i' \beta) \end{aligned}$$

gdzie F jest dystrybuantą składnika losowego u_i .

W przypadku logistycznego rozkładu składnika losowego wartości dystrybuanty oblicza się według formuły:

$$F(u) = \frac{\exp(u)}{1 + \exp(u)}$$

a zatem:

$$p_j = \frac{\exp(k_j - x_i' \beta)}{1 + \exp(k_j - x_i' \beta)} - \frac{\exp(k_{j-1} - x_i' \beta)}{1 + \exp(k_{j-1} - x_i' \beta)}$$

Otrzymany model nazywa się modelem logitowym dla kategorii uporządkowanych.

Parametry tego modelu szacuje się za pomocą metody największej wiarygodności (MNV) poprzez maksymalizację logarytmu funkcji wiarygodności:

$$\begin{aligned} \text{Log}L(y|x; \beta, k_1, \dots, k_6) &= \\ &= \sum_{i=1}^{648} \sum_{j=1}^7 d_j \log \left[\frac{\exp(k_j - x_i' \beta)}{1 + \exp(k_j - x_i' \beta)} - \frac{\exp(k_{j-1} - x_i' \beta)}{1 + \exp(k_{j-1} - x_i' \beta)} \right] \end{aligned}$$

gdzie d_j przyjmuje wartość 1, gdy w próbie zmienna y_i przyjęła wartość j i 0 w pozostałych przypadkach.

Ocenę dopasowania modeli przeprowadzono z wykorzystaniem następujących miar [Gruszczyński 2010]:

Współczynnik McFadden's Adj R2 (skorygowany pseudo-R2 McFaddena) wyrażony wzorem

$$R_{\alpha F}^2 = 1 - \frac{\ln L - p}{\ln L_0}$$

gdzie L , L_0 - oznaczają odpowiednio wartość funkcji wiarygodności badanego modelu i wartość funkcji wiarygodności dla modelu uwzględniającego jedynie progę, p jest liczbą szacowanych parametrów z wyłączeniem progów. Miara ta jest mniejsza od 1 i może przyjąć wartości ujemne. Analizowany model charakteryzuje się dobrym dopasowaniem do danych empirycznych, jeśli jej wartość jest bliska 1.

Test ilorazu wiarygodności, w którym wartość statystyki dana jest wzorem $LR = 2(\ln L - \ln L_0)$. Statystyka ta ma rozkład chi-kwadrat z liczbą stopni swobody równą p . Hipotezę zerową odrzuca się w przypadku wartości statystyki wyższych od wartości granicznej. Najmniejszy poziom istotności, przy którym można odrzucić hipotezę zerową, jest oznaczony jako p -value. Test ten zastosowano do badania łącznej istotności wszystkich zmiennych objaśniających. W tym przypadku hipoteza zerowa zakładała, że wszystkie parametry modelu oprócz progów są równe zero.

Wyniki badania

W dalszej części artykułu zbadano zależność pomiędzy wybranymi zmiennymi objaśniającymi, których zestawienie zawiera tabela 3, a zmienną objaśnianą opartą na ocenie stwierdzenia *Ogólnie jestem zadowolony z zajęć*. Wartości przyjmowane przez tę zmienną zostały zawarte w tabeli 2. Doboru zmiennych objaśniających do modelu dokonano na podstawie przeglądu literatury i był on uzależniony od danych udostępnionych autorce. Ze względu na sposób, w jaki były sformułowane oceniane przez studentów stwierdzenia, wyższa ocena (czyli im bardziej studenci zgadzali się z danym stwierdzeniem) oznaczała, że dane zajęcia były w sensie pozytywnym wyżej oceniane. Zmienne o numerach 1 i 2 zostały stworzone arbitralnie przez autorkę artykułu. Dla zmiennej 1 badane zajęcia podzielono na dwie grupy: na przedmioty, w których używa się wzorów i dokonuje obliczeń, nazywając tę grupę przedmiotami

ściłymi, oraz na grupę przedmiotów, w których wzory i obliczenia pojawiają się sporadycznie, nazywając je przedmiotami humanistycznymi. W przypadku zmiennej o numerze 2 dokonano arbitralnego podziału ocenianych przedmiotów na dwie grupy, w zależności od tego, czy są to przedmioty ogólne, takie jak na przykład matematyka, psychologia, socjologia i makroekonomia, czy związane z konkretnym kierunkiem studiów. Do drugiej grupy zaliczono również przedmioty do wyboru, seminaria dyplomowe i magisterskie oraz lektorat języka angielskiego, zakładając, że przedmioty te powinny cieszyć się większym uznaniem studentów jako te, które w przyszłości według nich mogą im się przydać w pracy zawodowej. I tak na przykład finanse przedsiębiorstw dla kierunku Finanse i rachunkowość zaliczono do pierwszej grupy, a dla kierunku Zarządzanie do drugiej grupy przedmiotów.

Charakterystyki liczbowe dotyczące poszczególnych zmiennych przedstawia tabela 4.

Tabela 3. Zmienne objaśniające wykorzystane w modelu

Nr zmiennej	Nazwa zmiennej	Sposób pomiaru
1	Specyfika przedmiotu I	1 – przedmiot ścisły, 2 – przedmiot humanistyczny
2	Specyfika przedmiotu II	1 – przedmiot ogólny, 2 – przedmiot kierunkowy
3	Zajęcia były prowadzone zgodnie z programem.	Wartości zmiennych są liczbami rzeczywistymi z przedziału $\langle 1;5 \rangle$, będącymi średnimi arytmetycznymi z ocen wystawionych przez studentów dla danego przedmiotu na skali od 1 (zdecydowanie nie) do 5 (zdecydowanie tak).
4	Wymagania dotyczące zajęć i zaliczenia były jasno sformułowane.	
5	Zajęcia prowadzone były w zrozumiały sposób.	
6	Zajęcia prowadzone były w interesujący sposób.	
7	Na tych zajęciach dużo się dowiedziałem.	
8	Prowadzący traktował studentów z szacunkiem.	

Źródło: opracowanie własne na podstawie danych WSB

Tabela 4. Charakterystyki liczbowe zmiennych objaśniających

Nr zm.		Średnia ogółem	Odchylenie standardowe	Średnia dla przedmiotów ścisłych	Średnia dla przedmiotów humanistycznych	Średnia dla przedmiotów ogólnych	Średnia dla przedmiotów kierunkowych
3	Zajęcia były prowadzone zgodnie z programem.	4,547	0,395	4,506	4,558	4,497	4,584
4	Wymagania dotyczące zajęć i zaliczenia były jasno sformułowane.	4,462	0,481	4,363	4,491	4,392	4,515
5	Zajęcia prowadzone były w zrozumiały sposób.	4,369	0,559	4,159	4,432	4,308	4,416
6	Zajęcia prowadzone były w interesujący sposób.	4,266	0,610	4,060	4,331	4,188	4,326
7	Na tych zajęciach dużo się dowiedziałem.	4,257	0,567	4,078	4,313	4,184	4,313
8	Prowadzący traktował studentów z szacunkiem.	4,672	0,368	4,623	4,687	4,659	4,682

Źródło: opracowanie własne na podstawie danych WSB

Tabela 5. Średnia i odchylenie standardowe dla wybranych zmiennych z podziałem na 3 wyróżnione grupy przedmiotów.

Nazwa zmiennej	Przedmiot					
	ściśle i ogólny		mieszany		kierunkowy i humanistyczny	
	Średnia	Odchylenie standardowe	Średnia	Odchylenie standardowe	Średnia	Odchylenie standardowe
Zajęcia były prowadzone zgodnie z programem.	4,492	0,299	4,506	0,434	4,595	0,380
Wymagania dotyczące zajęć i zaliczenia były jasno sformułowane.	4,350	0,438	4,405	0,519	4,541	0,449
Zajęcia prowadzone były w zrozumiały sposób.	4,203	0,516	4,293	0,604	4,479	0,510
Zajęcia prowadzone były w interesujący sposób.	4,055	0,579	4,205	0,642	4,382	0,568
Na tych zajęciach dużo się dowiedziałem.	4,086	0,556	4,190	0,606	4,365	0,514
Prowadzący traktował studentów z szacunkiem.	4,654	0,232	4,641	0,385	4,703	0,385
Ogólnie jestem zadowolony z zajęć.	4,214	0,490	4,298	0,574	4,450	0,522

Źródło: opracowanie własne na podstawie danych WSB

Ze wszystkich elementów składających się na całościową ocenę nauczyciela akademickiego najwyżej oceniany przez ankietowanych był stosunek wykładowcy do studenta (zm. 8) oraz organizacja zajęć (zm. 3 i 4). Najniżej studenci ocenili sposób przekazywania wiedzy przez nauczycieli akademickich (zm. 5, 6 i 7) i w tym zakresie należałoby rozważyć wprowadzenie stosownych działań korygujących. Można też zauważyć, że średnia ocena wystawiana przedmiotom ścisłym była niższa niż ocena wystawiana przedmiotom humanistycznym, niezależnie od rozważanej zmiennej. Podobnie przedmioty ogólne były oceniane niżej niż kierunkowe. Były to jednak różnice w wielkościach dziesiętnych, więc jak już zauważono wcześniej, powinno się zachować ostrożność w interpretacji tego typu wyników. Pomimo to wydaje się, że nałożenie się dwóch czynników, a więc tego, że przedmiot jest ścisły, a jednocześnie ogólny, może mieć już bardziej znaczący wpływ na ogólną ocenę wystawianą wykładowcy. Aby to sprawdzić, wszystkie zajęcia podzielono na 3 grupy. Do grupy pierwszej zaliczono te przedmioty, które wcześniej były sklasyfikowane jako ścisłe i ogólne, do drugiej – te, które były sklasyfikowane jako ścisłe i kierunkowe lub humanistyczne i ogólne, a do trzeciej – przedmioty humanistyczne i kierunkowe. W ostatnim wierszu dołączono zmienną *Ogólnie jestem zadowolony z zajęć*, przed podzieleniem jej na kategorie przedstawione w tabeli 2.

Jak można zauważyć, średnia ocena przedmiotu wzrasta wraz z przejściem od 1. do 3. grupy przedmiotów. Największe różnice występują dla zmiennej *Zajęcia prowadzone były w interesujący sposób*, a zaraz potem dla zmiennych *Na tych zajęciach dużo się dowiedziałem* i *Zajęcia prowadzone były w zrozumiały sposób*. Jednak biorąc pod uwagę odchylenia standardowe, należy zachować ostrożność w ich interpretowaniu. Wprowadzony podział spowodował, że średnie oceny dla przedmiotów ścisłych i ogólnych stały się nieco wyższe niż dla wszystkich przedmiotów ścisłych, poza tym w porównaniu ze średnią obliczoną osobno dla przedmiotów kierunkowych i przedmiotów humanistycznych wzrosły również średnie oceny dla przedmiotów jednocześnie kierunkowych i humanistycznych. Jedynymi średnimi, które były mniejsze, były średnie dla przedmiotów zakwalifikowanych równocześnie jako ścisłe i ogólne w porównaniu ze średnimi dla przedmiotów ogólnych. Można zatem wnioskować, że jeżeli przedmiot jest jednocześnie kierunkowy i humanistyczny, to następuje pozytywne połączenie tych dwóch czynników, co wpływa na wzrost średniej oceny. Natomiast sklasyfikowanie przedmiotu jako ścisły zaniża ocenę przedmiotów należących do grupy ogólnych. Na podstawie dokonanych analiz zdecydowano się na przedstawienie dwóch modeli, ze względu na to, że dołączenie do modelu zmiennych od 3–8 powodowało, że zmienne 1 i 2 stawały się nieistotne, a wydaje się, że zależność między tymi zmiennymi a zmienną objaśnianą jest

Tabela 6. Wyniki estymacji modelu logitowego zmiennej uporządkowanej dla zmiennej objaśnianej będącej oceną zadowolenia z zajęć.

Model 1			Model 2		
Nazwa zmiennej	Ocena	Błąd stand.	Nazwa zmiennej	Ocena	Błąd stand.
W.wolny 1	-4,936***	0,5024	W.wolny 1	-29,872***	2,7317
W.wolny 2	-4,371***	0,3813	W.wolny 2	-36,696***	2,9906
W.wolny 3	-3,299***	0,2295	W.wolny 3	-44,919***	3,3748
W.wolny 4	-2,517***	0,1631	W.wolny 4	-49,159***	3,5959
W.wolny 5	-1,227***	0,1057	W.wolny 5	-55,991***	3,9939
W.wolny 6	0,420***	0,0936	W.wolny 6	-62,863***	4,4006
1. Specyfika przedmiotu I (przedmiot ścisły)	0,377***	0,0876	4. Wymagania dotyczące zajęć i zaliczenia były jasno sformułowane	1,331***	0,4491
2. Specyfika przedmiotu II (przedmiot ogólny)	0,155**	0,0755	5. Zajęcia prowadzone były w zrozumiały sposób	3,833***	0,6493
			6. Zajęcia prowadzone były w interesujący sposób	1,909***	0,5478
			7. Na tych zajęciach dużo się dowiedziałem	4,551***	0,5718
			8. Prowadzący traktował studentów z szacunkiem	2,246***	0,5883
N	1802		N	1802	
LR	27,32516		LR	942,90243	
p-value	0,000		p-value	0,000	
R2Mcfaddena	0,01699		R2Mcfaddena	0,58712	

* $0.05 \leq p < 0.1$; ** $0.01 \leq p < 0.005$; *** $p < 0.01$

Źródło: Obliczenia własne z wykorzystaniem programu STATISTICA 8.0 na podstawie danych WSB

warta omówienia. Szacunki parametrów otrzymanych modeli przedstawiono w tabeli 6.

Test ilorazu wiarygodności dla otrzymanych modeli pozwolił w obu przypadkach na odrzucenie hipotezy zerowej o łącznej nieistotności wszystkich zmiennych objaśniających na poziomie $p=0,0000$. Wartość skorygowanego współczynnika pseudo-R² McFaddena dla modelu z 2 zmiennymi objaśniającymi wyniosła 0,01699, a dla modelu z 5 zmiennymi objaśniającymi – 0,58712, a więc model drugi jest dużo lepiej dopasowany do danych empirycznych użytych w jego estymacji. Dla obu modeli parametry statystycznie istotnie różnią się od zera.

Na podstawie ocen parametrów modeli można stwierdzić, że na ogólną ocenę zadowolenia z zajęć przez studenta wpływ mają następujące czynniki:

- Jest to przede wszystkim wiedza przekazywana na zajęciach studentom (zm. 7) oraz sposób jej przekazania, ale bardziej chodzi

tu o umiejętność zrozumiałego (zm. 5) niż interesującego (zm. 6) przekazu. Większy niż w przypadku ostatniego z wymienionych czynników wpływ na ogólną ocenę ma dla studentów to, że prowadzący zajęcia traktuje ich z szacunkiem. Najniższy wpływ ma zaś to, że wymagania dotyczące zajęć i zaliczenia są jasno sformułowane (zm. 4). Wszystkie omówione stwierdzenia dotyczące oceny zajęć były pozytywnie skorelowane z ogólną oceną zadowolenia z zajęć i wszystkie parametry stojące przy zmiennych odpowiadających tym pytaniom były istotne statystycznie na poziomie $p < 0,01$.

- Jeśli chodzi o stwierdzenia przedstawione studentom, to jedynym czynnikiem, który okazał się statystycznie nieistotny i został wykluczony z modelu, było *Zajęcia były prowadzone zgodnie z programem* (zm.3).
- To, że oceniane zajęcia należały do grupy przedmiotów humanistycznych, zwiększało

prawdopodobieństwo uzyskania wyższej ogólnej oceny (zm. 1), podobnie prawdopodobieństwo to było wyższe dla zajęć z przedmiotów kierunkowych w porównaniu z przedmiotami ogólnymi (zm. 2). Jednak większy wpływ na ogólną ocenę zajęć ma rozróżnienie, czy przedmiot jest ścisły czy humanistyczny, niż to, że jest ogólny czy kierunkowy.

Model 2 uzyskał zdecydowanie wyższy skorygowany współczynnik pseudo-R² McFaddena niż model 1, co oznacza, że czynniki charakteryzujące kurs, takie jak rodzaj kursu (ściśły/humanistyczny, ogólny/kierunkowy) mają jednak niewielki, choć istotny statystycznie wpływ na ogólną ocenę kursu w porównaniu z czynnikami, które są mierzone za pomocą narzędzia wykorzystywanego przez WSB. Oznacza to także, że stają się one nieistotne w przypadku wzięcia pod uwagę wyżej wymienionych czynników. Podobne wyniki uzyskano w Dalton, H. & Denson, N. (2009), gdzie indywidualne cechy kursu (takie jak powód jego wyboru, charakterystyki studenta) miały niski lub nieistotny wpływ na zmienną objaśnianą, natomiast największy procent jej wariacji (67%) wyjaśniały zadawane studentom pytania dotyczące sposobu prowadzenia kursu.

Umiejętność skutecznego komunikowania się od dawna przyjęta jest jako standard efektywnego nauczania. Uzyskane wyniki sugerują, że zdolność ta jest podstawą pozytywnego postrzegania wykładowcy przez studentów. Okazało się również, że największy wpływ na ocenę wystawianą przez studentów ma stosunek do studentów, a nie przestrzeganie wymagań formalnych, takich jak zgodność zajęć z programem czy jasne sformułowanie wymagań dotyczących zajęć i zaliczenia, choć ta kwestia nie była całkiem bez znaczenia.

Autorka sprawdziła również, czy na ogólne zadowolenie ma wpływ forma zajęć, jednak czynnik ten okazał się nieistotny statystycznie. Jeśli chodzi o liczbę osób obecnych na zajęciach, to ze względu na to, że w przypadku innych niż wykład form zajęć liczby osób z kilku grup były sumowane, gdy zajęcia prowadził jeden wykładowca, nie było możliwe zbadanie tego wpływu. Interesującym byłoby sprawdzenie, czy uzyskane wyniki różnią się w zależności od tego, z jakiej uczelni pochodzą studenci, np. czy studenci kierunków politechnicznych lub uczelni państwowych biorą pod uwagę inne czynniki przy wystawianiu ogólnej oceny zajęciom? Dodatkowych informacji mogłoby też dostarczyć połączenie oceny uzyskanej z danego przedmiotu z oceną wystawioną nauczycielowi przez studenta.

Można by to zrobić albo dołączając pytanie o taką ocenę, albo łącząc bazy danych w przypadku wypełniania przez studentów ankiety w formie elektronicznej. Ciekawym byłoby też sprawdzenie, czy semestr, na którym odbywają się zajęcia, ma wpływ na ocenę wystawianą przez studentów, i czy faktycznie jest tak, że studenci wyższych lat oceniają wykładowcę łagodniej.

Zakończenie

Z uzyskanych ocen wynika, że należy położyć szczególny nacisk na poprawę i usprawnienie sposobu komunikowania się nauczycieli akademickich ze studentami w trakcie prowadzenia zajęć dydaktycznych. Jest to bowiem czynnik, który był najniżej oceniany przez studentów, a jednocześnie jego wpływ na ogólną ocenę zajęć był największy. Wydaje się, że do wzrostu przeświadczenia studenta o tym, że dużo dowiedział się na danych zajęciach, mogłoby przyczynić się większe mobilizowanie studentów do samodzielnego myślenia i inspirowanie ich do większej aktywności intelektualnej. Wskazane byłoby prowadzenie zajęć dydaktycznych w sposób zachęcający studentów do brania w nich czynnego udziału i ułatwiający im opanowanie treści przewidzianych w programie.

Zauważono, że otrzymywanie przez nauczycieli informacji zwrotnych w postaci ocen studentów poprawia jakość nauczania, zwłaszcza gdy jest połączone z konsultacjami ekspertów zajmujących się kształceniem umiejętności nauczania (Marsh i Roche, 1993; McKeachie, 1997). Większość badań jest zgodnych co do kwestii wiarygodności tego typu ewaluacji. Wydaje się, że głównym zadaniem pozostaje obecnie to, aby regularnie dokonywać przeglądów instrumentów pomiarowych i kształtować je tak, aby odpowiadały potrzebom i celom danej instytucji, a jednocześnie uwzględniały aktualne wyniki badań w tej dziedzinie (Gravestock, Gregor-Greenleaf, 2008). Właściwie przeprowadzona ewaluacja może dostarczyć uczelniom informacji na temat tego, w jaki sposób prowadzić działalność, aby skupiać się nie tylko na kreowaniu wiedzy i umiejętności, ale przede wszystkim na samym nauczaniu i pomaganiu studentom w zdobywaniu wiedzy. Student jest obecnie uznawany za głównego „interesariusza” każdej uczelni, a więc ważne jest, aby jego potrzeby zostały zaspokojone w jak najwyższym stopniu. Należy jednak pamiętać o tym, aby uzyskanych ocen nie wykorzystywać w sposób krzywdzący dla nauczycieli, a więc nie interpretować niewielkich różnic rzędu jednego miejsca po przecinku,

gdyż istnieje szereg czynników niezależnych od prowadzącego zajęcia, które rozpatrywane pojedynczo mają niewielki lub nieistotny wpływ na

ocenę zajęć, jednak skumulowanie się ich może spowodować, że powstanie różnica w ocenie o wspomnianej wielkości.

Bibliografia

Abrami P.C. (2001), Improving judgments about teaching effectiveness using teacher rating forms. In M. Theall, P.C. Abrami, and L.A. Mets (Eds.). *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research* 109, 59–87.

Aleamoni L., & Hexner P. (1980), A review of the research on student evaluations. *Instructional Science*, 9, 67–84.

Algozzine B., Beattie J., Bray M., Flowers C., Gretes J., Howley L., Mohanty G., & Spooner F. (2004), Student evaluation of college teaching: A practice in search of principles. *College Teaching*, 52(4), 134–141.

Arreola R.A. (2007), *Developing a comprehensive faculty evaluation system* (3rd ed.) Bolton, MA: Anker.

Beran T., Violato C., & Kline D. (2007), What's the 'use' of student ratings of instruction for administrators? One university's experience. *Canadian Journal of Higher Education*, 17(1), 27–43.

Beran T., Violato C., Kline D. & Frideres J. (2005), The utility of student ratings of instruction for students, faculty, and administrators: A "consequential validity" study. *Canadian Journal of Higher Education*, 35(2), 49–70.

Bolton, MA: Anker. Marsh, H.W. (1997), The measurement of physical self-concept: A construct validation approach. In K. Fox (Ed.), *The physical self: From motivation to well-being* (pp. 27–58). Champaign, IL: Human Kinetics.

Cashin W. (1990), Students do rate different academic fields differently. In M. Theall & J. Franklin (Eds.) *Student ratings of instruction: Issues for improving practice*, pp. 113–121. San Francisco: Jossey-Bass.

Cashin W. (1995), *Student ratings of teaching: The research revisited* (IDEA Paper No. 32). Manhattan, KS: Kansas State University Center for Faculty Evaluation and Development.

Cashin W. E. (1988), *Student ratings of teaching: A summary of the research* (IDEA Paper No. 20).

Manhattan, KS: Kansas State University Center for Faculty Evaluation and Development.

Coren S. (2001), Are course evaluations a threat to academic freedom? In S.E. Kahn & D. Pavlich (Eds.), *Academic Freedom and the Inclusive University* (pp. 104–117). Vancouver: University of British Columbia Press.

Cramer J. S. (2001), *An introduction to the logit model for economists* (wyd. 2), Timberlake Consultants, London.

Cramer J. S. (2003), *Logit models from economics and other fields*, Cambridge University Press, Cambridge.

Crumbley, L., Henry B., Kratchman S. (2001), Students' perceptions of the evaluation of college teaching. *Quality Assurance in Education*, 9(4), 197–207.

Dalton H. & Denson, N. (2009), Student evaluation: what predicts satisfaction?, in *The Student Experience, Proceedings of the 32nd HERDSA Annual Conference*, Darwin, 6–9 July 2009: pp. 100–110.

Davidovitch N. & Soen D. (2006), Class attendance and students' evaluation of their college instructors. *College Student Journal*, 40(3), 691–703.

Elliott A. C. & Woodward W. A. (2007), *Statistical analysis quick reference guidebook. With SPSS examples*. Sage Publications, p. 40

Feldman K.A. (1984), Class size and college students' evaluations of teachers and courses: A closer look. *Research in Higher Education*, 21, 45–116.

Frey P., Leonard D., & Beatty W. (1975), Student ratings of instruction: Validation research. *American Educational Research Journal*, 12, 435–447.

Gravestock P. & Gregor-Greenleaf E. (2008), *Student Course Evaluations: Research, Models and Trends*. Toronto: Higher Education Quality Council of Ontario.

Greene W. H. (2003). *Econometric Analysis*, Prentice Hall, Upper Saddle River.

- Greene W. H., Hensher, D.A. (2009), *Modeling Ordered Choices*,
- Gruszczyński M. (red.) (2010), *Mikroekonometria. Modele i metody analizy danych indywidualnych*. Oficyna Wolters Kluwer business, Warszawa.
- Gruszczyński M. (2001), *Modele i prognozy zmiennych jakościowych w finansach i bankowości*, SGH, Warszawa.
- Heckert T.M., Latier A., Ringwald-Burton A. & Drazen C. (2006), Relations among student effort, perceived class difficulty appropriateness, and student evaluations of teaching: Is it possible to “buy” better evaluations through lenient grading? *College Student Journal*, 40(3), 588–596.
- Katiliūtė E. (2010), Students’ perception of the quality of studies: differences between the students according to their academic performance, economics and management: 2010. 15, issn 1822–6515
- Koh C. & Tan T. (1997), Empirical investigation of the factors affecting SET results. *International Journal of Educational Management*, 11, 170–178.
- Marsh H.W. (1987), Students’ evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11(3), 253–388.
- Marsh H.W. & Dunkin M. (1992), Students’ evaluations of university teaching: A multidimensional perspective. In J.C. Smart (Ed) *Higher education: Handbook of theory and research*, 8, 143–233.
- Marsh H.W., & Roche L.A. (1993), The use of students’ evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal*, 31(1), 217–251.
- Marsh H.W., & Roche L.A. (1997, November). Making students’ evaluations of teaching effectiveness effective: The critical issues of validity, bias and utility. *American Psychologist*, 52(11), 1187–1197
- McKeachie W. J. (1997), Student ratings: The validity of use. *American Psychologist*, 52(11), 1218–1225.
- Nasser F., & Hagtvet K.A. (2006), Multilevel analysis of the effects of student and instructor/course characteristics on student ratings. *Research in Higher Education*, 47(5), 559–590.
- Neumann R. (2001), Disciplinary differences and university teaching. *Studies in Higher Education*, 26(2), 135–146.
- Ory J.C. (2001), Faculty thoughts and concerns about student ratings. In K.G. Lewis (Ed.), *Techniques and strategies for interpreting student evaluations* [Special issue]. *New Directions for Teaching and Learning*, 87, 3–15.
- Pounder J. (2007), Is student evaluation of teaching worthwhile? An analytical framework for answering the question. *Quality Assurance in Education*, 15, 178–191.
- Scriven M. (1981), *Evaluation thesaurus* (3rd ed.). Pt. Reyes, CA: Edgepress.
- Shevlin M., Banyard P., Davies M. & Griffiths M., (2000). The Validity of Student Evaluation of Teaching in Higher Education: Love me, love my lectures?, *Assessment & Evaluation in Higher Education*, Vol 25:4, pp: 397–405.
- Theall M., & Franklin J. (2001). Looking for bias in all the wrong places: A search for truth or a witch hunt in student ratings of instruction? In M. Theall, P.C Abrami, & L.A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* [Special issue]. *New Directions for Institutional Research*, 109, 45–56.
- Wachtel H.K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education*, 23, 191–212.
- Williams W. M., and Ceci S. J. (1997). How’m I Doing? Problems with Student Ratings of Instructors and Courses. *Change*, 29(5), 13–23.
- Wright R.E. (2006). Student evaluations of faculty: Concerns raised in the literature, and possible solutions. *College Student Journal*, 40(2), 417–422.

FACTORS AFFECTING THE STUDENTS' OVERALL EVALUATION OF COURSES

Abstract

The main goal of this paper is to examine the factors that affect the overall evaluation of courses by students. The data used in the study comes from a survey: Evaluation of the course in the Schools of Banking in the winter semester of the academic year 2012/2013, which evaluated 1802 classes. Logistic regression model was used in the paper to identify factors affecting the overall evaluation of the course. The paper also includes a review of English-language articles on the evaluation of classes and the results obtained in them.

Keywords: evaluation of the course, logistic regression model